Post-mortem and ante-mortem 2D and 3D facial comparison for forensic identification

Isabel Denise Burton

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of

Doctor of Philosophy

This research programme was carried out in collaboration with

Texas State University, San Marcos TX USA

March 2023

Enough.

Contents

List of Figures

List of Tables

Abbreviations and Acronyms

2D	Two-Dimensional
3D	Three-Dimensional
ABC	Automated Border Crossing
AFR	Automated Face/Facial Recognition
AFTER	Australian Facility for Taphonomic Experimental Research
AM	Ante-Mortem
AMSL	Above Mean Sea Level
ARISTA	Amsterdam Research Initiative for Subsurface Taphonomy and Anthropology
ASA	Advertising Standards Authority
BKA	Bundeskriminalamt (German Federal Criminal Police Office)
BSI	Bundesamt für Sicherheit in der Informationstechnik (German Federal Office for Information Security)
CBCT	Cone-beam Computed Tomography
CDI	Cadaver Decomposition Island
CFS	Craniofacial Superimposition
CPS	Crown Prosecution Service
CT	Computed Tomography
DNA	Deoxyribonucleic Acid
DVI	Disaster Victim Identification
EI3D	Expression-Invariant 3D Face Recognition
ENFSI	European Network of Forensic Science Institutes
ERT	Emergency Response Team

FA	False Acceptance
FAC	Forensic Anthropology Center
FACTS	Forensic Anthropology Research Center at Texas State
FARF	Forensic Anthropology Research Facility
FBI	Federal Bureau of Investigation
FERET	Face Recognition Technology Program
FFP	Female Face Pool
FISWG	Facial Identification Scientific Working Group
FI	Facial Identification
FLO	Family Liaison Officer
FR	Facial Recognition
FR	False Rejection
FRONTEX	European Border and Coast Guard Agency [French: *Frontières extérieures*]
FRS	Facial Recognition Systems
FRVT	Face Recognition Vendor Test
FS	Facial Superimposition
FSTT	Facial Soft Tissue Thickness
GDPR	General Data Protection Regulation
HDF	Human Decomposition Facility
HTF	Human Taphonomy Facility
ICRC	International Committee of the Red Cross
ID	Identification
IHL	International Humanitarian Law
IHRL	International Human Rights Law

INTERPOL	The International Criminal Police Organization
LJMU	Liverpool John Moores University
Max	Maximum
MFE	Mass Fatality Event
MFI	Mass Fatality Incident
MFP	Male Face Pool
Min	Minimum
NIST	National Institute of Standards and Technology
PM	Post-Mortem
pPM	Pretend Post-Mortem (= living, pretending to be dead)
PMI	Post-mortem Interval
PS	Perceptual Sensitivity
REC	Research Ethics Committee
SCD	Subject-to-Camera Distance
SMAS	Superficial Musculo-Aponeurotic System
TA	True Acceptance
TR	True Rejection
TRF	Taphonomic/Taphonomy Research Facility
TXST	Texas State (University)

Terminology

Abrasion	Scrape; result of frictional denuding of skin or crushing of the epidermis
Albinism	Inherited, congenital condition; characteristic appearance of blue eyes, pale skin, and white hair due to reduced amount of melanin pigment in eyes, hair, and skin
Algor mortis	Cooling of the body after death
Ante-mortem	Before/prior to death
Anthroposcopic	Involving or based on anthroposcopy
Anthropometry	(Greek: *anthropós* = human; *metron* = measure) the scientific study of the proportions and measurements of the human body
Anthroposcopy	Determination of human bodily characteristics by inspection as opposed to exact measurements
Bertillonage	A system formerly in use for identifying persons by means of a detailed record of physical characteristics (syn. Bertillon system)
Biometrics	The measurement and analysis of unique physical or behavioural characteristics as a means of verifying personal identity
Caucasian	Of or relating to a group of people having European ancestry, classified according to physical traits (e.g., light/fair skin pigmentation
Craniometry	The study and measurement of skulls (e.g., to determine its characteristics as related to sex and ancestry/race)
Cranioscopy	The observation, examination, and description of the human skull
Cephalometry	The science of measuring the human head
Chelion	Corner of the mouth fissure
Contusion	Bruise - caused by blood flowing into surrounding tissue after blood vessels were traumatically disrupted
Edaphology	(Greek: *edaphos* = ground; *logos* = study) The ecological study of soil and its influence on living things, e.g., plants, animals

Endocanthion	Inner corner of the eye
Epidermis	The surface epithelium of the skin, overlying the dermis
Exocanthion	Outer corner of the eye
Inferotemporal cortex	Cerebral cortex on the inferior convexity of the temporal lobe, crucial for visual object recognition
Laceration	Tear; result of splitting or stretching the skin over a hard underlying surface (e.g., bone)
Livor mortis	Purple discolouration of skin after death, caused by gravitational pooling of blood into dilated capillaries after circulation has ceased
Melasma	Uneven facial skin pigmentation; possible causes are sun exposure, fluctuations in hormone levels (e.g., pregnancy, hormonal birth control)
Middle-aged	No uniform definition; generally referring to individuals around age 35-45 until 60-65
Morphological	Relating to the scientific study of the structure and form in biology
Nasion	Craniometric point, where the top of the nose meets the ridge of the forehead; the point on the skull corresponding to the middle of the nasio-frontal suture
Pedology	(Greek: *pedon* = soil; *logos* = study) scientific discipline concerned with all aspects of soils, incl. physical and chemical properties, formation, description, and mapping
Peri-mortem	Near or around the time of death
Phenotype	Observable characteristics of an organism (e.g., colour, shape, size)
Pogonion	Most anterior midpoint of the chin
Poikilothermic	Having a body temperature that varies with the surrounding temperature
Post-mortem	After death

Pronasale	Most protruding point at the top of the nose
Prosthion	Craniometric point that is the most anterior point in the midline on the alveolar process of the maxilla
Ptosis	Drooping of upper eyelid which can restrict or block vision
Rigor mortis	Stiffness of muscles and joints after death, resulting from chemical changes within muscle fibres
Selfie	Photograph taken of oneself, typically by using a smart phone or webcam and uploaded to a social media platform
Stomion	Midpoint of the incisive line of the lips
Taphonomy	The study of the events and processes, such as burial in sediment, transportation, and decomposition, that affect the remains of an organism after it dies
Vitiligo	loss of pigment cells in skin and/or hair; onset commonly prior to age 30; aetiology is genetic and/or autoimmune related

Acknowledgements and Contributors

Whilst the PhD journey is often perceived as a rather solitary endeavour, this section shall pay tribute to all without whom this project would ultimately not have been successful.

Firstly, I would like to give a special thank you to my director of studies and first supervisor, Professor Caroline Wilkinson, for adopting me into the Face Lab team and taking me under her wing as one of her PhD candidates. I very much enjoyed working on this interdisciplinary project and am grateful for your advice and support, given your vast knowledge and experience in the field of craniofacial identification. I really appreciated the opportunity to be involved in Face Lab projects and to be able to build on previous training and education in this unique environment. Thank you for supporting my research visit to Texas, which was undoubtedly one of the highlights of this journey and for not losing patience throughout, despite the many obstacles.

Secondly, I wish to thank Dr Julie Roberts for taking on the role of second supervisor mid-project, and contributing invaluable support, drawing from her substantial forensic anthropology and DVI field experiences, but also for the advice on careers and life post-PhD. Thank you for helping to navigate difficult aspects of this PGR journey and being so approachable and genuine throughout. Special thanks to Dr Sud Sudirman, who - after initial hesitation relating to the kind of visual data this project would inevitably entail - kindly agreed to be my third supervisor and provided guidance in all matters related to computer science and programming. It has been an absolute pleasure to work with you and I am truly grateful for your support. I would like to express my gratitude to my thesis examiners Professor Maryna Steyn (external) and Dr Mark Grabowski (internal), for their time and efforts in anticipation of the viva voce examination and their comments. I herewith apologise for the high word count of my interdisciplinary research.

I must thank Liverpool John Moores University for generously awarding me a full PhD scholarship, which covered four years of this project and without which this research would not have been possible. Thank you to Face Lab for partially funding my research visit to San Marcos, Texas. I am very grateful to Walter Schels for allowing me to include some of his truly moving and precious photographs from the *Nochmal Leben vor dem Tod (= life before death)* project in my thesis. And a special thank you also to Ralph Hutchings for giving his consent for his anatomical photographs to be used and the beautiful hand-written note that came with it. I would like to express my gratitude to Dr Lars Oesterhelweg at the Institute of Forensic and Legal Medicine, Charité Berlin, for allowing me to spend valuable time at his institute and for providing a sounding board and guidance throughout my endeavours. Thank you for introducing me to Hilja Hoevenberg and being such an approachable and genuine external advisor.

My sincere thanks go to Dr Daniel Wescott at Texas State University in San Marcos, for agreeing to this research collaboration with such a quick turnaround and for allowing me to spend inestimable time at your facilities in 2019. I shall treasure all the special memories and experiences, from visiting and working at my first taphonomic research facility to climbing into a muddy cave in your back yard with fellow researchers. I would also like to thank Laney Feeser for her support in collating the ante-mortem data for my research from the database at TXST, and all the other great students and researchers for showing me the ropes at FARF and ORPL and your kind companionship during this time. I am eternally grateful to have met and spend immensely valuable time with fellow researchers Dr Clara Alfsdotter, Nathan Blair, and Molly Kaplan whilst in Texas. Clara and I truly bonded over scooping water out of a flooded grave with nothing but a dustpan and taking well deserved and much needed breaks from life at Jo's Coffee. Thanks to Nate for the precious time we spent together, for Black Mirror, beef brisket, board games, dog walks, and fantastic conversations. Apologies to Molly for not inviting you around to hang by the pool sooner, but as you said, this was only the start of our friendship. I would love to share future adventures and projects with you all, no matter where in the world our paths may cross next.

Heartfelt thanks to Dr Hayley Mickleburgh, who – after several failed attempts to acquire data and related ethical approval in the UK and Europe on my part – pointed me in the direction of Dr Wescott and a collaboration with TXST University in the US, which ultimately saved this project and offered a fantastic opportunity for me to grow and branch out. I am thankful to Prof Charlie Frowd for being available at short notice when I urgently needed advice and guidance on my project, for inviting me to UCLAN to join your workshop, and for all your support with the psychological aspects of face recognition and possible statistical analyses.

Thank you to LJMU's Health & Wellbeing services, in particular Penny Bradshaw. The PhD Group has been a much-needed lifeline throughout my PGR journey. I am truly grateful for all your support and could never adequately put into words how much of a difference it made to me, knowing that you are in *our* corner and providing a safe space for shared frustrations, and all victories big and small. My sincere thanks to Julie Sheldon and Victoria Sheppard from the Doctoral Academy at LJMU, for being there when I needed support. Thanks for making sure I would keep up the resilience and understand that what we want and what we need in trying times are often two very different things, and for generally contributing to make the administration side of things a lot smoother.

A big and warm thank you to the current and former Face Lab crowd, especially Dr Sarah Shrimpton, Mark Roughley, Dr Jessica Liu, and Dr Eilidh Ferguson, for your support, guidance, encouragement, believing in me, and making me feel very welcome and as part of the team from the start. I am truly grateful for your comradeship and honesty. After having shared a lot of the same challenges and experiences, I hope to never lose touch with any of you in the future and wish you nothing but the very best in all you do and become. Special thanks to Sarah, for always being there for me. Your encouragement and support meant more than you will ever know. Stay fierce! Heartfelt thanks to Dr Stenton Mackenzie, for throwing some much-needed lifelines, being open and honest whilst always staying true to his unique and special self. I am immensely grateful for all your support and your friendship. Thanks to Dr Pietà Schofield for your patient and kind support with statistics, for significantly contributing to keeping me sane and making me feel very comfortable whilst discussing, tackling, and revising related approaches, and filtering out any inadequate and blatantly wrong advice from other sources.

A special thank you to Dr Constantine Eliopoulos for believing in and promoting my abilities by giving me the opportunity to join you in Cyprus twice for the 1974's conflict mass grave excavations and planting the seed of PhD life in my head during my forensic anthropology undergraduate years at LJMU. I could not have asked for a better mentor and friend during my early days in this field and sincerely hope we will be working together again in the future. A special thank you to Dr Dirk Rohde, my former biology teacher in Marburg and a true inspiration, who encouraged me early on to follow my forensic anthropology dreams, however niche they may be, and who sparked a lifelong curiosity for human anatomy, evolution, and the natural world around us.

I could not have undertaken this journey without the support of both my German and English families. I will be forever grateful to my parents Maria and Klaus Hengelhaupt for generously supporting my studies and life abroad and patiently believing in my abilities until the very end. A big thank you and hugsizzles to my brother David, who understands better than anyone what the constant battles with imposter syndrome, the instant gratification monkey, and our inner *Schweinehund* feel like, for our mutual encouragement throughout our academic and other endeavours. We can make it happen; we can make it real! To the Clan, the one and only, whom I've not been able to see nearly as much as I would have liked since moving to the UK in 2011. Knowing that you are always there for me, and coming home to open arms, warm smiles, and the best food in the world is simply wonderful. I immensely treasure your honesty and loving support. We are always stronger together and nothing could ever break this bond.

A big thanks to my English family, who have given me a home away from home and for all their manifold support over the years. Dave Burton, I cannot adequately put into words in either language how immensely grateful I am to have you in my life – bestest of all the best ones. We managed to pull off our wedding of the giants amidst all this madness and you have truly been the most patient, caring, encouraging, and supportive husband and companion during this long journey with all its highs and lows. Our Frau Katze has also been essential to my emotional support; scientifically proven to keep heartrates down during stressful times with her feline wonderfulness and loud purrs. I would like to thank my dear and loyal friends Clarissa Brierley, Josie Ide, Miriam Houle, Nina Robbers, and Laura Antonia Pohany (just to name a few), for providing support and distractions in equal measure whenever needed. For allowing me to switch between the studious hermit lifestyle and coming up for air in the shape of anything from weekends away to simple cups of tea.

This may be slightly unconventional, but as Calvin Cordozar Broadus Jr. aka. *Snoop Dogg* once said "*I wanna thank me. I wanna thank me for believing in me. I wanna thank me for doing all this hard work. […] I wanna thank me for never quitting. […] I wanna thank me for just being me at all times*" (Next Generation Media, 2018; online). My apologies for not becoming a *real* doctor (you know who you are), but I hope that one day you will understand that to me, the dead are (almost) equally as important as the living.

Finally, I would like to express my immense gratitude to all body donors, families and next of kin, who, with their selfless gift, enable invaluable research, and are allowing a multitude of disciplines to grow, develop, and evolve.

Declaration of Original Work

<u>Declaration by candidate:</u>

This thesis is a presentation of my original research work. I have consulted all references cited. Any contributions by others are noted, including collaborative discussion. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. The author/researcher has no conflict of interest related to this investigation.

Name: Isabel Denise Burton, BSc (Hons) MSc FRAI AFHEA

Signature:

Date: 25/02/2023

<u>Certificate of approval:</u>

As primary PhD supervisor/director of studies of Isabel D. Burton, I certify that the above declaration is true to the best of my knowledge.

Supervisor: Professor Caroline Wilkinson

Signature: Caroline Wilkinson Digitally signed by Caroline Wilkinson

Date: 2023.03.22 14:33:00 Z

Abstract

This research explores the application of manual facial comparison methods and offline (semi-) automated face recognition algorithms to human identification of recently deceased individuals, using 3D (N=3) and 2D (N=6) data. Current methods are generally designed for and applied to the living, are not commonly used in combination or hierarchical order, and are mostly tested under controlled, lab-based conditions. The aforementioned makes their applicability to real-life settings questionable, and research on their application to the deceased is scarce, although desperately needed. Further issues arise from the lack of uniformity and standardisation in methodological approaches, as well as feature descriptions and terminology.

This study investigates the applicability of combined, manual face comparison methods in hierarchical order (preliminary feature-based analysis, facial superimposition, detailed morphological comparison) and two (semi-) automated face recognition algorithms (MATLAB® and Picasa) to the recently deceased. Pilot and ancillary studies explore pretend-dead faces as a data source and evaluate geometry vs. texture in 3D face models. 3D data was obtained using the Artec handheld laser scanner.

Key findings indicate that human face matching ability is superior and more resilient to PM facial changes, non-standardised AM data, and limited data availability, compared to the automated methods tested. Results further suggest that a hierarchical approach to manual comparison is highly beneficial. MATLAB®'s algorithm is unreliable even as a filtering tool and Picasa struggled to detect PM faces in images; an issue not encountered with pretend-dead data. Both automated approaches utilised here are not suitable for application on similar datasets or in casework settings, unlike the manual approach – the latter requiring further validation. Limitations arise primarily from the small sample size and non-quantifiable approaches to manual facial superimposition.

In forensic casework and disaster victim identification scenarios, comparative AM data for primary and even secondary methods of identification are often lacking. However, facial photographs are almost always attainable and should be considered an important resource for post-mortem identification. The interdisciplinary nature of this field requires collaborative efforts to address remaining challenges in the future.

I. Introduction

1 Background and Context

The human face is a unique composition of morphological features (Wilkinson, 2004; Stephan, Taylor and Taylor, 2008; Lucas and Henneberg, 2015; Zuo, Saun and Forrest, 2019; Balazia *et al.*, 2021), by which we are able to recognise individuals that we are familiar with (Bruce, 2012; Evans, 2014; Young and Burton, 2017), and which also holds great biometric value for the identification of unfamiliar faces (Moreton, 2021b; White, Towler and Kemp, 2021). Facial identification is an umbrella term for several different approaches, including – but not limited to – face matching and comparison, craniofacial approximation, facial reconstruction, craniofacial superimposition, and (automated) face recognition. The human skull and/or face represent(s) the baseline for all such methods, and the aim is to either recognise, verify, or identify an individual. The chosen approach may also lead to an exclusion or rejection, if observed differences cannot be explained by naturally occurring changes (e.g., aging signs, weight differences, facial expression) or factors related to image capture (e.g., pose, illumination, distortions). Manual facial identification, which is generally practitioner-led and hence likely to be influenced by personal training, experience, and bias, always holds an element of subjectivity and can be somewhat difficult to quantify. Facial comparison or face matching is also rather time consuming, hence attempts to automate procedures by utilising machine-based systems have experienced a substantial influx in research interest, especially over the past two decades. This is also in part owed to the incredible technological advances that were made during this time. However, the application of automated face recognition or manual face matching and comparison methods are generally developed on and used for the identification of the living. Hitherto, the potential of such techniques to be applied to the identification of the deceased has only very tentatively been explored (Tillotson, 2011; Caplova, Obertova, *et al.*, 2017, 2018; Cornett *et al.*, 2019; Davis, Maigut and Forrest, 2019a; Khoo and Mahmood, 2020).

Human identification and the sub-discipline of facial identification are very much an interdisciplinary endeavour (Lee, Mackenzie and Wilkinson, 2011; Evans, 2014; Kealy *et al.*, 2014; Bennett, 2020; de Boer *et al.*, 2020), which is reflected in the extensive literature review section of this thesis. Facial identification is broadly divided into manual and automated approaches, although there is some overlap. Face familiarity plays an important role in facial identification, as facial recognition can only occur if we have encountered a face previously and features are known (i.e., familiar) to us (Kramer, Young and Burton, 2018), whereas face matching or comparison in a forensic context is mostly performed on unfamiliar faces (White *et al.* 2014; Dowsett and Burton, 2015). Automated approaches are generally labelled as face recognition systems, as they may be used to verify/recognise an identity by matching their features to a database of known individuals on system. Upon reviewing past literature and following current research, it becomes evident that the general lack of validation studies and field-applications (i.e., using non-standardised data) of proposed, to date mostly lab-tested methods, is a pervasive issue (Bacci, Houlton, *et al.*, 2021).

1.1 Manual Facial Comparison

Humans tend to be skilled in facial recognition on average, when a face is familiar to them (Young and Burton, 2017). There is, however, quite a significant inter-person variability, which – in the wider population – ranges anywhere from prosopagnosia (aka. face blindness) to super-recognisers (Balsdon *et al.*, 2018; Phillips *et al.*, 2018; Dunn *et al.*, 2020; Barton, Davies-Thompson and Corrow, 2021; de Haas, 2021). Unfamiliar face matching on the other hand has shown to be rather error prone, even under controlled laboratory conditions (Shapiro and Penrod, 1986; Megreya and Burton, 2008; Burton, White and McNeill, 2010; Bindemann, Avetisyan and Rakow, 2012; Megreya, Sandford and Burton, 2013; Havard and Memon, 2014). Contradicting results exist with regards to whether training and experience improve unfamiliar face matching abilities (Wilkinson and Evans, 2009) or not (Young and Burton, 2017). The terms facial recognition and face matching are often used interchangeably in the literature, but fundamentally, unfamiliar faces are compared or matched, whereas unfamiliar recognition is in itself an oxymoron, as one cannot recognise what has not been encountered before.

As manual methods in general are often practitioner-led, the lack of uniformity in approach and use of terminology has been criticised in the past. Ongoing efforts to standardise and provide guidelines are extremely important (ENFSI, 2018; FISWG, 2019a; OSAC, 2021). Current guidelines advocate for the use of morphological comparison or feature instructions in face matching, whereas anthropometry should not be utilised at all, and superimposition generally only in conjunction with other methods (Megreya and Bindemann, 2018; FISWG, 2019a; OSAC, 2021).

1.2 Automated Face Recognition

Automated face recognition is based on biometric features and/or their compositional structure, which is employed to establish, verify, and/or authenticate an identity (Oloyede, Hancke and Myburgh, 2020). Incredible advances and developments have been made since semi-automated pattern recognition systems and first fully automated face recognition system were first introduced in the 1960's and 1970's respectively. Now, machine learning algorithms are able to learn from and adjust to the information that is provided to them within a training database, a process which is also referred to as deep learning (Adjabi *et al.*, 2020; Bundesamt für Sicherheit in der Informationstechnik, 2021). Shortcomings in automated face recognition research frequently relate to the fact, that most studies are conducted under highly controlled, laboratory conditions, where excellent, near perfect recognition scores are achieved (Grother, Ngan and Hanaoka, 2022). Frontal view, standardised, high quality facial images are not the type of data that is usually encountered in real-world identification scenarios, making these systems mostly not suitable for this purpose. As a general rule, the larger and more diverse the underlying training dataset, the better the performance of the algorithm or system, even on somewhat unconstrained data, such as the *Labelled Faces in the Wild* dataset (Tolba, El-Baz and El-Harby, 2006; Schroff, Kalenichenko and Philbin, 2015).

Big data firms, such as Google Inc. and Meta Platforms (formerly known as Facebook) have an advantage in this regard, but image rights violations and involuntary data sharing are recurring issues with such cloud-based systems. Currently, considerable controversy is attached to Clearview AI's face recognition system being used to help identify deceased soldiers in the ongoing Russia-Ukraine war (Clayton, 2022; Romero Moreno, 2022). Identification is incredibly important not only as a right for the dead, but also as a means for next of kin to get answers and a chance at finding closure. An important, not to be ignored issue, however, is the path to reach this goal. Morally and ethically speaking, just because you can, does not necessarily mean you should. Whilst their algorithm has an incredibly high performance of up to 99% accuracy on living faces (Grother *et al.*, 2022) and identification success rates on the deceased are reported to be very good (Temple-Raston and Powers, 2022), the company has been sued in the past over collecting and using facial photographs from online sources (e.g., social media profiles) without consent, implying failure to adhere to GDPR regulations regarding biometric data, privacy and image rights, in order to expand their enormous facial recognition database (ICO, 2022; Ramage, 2022). Their technology is a black box system and very little is known about the functional set-up and operational background.

As image variations, such as pose, illumination, facial expression, and ageing, still present an issue to automated systems, the past two decades have seen an increase in studies investigating the use of 3D vs. 2D over 2D vs. 2D face data, which has the potential to overcome some of these hurdles (Bundeskriminalamt, 2007; Lynnerup *et al.*, 2009; Guo *et al.*, 2016). Another clear advantage of 3D facial models is the additional spatial and depth information present, which is not available in 2D images, making this data format more adaptable to the comparative 2D data in matching and recognition tasks, both automated and manual (Thomas, 1998; Cattaneo *et al.*, 2012; Gibelli *et al.*, 2017; Ulrich, 2017; Lee *et al.*, 2019). Some research even experiments with 3D vs. 3D data (Guo *et al.*, 2016; Gibelli *et al.*, 2017), a constellation that is unlikely to be encountered in the real world, hence potential transferability to casework is not (yet) given here. Equipment to obtain 3D data (e.g., portable hand-held laser scanners) have generally decreased in cost whilst simultaneously significantly improving in their performance over the past decade (Kersten, Przybilla and Lindstaedt, 2016; GoMeasure3D, 2017; Decker and Ford, 2017).

1.3 Facial Identification of the Deceased

The majority of research and real-life applications of facial identification to date have been focused on living individuals (e.g., Abate *et al.*, 2007; Frowd *et al.*, 2007; Lynnerup *et al.*, 2009; Moreton and Morley, 2011; Bindemann *et al.*, 2013; Strathie and McNeill, 2016; Phillips *et al.*, 2018; Kortli *et al.*, 2020) with the exception of facial depictions, approximations/reconstructions, and craniofacial superimposition (Taylor, 2001; Wilkinson, 2008; Yoshino, 2012; Kealy *et al.*, 2014; Wilkinson and Lofthouse, 2015; Damas, Cordón and Ibáñez, 2020). A considerable issue is the lack of suitable datasets, which is unsurprising, given the ethical and logistical implications, as well as the sensitive nature of post-mortem facial images. If well-funded, large-scale research, such as the FAST-ID project a decade ago are unable to pursue one of their main objective of comparing ante- and post-mortem facial images (Crabbe *et al.*, 2013), it is unsurprising that such a scarcity of theoretical background studies in this area persists.

Furthermore, facial post-mortem changes and their effects on identification methods are also considerably under researched (Wilkinson and Lofthouse, 2015; Cornett *et al.*, 2019), whilst the need to explore alternative methods of identification is evident from disaster victim identification efforts, both past and present, where primary identification methods (e.g., DNA, fingerprints, odontology) cannot be deployed due to a lack of comparative ante-mortem data (Soomer, Ranta and Penttilä, 2001; Morgan *et al.*, 2006; Tsokos *et al.*, 2006; Birngruber *et al.*, 2011; Khoo *et al.*, 2016; Caplova, Obertova, *et al.*, 2018; Bennett, 2020). To the best of the researcher's knowledge, there is no literature available to date that evaluated human unfamiliar face matching ability for deceased faces. Furthermore, no studies to date have explicitly investigated facial changes in the context of decomposition in water environments. Current guidelines and best practice recommendations for manual facial comparison focus almost exclusively on the living and would likely require adaptations to make them transferrable to the deceased face (ENFSI, 2018; FISWG, 2018a; OSAC, 2021).

The deceased face introduces its own set of potential obstacles, or "*non-idealities*" (Labati *et al.*, 2021; p.2). In addition to common challenges known from the identification of the living, such as illumination, pose, facial expression, or age, other phenomena like facial injuries, decompositional changes (e.g., bloating, corneal clouding), and a non-natural expression are often present. Such post-mortem obstacles are then possibly accompanied by further issues regarding availability, acquisition, and/or quality of comparative ante-mortem data. The latter may present as low-quality images, with a significant age or time gap between capture of ante- and post-mortem images or scans and consequently show significant variations in facial appearance. Ante-mortem data in real-world scenarios is also often uncontrolled in nature, whereas the majority of face recognition and comparison research utilises standardised, lab-based data.

There is much controversy surrounding the application of automated face recognition methods to the identification of the deceased, particularly with regards to cloud-based systems and the ethical and legal grey zones, in which they often operate (Khoo and Mahmood, 2020; Romero Moreno, 2022). Appropriate regulations and guidelines are still lacking; an oversight with potentially severe consequences (Kantayya, 2020). Research shows that some systems can achieve near perfect accuracy rates, especially under controlled conditions, one of which is Clearview AI's facial recognition system (Grother *et al.*, 2022). However, the latter made headlines in recent months due to accusations of unlawful acquisition of facial data from online platforms (Ramage, 2022; Romero Moreno, 2022). Nevertheless, this system is currently being applied to aid in the identification efforts of deceased soldiers in the Russia-Ukraine war (Clayton, 2022). It is somewhat questionable, whether the end really justify the means, or whether data protection and ethical frameworks should be at the forefront of such approaches.

2 Current Research

2.1 Summary of Aims, Objectives, and Hypotheses

Aims of the current project are to investigate to what extent manual and automated facial recognition and comparison methods – at present predominantly developed for and applied to the living – can be used to identify recently deceased individuals. Further, the aim is to test both human and machine face matching accuracy on unfamiliar, deceased faces and to evaluate whether approaches used in this research are potentially transferrable to real-world identification scenarios. This research strives to improve the understanding regarding application, accuracy, and reliability of facial identification methods and will scrutinise the practicalities of different data formats during collection, post-processing, and analysis.

This research explores whether several different manual face comparison methods are potentially more efficient and reliable when used in combination and hierarchical order, either as a filtering tool or to ultimately identify a post-mortem subject. Moreover, to determine whether offline, (semi-)automatic face recognition algorithms can accurately detect and match post-mortem faces with corresponding ante-mortem images from a face pool. An additional aspect is the use of both 3D and 2D face data, with the intention to study potential benefits and shortcomings regarding data formats and methods of comparison.

The main focus of this current research lies on the comparison of 3D facial surface scans (N=3) as well as 2D facial photographs (N=6) of recently deceased individuals against a face pool of 2D ante-mortem facial images (N=30). Subjects within the dataset are representative of the current Texan population (World Population Review, 2022), as data will be obtained from the Forensic Anthropology Center at Texas State University in San Marcos. Although post-mortem data will be collected in a standardised manner, all ante-mortem data used here is non-standardised by default. Data analysis is performed by combining three or two manual face comparison methods (depending on data format), namely a preliminary feature-based comparison, followed by a facial superimposition, and finally a detailed morphological comparison. For the (semi-) automated facial comparisons, a simple face recognition algorithm using MATLAB® and Google's Picasa 3 face recognition function shall be explored, and results compared to those of the manual, human observer analyses. Success for both manual and automated methods are determined by the ability to reach identification decisions that are ultimately verified as true match pairs.

Research hypotheses are, that existing methods may have to be adapted to be suitable for deceased faces, as the latter may introduce obstacles that are not present in living faces and hence not considered in the design of current guidelines and standards, and some facial features may have higher discriminating value than others. Furthermore, offline face recognition algorithms may struggle with non-standardised data and the *dead* face, as they are primarily developed for and trained on standardised living face images. It is tentatively expected that human comparison results will outperform chosen (semi-)automatic systems in this context, as post-mortem related changes in morphology might be easier to overcome for the human observer.

This research will focus on adult subjects only, the aspect of growth and development during childhood and adolescence will therefore not be included in the review section of this thesis.

2.2 Structural Outline of Approach

The main focus of this research is the comparison of 3D and 2D PM data against 2D AM face pools, firstly by means of a manual, combined methods approach (preliminary, feature based analysis; facial superimposition; detailed morphological comparison), and secondly through two (semi-) automated comparison techniques (MATLAB®, Google Picasa 3).

A short pilot study (Appendix A – *Pilot Study*) was conducted on living individuals, pretending to be deceased, to assess geometry vs. texture with regards to matching accuracy and hence inform the data collection modality and format for the main study of this project. This is a much-neglected aspect in the literature, although crucial in improving existing, and developing new methods of facial comparison, as it dictates which data formats (e.g., CT scan models vs. laser surface scan, photogrammetry models, or photographic images) are more suitable and where collection emphasis should lie – when- and wherever possible. Materials comprised six pretend-PM subjects (3 males, 3 females) and a face pool of 40 non-standardised facial photographs (20 males and 20 females respectively). Participants are middle-aged Caucasians from the UK, as data was collected by the director of studies, at the researcher's own institution in Liverpool, from individuals unfamiliar to the researcher. The analysis was conducted in a 2-step manual, computer-based, hierarchical approach: a preliminary, feature based comparison followed by facial superimposition. Results showed that 3D face models without texture information (geometry only) were misidentified in 50% of cases (3/6), whereas for 3D data with texture, a 100% true positive rate was achieved. As a result, 3D laser surface scanning and photogrammetry was chosen over CT scans for data collection of the main study, to have texture information available.

Unfortunately, there were some issues with the scanning equipment and AM data availability, resulting in low PM subject numbers for the main study. After ultimately only having a small sample available for the main study section (3D PM: N=3; 2D PM: N=6), some ancillary studies (Appendix B – *Automated Comparison – Ancillary Studies*) were conducted, in order to maximise the number of possible analyses and findings from this limited dataset. Those studies focused on assessing potential differences in (semi-) automated face recognition between actual PM and pretend-dead faces. Furthermore, they aimed to assess differences in matching accuracy between using only one AM image per person in the face pool vs. using multiple different images per person. The same baseline datasets as in the pilot and main study were used here. Findings show that Google Inc.'s Picasa 3 struggles to locate a deceased face within an image (face detection), which is then seemingly not considered in the automated face recognition/matching stage. This issue did not occur with pretend-dead faces. Further, using multiple different AM images appears advantageous over using artificially augmented versions of the same image. Matching results for MATLAB® were equally poor as in the main study, which can likely be attributed to the algorithm having been designed for and trained on highly standardised data only, and subsequent inability to overcome image variations as were present in the current data set.

Both the pilot study and ancillary studies were intentionally moved to the appendices section of this thesis. Although they are both deemed useful and informative with regards to the main research study and do show novel results, it would break up the narrative and distract from the main purpose of this research project, if those where to be incorporated into the main thesis.

II. Aims and Objectives

1 Research Questions

- Can face recognition and comparison methods, which at present are predominantly applied to the living, be utilised to aid in the identification of the recently deceased?
- Does human recognition ability outperform machine-based algorithms in an identification scenario with recently deceased individuals?
- Which of the materials and methods applied in this research are potentially transferable and useful to real-life identification scenarios involving recently deceased individuals (e.g., DVI), and which are not?
- If positive identification cannot reliably be established, are the approaches tested here still valid as a pre-screening tool to narrow down a face pool of possible matches, which would subsequently have to be followed up by other types of analyses (e.g., DNA, dental records) in real-life scenarios of forensic case work?
- What are the benefits and possible disadvantages of using 3D vs. 2D data?

2 Aims

The aims of this research are:

2.1 Main Study: The Deceased

- To investigate to what extent manual and automated facial recognition and comparison methods – at present predominantly developed for and applied to the living – can be used to recognise and/or identify recently deceased individuals.
- To test both human and machine face matching accuracy on unfamiliar PM vs. AM faces.

- To evaluate if the approaches used here are suitable for application to real-world identification scenarios, either fully, partially, or in an adapted manner.
- To explore the benefits and shortcomings of 3D vs. 2D data for facial comparison.

2.1.1 Part 1 – Manual Comparisons

- To investigate, whether several different manual face comparison methods are more efficient and accurate when used in combination, as a filtering tool, to narrow down a face pool and identify the matching target.

2.1.2 Part 2 – Automated Comparisons

- To determine whether freely available, offline, (semi-)automated face detection and recognition algorithms can accurately and reliably detect and match PM subject’s faces with their corresponding AM images from a face pool.

3 Objectives

The objectives of this research are:

3.1 Main Study: The deceased

- To obtain and utilise non-standardised 2D AM facial data and semi-standardised 2D and 3D PM data from recently deceased individuals, mimicking a real-life identification scenario with related challenges in data availability, acquisition, and analysis.

3.1.1 Part 1 – Manual Comparisons

- To employ several different manual facial comparison methods (preliminary feature-based analysis, facial superimposition, and morphological comparison) consecutively in a hierarchical approach.
- To employ a five-point rating scale in order to obtain quantifiable results for differences and similarities observed.
- The methods applied will be deemed successful if the final identification decisions are verified as true matches.

3.1.2 Part 2 – Automated Comparisons

- To test machine face recognition ability, by using two (semi-) automated face detection and recognition systems/algorithms: MATLAB® and Google Picasa.
- To operate both (semi-) automatic tools offline, to comply with related data protection, privacy, and ethics requirements for sensitive, identifiable data.
- The methods will be deemed successful by the algorithm's ability to produce true positive identification results through correct matching decisions.

4 Hypotheses

- Manual face matching accuracy will outperform (semi-) automated systems in this current study, given the PM element and non-standardised AM data.
- Using several manual comparison methods in a combined, hierarchical approach will lead to reliable identification decisions.
- Manual facial comparison methods developed for and tested on the living are expected to be easily transferrable and adaptable to PM data, without losing their potential success rate.
- Automated methods will struggle with PM and non-standardised AM data, likely resulting in a negative impact on accuracy rates.

- 3D PM facial data will be more beneficial compared to 2D PM data, as various facial angles and detail are available and alignment with head pose in 2D AM images can be achieved.
- Facial laser surface scans, collected with the handheld Artec Spider model, will be a fast and efficient way to capture high detail and create accurate 3D PM face models.

5 Novelty of Research

Individual methods used in the manual comparison section of this research are not entirely new, but some (e.g., morphological analysis following FISWG guidelines) are severely lacking in validation studies (Bacci, Houlton, *et al.*, 2021). Furthermore, using these methods in combination is a novel approach, that may be able to balance out minor individual shortcomings of both the methods and potential practitioner bias. The application both manual and automated analyses to the identification of the deceased is an extremely neglected but important aspect in the literature and research in general is sparse, likely due to the difficulty in obtaining suitable datasets. Given the extensive ethical challenges involved in collecting and utilising identifiable (face) data from recently deceased individuals, there is no open-access database for facial photographs or 3D face models of recently deceased individuals, unlike the several databases of living individual's facial photographs that can be utilised for face recognition research (see e.g., Chellappa, Wilson and Sirohey, 1995; Alashkar *et al.*, 2014; Chen *et al.*, 2018; Clifford, Watson and White, 2018).

Most facial identification methods are developed for the purpose of, and are predominantly applied in, scenarios involving living individuals. This is especially true for automated face recognition algorithms. Therefore, the novelty of this research lies in its application to the deceased, and in comparing human and machine face matching ability. It utilises realistic, *wild* data that incorporates a variety of factors still considered challenging to recognition algorithms and the human recognition ability alike, namely variations in lighting, pose, age, facial expressions, occlusion, weight loss or gain, and image quality. Although neither the manual facial comparison methods (i.e., feature based comparison, facial superimposition, morphological analysis) individually nor Google Picasa or MATLAB® face recognition algorithms are new, using them in combination and applying this approach to the deceased is.

This research strives to improve the understanding regarding application, accuracy, and reliability of facial identification methods, both manual and automated. It aims to contribute to the body of knowledge, particularly through testing applications of methods on recently deceased faces and scrutinising the practicalities of different data formats during collection, post-processing, and analysis. This project will show whether manual facial comparisons and/or automated face recognition algorithms could be applied in PM identification or verification scenarios, either on their own or alongside other methods. It will also show, whether automated methods could help to reliably narrow down a large face pool in a more time-efficient manner, as opposed to manually, which can be an extremely long and arduous process (ICRC, 2008; Blau and Hill, 2009; Khoo *et al.*, 2016).

Results will give an indication of whether human face recognition ability for AM vs. PM faces is still superior to automated methods, when confronted with unfamiliar, unconstrained, uncontrolled data – as would be the case in most real-life scenarios. Approaches included in this current research should be validated through further studies with larger sample sizes and more diverse population groups in the future, prior to possibly applying methods to real-life casework, both nationally and internationally, especially in DVI. When time is of the essence, resources are limited, and circumstances do not provide access to AM data for primary identifiers, photographs may be the only comparative AM data available (Caplova, Obertova, *et al.*, 2017; Olivieri *et al.*, 2018; Wilkinson and Castaneyra-Ruiz, 2021). Therefore, visual identification methods that explore this type of data as a baseline must be further investigated, improved, and expanded.

III. Literature Review

1 Human Identification - Identity vs. Identification

According to Marshall (2014), elements contributing to who we are (i.e., identity) can be broken down into three main groups: the things we all have in common with others, those we share with only a few, and elements that are unique to us as an individual. Attributes we have in common can be referred to as universal, pre-social, and essential, such as the human potential for rational thought. Characteristics such as sex, national identity, ethnic affiliation, and religious beliefs are often linked to a social and/or political identity, that we only share with a few other people and most of these elements are not defined by choice or arrangement. Finally, uniqueness in an individual can be one's origin, past experiences, personal choices, and roles (e.g., being a parent, nurse, friend, etc.), but also comprise an individual's intellectual aptitudes and talents. Identity in the sense of a social, emotional, and/or psychological construct, is generally associated with how an individual sees themselves and is portraying themselves to the outside world through their unique character and personality. It can also mean *who a person is*, and the claim of identity can then be verified or rejected by process of identification, which is synonymous with recognition by means of comparison (Cambridge Dictionary, 2021a). Identification (from Latin *idem* = the same, identical) is usually achieved through verification, evaluation, comparison, and/or analysis of two sets of biological or other data that is ideally non-transferable (Black, 2006; Bikker, 2014).

From a forensic human identification standpoint, the identity of an individual is established by narrowing down broader, shared characteristics to unique elements that are either extremely unlikely to be shared or are not shared with anyone else. Referring back to Marshall (2014), the questions behind this approach would broadly be *is it human?* (i.e., shared with many/all others); *what is their sex, age, and ethnic affiliation?* (i.e., shared with some/few others); *is it possible to obtain and match DNA, dental records, fingerprints, etc.?* (i.e., unique). Many human hard and soft tissue features are unique and can therefore be used for biological and personal identification (Cappella *et al.*, 2019). Unique in this sense means that the probability of someone else sharing the same trait or marker is so unlikely, that it is assumed no one else or only very few would have the same. An example would be a calculated DNA match probability of one in 24 billion, three times the world's current population (7.9 as of January 2022; see https://www.worldometers.info/world-population/ [accessed 28.01.2022]).

Biological identification or biological identity is defined through characteristics such as age, biological sex, ancestry, and genetic profile. Some of those are shared (e.g., sex, age), and some are unique or at least extremely rare amongst individuals, meaning they hold great statistical power (e.g., DNA, iris pattern, fingerprints). This power increases with the number of traits that are analysed for one particular method. Personal identification is based on features or characteristics that result from any given biological trait. It is more concerned with the development and growth aspect and resulting changes during a person's lifetime, whilst considering and evaluating the similarities or differences between shared traits, the absence or presence of features, any modifications and whether they can be explained by natural changes or body modifications. There are several different techniques that have been applied in personal identification, such as dactyloscopy, dental comparison and bite mark analysis, frontal sinus shape or palatal rugae comparison, palatal rugae matching, and facial identification and recognition. All of those are constantly evolving, as they are being refined and revised, as important research findings are added to the body of knowledge. The following sections will focus predominantly on the identification of unfamiliar deceased individuals, as this is also the content of the current research. Covering all identification methods in depth would go beyond the scope of this thesis.

1.1 The Importance of Identification

In human identification, the first question to be addressed should not be *how* to identify someone but *why* (Cook, 2020). Whenever an individuals' identity is assessed or described by means of scientific techniques, the context is crucial and the need for identification may arise from a variety of possible scenarios, e.g., the individual is deceased and their identity unknown (archaeological or forensic), or the individual is alive but unable or unwilling to provide the required information (e.g., sub-adult, incapacitated individual, or suspected case of identity fraud). In bioarchaeology, identification - as far as it is possible - provides a fundamental contribution to understanding human interactions, the structure of historic societies, as well as past events and environments. Within a forensic context, identification is relevant to establish both the identity of the perpetrator and the victim. More recent applications of human identification focus on issues relating to national security (e.g., terrorist threats or identity theft) (Bisgaard Munk, 2017; Anwarul and Dahiya, 2019). Identification efforts from varying disciplines (including but not limited to pathology, medicine, anthropology, and archaeology) are aiding in a dignified management of the deceased, by providing required documentation and information. A death certificate is necessary from a legal perspective for insurance policies, testaments, and pensions, but from a criminological and social point of view, it provides information on cause and manner of death.

An identification is a much-needed answer to the *where* and *what* of someone's personal circumstances. The humanitarian element prevents the "*social death of the dead*" (Parra *et al.*, 2020; p.86), which occurs when the missing are not found and/or the unidentified remain nameless, when surviving relatives do not get answers, and closure is ultimately denied (Baraybar, Caridi and Stockwell, 2020). The majority of relatives will never stop searching for missing loved ones, and the continuous uncertainty of ambiguous loss will likely lead to complicated grief (Boss, 2010; Lenferink *et al.*, 2018; Cook, 2020; Mazzarelli *et al.*, 2021). This creates a state of limbo between clinging to hope and anticipatory grief (Wayland *et al.,* 2016 in Lenferink *et al.*, 2018). Relatives have the right to know what happened to their missing family members, and for their human remains to be returned to them (ICRC, 2008; IHRL AP I, Art 32; ICPPED, Art. 24). However, this may not always be possible, as not all missing persons are found, not all that are found are identified, and repatriation can have its own challenges (Jensen, 1999; Wright *et al.*, 2015). Should the latter not be feasible, human remains should be buried in a dignified manner and families have the right to visit their relatives' grave (European Court of Human Rights, *Sabanchiyeva and Others v. Russia,* application No. 38450/05, Judgement of 6 June 2013 in UN, 2017).

Identification of human remains is relevant for legal, ethical, social, religious, and cultural contexts (Ranson, 2016). According to the International Criminal Police Organization (INTERPOL), "*human beings have the right not to lose their identities after death*" (Interpol GA-1996-65-RES-13 DVI, 1996; p.1). The missing and the deceased are protected under national and international law. Most relevant for this context are arguably the international humanitarian law (IHL) and the international human rights law (IHRL) (Londoño Romanowsky and Silva Chau, 2020). In IHRL, the primary duty is to search for, recover, and evacuate the dead in or following armed conflict (ICRC, 1949). IHRL applies at all times and is not just linked to armed or other conflicts and mass fatality incidents (MFIs; aka. mass fatality events / MFEs) (Vega Dulanto, 2020). Humanitarian forensic action, as developed by the International Committee of the Red Cross (ICRC), entails the implementation of forensic science to "*address the needs of victims of armed conflicts and other catastrophes for humanitarian, rather than criminal, purposes*" (Tidball-Binz, 2020; p.3). One of those needs is identification, which still poses a significant challenge for a number of reasons, such as the location of an event and accessibility, limited available resources, a high victim count, state of the remains (e.g., fragmented, comingled) and degradation thereof, the availability of AM data for comparison and quality thereof (Alonso *et al.*, 2005; Cornett *et al.*, 2019). Nevertheless, identification of the dead is regarded as a responsibility of the living (ICRC, 1949, 2008; Interpol, 2018).

1.2 Disaster Victim Identification

Disaster Victim Identification (DVI) is the formal process of recovering, processing, and identifying deceased persons in an MFI (Alonso *et al.*, 2005; Bikker, 2014; Brough, Morgan and Rutty, 2015). In the UK, the following criteria are considered when declaring a DVI response necessary: the significant number of deceased (actual or potential), the nature of the incident (whether it is likely to make identification of the deceased difficult), the location/s and ease of access to the deceased, the condition of the remains, a likely terrorist or criminal cause of the event, the potential hazards at the scene and the mortuary suitability and capacity (Home Office, 2013). On an international level, INTERPOL's DVI guidelines (Interpol, 2018) offer general recommendations on procedures, best-practice, documentation, etc. and aim to advocate international standards to aid in a consistent approach to disaster management and identification efforts by transnational response teams. First published in 1984, it has since been reviewed and updated every five years by specialised DVI working groups and the DVI standing committee of INTERPOL (de Boer *et al.*, 2020).

It has become evident over the past decades, that MFIs and natural disasters are inevitably occurring on a global scale. As climate change progresses natural disasters especially will become more frequent (Seneviratne *et al.*, 2012; UN, 2021). The migrant crisis, predominantly in the Mediterranean, in Africa, and at the US-Mexico border, calls for an ongoing humanitarian and forensic response effort (e.g., Ellingham, Perich and Tidball-Binz, 2017; Cattaneo *et al.*, 2020; Spradley and Gocha, 2020; Zorba *et al.*, 2020; Wilkinson and Castaneyra-Ruiz, 2021). As Hinchliffe (2007; p. 493) rightly pointed out, "*it is not a question of 'what if this happens again' – but 'what shall we do when this happens again*'". INTERPOL differentiates between *open* and *closed* disasters (Interpol, 2018). The former refers to any major catastrophic incident or attack, which results in an unknown number of casualties (e.g., attack at public gathering: 2016 Berlin truck attack on Breitscheidplatz). Closed disasters are defined as devastating events, claiming the lives of a known and identifiable group of people (e.g., plane crash: 2015 Germanwings flight 9525). A combination of both is theoretically possible, and approaches need to be tailored accordingly. In the most recent version of INTERPOL's DVI guidelines, four stages of the DVI process are specified (Interpol, 2018):

1. Scene (location, recovery, processing of human remains and property/belongings)
2. Post-mortem (detailed PM examination of human remains, collection of PM data)
3. Ante-mortem (collection of missing person and AM data from various sources)
4. Reconciliation (matching of AM and PM data, returning remains to families if possible)

Many countries have highly trained, experienced, and deployable emergency response teams (ERTs), which are usually interdisciplinary in nature, and each role is regarded equally valuable to the overall mission (e.g. UK DVI Unit, Blake Emergency Services, Kenyon International Emergency Services) (Brough *et al.*, 2015; Sledzik and Mundorff, 2016). Teams are deployed either after the government of the country in which the MFI occurred has requested support, or their offer to get involved has been accepted. It is important to understand that in the majority of cases, any operational framework is subject to laws, rules, regulations, and accepted approaches (i.e., religious, cultural beliefs) of the affected country, as a foreign response team is primarily in the role of the supporter. This may shift if the respective local government transfers responsibilities. It is understood however, that identification of foreign nationals is not possible without the cooperation of the respective home country (Robertson, 2008). The first efforts are always to secure the scene, save lives (if possible), and if required, (re-)establish infrastructure, before recovery and identification commences. Another main task following MFIs is incident and scene investigation, as this commonly brings about relevant information for the DVI teams and may aid in manner and cause of death determinations (Ranson, 2016). All aforementioned phases can run concurrently, depending on the scene and circumstances (Sledzik and Mundorff, 2016).

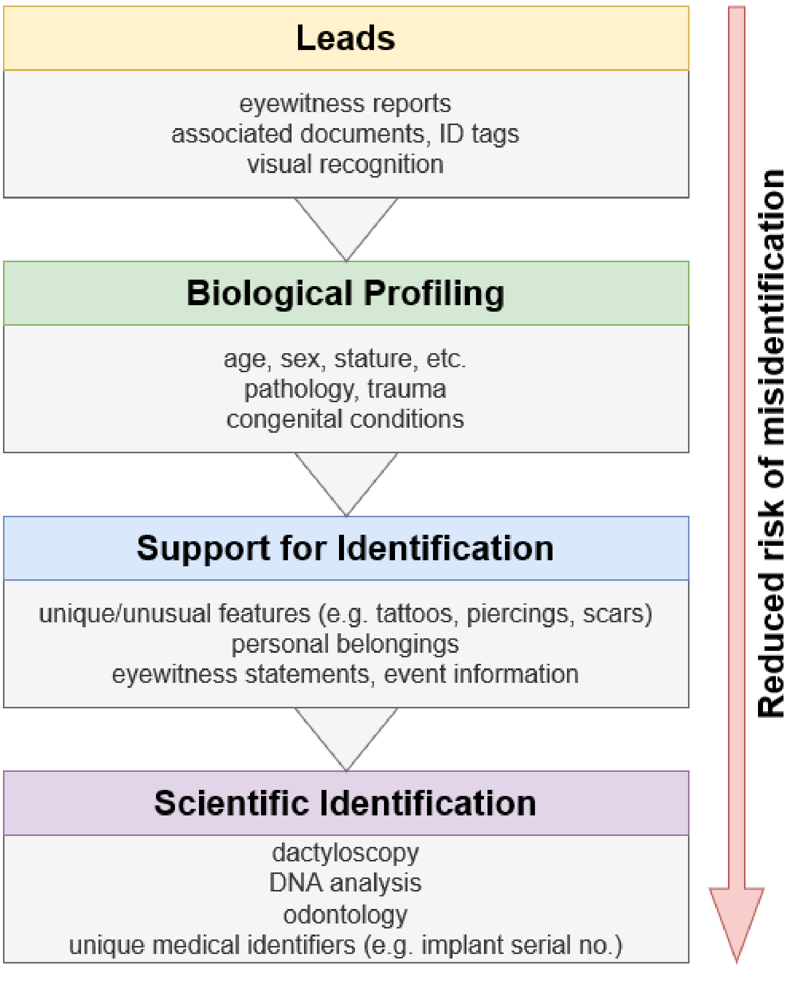

Figure 1: Diagram showing an ideal scenario to be followed for identification efforts, with the weakest or most circumstantial evidence at the top and strongest, most scientifically valid at the bottom, as recommended by the ICRC. The more information collected and compared, the better, and the stronger the evidence for positive identification or rejection decisions, which in turn aims to avoid misidentifications (adapted from ICRC (2008); p.13).

1.2.1 Primary and Secondary Methods of Identification

Victims of MFIs, armed conflicts, migration, and crime should be searched for, recovered, treated with dignity and respect, and ultimately identified. Forensic human identification approaches are all based on the principle of comparison of at least two data sets and finding compatible, matching structures, when excluding differences are absent (Campomanes-Álvarez *et al.*, 2018). However, not all identification methods have the same discriminating power. In a forensic context, a distinction is therefore made between primary and secondary methods of identification (aka. primary and secondary identifiers) (ICRC, 2008; Interpol, 2018).

Primary Identifiers

- Comparative dental analysis (forensic odontology)
- DNA analysis
- Fingerprint analysis (dactyloscopy)

Secondary Identifiers

- Medical record comparison
- Personal effects
- Visual identification and/or description
- Tattoos and other body modifications

Although most authorities accept a single primary identifier as sufficient evidence for positive identification, the recommendation is always to corroborate this with other primary or secondary identifiers (de Boer *et al.*, 2020). Primary identifiers have limitations, as they rely on existing, available, and reliable AM data for comparison, which are not always available. Applicability of methods is also dependent on the state of the remains and survival of specific body parts, as some withstand decompositional changes and extreme temperatures better than others, which deteriorate quickly and are at risk of cross-contamination in cases of comingled remains (e.g., DNA). Lack of standardisation and terminological differences in each of the methods listed above can also cause issues, when comparing AM to PM records becomes confusing at best (Hinchliffe, 2011). Whilst DNA is still perceived as the magic answer to all identification needs, it is not always possible to implement this in all identification scenarios (Farrell, 2018; Bennett, 2020). Cultural or religious beliefs, and/or the state of the remains may not allow for PM DNA samples to be collected, relevant AM data may not exist for comparison, or the required resources to implement DNA analyses may not be available (ICRC, 2008; Khoo and Mahmood, 2020). DNA can also quickly degrade, due to environmental conditions and/or incorrect sampling, transport, and storage, causing an issue for PM data collection (Alonso *et al.*, 2005; Hartman *et al.*, 2011). If more members of the same family are amongst the deceased, other methods should corroborate findings from DNA comparison (Prinz *et al.,* 2007 in Wright *et al.*, 2015).

Forensic odontology has shown to be very reliable as an identification method, as teeth are not only very individualistic but also stable over a long period of time and in varying conditions after death (Black and Bikker, 2016). In fact, 61% of the 2004 Tsunami victims were identified through this approach, as hot and humid weather conditions and subsequent rapid decomposition and DNA degradation rendered the latter almost unusable (only 1.3% of victim ID through DNA) (Morgan *et al.*, 2006). However, this method is not feasible without available and accurate AM dental records for comparison (Hinchliffe, 2007). Similar issues arise with dactyloscopy. Even if fingerprints are collected amongst the living population for identification purposes (e.g. ID cards), not all countries have a central fingerprints database for their citizens and required records may therefore not be attainable (Kobus, Kirkbride and Raymond, 2016). Children may not have been old enough to be registered via fingerprints (Wright *et al.*, 2015). The skin and friction ridge pattern on the hands of the deceased can be compromised by wounds, missing parts, desiccation and/or decomposition (Dror, Charlton and Péron, 2006; Chaikunrat, Pongpanitanon and Petiju, 2011; Mulawka, 2014; Kobus *et al.*, 2016). Nevertheless, dactyloscopy is "*often the fastest primary identification method in DVI work*" (Johnson and Riemen, 2019; p.294), and can be incredibly useful and valuable in some situations, e.g. 2014 MH17 plane crash: 151/283 victims identified through dactyloscopy (Johnson and Riemen, 2019); 2016 LaMia Flight 2933: 64/73 victims identified through dactyloscopy (de Souza *et al.*, 2021)).

The applicability of methods is majorly dependent on available resources and the state of preservation of human remains. If soft tissue is present, secondary methods, such as latent print analysis (e.g., finger, ear, palm, foot, lip), comparison of body modifications (e.g., tattoos, implants) and physical appearance in general (e.g., facial features, scars, moles, vein pattern), or certain pathologies and medical procedures, may be possible. If only skeletal elements remain, a biological profile is generally established (incl. age, sex, stature, hard tissue trauma and pathology). A stable isotope analysis using remaining hard tissues such as hair, teeth, bone, and/or nails may provide information on broad dietary background and likely geographical origin, which can also aid as a screening tool to narrow down AM data pools for further comparisons, which are then conducted through other methods (Chesson *et al.*, 2020). INTERPOL has recognised the importance of secondary identifiers, due to the known limitations of primary identifiers (Interpol, 2018; Cornett *et al.*, 2019). Secondary identifiers are often used as supporting evidence in combination with primary methods, as the former hold less discriminating power that – in isolation – is insufficient for positive identification. Cumulatively however, secondary identifiers may reach or surpass a threshold that justifies this (Black and Bikker, 2016). Identifying someone based on analysis and comparison of elements that can be easily lost, stolen, or exchanged between different individuals should be avoided (Blau and Hill, 2009; Cattaneo and Gibelli, 2013). To prevent cognitive bias (e.g., confirmation bias), primary and secondary identification efforts are kept separate during the compiling and analysis stage (Wright *et al.*, 2015).

1.2.1.1 Visual Identification Methods in DVI

The terminology in the literature is not particularly uniform, as *visual identification* and *recognition* are used interchangeably (Caplova, Obertova, *et al.*, 2018), and error rates are often not mentioned (Soomer *et al.*, 2001; Morgan *et al.*, 2006). Visual identification has been used as supporting evidence alongside primary identifiers, but its applicability is very much dependent on the type of MFI, environmental conditions, and state of the remains (Soomer *et al.*, 2001; Morgan *et al.*, 2006; Tsokos *et al.*, 2006; Chaikunrat *et al.*, 2011). Tattoos and other body modifications can be valuable in this context (Birngruber *et al.*, 2011; Starkie *et al.*, 2011; Byard, 2016), and have been successfully used in DVI scenarios (Beauthier, Lefèvre and De Valck, 2011). In past MFIs, visual identification – including from facial features – has shown to be highly inaccurate and error prone (Black and Bikker, 2016), especially in cases of advanced decomposition and when identification is not corroborated by other methods (Tsokos *et al.*, 2006).

Lain, Griffiths and Hilton (2003) stated that 10% of the 2004 Tsunami victims and 50% of the 2002 Bali bombing victims were misidentified by means of unaided (i.e., not conducted by trained experts) facial recognition. Family members are not reliable at visual identification or familiar face recognition when in an emotional state of emergency, and exposing relatives to this task will further negatively impact their ordeal (Ranson, 2016). PM and injury related changes to the face further render an objective analysis or identification through relatives impossible (Black and Bikker, 2016). By contrast, other sources emphasise the immense value of visual identification. According to Chaikunrat, Pongpanitanon and Petiju (2011), approximately one third of the 2004 Tsunami victims in Thailand were identified based on physical appearance in the early stages of DVI efforts, when decompositional changes had not affected facial appearance beyond comparatibility to AM records. Physical descriptions (95%) and visual identification (48%) were also successfully applied in the 1994 Estonia ferry accident, although one crew member was misidentified which was later rectified (Soomer *et al.*, 2001).

Misidentification can have several underlying causes. If relatives are in denial and cannot accept that their loved one has died, they may not recognise or be able to confirm the identity. If they are desperate to get answers and solve their very personal missing person's case, they may claim a body as their relative, even though it is in fact someone else. The remains are likely not resembling the mental image of their loved one, as death and decompositional processes (and possibly injuries) changed the features, rendering them unrecognisable. Potente *et al.* (2021; p.1) developed and proposed a "*bubbling procedure*", by which injured or decomposed facial features are obstructed digitally, to spare relatives some distress and hide distracting elements. As humans tend to possess the ability of perceptual filling-in (i.e., missing information not directly provided by sensory input is automatically, cognitively filled in to create a complete picture/face), the bubbling does not hinder familiar face recognition. In cases where larger portions of the face are compromised, identification may no longer be facilitated through this method, and facial depictions or other means of visual presentation may be more suitable (Wilkinson, 2014).

In cases where the main primary and secondary methods of identification cannot be applied, other means become necessary (Khoo *et al.*, 2016; Caplova, Obertova, *et al.*, 2018; Bacci, Briers and Steyn, 2021). Facial identification efforts by qualified, trained, and experienced experts through morphological comparison (FISWG, 2018a), craniofacial superimposition (Jayaprakash, Srinivasan and Amravaneswaran, 2001; Al-Amad *et al.*, 2006), face matching (Fysh and Bindemann, 2017b; Moreton, 2021b), craniofacial reconstruction (Taylor, 2001; Rynn, Balueva and Veselovskaya, 2012), and/or facial depiction (Wilkinson, 2014) have been and should be considered (Black and Bikker, 2016; Ranson, 2016). The method of choice is largely dependent on the state of the remains, as visual identification can also be an impossible task for the expert, if features are no longer comparable (Ranson, 2016).

Addressing the common issue of availability of AM data, it can be argued that, with the popularity of smart phones and social media, even in developing nations, facial images as comparative AM data are almost always available, and often more accessible (digital format, e.g. social media) than AM DNA or dental records (Yoshino, 2012; Broach *et al.*, 2017; Nuzzolese, Lupariello and Di Vella, 2018). Cloud-based AM data is also less likely to vanish compared to other AM data that may be destroyed during an MFI, or lack of comparative DNA samples when entire families fall victim to the event (Farrell, 2018). People in developing countries are taking selfies, not necessarily with the intention to post them on social media, but more so because they are aware this may be the only way to identify them after a MFI (E. Untoro 2020, personal communication, 23rd July). Olivieri *et al.* (2018) endorse the use of selfies and other face data in DVI, especially in relation to the migrant crisis (case study on 2013 Lampedusa shipwreck).

Most scientific approaches regarding visual and especially facial identification of the deceased, have not yet made a successful transfer from promising laboratory-conditions-based results to real-life scenarios, and their applicability in field conditions remains to be explored (see Caplova, Obertova, *et al.*, 2018 for review). Broach *et al.* (2017) proposed using commercial face recognition software in DVI. Whilst the general idea is auspicious, their study demonstrated this method on different "*victim moulage[s]*" (Broach *et al.*, 2017; p. 569) derived from frontal view images of living individuals, with artificial injuries digitally patched onto the face. The visual cues are not remotely comparable to deceased individual's faces, hence the transferability of the method to real-life scenarios is not given. The small sample size used is another major limitation that was acknowledged by the authors. Khoo and Mahmood (2020) cautiously suggest artificial intelligence (AI) and facial recognition as a potential fourth primary identification method, proposing to compare PM facial photographs taken immediately following an MFI against government databases of facial images captured by surveillance or security cameras (e.g., passport gates). This in turn opens up heated debates and concerns regarding data protection, identity theft, and surveillance states (e.g., Giroux, 2015; Corcoran and Costache, 2016).

Many human remains have not been found and/or recovered, many have not been formally identified and remain in unmarked (mass) graves, and the next disaster will inevitably occur and claim further victims (Farrell, 2018). This demands an ongoing humanitarian and scientific engagement. Therefore, DVI team trainings are essential, and best practice recommendations, guidelines, and standardised procedures need to be continuously updated, improved, and expanded. Digital databases for AM and PM data collation are extremely valuable and hold great potential, as either dataset may be collected at very different locations (e.g., deceased migrants with long travel routes) (Interpol, 2018; Kobelinsky and Furri, 2020; Ramírez Páez, 2020). However, data sets are only comparable, when certain standards for collection are followed and more effort needs to be invested into establishing such databases, ideally merging them into a global database (UN, 2017). More research and case reports are urgently needed, to establish a clearer picture of what approach can be successful under which circumstances. Methods that rely on morphological comparisons are harder to quantify compared to DNA or dactyloscopy, and the question then remains: “*how much is enough for positive identification?”* (Cattaneo and Gibelli, 2013; p.158).

2 Forensic Anthropology

Anthropology, the study of humans (Greek: *anthropos* = human; *logos* = ground / reason / word), is comprised of many sub-disciplines, such as social, cultural, linguistic, and physical or biological anthropology. The latter is focused on the behavioural and biological aspects of individuals, in both modern humans and their extinct ancestors (hominins), which entails their evolutionary biology and developmental history, and can be considered the "*the parent discipline of forensic anthropology*" (Sauer, Wankmiller and Hefner, 2016). In the past, forensic anthropology was considered a sub-field of biological or physical anthropology, but more recently it has been argued that it should be regarded as an independent discipline (Burns, 2007; Passalacqua, Pilloud and Congram, 2021; p.243). It is often defined as the search, recovery, analysis, and collection of evidence from human skeletal remains for medico-legal purposes (Burns, 2007; Ubelaker, 2018). However, definitions and particulars of the scope of duties does vary depending on region or country, position, and operation (Blau and Ubelaker, 2016b; MacKinnon and Harrison, 2016; Obertová *et al.*, 2019; Passalacqua *et al.*, 2021). Anthropologists and archaeologists alike have collected skeletal material, conducted analyses, and drawn their conclusions for many hundreds of years. Historically, forensic anthropology was not so much an independent discipline as it was a scientific side-line interest of biologists, medical doctors, and anatomists.

It was not until the mid-19th century that their approaches and findings were applied to and used in medico-legal matters, e.g., the Webster/Parkman murder trial in 1850 (Webster and Bemis, 1850; Burns, 2007; Ubelaker, 2018). Internationally, those particular skills were called upon on a large scale in mass grave identification efforts, following genocide and other war crimes or casualties, notably in the former Yugoslavia and Rwanda. These deployments of Argentine, US, and UK personnel led to a new-found appreciation for this 'new' field of expertise. Forensic anthropology has since gradually become a truly independent discipline in the UK, with appropriate training and education frameworks in place (MacKinnon and Harrison, 2016). Although forensic anthropology was mainly developed in North America and Europe, it has since expanded globally, which simultaneously brought about overall developments, expansions, and improvements of the practice, methodologies, and standards used, as more variation in population specific traits are now being taken into account and new areas of human identification are being explored (Ubelaker, 2018). However, it remains a relatively young field and therefore standards and methodologies are constantly being revised and supplemented by new research findings. The field has grown rather substantially over the past few decades and is now overlapping with a large number of other fields, mainly areas within anthropology, medicine, pathology, computer science, criminology, and biology.

Significant developments in the UK over the past decade have generated a much-needed standardisation and professionalisation of forensic anthropology as an independent discipline, with the introduction and implementation of a certification scheme (MacKinnon and Harrison, 2016). The Royal Anthropological Institute (RAI) of Great Britain and Ireland, in collaboration with the British Association of Forensic Anthropology (BAFA) and the Home Office's Forensic Science Regulator (FSR), have established a three-level certification framework, that evaluates and accredits appropriate and sufficient training, experience, and competence as a new professional standard within the field (for further information, please see https://www.therai.org.uk/forensic-anthropology [accessed 09.08.2022]).

Certification on any of the three levels is now considered a prerequisite for UK-based forensic anthropologists who wish to be involved in forensic casework and international deployments. The Royal Society and The Royal Society of Edinburgh have very recently developed and published a judicial primer for forensic anthropology in the UK (Hackman *et al.*, 2022). This is a document specifying and explaining the science, making it understandable to judges and juries, but also defining the possibilities and limitations of the field. Relating to this is undoubtedly the *CSI effect*, which poses a real issue, in the sense that many people tend to get overly excited and have unrealistic expectations about what forensics can and cannot achieve, broadly influenced by popular fiction and true crime formats (Foltyn, 2008; Adams, 2016). Many students find themselves drawn to anything forensic as part of a degree, without really grasping what case work later on will truly entail and ask of them, and often glamorous and miraculous portrayals of the discipline on TV shows are very far removed from the real world.

Tasks of the forensic anthropologist commonly include establishing a biological profile by means of sex, age, stature, and population affinity estimations (Forbes and Nugent, 2016; Sauer *et al.*, 2016), but can also entail – and are not limited to – determination of time since death, ante-, peri-, and post-mortem trauma, expert witness work, crime scene attendance, human rights investigations, managing the appropriate response or approach to an event, and interaction with next-of-kin, other practitioners, and government authorities (Sledzik and Mundorff, 2016; Márquez-Grant and Roberts, 2021; Passalacqua *et al.*, 2021). More recently, forensic anthropologists are also often tasked with examination of living individuals (e.g., age estimation from dental records or epiphyseal fusion patterns; aka. *judicial anthropometry*), and less skeletonised human remains (e.g., in contexts of burning, drowning, or dismemberment) (Blau and Ubelaker, 2016a). Forensic anthropology and identification efforts always have an ethical dimension, in the sense that human remains are not merely an object, they are evidence and should be treated with dignity and care, whilst remaining objective. Furthermore, the scope of the work encompasses other anthropological aspects, and therefore requires cultural awareness (e.g., religious beliefs, burial practices), and a consideration for potential political, social, and judicial implications (Blau, 2016). Forensic anthropology is now commonly requested and implemented in criminal case work in the UK and in humanitarian identification efforts internationally (MacKinnon and Harrison, 2016). There are many complimentary disciplines, such as forensic archaeology, entomology, criminology, odontology, or environmental geoscience, and an interdisciplinary approach should be encouraged (Sledzik and Mundorff, 2016; Márquez-Grant and Roberts, 2021).

3 Facial Anthropology and Identification

Facial anthropology is a branch of forensic anthropology, which is focused on the identification of living and deceased individuals through assessment of craniofacial anatomy and biometrics (Wilkinson, 2015; Wilkinson and Lofthouse, 2015). Forensic craniofacial identification requires solid knowledge of facial anatomy and anthropology. It is facilitated through a variety of different methods of facial imaging and comparison that are generally applied by different experts from a broad range of fields, such as forensic odontology, law enforcement, anatomy, forensic art, forensic anthropology, medical imaging, computer science, genetics, and psychologists (Stephan *et al.*, 2019). Methods include facial image comparison and analysis, craniofacial superimposition, facial depiction, age progression and regression, facial composites, automated and manual face recognition, and facial reconstruction or approximation. This section will focus predominantly on facial image comparison, as other techniques of facial anthropology and identification are not relevant for the current research. For a comprehensive overview of current facial imaging methods, please also see Stephan *et al.* (2019).

3.1 Facial Identification

The face has been deemed the third most individualistic and diverse feature bearer for personal identification, after DNA and fingerprints (Kreutz and Verhoff, 2004). It is indeed a matrix of seemingly endless possibilities for variation. It changes shape throughout development, through facial expression, intrinsic and extrinsic factors, but also in death, through loss of muscle tone and eventually decompositional changes. The unique configuration of the human face plays an important role in human identification (Stephan *et al.*, 2008). Visual information, such as facial data, has a very high recognition value (Alt, 2012) and faces hold and provide a multitude of social, emotional, and biological cues, which are extremely important for everyday interactions (Rule *et al.*, 2013; Hehman *et al.*, 2017; Verosky *et al.*, 2018). As we define fellow humans primarily by their face, its perception becomes a key factor in identification. Therefore, in order to recognise or identify someone, the instinctive feature or composition thereof would first of all be the face. Faces are used in the context of identification of the living and the dead, however, there is very little research and even less validation studies published to date on the latter. The sections below will provide a brief overview of the main facial comparison methods, their applications, accuracy, limitations, evidentiary value, and legal requirements, as well as current best practice guidelines and standards.

3.1.1 Facial Identification Methods

Facial identification methods can be subdivided into manual and automated methods, with some overlap. Morphological analysis is entirely manual and reliant on human experts (e.g., Stavrianos *et al.*, 2012; Towler *et al.*, 2019; Schüler and Obertová, 2020; Bacci, Houlton, *et al.*, 2021), whereas craniofacial superimposition used to be manual in the past and is now commonly conducted as semi-automated and/or computer assisted (Yoshino *et al.*, 2000; Iino *et al.*, 2016; Damas *et al.*, 2020). Electronic face recognition systems are mostly automated, but most require human input in the verification stage (Hassan *et al.*, 2015; Jain, Nadnakumar and Ross, 2016; Fysh and Bindemann, 2018; Kaur *et al.*, 2020). Data formats also differ, as most facial comparisons are performed between two sets of 2D images, but 3D-to-2D or even 3D-3D comparisons are also possible in certain circumstances, although less well researched and validated (e.g., Bowyer, Chang and Flynn, 2004; Lu, Jain and Colbry, 2006; De Angelis *et al.*, 2009; Senthilkumar and Gnanamurthy, 2013; Gómez *et al.*, 2018; Adjabi *et al.*, 2020). All facial identification procedures that utilise facial images can be summarised under the term facial imaging (Stephan *et al.*, 2019).

Facial matching or facial comparison is generally used as an umbrella term for several different approaches that can be applied when comparing facial images of an unknown or unidentified individual to a known person, to determine whether or not they are the same (Stavrianos *et al.*, 2012; Stephan *et al.*, 2019; Fysh, 2021). Side-by-side morphological or feature-based comparisons of 2D photographs and/or video stills from CCTV footage, and superimposition of 2D and 3D data (e.g., face scan or photogrammetry models) are currently the most common approaches (Cattaneo *et al.*, 2012). Anthropometry or photo-anthropometric analysis has also been used rather widely over the past 150+ years, but is no longer recommended (İşcan, 1993; Davis, Valentine and Wilkinson, 2012; FISWG, 2019a). It is however still important to discuss anthropometric approaches, in order to fully understand the history and developments in facial identification to date. The FISWG also include anthropometry in their summary of the four main categories of facial comparison approaches (FISWG, 2019a), which are briefly summarised hereafter. Automated face recognition and verification systems are discussed separately under section 6.2 Automated Face Recognition Systems below.

3.1.1.1 Holistic Comparison aka. Facial Review

Holistic face processing is an automatic and natural way for humans to recognise and identify other people, by considering the overall appearance and a holistic composition of individual features. This has been found to be very accurate in familiar face recognition (e.g., friends, family members, colleagues), considerably less so for mentally matching unfamiliar faces (Bruce, Henderson and Newman, 2001; Maurer, Le Grand and Mondloch, 2002; Sinha *et al.*, 2006). FISWG does not consider this a method of facial comparison (FISWG, 2019a), and it is known that large inter-person variability regarding holistic, unfamiliar face matching ability exists (see section 6.3.2 Face Recognition and Matching Ability below for further details). Holistic face matching however is still employed in isolation when fast identification decisions have to be made (e.g., border control) (Megreya, 2018), and is also inherently part of other facial comparison approaches, as it appears to be extremely difficult to by-pass this automatic process (Michel, Caldara and Rossion, 2006; Richler and Gauthier, 2014; Meinhardt-Injac *et al.*, 2017). However, the European Network of Forensic Science Institutes (ENFSI) only recommends holistic comparisons in scenarios where no other method is feasible or available, but facial examiner's unfamiliar face recognition baseline abilities should be tested and taken into consideration (ENFSI, 2018).

3.1.1.2 Anthropometry

As a subcategory to biometrics (see section 6.1 Biometrics below), anthropometry is the scientific study of proportions and measurements of the human body (Farkas, 1996; Kleinberg, Vanezis and Burton, 2007; Caple and Stephan, 2016). Alphonse Bertillon, a French police officer and biometric researcher in the late 19th century, developed a personal identification system based on frontal and profile view mugshots, a defined set of facial anthropometric measurements, and individual features such as eye colour, scars, and other marks (Bertillon, 1890; Galton, 1896; Fosdick, 1915; Jain *et al.*, 2016). Differences in measurements and morphology between individuals were used to record, catalogue, and subsequently verify the identity of criminals, which is sometimes referred to as *Bertillonage* (Bertillon, 1890; Grüner, 1993). Although the approach initially appeared very promising and was rapidly adopted across other countries, it was ultimately deemed too time consuming and laborious, the element of anthropometry too error-prone, due to carelessness, inconsistency, intra- and inter-observer errors, non-standardised descriptions, and incorrect readings by oftentimes untrained staff (Galton, 1896; Jain *et al.*, 2016). The Bertillon system was mostly replaced by the manual comparison of fingerprints (aka. dactyloscopy) in the early 20th century, largely influenced by Henry Faulds, William Herschel, and Sir Francis Galton (Jain *et al.*, 2016). Craniofacial anthropometry still forms the basis for facial approximations (Stephan, 2003; Wilkinson, 2008; Duan *et al.*, 2014; Maltais Lapointe, Lynnerup and Hoppa, 2016), and is widely used in aesthetic medicine and clinical treatment (e.g., craniofacial surgery and orthodontics) (Farkas, 1994, 1996).

Photo-anthropometry, which relies on and utilises measurements, dimensions, and angles of facial landmarks and inter-feature relationships from images, has been applied in facial comparison and identification scenarios in the past. A frequent criticism however lies in the difficulties of locating and consistently marking landmarks on photographs, and drawing reliable conclusions from non-standardised images (İşcan, 1993; Campomanes-Álvarez *et al.*, 2015; Lee *et al.*, 2019). Any image variations, such as distortion, differences in facial angle, low resolution, obstructed features, or lighting issues can significantly interfere with this type of analysis, and it becomes extremely difficult to clearly define the threshold for differences and commonalities (Moreton and Morley, 2011; Gibelli *et al.*, 2016). Although appealing, as it has quantitative and therefore seemingly more objective properties, photo-anthropometry is no longer recommended as a reliable method in this context, not even when used in conjunction with other methods (FISWG, 2019a; OSAC, 2021). A distinct lack of discriminating power, a large non-negligible component of subjectivity, and inter-observer differences have been discovered in past research, deeming photo-anthropometry misleading and unreliable (Kleinberg *et al.*, 2007; Evison and Vorder Bruegge, 2010; Moreton and Morley, 2011). The applicability of photo-anthropometry to real-world identification scenarios is extremely limited, as most images (both ante- and/or post-mortem) are generally non-standardised, and the visibility and hence availability of landmarks is likely diminished or obstructed.

Obtaining both frontal and profile view images as mugshots has been discovered and appreciated over a century ago (Bertillon, 1890). It was temporarily adapted in multiple countries in a forensic context, yet research on profile view facial comparisons is still very sparse (Abdel-Mottaleb and Zhou, 2006; Kramer and Reynolds, 2018). Arguably, a second or any additional perspective can add considerable value, as it provides additional information on the morphology of an individual, a wider context and therefore more possibilities for comparison. In their unfamiliar face matching study on frontal and profile view images, Kramer and Reynolds (2018) asked participants to a) undertake the Glasgow Face Matching Test (GFMT) with frontal view images only, b) with lateral images only, and c) with both frontal and lateral view image stimuli available for the face pool and target images. No advantage of frontal over lateral view or vice versa, nor an additional benefit of having both views available was found. This challenges other theories of face matching (Robertson, Middleton and Burton, 2015; Ritchie and Burton, 2017; Fysh, Stacchi and Ramon, 2020). A scenario in which one might have to compare a frontal view (AM) to a lateral view (PM) image (or vice versa), as can be the case in real life scenarios, was not explored. Nevertheless, this study has raised important questions that need to be addressed in future research, namely, what are the differences – if any – in the information humans can draw from frontal as opposed to lateral images and why do we seem unable to extract from and utilise both viewpoints in conjunction? One possible explanation might be that we are simply more used to frontal face-to-face (as opposed to lateral ear-to-ear) interactions and hence subconsciously place more importance on frontal views and are less used to and hence versed in extracting and interpreting cues from profile views. Although this would not explain the similar performance in frontal and profile view image stimuli, when observed in isolation (Kramer and Reynolds, 2018).

3.1.1.3 Superimposition

By creating an overlay of two faces from either 2D images and/or 3D models, the best possible alignment is sought, which serves as an aid for visual comparison of the facial structures and overall morphology. Several different approaches exist, such as manual photographic or video-based comparisons, as well as computer-aided and fully automatic alignments (Damas *et al.*, 2020). Current recommendations and best-practice guidelines do not promote this as a stand-alone method, as it has severe limitations in its applicability and is considered subjective and often unreliable (FISWG, 2019a; OSAC, 2021). The ENFSI (ENFSI, 2018) does not recommend to use this technique for forensic facial comparisons at all. However, new developments from Panacea Cooperative Research (2021) seem highly promising. Their recently launched *Skeleton ID* software package entails a craniofacial superimposition module, which relies on cephalometric and craniometric landmarks. This function can semi-automatically produce an overlay through a computer aided CFS approach, which provides quantitative results in form of likelihood ratios for decision making and statistical proof of uniqueness, ultimately creating a matching score. To date, validation studies thereof have only focused on skull-face overlays and face-to-face studies are still required to assess this method further.

Please refer to section 3.2 Craniofacial Superimposition below for a comprehensive background of history, applications, and limitations of craniofacial superimposition.

3.1.1.4 Morphological Comparison

Morphological comparison (aka. morphological facial examination, feature based analysis) is a systematic method in which individual features and other (facial) components are compared one-by-one. The differences and similarities of each characteristic are noted, which are then reviewed and translated into a rating scale, similarity score, or levels of support for identity verification or exclusion (İşcan, 1993; FISWG, 2018a; Bacci, Houlton, *et al.*, 2021). This is by no means a novel approach, as it has been used by anthropologists and medical professionals over the past two centuries (İşcan, 1993). Analysis will generally focus on overall features (e.g., face shape) first, and subsequently become more detailed, by analysing individual features (e.g., ears, nose, eyes) and their structure or components (e.g., eyebrow or beard hair growth pattern, ear tubercles, nostril shape and size, skin marks, etc.). It is not only individual features but also their spatial relationship to one another that are analysed. Body modifications, such as scars and tattoos, can also be very useful as they often present discriminating characteristics (FISWG, 2018a; OSAC, 2021). Despite morphological comparison being the recommended approach for facial identification through face matching, only a limited number of validation studies exist (Steyn *et al.*, 2018; Bacci, Briers, *et al.*, 2021; Bacci, Houlton, *et al.*, 2021). Nevertheless, Schüler and Obertová (2020) highlight the value of this approach, as trained experts will outperform untrained individuals and automated face recognition systems in this task. This view is *inter alia* based on research by Towler, White and Kemp (2017), who found that by providing novices and experts with a simplified feature list, experts outperformed novices and showed a better understanding for the varying discriminatory power of certain features. In their study, ears, scars, and other marks were deemed the most useful features for comparison in this context.

In an effort to standardise which facial features are considered for morphological analysis, as well as terminology and descriptions thereof, several summaries, studies, guidelines, and standards have been published to date (e.g., Knussmann, 1988; Schwidetzky and

Knussmann, 1988; Vanezis *et al.*, 1996; ASTM, 2018; ENFSI, 2018; FISWG, 2018a; OSAC, 2021). Two atlases on male and female facial features have been compiled, which entail a feature list and feature descriptions with visualisations and examples of rating sheets (Ohlrogge *et al.*, 2008, 2009). In Germany, the *Arbeitsgruppe Identifikation nach Bildern* (AGIB; = working group for identification from images) has their own standards, that are still currently being applied for image comparison and expert witness work (AGIB, 2011). The FISWG published a *Facial Image Comparison Features List for Morphological Analysis* in 2014 and a slight adaptation of said list was published by the American Society for Testing and Materials (ASTM) shortly after (ASTM, 2018). The FISWG have since updated their feature list FISWG (2018a), and the application thereof has been found to be useful in recent research studies (Bacci, Briers, *et al.*, 2021; Bacci, Davimes, *et al.*, 2021; Bacci, Houlton, *et al.*, 2021). Diagrams for characteristic descriptors of facial features can also be found in the EFNSI *Best Practice Manual For Facial Image Comparison* (ENFSI, 2018). Despite this, there is no one document that practitioners adhere by, and there is still an element of practitioners relying on different standards and adapting their methodologies as they see fit for casework.

3.1.2 Applications of Facial Identification

Most research and real-life applications in facial identification to date have been focused on the living (e.g., Bromby, 2006; Abate *et al.*, 2007; Frowd *et al.*, 2007; Lynnerup *et al.*, 2009; Moreton and Morley, 2011; Bindemann *et al.*, 2013; Solomon and Gibson, 2014; Strathie and McNeill, 2016; Phillips *et al.*, 2018; Kortli *et al.*, 2020; Schüler and Obertová, 2020) with the exception of facial reconstruction/approximation, craniofacial superimposition, and facial depiction (e.g., Taylor, 2001; Yoshino, 2012; Kealy *et al.*, 2014; Wilkinson, 2014; Joukal and Frišhons, 2015; Wilkinson and Lofthouse, 2015; Damas, Cordón and Ibáñez, 2020). Some applications of facial identification, recognition, and verification include age verification (e.g., in a supermarket or night club) (Kemp, Towell and Pike, 1997; White, Burton and Kemp, 2014; Osman and Viriri, 2018; Papesh, 2018), identity cards (Jain, Ross and Pankanti, 2006), passport control (borders), missing persons investigations (e.g., Blythe and Woodforde, 2007; Cattaneo and Gibelli, 2013; Bikker, 2014; Nuzzolese, Lupariello and Di Vella, 2018; NCA, 2021), and CCTV surveillance (e.g., Davis and Valentine, 2009; Bindemann *et al.*, 2013; Moreton, 2021). There is a notable imbalance in terms of research focus, as some methods like facial composites, craniofacial superimposition, and craniofacial approximation have been studied considerably more in depth and volume (Stephan *et al.*, 2019) compared to facial comparison. This is despite the latter increasingly gaining relevance in casework (Steyn *et al.*, 2018; Bacci, Houlton, *et al.*, 2021). Even less research and published case studies are available on face identification of the deceased (Anderson, 2008; Lee *et al.*, 2011; Cornett *et al.*, 2019; Khoo and Mahmood, 2020; Potente *et al.*, 2021).

3.1.3 Reliability, Limitations and Challenges of Facial Identification

Most facial imaging approaches are generally used in combination with other methods of identification but are in their own right truly valuable, as they can help to direct investigations, narrow down face pools of possible matches, and in some cases also facilitate positive identification in isolation. It is understood that rejection or exclusion decisions are easier compared to establishing positive identification, as "*a single morphological difference […] will lead to an exclusion of identity, unless […] the difference in the feature expression can be plausibly explained*" (Schüler and Obertová, 2020; p.326) e.g., through ageing. Past research has shown that unfamiliar face matching can be highly error prone for a number of different reasons (see e.g., Bruce *et al.*, 1999; Megreya and Burton, 2008; Bindemann, Avetisyan and Rakow, 2012; Megreya, Sandford and Burton, 2013; Fysh and Bindemann, 2017b; Balsdon *et al.*, 2018; Megreya, 2018; Megreya and Bindemann, 2018). Most of the aforementioned studies have focused on the identification of the living, but findings are still relevant for the deceased, as existing issues in facial identification (e.g., illumination, changes in appearance, distortion, image quality) would presumably only increase or worsen when introducing the added factor of loss of life and associated facial changes (see section 5.3 Post-Mortem Changes to the Face below).

The seemingly easiest, most straight-forward, and highly favourable scenario for facial identification – namely being presented with same-day, very similar, standardised photographs – has shown to be very difficult and even then, related face matching results are not as accurate as one would assume (Megreya and Burton, 2006; Burton *et al.*, 2010). Therefore, it is only logical that performance further declines under more challenging conditions, e.g., large time gaps between image capture dates and/or unfavourable imaging conditions (e.g., illumination, pose, angle, facial expression, distortion). A considerable issue with facial comparisons overall is that it entails judgement under uncertainty. On the one hand, the face of the same individual can appear very different in two photographs, or when a live person or video footage is compared to a still photograph. This is due to factors such as age, body modifications, changes in lighting and/or pose (Schüler and Obertová, 2020). On the other hand, two different individuals' faces may look rather similar. An extremely limiting factor in most research to date in unfamiliar face matching and recognition, is the utilisation of mostly standardised, frontal view images, that are acquired under laboratory conditions (Bruce *et al.*, 1999, 2001; Megreya and Burton, 2006, 2008; White *et al.*, 2013; White, Burton, Jenkins, *et al.*, 2014; Dowsett and Burton, 2015; Kemp *et al.*, 2016). However, image variability is a key aspect of understanding and improving face matching, therefore more natural, wild, or ambient facial images must be considered in future experimental studies. Although this limits the control researchers have over certain aspects or variations, Young and Burton (2017; p. 216) argue that it is necessary "*in allowing us to find consistent cues for recognizing face identity and for other aspects of face perception*".

The quality of photographs or other facial representations (e.g., video stills, 3D face models) is arguably one of the most significant components in facial identification. The higher the quality (i.e., resolution, angle, illumination), the more features and details are likely visible, which can be used in morphological analyses. The more features are compared, the less random and statistically more powerful a comparison becomes, hence increasing its evidential value (Schüler and Obertová, 2020). Ultimately, the aim is to determine uniqueness or singularity by comparing a high enough count of morphological features (Lucas and Henneberg, 2015). In reality however, image quality is often not ideal and other factors, such as time elapsed between when two photographs were taken, can also introduce difficulties and limitations (see section 3.2.1 Issues and Best Practice Recommendations for more details). As per the current *Best Practice Manual for Facial Image Comparison*, sufficient image quality and feature visibility should therefore be considered a necessary prerequisite for morphological and anthropological comparisons

(ENFSI, 2018). Schüler and Obertová (2020; p.325) suggested a six-point scale for evaluating image quality prior to conducting facial comparisons.

There appears to be a lack of consideration regarding quantifiability of decisions and related results (Gibelli *et al.*, 2016). This may also be due to the very nature of facial comparisons and the fact that unique, individualising morphological features are extremely difficult to reliably quantify and measure (Komar and Lathrop, 2006 in Sledzik and Mundorff, 2016). To date, research on morphological comparison has been extremely scarce (Stephan *et al.*, 2019), and in those few available publications, error rates are often not stated.

Previous research on facial identification of the living did often not incorporate an extensive morphological feature list, but rather focused on improving unfamiliar face matching by providing feature instructions (Megreya and Bindemann, 2018) or adding feature classifications and categorisations. This potentially allows to differentiate more reliably between an oftentimes mere holistic matching decision and a feature-based approach. For example, Towler, White and Kemp (2014) tested whether face shape classification (i.e., oval, oblong, round, etc.) in face matching tasks (using GFMT) would be "*diagnostic of identity*" (p.215) and improve matching accuracy, which it failed to accomplish. Perceived face shape appeared to vary widely depending on the observer and matching based on face shape was not consistent. Adding this tool did not improve matching task results (mean accuracy = 82% for both scenarios, with and without face shape classifications).

More recent studies (published after the practical part of this current research project had already been completed) aimed to validate the feature list approach as recommended by the FISWG (2018a) on the living. Bacci, Briers, *et al.* (2021) compared CCTV stills vs. standardised photographs, with and without obstruction (i.e., cap, sunglasses) using morphological analysis. Results indicate that brimmed caps were more detrimental to matching accuracy (61.8%) and reliability (κ=0.639), compared to sunglasses (90.4%; κ=0.798). This is in agreement with other research, such as Henderson, Bruce and Burton (2001) and Lee *et al.* (2009), who also found brimmed caps to severely affect accuracy rates (57.1%; 40-43%). In another validation study, applying FISWG recommended protocol for morphological comparison, Bacci, Houlton, *et al.* (2021) compared wild-type photographs and suboptimal CCTV stills against a face pool of standardised images. The lower quality of the suboptimal CCTV data had a bigger negative impact on accuracy (82.6%), compared to the non-standardised photographs (99.1%). The authors also state that, although this is in agreement with previous findings, accuracy was still higher for a similar set-up, suggesting that the FISWG recommended morphological, feature-based approach leads to more reliable results. Nevertheless, although morphological comparison is applied in casework, case studies and especially validation studies are still severely lacking, making efforts to place research findings within the body of knowledge and greater context extremely difficult.

Further limitations persist, as Moreton (2021; p. 162) points out that "*[i]t is currently unknown how applicable morphological analysis is to matching different sources of imagery, such as low-quality CCTV footage to high-quality police mugshots*". This approach is also highly dependent on the quality and quantity of the data, and judgements or evaluations are and will always entail an element of subjectivity or investigator bias (Sauer, Michael and Fenton, 2012). It is extremely difficult to quantify which features (or components thereof) are truly unique, or distinctive. The FISWG guidelines (FISWG, 2018a, 2019a) list an extensive number of morphological features, but do not specify a minimum number required to make an identification or exclusion decision, and also do not acknowledge the fact that not all features are applicable in most comparison scenarios. The latter being influenced by data quality, head angle, obstructions of facial features and other limiting factors. Furthermore, the FISWG guidelines provide a list of features but no terminology for descriptions thereof, hence the issue of non-standardised terminology and subsequent inter-observer variations of its use persist.

Feature descriptions do exist (İşcan, 1993; Vanezis *et al.*, 1996; Dunn and Harrison, 1997; Allanson *et al.*, 2009), but are not universally standardised nor uniformly applied, neither in theory (research) nor in practice (field) (Dunn and Harrison, 1997; Caple and Stephan, 2016; Steyn *et al.*, 2018). Ritz-Timme *et al.* (2011) found that individual intra-observer feature descriptions are highly subjective, which is likely *inter alia* a result of the aforementioned issues. Whilst one might argue that facial comparison may still be effective and reliable, as long as the observer applies their approach uniformly and consistently, it still hinders collaborative efforts and complicates interpretation and potential transferability of research or case study findings. Clear, uniform terminology is desperately required for this very specialised field and area of research (Bacci, Davimes, *et al.*, 2021).

The most common challenges encountered in unfamiliar face matching, when applied in a real-world context, are listed below (Robertson *et al.*, 2015; FISWG, 2021; Fysh, 2021):

- Face variation over time (Jenkins *et al.*, 2011; Erickson, 2016; Stephan *et al.*, 2019)
- Time pressure (Fysh and Bindemann, 2017a; Wirth and Carbon, 2017)
- Impostors, identity fraud (Balazia *et al.*, 2021)
- Low image quality (Hancock, Bruce and Burton, 2000; Bourlai, Ross and Jain, 2011; Ritchie *et al.*, 2018)
- Viewpoint and facial angle differences (2D-2D) (Estudillo and Bindemann, 2014; Favelle, Hill and Claes, 2017)
- Comparing 3D person (real life) to 2D image (Kemp *et al.*, 1997; Megreya and Burton, 2008)
- Disguises, obstructions (Ramanathan, Chellappa and Roy Chowdhury, 2004; Ghiass *et al.*, 2014; Suri *et al.*, 2019)
- Facial expression (Bruce *et al.*, 1999; Drira *et al.*, 2013; Lei *et al.*, 2014; Jin and Tan, 2016)
- AM to PM changes / living vs. deceased (Tsokos *et al.*, 1999; Bolme *et al.*, 2016; Caplova, Obertova, *et al.*, 2017; Cornett *et al.*, 2019; Davis, Maigut and Forrest, 2019b)

In order to control for certain factors and establish their individual impact on face matching or recognition, most studies use standardised data that only introduces one or a select few of the challenges listed above. While this approach is comprehensible, it also entails that methods used and resulting accuracy levels cannot be directly applied to non-standardised data in real-world settings, where it is very rare to only have a single limiting factor influencing the analysis. Importantly, Bacci, Davimes, *et al.* (2021) state that data given to or held by law enforcement agencies generally reflects that these known limitations are oftentimes not taken into account and available guidelines, standards, and recommendations are evidently disregarded. This can result, for example, from image capture by outdated systems, or acquisition by untrained staff. Compared to police line-ups, or passport control scenarios, some important information about an individual is lost when only a 2D or 3D representation of the face is available for comparison. A passport contains information about height and exact eye colour, therefore seeing someone in person and comparing more than just facial features can be beneficial (Stevens, 2021). However, facial identification of the living has shown to be challenging, even under ideal conditions (see section 6.3.2 Face Recognition and Matching Ability below for further details). Going even further, neither average humans nor machines are capable of accurately matching unfamiliar faces (Jenkins and Burton, 2008). Misidentification can have a wide range of implications, ranging from mistaking someone on the street for your neighbour and resulting awkwardness to wrongful convictions and a death sentence (McNeill, McNeill M. and Strathie, 2015; Innocence Project, 2022).

3.1.4 Legal Requirements and Evidentiary Value

The requirements for presenting face comparison evidence and respective levels of confidence or other result formats for positive or potential matches vary, depending on the country, in which a case is heard. In England and Wales, evidence is generally presented on a scale of or levels of support, but it has been argued that this format, e.g., Bromby scale (Bromby, 2006), is problematic, as the scale itself does not provide transparent guidance for differentiation and decision making, which ultimately then lies with the expert, introducing subjectivity and not necessarily uniformity (please refer to Steyn *et al.*, 2018 for details). The *Anthropological Atlas of Male Facial Features* (Ohlrogge *et al.*, 2008) and the *Anthropological Atlas of Female Facial Features* (Ohlrogge *et al.*, 2009) were created as part of the 'Optimisation of Methods for Identification of Persons from Photographs (Photoidentification)' project. An important component of those atlases are frequency lists, which show the prevalence of certain characteristics within the population. However, it is unclear as to which population groups these include. Assuming it relates to White/Caucasian Europeans only, the frequencies would not be transferable in other contexts and therefor hold limited value and emphasise the need to collect more population data on an ongoing basis, to determine global frequencies as accurately as possible, which could then be used for quantifying evidentiary observations.

A decade later, this issue still persists. The ENFSI state that "*at present, it is not possible to calculate the probability of a match between two faces. This is due to the absence of relevant population data […]*" (ENFSI, 2018; p.16).

Until quantifiable alternatives are found, existing methods are validated, and sufficient population data is collected and collated, some forensic sciences remain reliant (in part) on subjective probabilities or qualitative approaches (Schüler and Obertová, 2020). However, this "*justified subjectivism*" (Biedermann *et al.*, 2017; p.477) has limits and the need to use precise, uniform terminology and follow best practice guidelines and standards, as well as ensure that experts have the relevant training and accreditation should be paramount (Taroni *et al.*, 2018). Even then, a subjective component remains in practitioner-led facial identification, but according to current guidelines, "likelihood […] can be informed by subjective probabilities using expert knowledge" (ENFSI, 2018; p.16). When automated methods are sought, these also currently hold an element of human bias introduced at the programming stage (Fysh and Bindemann, 2018; Crumpler, 2020).

Forensic techniques and approaches are required to be replicable, validated, their reliability proven, and accuracy rates known, when applied in human identification scenarios and therefore potentially having to hold up in court (Imwinkelried, 2016; Milroy, 2017). In most states of the U.S., the Daubert criteria are an established reliability standard for all expert testimony and scientific methodology used in a court of law (Daubert v. Merrell Dow Pharmaceuticals (92-102), 509 U.S. 579 (1993)) (Lesciotto, 2015). Factors that are taken into consideration here are whether the scientific method has been tested, been subject to peer review, the potential error rate, conformity with existing standards, reliability, and whether the method is accepted by the scientific community in question (Kadlec, 2021). Prior to Daubert, U.S. courts followed the Frye standard (some states still adhere to this), which only follows one of the Daubert aspects, namely that scientific evidence is only admissible if the technique used to obtain it is reliable and generally accepted in the respective scientific community (Frye v. United States, 293 F. 1013 (D.C Cir. 1923)). The integration of the Daubert criteria into the UK legal system has been debated, but to date neither this nor other similar U.S. standards have been adopted (Law Commission, 2009; Ireland and Beaumont, 2015). The debate arose from a need for standardisation and best practice guidelines, as well as accreditation and practitioner certification for all forensic disciplines (Forensic Science Regulator, 2016a, 2016b).

As "*science is a fluid concept*" (Ireland and Beaumont, 2015; p.4), standards and guidelines are constantly evolving and need to be updated and supplemented regularly. The same holds true for legal requirements with regards to expert testimony (Lesciotto, 2015). It could be argued that "*any kind of uncertainty is assessed in the light of the knowledge possessed at the time of the assessment*" (Taroni *et al.*, 2018; p.244). Or more precisely, as Schrödinger (1947; pp.53) stated "*[s]ince the knowledge may be different with different persons or with the same person at different times, they may anticipate the same event with more or less confidence, and thus different numerical probabilities may be attached to the same event. […] Thus, whenever we speak loosely of the probability of an event, it is always to be understood: probability with regard to a certain given state of knowledge.*" It should be mentioned that there is currently no uniform certification or accreditation required nor available for professionals (e.g., border control officers, police personnel, facial imaging experts, academics) conducting facial comparisons.

3.1.5 Standardisation and Best Practice Guidelines

Several guidelines, case reports, standards, and best practice manuals in relation to facial identification have been published in recent years, by scientific working groups, industry experts, and law enforcement agencies (e.g., ACPO, 2009; ENFSI, 2018; FISWG, 2019; OSAC, 2021). These documents require ongoing updates and improvements, in line with new research findings and developments. To date, there are no standardised, international guidelines, that are being followed by all experts tasked with facial identification in some capacity. Best-practice approaches ideally reflect the current state of the art, with the limitation, that not all recommendations can be followed or are applicable in all scenarios. Better or more data is almost always desirable, but more often than not, the practitioner or expert will have to work with what is available to them.

A non-exhaustive list of said standards and best practice manuals is given below:

- *Facial Identification Guidance*; National Policing Improvement Agency (NPIA) & Association of Chief Police Officers (ACPO) (ACPO, 2009)
- Best Practice Manual for Facial Image Comparison; European Network for Forensic Science Institutes (ENFSI) (ENFSI, 2018)
- *Law Enforcement Facial Recognition Use Case Catalogue*; Integrated Justice Information Systems (IJIS) Institute + International Association of Chiefs of Police (IACP) = Law Enforcement Imaging Technology Taskforce (IJIS Institute and IACP, 2019)
- *Forensic Image Comparison and Interpretation Evidence: Guidance for Prosecutors and Investigators Issue 1*; National Crime Agency & CPS & Metropolitan Police & UK Forensic Science Regulator (Forensic Science Regulator, 2020)
- *Standard Guide for Facial Comparison Overview and Methodology Guidelines*; Organization of Scientific Area Committees for Forensic Science (OSAC) (OSAC, 2021)
- *Artificial Intelligence Act*; European Commission (European Commission, 2021)
- Inter alia: *Standard Guide for Postmortem Facial Image Capture* (FISWG, 2018b)*; Facial Comparison Overview and Methodology Guidelines* (FISWG, 2019a)*; Physical Stability of Facial Features of Adults* (FISWG, 2021)*; Facial Image Comparison Feature List for Morphological Analysis* (FISWG, 2018a); Facial Identification Scientific Working Group (FISWG). (For all current FISWG guidelines and documents, please see https://fiswg.org/documents.html [accessed 18.02.2022])

Although to date it is not a legal or otherwise regulated task to undertake facial comparison case work, the FISWG and other groups and organisations strongly advocate for thorough mentorship and training (ENFSI, 2018; FISWG, 2019b, 2020b). Existing training courses vary in content and duration, despite existing guidelines and best practice recommendations. Research has shown that short training courses are less effective, as no improvement or if so, no sustainable and replicable improvement in accuracy levels has been observed (Woodhead, Baddeley and Simmonds, 1979; Towler *et al.*, 2019). Considerably longer training or mentorship programmes, that allow an individual or trainee to gain varied exposure and professional experience in addition to substantial training and background knowledge (e.g., anatomy, ageing), could be the most beneficial and promising route for facial examiners (Towler, Kemp and White, 2021).

However, the aforementioned studies on training effectiveness have yet to be validated. More research on training and face matching expertise and possible avenues leading to improvement are needed. Focusing on the baseline face recognition ability appears to be promising, but feature-based, morphological comparison training, although only having been tested in a few studies to date, could help to slightly level the field in this regard. Results show improvements in accuracy for otherwise low performing trainees, but less changes in already high(er) performing candidates (Towler *et al.*, 2017; Megreya and Bindemann, 2018). For further reading and a comprehensive review on face matching training and face recognition expertise, please refer to Moreton (2021a).

Beyond this, there are also ethical, political, and societal responsibilities that need to be considered. With regards to validation studies and publications, it is ethical concerns and privacy rights that are a major factor and often prevent identifiable data (which faces are, by default), to be published and shared, especially in cases of deceased individuals (Ulguim, 2017; Van Noorden, 2020). Although the material may be highly sensitive, explorative and validation studies are nevertheless important if not mandatory to establish and improve best practice approaches and accuracy levels. The aim being, that existing facial identification methods are validated and improved, so they can become more reliable and replicable, and new approaches may be developed, ultimately broadening the field of human identification with more options and possible alternatives to existing methods.

3.2 Craniofacial Superimposition

Craniofacial superimposition (CFS) is a method that involves the overlay and comparison of an unidentified skull to the face of a known individual, in order to establish whether they belong to the same person, based on morphological traits, landmark alignment, and/or the correspondence of craniofacial features (Ubelaker, Wu and Cordero, 2019). The current literature differentiates between three different approaches, depending on equipment and data format used: still photographic, video, and computer-assisted superimposition (Campomanes-Alvarez *et al.*, 2018). All of these approaches have been revised and modified multiple times over the last century, and there were and still are several variations within the three main approaches (Glassman, 2001; Damas *et al.*, 2020). The umbrella term CFS on its own does not implicate a specific data format (i.e., 2D photo vs. 3D model), the acquisition and analysis method/device/equipment (i.e. photo or video camera, scanner, software), nor does it define the means of comparison (i.e. holistic, morphological, landmark-based, anthropometrical) (Damas *et al.*, 2011). Therefore, whenever CFS is utilised, all those factors need to be clearly specified.

When discussing CFS in the past and present literature, the terminology used varies amongst practitioners. It can be somewhat misleading as to which part of the process authors are referring to, but even defining the overall CFS umbrella term is lacking in consistency. Additional difficulties may arise from language barriers and translation problems, like the following examples derived from German native speakers that either pioneered or contributed significantly to CFS developments. Hermann Welcker (1883) refers to his death mask-to-portrait comparisons of Kant and Schiller as *Ineinandersetzung* (German, = interlocking, here: of images), whereas Hermann Edelmann (1938) denotes his approach of radiographic superimposition of skull and head as *Durchdringungsbilder* (German, = permeating images). Richard Helmer (1984) uses the terms *Ineinanderstellen* (German, = stacking/stashing inside each other, here: of images) for CFS and *elektronische Bildmischung* (German, = electronic image mixing) when talking about the skull-face-overlay process in video superimposition. Other authors use yet another rather unspecific term of *photographic supra-projection* (İşcan, 1993; Bronkhorst, Kenter and Stratmann, 1998; Ibáñez *et al.*, 2009; Santamaría *et al.*, 2009). CFS should not be confused with or mislabelled as forensic image comparison (Sauer *et al.*, 2012), as this describes a wider field of facial image analysis and stands above CFS for an array of methods, for which CFS is merely one (see section 3.1.1 Facial Identification Methods above).

It has previously been shown (Kleinberg *et al.*, 2007; Lynnerup *et al.*, 2009; Sauer *et al.*, 2012; Strathie and McNeill, 2016; Gibelli *et al.*, 2017) that CFS approaches can also be applied to face-to-face rather than skull-to-face comparisons, which would then be referred to as facial superimpositions (FS). This has been suggested to make comparisons easier and the outcome potentially more accurate, as it is no longer a comparison between two different objects: skull and face (Duan *et al.*, 2014; Ibáñez, Cavalli, *et al.*, 2015; Campomanes-Alvarez *et al.*, 2018). Both CFS and FS can be applied as an identification method, whenever there is a skull or face present and (ante-mortem) facial photographs are available for comparison. If dental records, DNA samples, or fingerprints of the presumed individual are also accessible, it is likely that primary identification methods would be prioritised - depending on available resources - or that CFS/FS would be applied in combination with a primary identification method (Bilge *et al.*, 2003). Oftentimes however, CFS/FS can be the only method possible, due to a known and widely recognised lack of AM data especially in the developing world, whereas facial photographs - in the age of affordable digital cameras and smart phones - appear to be accessible in almost any identification scenario (Yoshino *et al.*, 1995; Al-Amad *et al.*, 2006; Birngruber *et al.*, 2010; Jayaprakash *et al.*, 2010; Wilkinson and Lofthouse, 2015; Ibáñez, Valsecchi, *et al.*, 2016). CFS and FS are usually conducted by forensic anthropologists, anatomists, forensic odontologists (Whittaker, Richards and Jones, 1998; Cattaneo *et al.*, 2012), or by an interdisciplinary team including one or more of the former as well as pathologists, and/or computer scientists (Bilge *et al.*, 2003; Campomanes-Álvarez *et al.*, 2018). A solid understanding and knowledge of facial anatomy and anthropology is a prerequisite (Scully and Nambiar, 2002; Wilkinson and Lofthouse, 2015) and experience in facial identification desired. Moreover, since the technique relies on photographic images, an understanding of basic photographic principles is also indispensable (Taylor and Brown, 1998).

The three main stages of any CFS or FS process are essentially comprised of the same or similar elements, irrespective of which technological approach is being applied by the expert (photo, video, computer-aided) (Damas *et al.*, 2011; Oscar Ibáñez *et al.*, 2012; Campomanes-Álvarez, Cordón, *et al.*, 2014): Face/image enhancement and skull/face modelling , skull-face or face-face overlay, and decision making.

Stage 1: Face/image enhancement & skull/face modelling

The most important factor in this stage is the acquisition of the most suitable AM and PM data, i.e., medium to high resolution images in which the face is clearly visible, and a physical skull or accurate skull/face model, whenever possible. If the skull is fractured, the assembly and subsequent photographing, radiography, or scanning (CT, laser scanner) all fall within the skull modelling stage. The same applies for a skull/face scan if it is fragmented or several sweeps are combined, and the skull model requires processing prior to being used in the overlay. Enlargement, sizing, and/or scaling of the images and objects to match the comparison data is also part of this first stage.

Stage 2: Skull-face overlay (SFO) or face-face overlay (FFO)

The aim of the SFO/FFO is to find the best fit in the overlay of the two 2D images or the 2D image and a 3D object (Damas *et al.*, 2011). This is achieved through orienting and scaling the face or skull to match the face in the comparison image (Oscar Ibáñez *et al.*, 2012) or vice versa. Depending on the approach and equipment used, this second stage is arguably the most time-consuming (O. Ibáñez, Cordón and Damas, 2012). According to Valsecchi, Damas and Cordon (2018), most forensic laboratories are still relying on a manual overlay process and the required sizing, orientation, and other adjustments are achieved by rather difficult trial and error approaches (Campomanes-Álvarez, Ibáñez, *et al.*, 2014). However, there has been a focus on automating this step in a computer-based automated procedure by utilising evolutionary algorithms over the past decade (Ibáñez *et al.*, 2009; Lee *et al.*, 2011).

Usually the expert guides and judges the SFO/FFO by a large variety of anatomical criteria, such as morphological features, proportions, craniofacial landmarks, tissue depth, and asymmetries (Ibáñez, Vicente, *et al.*, 2016), but there is no set approach that needs to be followed and the weight attributed to different criteria is also unregulated (Damas *et al.*, 2015). Hence, both the analysis and interpretation are somewhat prone to bias and subjectivity. Computerised methods are able to provide an accuracy index (score on a scale from zero to one, with one suggesting a perfect match) (Campomanes-Alvarez *et al.*, 2018), which allows for a quantifiable result of the SFO/FFO stage that can guide the expert in a more objective manner when deciding upon the final result in the next stage.

Stage 3: Decision making (DM)

Irrespective of the approach and equipment involved, the expert has to make the final identification decision (Campomanes-Alvarez *et al.*, 2018). There is no common guideline as to the approach on how to reach that decision and possible classifications (Lee *et al.*, 2011). Examples of how identification decisions are expressed in the literature are as follows:

1. Positive identification vs. exclusion (Cattaneo and Gibelli, 2014);
2. Grade 1-4: close match, reasonable match, cannot exclude match/unlikely match, definite exclusion (Glassman, 2001);
3. Positive, likely positive, likely negative, negative identification (Ibáñez *et al.*, 2009).

As part of their recent best practice recommendations, Damas *et al.*(2015) suggest a sliding scale, i.e. strong support, moderate support, or limited support for not being a match; undetermined; limited support, moderate support, or strong support for being a match. The identification decision should always be accompanied by an explanation and/or a description of how this conclusion was drawn and what it was based on (Ibáñez, Vicente, *et al.*, 2016). Factors guiding the final decision can be craniofacial landmark alignment, soft tissue thickness and/or morphological feature correspondence, as well as a consistency between symmetries and asymmetries observed in both objects of comparison (Ibáñez, Valsecchi, *et al.*, 2016). The impossibility to align points or features of orientation utilised for the overlay may indicate that the compared data are not a match (Sauer *et al.*, 2012), and inexplicable differences are also leading to an exclusion rather than a positive identification (Birngruber *et al.*, 2010). Although considered somewhat subjective, irrespective of the amount of computational support in the overall procedure, the manual shape analysis and visual evaluation of the comparisons and overlays is still necessary and the expert's input based on their knowledge and experience is always required to achieve an accurate outcome (Ibáñez, Cavalli, *et al.*, 2015; Campomanes-Álvarez *et al.*, 2018). Historically and in line with technological advances over the past decades, there are three different underlying approaches to CFS: Photographic, video, and computer-based/-aided superimposition.

To date, only very few studies have focused on a face-to-face comparison in 2D-3D format. One of those, focusing on the identification of the living from surveillance footage, is the pilot study by Lynnerup *et al.* (2009). A laser surface scanner was used to obtain 3D face models from potential 'suspects', which were then superimposed over still frames from the surveillance tape. De Angelis *et al.* (2009) carried out a similar study, attesting to the high quality and reliability of the face-to-face superimposition, but stating very high acquisition costs for laser scanning equipment, which has undergone considerable improvements over the past 10 years and is now more affordable. Another face-to-face comparison was conducted by Cattaneo *et al.* (2012), again using a laser surface scanner for the 3D face model and 2D photographs. The shortfall of this study in relation to real-life applicability is the evaluation of correspondence based on lateral facial profiles only. It is unlikely that standardised lateral view AM photographs are widely available to replicate this approach in real case scenarios. For a more comprehensive review of the existing literature, descriptions, and applications of the three main methods applied by experts in the past and case studies thereof, please refer to Damas, Cordón and Ibáñez (2020).

3.2.1 Issues and Best Practice Recommendations

3.2.1.1 Landmarks

Landmarks are used as one of the tools in CFS and FS during the SFO/FO stage, and as a basis for evaluation of the level of correspondence. Craniofacial landmarks were identified as a source of uncertainty in the MEPROCS (New Methodologies and Protocols of Forensic Identification by Craniofacial Superimposition) project (Damas *et al.*, 2015). Main issues are related to incorrect and often non-replicable manual placement of landmarks and the influence of variations in soft tissue depth, the latter affected by facial expression. Landmark location is also affected by photographic conditions and image quality (Cummaudo *et al.*, 2013), as well as the position of the face and angle of head (Lee *et al.*, 2019). Best practice recommendations drawn from the MEPROCS project recommend the placement or marking of landmarks on the skull prior to scanning or otherwise acquiring the 3D model. However, this refers to skull-face superimpositions and no guidance is given on face-to-face comparisons within those guidelines. It is to be expected that the discrepancies between landmarks would be considerably smaller when comparing the same object and using the same category landmarks (e.g., facial), but the subjective factor of landmark location by the expert remains. A validation study of the MEPROCS best practice recommendations (Ibáñez, Vicente, *et al.*, 2016) has shown that not all experts employ landmarks during CFS/FS. One of the reasons being that not all software allows for landmark placement/marking, and a definitive lack of and need for specialised software for CFS.

Cummaudo *et al.* (2013), conducting the first accuracy study on landmark positioning in 2D images from frontal and lateral view, found that only a very small number of landmarks seem reliable - with the least dispersion in inter- and intra-observer error studies - therefore limiting the amount of information for meaningful conclusions. The follow-up study focused on the statistical evaluation of landmark dispersion. Findings included no difference in positioning accuracy between experts and students, and again only a limited number of landmarks were deemed reliable (incl. pupil, chelion, endocanthion, stomion, pronasale, subnasale). A study by Campomanes-Álvarez *et al.* (2015), examining inter- and intra-observer variations in facial landmark identification on 2D photographs, found that anatomical landmark location - especially for type 3, as defined above - are "*highly variable*" (p. 235). Those studies only used frontal and lateral 2D images but did not consider other facial angles or viewpoints. As landmarks are defined and dependent on a 3D surface and viewpoint, the mere application to 2D surfaces will inevitably introduce error, as some points are extremely difficult if not impossible to locate and define in a 2D image of the 3D object (Caple and Stephan, 2016). Lee *et al.*, (2019) used 2D images derived from 3D facial surface scans, showing the face from six different angles. Those where then further edited, presenting the expert with three different levels of image quality in terms of resolution. Their findings showed that landmarks were particularly inaccurate (high dispersion) in 30° lateral downward view, remarkable since this represents the typical CCTV footage angle. There is also a significant link between image resolution and landmark placement accuracy. For frontal view images, Lee et al.'s results were consistent with the previous accuracy studies mentioned above.

Computer science studies, especially those linked to facial recognition and face detection, aim to develop and improve automated landmark detection. However, most are conducted under controlled conditions and require an interdisciplinary team (Azouz, Shu and Mantel, 2006; Asi *et al.*, 2014; Valsecchi *et al.*, 2018). The study by Valsecchi, Damas and Cordon (2018) has failed to quantify erroneous landmark location. MEPROCS recommends a list of landmarks deemed most reliable and specifies advantageous software capabilities for employing landmarks for CFS (Damas, Cordón and Ibáñez, 2020;

pp.145-147). Automatic landmark detection for CFS/FS has the potential to eliminate the subjective factor of manual landmark placement by a human, as well as saving considerable time in the overall process. Unfortunately, at this stage, more studies and further developments are required in this area to make the approach applicable to possibly non-standardised real-life datasets.

3.2.1.2 Using More than One AM Photograph

As a future improvement of their technique, Glaister and Brash recommend the comparison of a skull with three rather than two portrait photographs, stating CFS could then "*possibly lead to certainty in identification*" (Glaister and Brash, 1937; p.161). Almost 80 years later, Wilkinson and Lofthouse (2015) confirm this and also emphasize that passport style images, such as those found on ID cards and drivers licences, and stored by national and international agencies (i.e., INTERPOL's Facial Recognition database) are not suitable, as such photographs "*do not provide enough distinguishing facial detail*" (p. 187.e5). The quality of the AM facial photograph is very important, and poor quality images often pose a serious issue in CFS identification scenarios (Damas *et al.*, 2011; Gordon and Steyn, 2016).The recommendation from previous studies and as part of the MEPROCS guidelines is to use more than one AM facial photograph for comparison, ideally ones in which the face is visible from more than one angle (Webster, 1955; Yoshino *et al.*, 1995; Al-Amad *et al.*, 2006; Stephan *et al.*, 2008; Damas *et al.*, 2015; Wilkinson and Lofthouse, 2015; Gordon and Steyn, 2016). This has shown to significantly improve accuracy and reliability of the technique and subsequent results (Austin-Smith and Maples, 1994; Ibáñez, Vicente, *et al.*, 2016). The misidentification rate in the Austin-Smith and Maples study was reduced from between 9.6% (lateral) and 8.5% (frontal) to 0.6% when more than one photograph was used in the comparison, enabling the face to be analysed from varying angles.

However, it is not always possible to obtain more than one AM photograph in a casework scenario. When working with a single image, oftentimes it does not fulfil the ideal requirements, but the more detail it entails, the higher the informative value and the potential validity of the CFS comparison (Stephan *et al.*, 2008). It is the experts task to evaluate the available AM data, and when deeming self-same not sufficient to facilitate an accurate comparison, CFS/FS should neither be the chosen identification method, nor should it contribute to identification by multiple methods (Campomanes-Alvarez *et al.*, 2018).

3.2.1.3 Distortion and Image Quality

Ante-mortem records are the standard on which CFS and FS have to rely on. The better the AM image, the more comprehensive and precise the analysis and subsequent results can be. Nevertheless, in most cases the quality of the PM data is much higher, as their acquisition and capture conditions receive more attention and a higher level of detail (Dorion, 1983), whereas AM photographs are often merely a result of happenstance, as the majority of AM photographs are captured under unknown and imperfect, or at least uncontrolled, conditions (Damas *et al.*, 2011; Gordon and Steyn, 2012). Nevertheless, image quality is arguably one of the most crucial factors in CFS/FS, hence photographic conditions and settings become all the more important. Image quality can be affected by compression, lens and/or perspective distortions, blurring, low resolution/pixel count. Unknown elements can be some or all of the following: camera model, type of lens, focal length and aperture, and subject-to-camera distance (SCD) - the latter greatly influencing perspective distortion. An example of this is highlighted in a study by Ward *et al.* (2018), who state that relative nasal size increased by 30% in males and 29% in females when the photograph was taken at a 12 inch distance (typical selfie distance), compared to no size difference to the original 3D shape within the face when taken at a 5 feet or 1.5m distance. In addition, differences in facial expression, lighting conditions, flexion/extension and rotation/angle of the head, as well as style or age-related changes between the reference material, need to be considered. All those factors can potentially generate errors, misinterpretations, and uncertainty (FISWG, 2012). Unfortunately, in the majority of cases, it is not possible to reconstruct the initial conditions and configurations of an AM image and/or consult with the photographer who took the photograph.

Distortions, in optical terms defined as a lack of proportionality in an image, is a major source for error in CFS/FS (Sauer *et al.*, 2012) that can be introduced in a number of ways.

Depending on photographic equipment used, there is the issue of lens distortion. The more the lens is curved (wide-angle), the more the mid-section of the image/face will be enlarged, which is known as barrel distortion. In telephoto lenses, the opposite effect occurs, and the mid-section appears smaller, causing pin cushion distortion. In any picture, if the face in question is located on the edge of the photograph rather than the middle, pin and/or barrel distortion can occur (Scully and Nambiar, 2002).

However, the most important factor is subject-to-camera distance (SCD). Stephan (2015) criticises that most CFS validation studies fail to mention the respective SCDs, therefore not accounting for potential distortions and resulting effects. He further raises the issue of lack of agreement in terms of a specific distance at which perspective distortion is definitely affecting CFS/FS. Perspective distortion is subdivided into compression distortion (i.e., long-lens or telephoto distortion) and extension distortion (= wide-angle distortion). Selfies are an example of a SCD that is too short, hence causing warping that makes the middle of the image appear larger than it actually is, whereas features on the outside appear smaller or are not visible at all. Although SCD should always be considered in any CFS/FS scenario, if reproducibility of the original photographic conditions is not possible, what remains is a trial and error approach (Stephan, 2015) whilst keeping the distortion issues in mind. If the SCD is < 1m or 40" (i.e., selfie), then such an image should ideally be compared to another picture taken under similar conditions. Several published guidelines (FISWG, 2018b; Ward *et al.*, 2018) recommend SCDs between 1.2 to 3.5 metres (4'-8' respectively), and at eye level.

In cases of distorted lower resolution images, there are possibilities to improve quality using computer algorithms (Ibáñez, Cavalli, *et al.*, 2015). This needs further development and validation, and might not be available to all practitioners, depending on software and background training. In any case, fundamental knowledge of the equipment (hard- and software), basic photographic principles, as well as possible limitations and obstacles is indispensable. For further details on, and examples of image factors regarding facial image comparisons, please consult the following guidelines: ENFSI (2018) and FISWG (2019).

3.2.1.4 3D-2D Data Comparison Preferred Over 2D-2D

In reference to taking a photograph of the skull or face, aiming to replicate the exact face and head position in the AM comparison image, Al-Amad et al. (2006) claimed that reproducing the same camera angle is feasible through a trial-and-error approach. Birngruber et al. (2010) on the other hand stated that taking the same photograph twice is not possible. Others have also commented on the difficulties of replicating photographic conditions and face angles, whilst simultaneously emphasising the importance of achieving this to enable a valid comparison (Kleinberg *et al.*, 2007; Cattaneo *et al.*, 2012). A lot of significant information is lost in the process of converting a 3D object into a 2D image, which can make facial analysis of any kind rather difficult (Stephan *et al.*, 2008; Cummaudo *et al.*, 2013). Minor discrepancies in alignment and orientation are known to potentially cause serious misjudgements (Cattaneo and Gibelli, 2014). Moreover, it is very time consuming to try and replicate the angle of the head and face in 2D, potentially needing hundreds of photographs in an attempt to find the perfect PM match for the AM photograph. This can be overcome by using 3D models of the skull or face instead (Thomas, 1998). Nickerson *et al.*, (1991) were the first to suggest the use of a 3D model in CFS, as it provides the expert with a more detailed depiction of the skull or face (Damas *et al.*, 2011). The 2D-3D approach has since been stated to be more accurate and reliable for facial identification cases (Cattaneo *et al.*, 2012), not just for CFS/FS but also for facial recognition.

Whereas this possibility seemed somewhat out of reach due to lack of suitable and affordable equipment some 10-20 years ago, laser surface scanning devices have undergone rapid and rather impressive developments since then and are now readily available and more reasonably priced. Another 3D source method is photogrammetry, which can be employed by merely using a mobile phone camera and Freeware (e.g., Agisoft Metashape) or open-source software (e.g., Meshroom) to convert 2D images into 3D models. A recent study by Gibelli *et al.* (2017) has explored the possibility of 3D-3D comparisons for superimposition. However, the data (stereo-photogrammetry) was captured under ideal conditions and the small sample size resulted in the inability to quantify this approach. It is also unlikely that this type of data would be readily available in real life identification scenarios.

The advantages of 3D over 2D data are evident. It allows for the model to be freely positioned to replicate a 2D face angle, and to enable a more accurate alignment of the comparative data (Lee *et al.*, 2019). If landmarks are used in the CFS/FS, their location is easier and more reliable on a 3D object (Caple and Stephan, 2016; Lee *et al.*, 2019), The overall process of using 3D-2D comparisons - also in the identification of the living from surveillance footage - is said to be more reliable and better suited than a 2D-2D analysis, even when the head or face angle in the AM/comparison image is less than ideal (Cattaneo and Gibelli, 2014).

Nevertheless, there are some limitations and aspects that need to be taken into consideration here. An inaccurate 3D model acquisition has definite negative potential to render the superimposition erroneous. Misalignments and inaccuracies can occur in any or all stages of data capture and post-processing. O. Ibáñez *et al.*(2016) noted in their validation study that not all software available to and employed by experts is capable of displaying the model with texture information, in addition some software in use does not correctly display and allow for interacting with both 2D and 3D images simultaneously. When CT or MRI data is used to capture and create the 3D model for CFS/FS, facial texture is not available. This is considered an issue, as texture provides valuable clues and information that can aid in the analysis. The effect of textured vs. non-textured face scans on identification outcomes have been tested in a pilot study and can be found under Appendix A – *Pilot Study*. Another potential shortfall of CT or CBCT (Cone-beam Computed Tomography) data is that it is often an incomplete model of the skull/head, which is deemed to prevent a reliable conclusion on match compatibility (Ibáñez, Cavalli, *et al.*, 2015).

3.2.1.5 Exclusion Only or Positive ID

Traditionally, it has been argued by most that CFS should be used for exclusion of individuals rather than as a stand-alone method for positive identification (Yoshino *et al.*, 1995; Jayaprakash *et al.*, 2001; Oxlee, 2007; Birngruber *et al.*, 2010; Gordon and Steyn, 2012; Sauer *et al.*, 2012; Gaudio *et al.*, 2016). However, some argue that it is possible to establish a positive identification through CFS alone, potentially useful in cases where other means of identification cannot be applied (i.e., lack of AM DNA/fingerprint/dental data for comparison and/or no access to required equipment and resources for such analyses). CFS has shown to be sufficiently accurate and reliable as an identification method when applying and following the most recent recommended best practice guidelines as much as possible (Ibáñez, Vicente, *et al.*, 2016) and using more than one AM photograph for comparison (Austin-Smith and Maples, 1994; Taylor and Brown, 1998; Wilkinson and Lofthouse, 2015). In forensic casework, there can be an element of identification through *failure to exclude*. One such case was described by Fenton, Heard and Sauer (2008), who noted that in closed disasters (i.e., where a skull is known to belong to one of two individuals in question) the identification of one individual leads to the "*circumstantial identification*" (p. 40) of the other.

Nevertheless, it has not yet been specified how many, and which features and details, need to align or match in CFS/FS to imply a positive or likely identification (Ubelaker *et al.*, 2019). Nor "*how many areas of non-conformity are necessary to make an exclusion decision*" (Glassman, 2001; p. 496), bearing in mind any differences caused by distortion or image quality, non-permanent features etc. For now it is still for the expert to decide in the final step whether the points of comparison provide a reliable basis for potential identification or certain inconsistencies warrant an exclusion (Sauer *et al.*, 2012).

Future research will undoubtedly aim to address this issue, and eventually try to further quantify if not possibly automate the last stage of CFS/FS. In the meantime, it is advisable to include or keep an individual in the face pool for further analysis, should some uncertainties linger, rather than exclude somebody prematurely with potentially devastating repercussions. Both MEPROCS and FISWIG guidelines recommend the use of craniofacial or facial superimposition only in conjunction with other techniques (i.e. morphological comparison), but not as sole method for establishing positive identification (FISWG, 2012; Damas *et al.*, 2015).

3.2.1.6 Accuracy, Reliability and Validity

Although used in several court cases in the past (Glaister and Brash, 1937; Helmer, 1987), very few publications and case studies address the potential difficulty of presenting CFS evidence and imagery to lay people (e.g., jury) for decision making (Mallett and Evison, 2013; Moreton, 2021b). It has been shown that the general public struggle to recognise a chimeric image (i.e., an image composed of left and right half faces from two different individuals), as most assume it shows the same individual (Strathie, McNeill and White, 2011). This demonstrates a biasing effect, that has also found to be present, when using video wipes (Strathie and McNeill, 2016).

With their ultimate accuracy, validity, reliability, and quantifiability still somewhat in question (Gordon and Steyn, 2012; Ibáñez, Vicente, *et al.*, 2015, 2016; Strathie and McNeill, 2016; FISWG, 2019a), CFS and FS remain controversial and much discussed topics in the field of facial identification (Ubelaker *et al.*, 2019). To date, many have criticised the lack of a uniform approach and absence of common guidelines (O. Ibáñez *et al.*, 2012; Campomanes-Álvarez, Ibáñez, *et al.*, 2014; Huete *et al.*, 2015), which has so far lead to every expert applying their own approach based on individual professional training and experience, but also dependent upon resources and equipment available (Ibáñez, Vicente, *et al.*, 2016).

Although several authors acknowledge that accuracy, speed, and reliability continue to improve through validation studies and progressive advances in technology (Aulsebrook *et al.*, 1995; Damas *et al.*, 2011; Gordon and Steyn, 2016), the MEPROCS project's interdisciplinary efforts were focused on establishing best practice guidelines that should be followed in any CFS scenario (Damas *et al.*, 2015; Ibáñez, Vicente, *et al.*, 2015). These guidelines are based on CFS rather than FS, and a validation study involving 12 practitioners with varying levels of training and experience found that not all recommendations could be implemented, due to software and/or hardware restrictions, data availability, limitations due to study design and/or practitioner skills, issues resulting from misunderstanding or simply reluctance to follow some of the recommended guidelines (Damas, Cordón and Ibáñez, 2020; p.168). However, the study did show a definitive improvement of the practitioner's performances linear to the level of fulfilment of the proposed guidelines, suggesting that - almost irrespective of experience - the more an expert follows the guidelines, the more reliable the outcome and subsequent identification decision can be expected to be (Ibáñez, Vicente, *et al.*, 2016).

3.2.1.7 Time Requirements

A manual approach to CFS/FS can be rather time consuming, and when confronted with the task of analysing a large number of comparisons, even more so. The required time frame is also dependent on the chosen approach (photo, video, computer-aided, non-automatic vs. automatic elements, etc), but the general guideline for a manual comparison would be approximately 24 hours for one comparison (Lee *et al.*, 2011), depending on hard- and software, experience and the size of the team. Especially in scenarios with large face-pools of potential suspects or victims, being able to automate some or most of the steps required in CFS/FS is not only tempting but seeing a fast-increasing need, due to the growing demand of facial identifications of the dead as well as the living. As an example of technological advancements, Ibáñez *et al.* (2009) achieved a fully automated algorithm-based SFO in less than 18 seconds. Only nine years later, the required time has been reduced to under 0.2 seconds by (Valsecchi *et al.*, 2018). Those studies are primarily conducted under controlled conditions, with ideal data and they are not yet transferrable to real life forensic case work. In order to thoroughly validate these concepts, more realistic - ideally forensic - data is required for extensive validation studies. This does not take away from the incredible developments, which are not only focused on acceleration, but also support the objective to further standardise CFS and FS procedures, to make them more quantifiable and reliable (Ubelaker *et al.*, 2019).

3.2.2 Current and Possible Further Applications of CFS

Since not all practitioners are able or willing to publish their casework studies, a group of researchers and experts have conducted a large scale survey with the aim to establish most common applications of CFS in the field (Damas *et al.*, 2020). However, the response rate to the questionnaire was rather low: Approximately 600 practitioners were approached and only 97 replies received. Out of those 97 only 55 expressed they use CFS on a regular basis and reported a total of 3854 cases, 1246 of which resulted in positive identifications. In the majority of processed CFS cases, there was a presumed identity, limiting the number of required comparisons considerably in comparison to larger face-pool based identification scenarios. From most to least common scenario the respondents reported and specified, the nature of investigation was related to missing persons, mass casualties, mass graves, and terrorist attacks.

Examples of published case studies for those and other possible CFS/FS scenarios listed below (non-exhaustive list):

- Missing Persons (Dorion, 1983; Mckenna, Jablonski and Fearnhead, 1984; Fenton *et al.*, 2008; Santamaría *et al.*, 2009; Jayaprakash *et al.*, 2010; Sauer *et al.*, 2012)
- Disaster Victim Identification (Al-Amad *et al.*, 2006)
- Counter-Terrorism (Indriati, 2009)
- Unidentified human remains - other (Glaister and Brash, 1937; Webster, 1955; Sekharan, 1971; Dorion, 1983; W.R., 1985; Helmer, 1987; Whittaker *et al.*, 1998; Bilge *et al.*, 2003; Ishii *et al.*, 2011; Khudomoma, 2017)
- Surveillance footage identification (Vanezis and Brierley, 1996; De Angelis *et al.*, 2009; Lynnerup *et al.*, 2009; Gibelli *et al.*, 2017)
- Plastic, reconstructive and/or maxillofacial surgery (Sforza *et al.*, 2018)

The two MEPROCS validation studies (Ibáñez, Valsecchi, *et al.*, 2016; Ibáñez, Vicente, *et al.*, 2016) found that, although recommended best practice was understood, the application of all requirements is not always possible, mostly but not exclusively due to limitations of available hard- and software. Further studies are needed to quantify and corroborate, as well as further develop CFS and FS (Ibáñez, Valsecchi, *et al.*, 2016). Superimposition is still applied in casework on a regular basis and research interest is very strong at this time (Ubelaker *et al.*, 2019; Panacea Cooperative Research, 2021). Most research to date has focused on skull to face 3D-2D superimposition, and best practice guidelines and recommendations also fail to include face-to-face approaches.

4 Ante-Mortem Changes to the Skull and Face

4.1 Facial Ageing

This section will exclusively focus on facial ageing in the adult face and will not discuss the developmental growth and age-related changes in infants, children, and adolescents. For further reading on the latter, please refer to Hans and Enlow (2008) and Mullins (2012). At first glance, faces of young and old adults seem to be very different, but when studied at length, both can display similarities as well as differences. The older or ageing face is characterised by deep lines, wrinkles, and folds, a thinner, paper-like skin with numerous discolourations and marks. A younger face may show first signs of lines, but those are mostly dynamic and not static, the younger skin therefore appears smoother and fuller (Cotofana, 2018). The aged face often seems thinner, devoid of most of its volume, slack, and as if having lost its retention towards gravitational forces that impacted the individual throughout a lifetime. This chapter aims to provide an overview of the literature from the past 20 years and the current state of science on facial ageing, its impact on the various tissue types, and manifestations of facial ageing in different parts of the head and neck, to illustrate how changes from young to old occur and why.

Facial ageing is caused by an interplay of several complex factors and processes that manifest themselves in the clinical signs of ageing that are visible on the surface of the head and neck (Guisantes, 2019; Venkatesh, Maymone and Vashi, 2019). Contrary to what had been assumed in the past, it is not only gravitational forces and an age-related loss of elasticity in the skin that leads to uneven skin, pigmentation changes, lines, wrinkles, and skin sagging (Ricanek Jr. and Tesafaye, 2006). Currently, the most accepted theory attributes facial ageing to being the result of an interplay between a multitude of factors, such as continuous skeletal remodelling and resorption, displacement and redistribution of subcutaneous fat and other tissues, a decrease in tissue elasticity, hormonal imbalances, as well as environmental factors (Guisantes, 2019; Venkatesh *et al.*, 2019; Kyllonen and Monson, 2020). However, the inferior displacement of facial soft tissues over time suggests at least some degree of gravitational impact. Ageing is severely influenced by both intrinsic and extrinsic factors. Most of the factors in either category are independent from one another. The majority of the research conducted to date has focused on extrinsic ageing factors, as those are easier to measure and quantify and have been shown to be applicable to, and comparable between, a variety of populations (Kyllonen and Monson, 2020).

The majority of the published literature points out that a direct correlation between an individual's chronological age and the perceived age from facial features and/or ageing skin is extremely unreliable and therefore not recommended as an age assessment tool (Taister and Holliday, 2000; Albert, Sethuram and Ricanek, 2011; Kaur, Garg and Singla, 2015). What is clearly evident from reviewing the related publications is that facial ageing research is very much an interdisciplinary effort (Albert, Ricanek Jr. and Patterson, 2007; Panis *et al.*, 2016; Al-Meyah, Marshall and Rosin, 2017; Flament, Amar and Bazin, 2018; Jdid *et al.*, 2018). Ageing affects all tissue types and layers of the face and most ageing signs manifest in mid-adulthood, showing a progression, deepening, and/or increased pronunciation with age (Avelar *et al.*, 2017; Cotofana, 2018). The onset, rate, and degree of age related changes is extremely variable across males and females, different ancestry or ethnic groups and even individuals within those groups (Schmidlin *et al.*, 2018).

4.1.1 Intrinsic Factors of Facial Ageing

Intrinsic factors of skin ageing are linked to individual genetic makeup, hormone levels, anatomical variations, ancestry, and simply the passage of time (De Rigal *et al.*, 2010; Mullins, 2012; Venkatesh *et al.*, 2019; Kyllonen and Monson, 2020). Those innate changes entail a "*progressive functional decline*" (Vashi, Maymone and Kundu, 2016; p.32) of both the hard and soft tissues of the human body. Intrinsic ageing is unpreventable and inevitable, the pattern and speed of change is largely genetic, and presents through "*cellular senescence, decreased proliferative capacity, decreased cellular DNA repair capacity, oxidative stress, and gene mutations*" (Ichibori *et al.*, 2014; p. 160). On a cellular level, intrinsic ageing is programmed into the genetic makeup of an individual by what is known as replicative senescence or the Hayflick limit (Hadi, 2014). In 1961, Leonard Hayflick and Paul Moorhead discovered that the number of cell divisions is limited and will eventually stop (Hayflick and Moorhead, 1961). The length of DNA repeats at the end of each chromosome (aka. telomeres) are reduced after mitosis (cell division) and will eventually become too short to allow for further cell divisions, slowing down and eventually bringing cell division to a halt and further inducing potential DNA damage or cell death (Freeberg, Kallenbach and Awad, 2019).

It has been suggested by several researchers over the past decade to turn to progeroid syndromes (Greek: *pro* = before, premature; *gēras* = old) and related triggers and genetic mutations for a better understanding of intrinsic factors affecting facial ageing (Makrantonaki, Bekou and Zouboulis, 2012; Makrantonaki, Pfeifer and Zouboulis, 2016; Nikolakis *et al.*, 2016). The general pattern of this disorder is characterised by premature aging, which manifests in accelerated skin ageing, alopecia, skin atrophy, sclerotic skin changes, loss of subcutaneous fat, and musculoskeletal degeneration (Coppedè, 2018). Causes for these rare genetic disorders are attributed to mutations in DNA repair genes and particularly the Lamin A gene, which is responsible for holding the nucleus of a cell together by producing a protein. In this mutation, the protein is defective and causes for the cells to become unstable (Vidak and Foisner, 2016; Ahmed *et al.*, 2017). How exactly this accelerated ageing process in progeroid syndromes translates to normal human ageing remains yet to be established through further research.

In the literature, hyper-dynamic or habitually repeated facial expressions are inconsistently assigned to either intrinsic or extrinsic factors, with some overlap (Ricanek Jr. and Tesafaye, 2006; Gowland and Thompson, 2013a). Although genetics might contribute to a more active mimicry overall, facial expressions are often a result of, or response to, environmental factors, such as stress (e.g. transverse frontal lines/horizontal forehead creases; vertical glabella lines), sleep deprivation, lighting conditions, or habitual facial movements e.g., when smoking (cirumoral striae), all of which contributing to an accelerated manifestation of premature clinical ageing signs (Taister and Holliday, 2000; Ilkankovan, 2014).

4.1.2 Extrinsic Factors of Facial Ageing

Extrinsic factors contributing to or inducing advanced facial ageing are mostly environmental and can be influenced and controlled to a much larger extent than intrinsic factors. The former accumulate over time, are mostly irreversible and can cause severe pre-mature ageing of facial features (Farazdaghi and Nait-Ali, 2017). Extrinsic ageing predominantly affects the head and neck area, as those are commonly exposed to the environment (Avelar *et al.*, 2017). To determine differences between intrinsic ageing and the impact of extrinsic ageing in an individual, the facial skin is often compared to that on the inside of the upper arm or upper thigh region (Makrantonaki *et al.*, 2012, 2016; Trojahn *et al.*, 2015).

Current research on facial ageing is focused predominantly on the effects of extrinsic ageing, as they are "*easier to quantify, measure, and compare across different populations*" (Kyllonen and Monson, 2020; p. 2). Examples of extrinsic ageing factors are smoking, sun/UV-radiation exposure, diet, lifestyle and living conditions, exposure to elements, drug and/or alcohol abuse, air pollution, and trauma, accidents or injuries (Ricanek Jr. and Tesafaye, 2006; Martires *et al.*, 2009; Ichibori *et al.*, 2014; Schmidlin *et al.*, 2018; Venkatesh *et al.*, 2019). Those factors act independent of one another and cause progressive, accumulative deterioration of the soft and hard tissues on a molecular and cellular level (Rowe and Guyuron, 2010). In relation to diet and lifestyle, Kyllonen and Monsoon (2020) found, that individuals with a higher body mass index (BMI) have an older facial appearance prior to the age of 55, but often appear younger than their chronological age after that. Not much is known about how weight loss or gain around the head and neck region affect ageing or the clinical signs thereof (Albert *et al.*, 2011; Kaur *et al.*, 2015). General exposure to elements (e.g. wind, arid air) will contribute to a change in skin colour and texture, with skin becoming tighter and either leather-like or loose, with yellow and freckled skin that might display enlarged superficial blood vessels (Albert *et al.*, 2007).

Photoageing, as a result of repeated and prolonged exposure to sunlight and therefore ultraviolet (UV-A and -B) radiation, is currently considered to be the most severe extrinsic component of facial ageing, as illustrated by Figure 2 (Ricanek Jr. and Tesafaye, 2006; La Padula *et al.*, 2019; Kyllonen and Monson, 2020). In this context, lighter skin is more susceptible to sun damage and hence more affected by photoageing, with some research indicating an impact of up to 80-90% towards facial skin ageing for Caucasians (Farage *et al.*, 2008). Facial ageing in non-white populations appears to be more driven and affected by intrinsic factors, as their higher collagen and melanin percentage, as well as a thicker dermal layer makes them less prone to UV radiation damage (Venkatesh *et al.*, 2019).

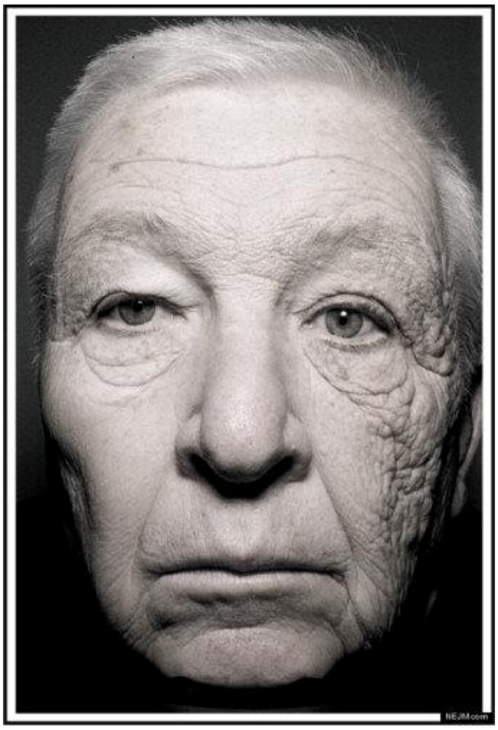

Figure 2: William McElligott, a 65-year-old truck driver, presenting with unilateral dermatoheliosis (aka. Favre Racouchot syndrome) due to one-sided exposure to sunlight whilst on the road, over a period of 25+ years (Gordon and Brieva, 2012).

Martires *et al.* (2009) and Ichibori *et al.* (2014) were able to draw valuable conclusions from two independent twin studies, investigating intrinsic and extrinsic factors related to facial ageing. Their results further corroborated the general assumption that there are significant and quantifiable differences in facial ageing, not only between population groups and between individuals within those groups, but also between very close relatives with almost the same genetic makeup. The differences in severity or rate of progression of facial ageing signs between monozygotic twins are entirely attributed to extrinsic ageing factors. Correlations were found to be significant between the use of sunscreen and/or smoking and wrinkle scores (Ichibori *et al.*, 2014).

Smoking is considered the second most significant extrinsic factor for premature skin ageing (Farage *et al.*, 2008). Nicotine is known to have a vasoconstrictive effect, which in facial skin over time translates to a diminished capillary blood supply, causing oxygen and nutrient deficiency, which contributes considerably to wrinkle formation, due to loss of skin elasticity and hardening of the dermis (Farage *et al.*, 2008). Long-term, heavy smokers may also present with a gray-ish complexion and fine line formation around the mouth (circumoral striae), which are the result of repetitive constriction of the orbicularis oris muscle (Albert *et al.*, 2007). Furthermore, Martires *et al.* (2009) state that smoking tobacco correlates with increased photo-damage to the skin. There is a lack of research on how other drugs influence facial changes, and other than smoking, the types of drugs that could potentially have or are having an impact are not specified in the literature (Albert *et al.*, 2007). For side-by-side mug shot images of individuals affected by substance abuse and the rather severe effects thereof on facial appearance, please visit the following website: https://www.rehabs.com/explore/faces-of-addiction/ [accessed: 16th June 2019].

4.1.3 Ageing of the Hard and Soft Tissues in the Head and Neck

4.1.3.1 Skull

The skull and hard tissue structures of the head and neck provide the scaffolding and framework for soft tissues in this area. Facial soft tissue structures are connected to the viscerocranium via adhesions and ligaments, which provide stability and support (Cotofana, 2018). Therefore, any changes that occur in the bony structure will translate to alterations of the overlying soft tissues (Rhine, 1990; Kyllonen and Monson, 2020). With fairly recent availability and increasing affordability of CT scanning, it has become possible to analyse changes within the skull in more detail.

Not only has it been found that the skull continues to grow and change throughout adulthood, but bony resorption and remodelling have also been recognised as an important contributing factor for clinical signs of ageing (Pessa, 2000; Albert *et al.*, 2011; Kaur *et al.*, 2015). Bone is progressively changing, with more remodelling and growth in younger individuals and an increase in resorption and atrophy with age (Avelar *et al.*, 2017). Primary sites for age-related bone resorption are the inferior-lateral and superior-medial rims of the orbits, the pyriform region, the maxilla, and mandible, as shown in Figure 3 below (Mendelson and Wong, 2012; Guisantes, 2019).

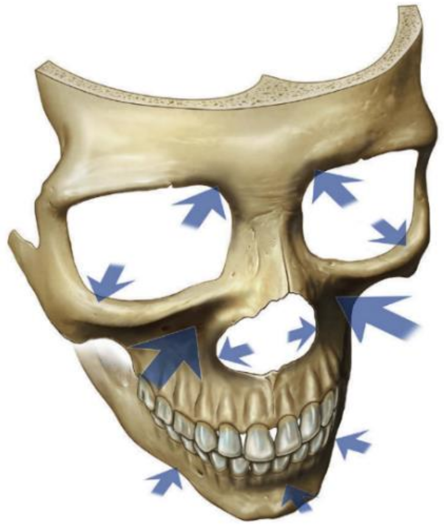

Figure 3: Arrows marking the main sites of skeletal resorption and remodeling as a result of ageing, with direct correlation between arrow size and amount of change; (right) darker discolourations on the skull indicating areas of major skeletal tissue loss with the effects on facial soft tissues shown alongside it (Mendelson and Wong, 2012).

Throughout adulthood and into senescence, an increase in cranial circumference and width, bi-zygomatic breath, as well as anterior and posterior skull height has been shown to occur (Albert *et al.*, 2011). The skeletal profile appears to flatten, due to the above mentioned alterations through resorption (Taister and Holliday, 2000) and a rotational change in the mid-face region, the latter also known as Lambros's theory. Described by Pessa (2000) and verified within the same study, this model proposes a clockwise rotational change within the mid-face relative to the cranial base from a lateral view, caused by an age-related decrease of the pyriform, maxillary and glabella/fronto-nasal angles due to bone remodelling and resorption (see Figure 4). This leads to the appearance of a less protruded maxillary profile, which ultimately causes a sliding effect of the overlaying soft tissue (Avelar *et al.*, 2017).

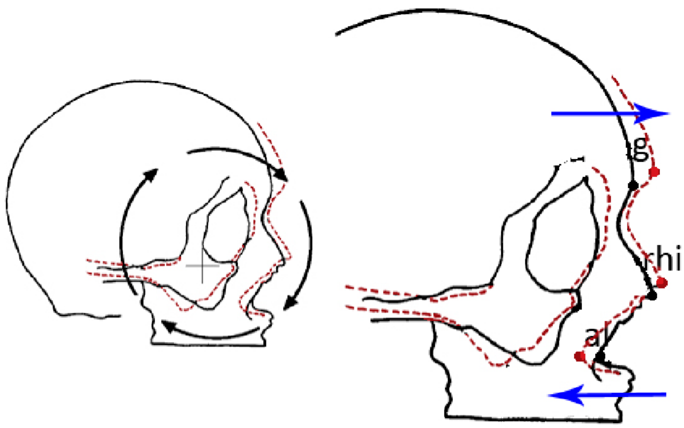

Figure 4: Lambro's theory of skull ageing as a clockwise rotation within the mid-face region, with red dotted line indicating age-related shape changes. Arrows marking the shift in skull features, with glabella (g) and rhinion (rhi) protruding anteriorly, and alae (al) shifting towards superior (Hadi, 2014).

Tooth loss in the maxilla and/or mandible will lead to additional and sometimes accelerated resorption, and in late senescence can cause a decrease in vertical facial height dimensions (Albert *et al.*, 2007). Further retrusion of the mid-face region in edentulous individuals has also been shown (Mendelson and Wong, 2012) and a decrease in alveolar height due to resorption in the mandible enhances the labio-mental fold and may lead to what is commonly referred to as ptosis of the chin or witches chin (Ilkankovan, 2014). Most of the skeletal changes within the skull are attributed to either a lack of or increased functional demand (Avelar *et al.*, 2017). The older the individual, the less stress is applied on the bony tissues through muscle and ligament pull, as facial expressions and functions tend to decrease (Mendelson and Wong, 2012; Guisantes, 2019). Individuals that have a more prominent bone structure may display a later onset of age-related changes within the face, as the resorption of the supporting bony elements will not be immediately visible on the exterior. The opposite applies to individuals with a less distinctive bone structure, who may already display premature facial aging signs in their 20s (Mendelson and Wong, 2012). The often typical, somewhat skeletal appearance of the elderly is attributed not only to skeletal resorption and tooth loss, but also to a loss of volume in adipose and muscular tissue due to ageing (Cotofana, 2018).

4.1.3.2 Muscles

Effects of ageing on the muscles of the head and neck are primarily causing muscle fibre atrophy (Albert *et al.*, 2011) and thus a decline in muscle strength and performance. The effects of the facial muscles on the clinical signs of ageing within the face, however, are caused by long-term, repetitive contractions and constrictions (Le Louarn, 2009). Those contractions cause the formation of wrinkles and lines that are related to facial expression, e.g. horizontal forehead creases, crow's feet, or vertical glabella lines (Guisantes, 2019). The most affected muscles of facial expression (aka. mimetic muscles) in this context are the mentalis, depressor anguli-oris, orbicularis oris, depressor septi nasi, nasalis, orbicularis oculi, levator labii superioris alaeque nasi, procerus, frontalis and the corrugator muscles (Ilkankovan, 2014). Adipose tissue compartments are located both underneath and above the facial muscles. In the younger face, the muscle movement occurs over a more convex and padded structure. With increasing age, there is a shift in adipose tissue (further discussed in the next section below), which causes the muscle to eventually become shorter and straighter due to flatter surrounding soft tissue structures (Le Louarn, 2009; Kaur *et al.*, 2015; Guisantes, 2019). Moreover, the muscles and their structural insertion into the skin divide certain adipose compartments, and it is at those junctions that prominent lines and folds will form (e.g. nasolabial crease) (Cotofana, 2018). Other lines and wrinkles are likely to form "*perpendicular to the stretch of a muscle*" (Taylor, 2001; p.51).

4.1.3.3 Adipose tissue

Adipose tissue compartments are positioned above and below the superficial musculo-aponeurotic system (SMAS). The SMAS is a superficial facial fascia complex, separating the sub-cuntaneous adipose tissue from deeper fat compartments (Cotofana, 2018). The latter lie between the SMAS and periosteum. This distribution allows for the muscles to move and slide relatively freely for facial expression or mastication (Guisantes, 2019). In younger faces, fat is more evenly distributed, and hence attributing to a fuller, smoother appearance. During the facial ageing process, the superficial adipose tissue will slowly descent, whilst the deeper compartments remain relatively steady in relation to the hard tissue structures. However, as the skull also undergoes changes, those will translate to shifts within the deeper adipose tissues as well (Kaur *et al.*, 2015; Guisantes, 2019). A descent and drooping of adipose tissue around the jaw line will eventually contribute to the formation of jowls (Albert *et al.*, 2011). Factoring out any weight gain or loss, adipose tissue within the face will continue to lose volume and atrophy over time, causing a more emaciated and hollow appearance on the facial surface, particularly in the cheek and temple regions of the elderly (Vashi *et al.*, 2016). The currently favoured volumetric theory also entails the above-mentioned changes, taking into account that some fat compartments or pads will decrease or shift sooner than others. With the loss of volume in one region, lines and eventually folds will appear at the intersection to the neighbouring facial part (Guisantes, 2019).

4.1.3.4 Ligaments

As age-related changes of the skull are now better understood, a relationship between bone resorption and remodelling and its effect on ligament and appendage tension and/or laxity and positional changes thereof is widely acknowledged (Kaur *et al.*, 2015; Cotofana, 2018; Guisantes, 2019; Okuda, Yoshioka and Akita, 2019). Major ligaments show less age related changes, whereas finer ligaments (e.g. orbital retaining ligament, mandibular retaining ligament, zygomatic ligament) are more affected by repetitive movements, causing a weakening and eventual bend or increased laxity (Guisantes, 2019). Okuda, Yoshioka and Akita (2019) also found a correlation between SMAS laxity increase and ageing, meaning the SMAS is less able to support other facial soft tissue structures over time. Appendages tend to become thinner (Albert *et al.*, 2011) and the progressive weakening of parotid and masseteric ligaments results in an inferior shift of the malar and buccal fat pads, promoting the nasolabial fold and jowl formation.

4.1.3.5 Skin

Venkatesh, Maymone and Vashi (2019; p.351) describe the skin as "*the most perceivable indicator of the aging process*". The skin is the largest organ of the body and fulfils a barrier function between the internal structures of our body and external influences (Sharma, Arora and Valiathan, 2014; Marionnet, Tricaud and Bernerd, 2015). Skin thickness varies throughout life, with an increase of thickness from birth to approximately age 15, remaining relatively stable until the mid-60's, at which point a continuous reduction of skin thickness follows, so that skin will eventually become thinner than compared to one's five year old self, if the individual lives to be 90+ years old (Escoffier *et al.*, 1989). According to (Calleja-Agius, Muscat-Baron and Brincat, 2007) age-related changes to the skin can be divided into genetic, chronological, endocrine, gravitational, environmental, behavioural, and catabolic causes. During adulthood and into senescence, the human skin progressively loses elasticity, displays decelerated cell division, disintegration of the collagen matrix within the dermis, and thinning of the epidermal layer (Hadi, 2014; Guisantes, 2019). The latter correlates to wrinkles, in that the deepest part of any line or wrinkle corresponds to the thinnest area of the epidermis (Ilkankovan, 2014). These physiological changes result and manifest in an increasingly uneven skin surface, with the formation of lines and wrinkles, skin sagging, as well as dryness, increased susceptibility to trauma and slower tissue repair abilities.

Intrinsic factors will cause thinning and atrophy of the epidermal layer of the skin, decreased circulation and collagen synthesis, and degradation of elastic fibres. With a decline in melanocyte function comes reduced natural photo-damage protection (Venkatesh et al., 2019). The elasticity of the skin becomes more and more diminished, which contributes to wrinkling (Nikolakis et al., 2016). Through loss of elasticity and facial atrophy and inferior shifts of soft tissue, skin sagging becomes increasingly prominent, with the formation of nasolabial folds, jowls, and sub-mental platysmal bands (Ilkankovan, 2014). The primary extrinsic ageing factor influencing skin ageing is photoageing due to UV radiation. While it does affect all skin layers, the visible changes to the dermis cause further loss of skin elasticity, pigmentation, keratosis, and fine line and wrinkle formation, with overall increased and accelerated manifestation of facial ageing signs (Farage et al., 2008; Cotofana, 2018). Smoking induced changes may also include changes to the pigmentation, decreased radiance, and formation of circumoral striae around the mouth through habitual constriction of the orbicularis oris muscle. Perceived age is mostly linked to texture changes and an overall facial atrophy, due to reduction, shift, weakening, and/or decrease of underlying tissues (Kaur et al., 2015).

Figure 5 below illustrates the main facial lines, creases, and wrinkles within the human face. It would go beyond the scope of this thesis to describe those individually at this stage. For a comprehensive list and description of facial lines, folds, and wrinkles, please refer to the following references: Hadi (2014); Kaur, Garg and Singla (2015); Hadi and Wilkinson (2017); FISWG (2018a).

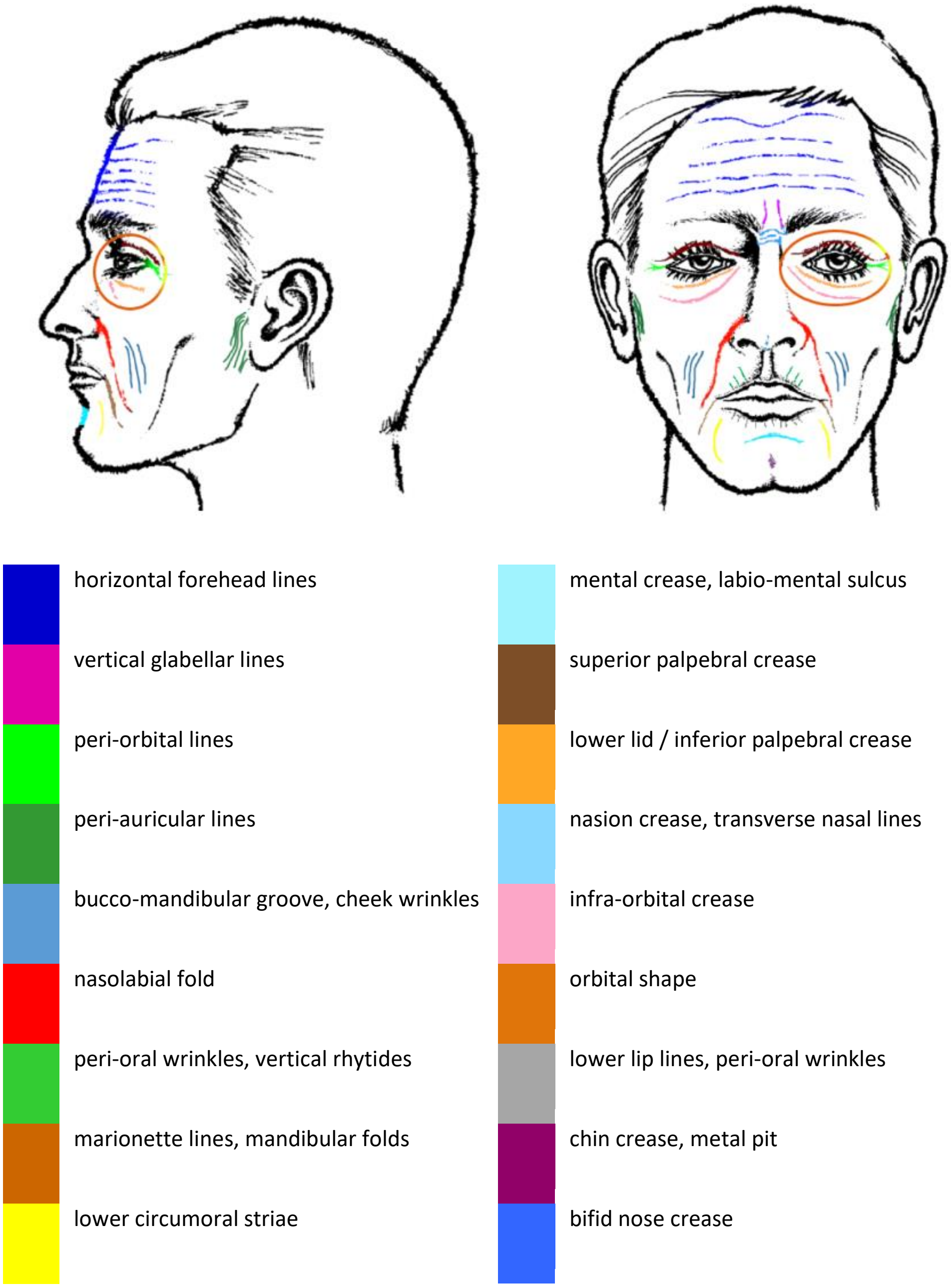

Figure 5: Lateral and frontal view of the face with colour-coded facial lines and wrinkles, modified from Hadi (2014) with changes in terminology following FISWG guidelines (FISWG, 2018a).

4.1.4 Age-Related Changes by Facial Regions

The following section describes the three main facial portions (see Figure 6), each of which are further sub-divided into the relevant facial regions. The aim is to provide a more comprehensive description of age-related changes within the face, underlying structural changes, tissue relationships, and clinical manifestations on the facial surface.

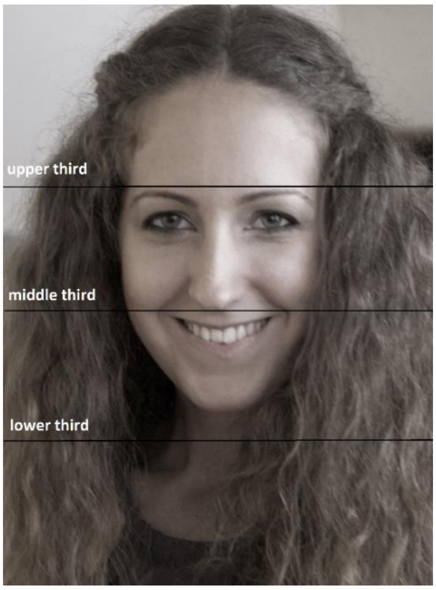

Figure 6: Frontal view image of the researcher divided into upper, middle, and lower facial thirds.

4.1.4.1 Upper Third of the Face

4.1.4.1.1 Hair

Hair growth and ageing is also subject to intrinsic and extrinsic factors. Intrinsic hair ageing is related to genetics and can include familial patterns of pigment loss onset (greying) or baldness. Extrinsic factors, such as stress, malnutrition, smoking or pollution can increase the speed of progression and severity of ageing signs (Ilkankovan, 2014). An interplay of intrinsic and extrinsic factors over time leads to weakening and damage of the hair shafts and follicles (Ilkankovan, 2014). Hair will eventually lose its colour, as melanocytes (pigment cells) are depleting within the hair bulb, causing hair strands to become more transparent and subsequently appear grey, silver, or white over time (Tobin, 2017). As for other hair within the head region, it has been found that in males, eyebrows, nose, and ear hair continue to grow.

Visible changes to hair in relation to an individual's age are greying and loss of hair density, as well as pattern hair loss and/or senescent alopecia. Pattern hair loss or pattern baldness, also referred to as androgenetic alopecia, is caused by a variety of genetic and environmental factors, but the exact causes are not completely understood (Vary Jr., 2015). In men, the clinical signs progress from hair loss at the temples, to receding hairline into an M-shape, hair thinning or loss at the crown of the head, and often resulting in partial or complete baldness (Tobin, 2017). Senescent alopecia of unknown aetiology – commonly characterised by an overall thinning of hair (Trüeb, Rezende and Gavazzoni Dias, 2018) – affects both sexes, with 50% of men being affected by the age of 50 and 50% of women once in their 60s (Whiting, 1998).

Several classification systems of patterned hair loss for both males and females have been proposed over the past 60 years, but most present limitations by either being too detailed to be practical and universally applicable, or by not taking all stages of hair loss and the inter-individual variations thereof into account (Gupta and Mysore, 2016). For males, the most commonly used is the Norwood classification system (Norwood, 1975; see Figure 21), whereas for females, the Sinclair self-reporting photographic measure scale has been shown to be an easy and valuable tool for both patients and clinicians to assess hair loss (Sinclair *et al.*, 2004). For a comprehensive review and more detailed descriptions of the different scales, please see Gupta and Mysore (2016).

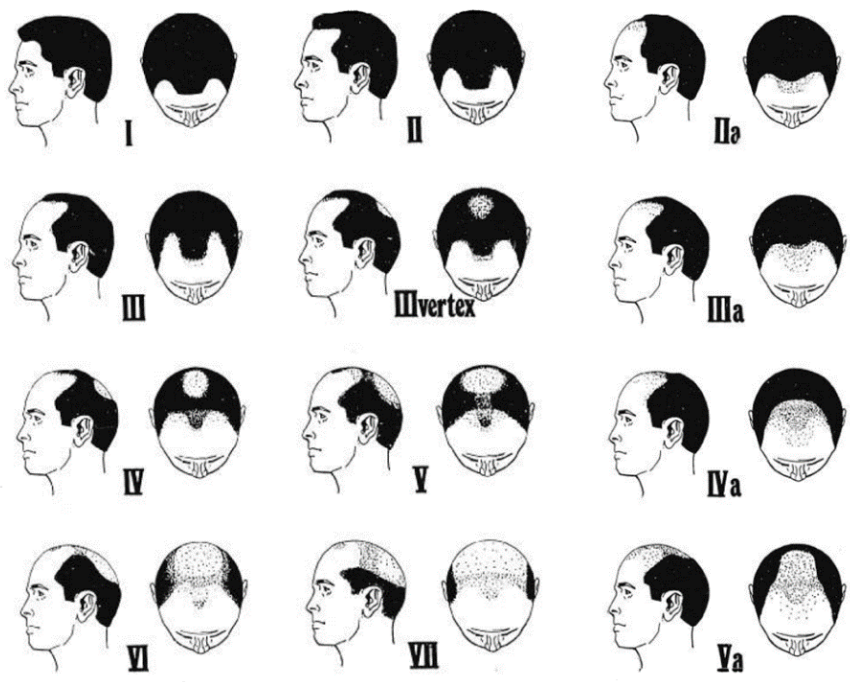

Figure 7: Norwood classification scale for male patterned hair loss (Norwood, 1975; used with permission).

4.1.4.1.2 Forehead, Temples and Glabella

One of the earliest signs of facial ageing are the horizontal forehead creases, a result of repeated contractions of the frontalis muscle and atrophy of subcutaneous adipose tissue. Even though lines may begin to show in the 20s, they remain a dynamic rather than static feature until one or two decades later (Ilkankovan, 2014). There are contradicting statements regarding the glabella angle to be found in the literature. Shaw and Kahn (2007) refer to a decrease in glabella angle with age, whereas Ilkankovan (2014) states an increase of said angle, the latter leading to a downward shift of the eyebrow and formation of lines in the glabella region. Those lines are further emphasised and deepened by repetitive contractions of the medial depressor supercili and corrugator muscles. The adjacent temporal region is affected by loss of subcutaneous fat, slight muscle atrophy, as well as thinning of the epidermis and therefore appears increasingly sunken and more flat or hollow as an individual ages (Kaur *et al.*, 2015; Sandulescu *et al.*, 2019).

4.1.4.1.3 Orbital Region

Orbital aperture size increases with resorption and remodelling occurring at the skeletal inferior-lateral and superior-medial rim respectively (Shaw and Kahn, 2007; Kahn and Shaw Jr, 2008; Sforza, Grandi, Catti, *et al.*, 2009). Sforza *et al.* (2009) reported an increase of the inter-canthal and bi-ocular widths, and orbital height with age, as well as a change in the angle of the orbits relative to the Frankfurt horizontal plane (FHP) of up to 10°. The same study found an increase in the length of eye fissure until the 30s or 40s, followed by a continual decrease into old age. This leads to an anterior shift of the orbital septum and a protrusion of the intra-orbital fat pads, the latter becoming more prominent throughout mid and late adulthood (Mendelson and Wong, 2012). The orbicularis retaining ligament (ORL) drifts from a horizontal orientation into an inferior inclined position, which causes a loss of support for the orbicularis oculi muscle and therefore promotes drooping of the sub-orbicularis oculi fat (SOOF) and the retro-orbicularis fat (ROOF) (Guisantes, 2019). A progressive weakening of the orbital septum leads to emergence of the medial upper lid fat. In combination with loss of elasticity in the skin, repetitive muscle contractions of the orbicularis oculi and surrounding muscles, as well as gravitational pull, this causes lateral hooding of the upper eyelid and eyebrow ptosis (Albert *et al.*, 2011).

On the skin surface, the aforementioned changes to the orbit structure also translate to age-related formation of tear through deformities (lower lid or inferior palpebral crease and infra-orbital crease) between the palpebral and orbital sections of the orbicularis oculi (

Figure *5* above), as well as the naso-jugal groove/fold at the inferior border of the orbicularis oculi (Lee and Hong, 2018). Those in combination with the upper lid and eyebrow drooping, will cause for the eyes to appear more deep set and smaller with age (Taister and Holliday, 2000). A progressive hypertrophy of the muscles in and around the eye and orbit, result in additional wrinkles in this region (Ilkankovan, 2014). It has also been found that the curve of the lower lid tissue changes with age (Hamra, 1995). According to Sharma, Arora and Valiathan (2014), increasingly darker shades below the eyes can be attributed to a thinning of the epidermis and loss of subcutaneous tissue in the infra-orbital region, making the vascular system more visible.

The orbital region and any changes or modifications thereof are of particular interest within the field of biometrics, especially facial recognition (Albert *et al.*, 2011), which is further addressed in section 6 Face Recognition below.

4.1.4.2 Middle Third of the Face

4.1.4.2.1 Maxilla and Cheek

The maxilla is affected by bone resorption and remodelling, as previously mentioned, and is also subject to Lambro's mid-facial rotation as part of the facial ageing process (Shaw and Kahn, 2007). The resulting deficit in hard tissue support structures contributes to the formation of a tear-through deformity (Mendelson and Wong, 2012), which is commonly referred to as bucco-mandibular groove (FISWG, 2018a). Due to the changes in the orbital region and retrusion of the maxilla (Ilkankovan, 2014), the malar fad pad and other soft tissues of the cheek drift towards inferior, leading to the formation of the nasolabial fold/crease (Albert *et al.*, 2011; Cotofana, 2018), a process that Pessa (2000; p.483) describes as "*pathogenesis of the mid-facial soft-tissue mal-position*". In case of tooth loss in the maxilla (and/or mandible), and subsequent alveolar remodeling and resorption, the cheeks can appear hollow or more concave (Pessa, 2000). The latter becoming increasingly prevalent with age, as more tissue atrophy and vertical downwards movement thereof occur (Forsberg *et al.*, 2009).

4.1.4.2.2 Nose

The cartilaginous nose, not unlike the ears, appears to show continuous growth throughout adolescence and senescence, as was found by several studies (Zankl *et al.*, 2002; Sforza, Grandi, De Menezes, *et al.*, 2010; Kaur *et al.*, 2015). Changes in shape in the nose and nasal region are primarily caused by the previously mentioned age-related alterations to the skull, with an increasing lack of hard tissue support, a widening pyriform aperture and decrease of projection within the maxilla (Mendelson and Wong, 2012). The lack of support of the alar base contributes to a development and deepening of the nasolabial fold (Albert *et al.*, 2007). As the hard tissue anterior nasal spine is receding, the soft tissue columella appears to retract (Guisantes, 2019). Sforza, Grandi, De Menezes, *et al.*, (2010) and Kaur, Garg and Singla (2015) both conducted quantitative studies on an Italian and a Punjabi Sikh population respectively. Both found that the nasal tip angle decreased with age as well as finding an increase in nasal tip protrusion and nose length, all of which are in agreement with, and add to previous research by Zankl *et al.* (2002), who relied on a Swiss population.

4.1.4.2.3 Ears

The cartilage within the human ear continues to grow very slightly throughout adulthood and into senescence (Sforza, Grandi, Binelli, *et al.*, 2009; Bhanu and Govindaraju, 2011; Guisantes, 2019). In their cross-sectional study on 843 Italians, Sforza, Grandi, Binelli, *et* *al.* (2009) verified and confirmed findings from previous studies (Ito *et al.*, 2001; Brucker, Patel and Sullivan, 2003; Meijerman, van der Lugt and Maat, 2007), namely the general trend of increased ear width and length with age. The drooping can be explained by increasingly weakened collagen fibres and thinning of the dermis (Wong *et al.*, 2016), a process potentially emphasised by heavy jewellery. The same study also confirmed the commonly accepted observation that – on average – males have larger ears than females. The overall characteristic shape of the ear however does not change significantly with age, and so remains a good morphological feature for facial comparison and other biometric analyses (Rahman *et al.*, 2007; Hurley, Arbab-Zavar and Nixon, 2008; Nixon *et al.*, 2010). Figure 8 below emphasises this point, but also shows one of the primary difficulties of using ears for identification, namely partial or complete obstruction of the feature within an image or other footage.

Figure 8: Comparison of the left ear of Prince Philip, Duke of Edinburgh and Queen Elisabeth II in age-different images, with respective enlargements of the relevant feature (adapted from Deigin, 2019).

4.1.4.3 Lower Third of the Face

4.1.4.3.1 Mouth and Lips

As the mandible becomes less firm, circumoral striae and vertical rhytides develop around the mouth but primarily above the upper lip vermilion border (Kaur *et al.*, 2015). Ageing does also appear to have a significant effect on the mouth and lips, causing overall lip thickness reduction, a decrease in vermilion dimensions, and increased mouth width (Albert *et al.*, 2011). Sforza, Grandi, Binelli, *et al.* (2010) confirmed age-related changes regarding lip thickness reduction, an increase in lip length and philtrum length for both males and females. As the findings were based on an Italian population, it is interesting to note that lip thickness did appear to be less affected in South African males, who on average have thicker lips compared to Caucasians (Schmidlin *et al.*, 2018). Wrinkle formation in the oral region and onset of loss of volume in the cheek were delayed when compared to European and Caucasian data. Another study by Albert, Ricanek Jr. and Patterson (2007) found that lip height increased in the early 20s, with an almost simultaneous reduction in lip thickness, followed by retrusion of both upper and lower lips between the ages of 25 and 45. Furthermore, lips were observed to shift downwards, lengthen and flatten between the ages of 30 and 45. Considering the phenotypic variation with regards to lip morphology and age-related changes, the existing studies are not sufficient as of yet to provide a reliable age assessment tool or timeline of events.

4.1.4.3.2 Chin and Jawline

The jawline becomes increasingly less well defined throughout the ageing process, due to a downward drift of soft tissue. Jowls and a double chin are likely to develop in most individuals in later adolescence, although the chin may be more defined, through the mandibular bony changes enhancing the mento-labial sulcus (Albert *et al.*, 2007). Pecora, Baccetti and McNamara Jr (2008) conducted a longitudinal study analysing cephalometric radiographs and found that the mandibular length increases with age in both males and females. Width of the gonial angle was shown to increase with age (up to 6° for edentulous individuals), but was deemed rather unreliable as no significant pattern could be established between different age groups (Upadhyay *et al.*, 2012). Venkatesh, Maymone and Vashi (2019) state that volume loss at the temples, cheeks, lateral chin cause an overall narrowing of the forehead and widening of the lower face, similar to a V-shaped triangle that eventually inverts with advanced age.

4.1.4.3.3 Neck

Only very little is reported in the literature on age-related changes of the neck, as most studies on facial ageing focus either on the face alone or on a particular facial region. Ilkankovan (2014) states that the exact causes for clinical signs of ageing in the neck area remain unknown. It is assumed that gravitational impact causes ptosis of the skin and loss of volume in the head and neck area leads to loose, sagging skin. The onset of the latter is often found in people in their 60s, but lines and wrinkles on the neck can form much earlier and will become static and deeper over time (Albert *et al.*, 2007). The most dominant sign of ageing in this region are the platysmal bands that form inferiorly on either side of the chin. A rather recent study by Sandulescu *et al.* (2020) describes the platysma as a mimetic muscle with several anatomical variations, that often display the first ageing signs at the neck. The development of platysmal bands is attributed to repetitive contractions and/or skin sagging with the exact aetiology not yet understood. With atrophy of subcutaneous fat and thinning of the skin, the sternocleidoid muscles may become more visible and therefore might appear more pronounced. The thyroid gland does enlarge in some individuals, especially those with a lower BMI (potential hyperthyroidism) and can lead to an increase in neck circumference.

4.1.5 Sexual Dimorphism in Facial Ageing

There are some differences repeatedly mentioned in the literature in relation to sexual dimorphism in facial ageing. Hormonal influences often lead to a later onset of clinical ageing signs in women, but post-menopause those changes will accelerate, whereas the manifestations of facial ageing within the male features are slower but more regular (Mealey, 2000b, 2000a; Albert *et al.*, 2007). However, males are more likely to be affected by hair loss and/or baldness and an earlier onset thereof. The differences in perceived age, as well as associated societal standards and pressures leading to the boom in anti-ageing treatments, is further discussed under section 4.2 Facial Aesthetics and Body Modifications below.

4.1.6 Considerations and Limitations

Although the published literature on facial ageing, skin maturation, and age-related changes to the craniofacial hard tissue structures appears extensive, there is still a lack of standard guidelines and validated uniform methods to assess age from facial appearance. Understanding and being able to quantify facial ageing is also relevant for biometrics in relation to automated facial recognition systems (Kaur *et al.*, 2015). Studies on human ageing are complicated by the fact that most rely on, and are ultimately limited by, small sample sizes, cross-sectional data, and examining different individuals at different ages, rather than being able to draw results from longitudinal observations of the ageing process from the same individual (Lynnerup, 2001; Sforza, Grandi, Catti, *et al.*, 2009). In the literature and for assessing patients for potential treatment, ageing features, their characteristic appearances, and severity are rated or categorised based on:

- a numerical and/or photographic scale (e.g. 0-4;0-6; 0-9) (Jdid *et al.*, 2018)
- a scale of pre-defined terms (e.g. none - minimal - fair - marked - prominent) (Kaur *et al.*, 2015; La Padula *et al.*, 2019)
- or by morphological assessment and comparison (Neave, 1998)

These approaches are all somewhat limited, as they attempt to quantitatively assess qualitative and extremely variable clinical signs of facial ageing (Kyllonen and Monson, 2020). Although the general sequence of ageing is fairly predictable, the exact onset, timings, progression, and severity remain uncertain, due to human variation and the multitude of factors (potentially) influencing the process (Albert *et al.*, 2007). There is a common consensus that the timing of age-related changes varies greatly not only between different population groups, but also "*among different anatomical sites even within a single individual*" (Farage *et al.*, 2008; p.87). When analysing and assessing a face, face photograph or facial model, it is not sufficient to state that the person in question appears to be young, middle aged, or old. The importance lies in a quantifiable approach to define *why* someone appears to be of a certain age. Age estimation from visual cues alone can be quite subjective and dependent on one’s own perspective in daily life, as a teenager would likely describe the face of a middle-aged person as old, whilst a senior citizen might describe the same face as young.

As for scientific purposes, this matter does require a certain tool to assess the visible ageing sign, but most researchers and clinicians develop their own scales or classification system. This means the existing ones are almost never validated by an additional study and there is huge variety and a definite lack of consensus and standardisation with regards to not only scales and categories, but also terminology (Hadi, 2014; Kaur *et al.*, 2015; Guisantes, 2019). As an example, Sandulescu *et al.* (2019; p.) mentions *'horizontal forehead creases*', '*horizontal forehead lines'*, '*forehead lines*', and *'transverse frontal lines'*, all referring to the same feature. The same applies to the terms *'wrinkle'*, *'line'*, *'fold'*, and *'crease'*, which are also used rather interchangeably in the literature (Hadi and Wilkinson, 2017). FISWG (2018a) have included facial lines, wrinkles, and folds for their guidelines on facial comparisons and presented a rather comprehensive list of useful terminology with the aim to provide a more standardised framework.

Terminology issues aside, it is incredibly difficult, if not impossible, to correctly judge an individual's age based on facial appearance alone, and even more so if this has to be achieved based on a photograph rather than the actual person (Kyllonen and Monson, 2020). In the absence of birth certificate and passport (or other legal documents containing the date of birth) and with no additional means and sources of age estimation available, facial ageing can only ever be an estimate of actual chronological age (Aggrawal *et al.*, 2010) and should therefore be presented as an age-range rather than a fixed number. Every publication seems to adapt what suits the particular research or case, and every diagram or figure in textbooks is labelled differently or has different terms and descriptions for the respective feature (Albert *et al.*, 2011; Hadi, 2014; Kaur *et al.*, 2015; FISWG, 2018a).

This inconsistency has been criticised as inevitably leading to miscommunication and misunderstanding, hence a standardised classification is required (Avelar *et al.*, 2017; Lee and Hong, 2018; Sandulescu *et al.*, 2019). The lack of standardised terms and procedures in this context results in studies that are not comparable to one another as well as non-validation of the scales or different approaches, as there is no one method being applied to different populations. Resembling a large puzzle of valuable results, all the pieces of individual research in this area are of a different size and shape, but no congruent picture has yet been produced. In an area like facial ageing, this may remain to present an issue in the future, due to the variety of research interests in the matter, originating from fields such as biology, genetics, dermatology, plastic surgery, aesthetic medicine, computer science, orthodontics, law enforcement, biological anthropology, and forensic science (Zankl *et al.*, 2002; Coleman and Grover, 2006; Ricanek Jr. and Tesafaye, 2006; Martires *et al.*, 2009).

Another limiting factor of the existing classification and grading scales is the fact that most are based on Caucasian populations, in addition, comparative data and evaluations are regularly obtained under ideal and controlled laboratory conditions with perfect illumination, facial expressions are mostly neutral, and all ageing scales are based on living individuals (Jdid *et al.*, 2018; La Padula *et al.*, 2019). While controlled conditions are important to develop such scales for reference, it might be difficult to observe the example features in real life comparisons. In forensic scenarios, the available facial photographs often show non-neutral facial expressions and pictures are mostly taken under non-controlled conditions with variations in photographic settings and influencing factors (e.g., lighting, object-to-camera distance, occlusions).

When comparing or assessing deceased individual's faces, there is currently a large knowledge gap with only a handful of studies addressing whether existing facial ageing scales or similar approaches would be applicable to the lifeless face. The face of a deceased individual may potentially be exhibiting positional/gravitational and other post-mortem changes, which can distort or conceal the described ageing signs. The only research in this direction has been conducted by Hadi and Wilkinson (2014) on the post-mortem resilience of facial features, and by Wilkinson and Tillotson (2005) on the post-mortem prediction of facial features. This matter will be discussed in more detail under section 5.3 Post-Mortem Changes to the Face below.

4.1.7 Facial Ageing in Human Identification and Recognition

A good understanding of the facial ageing process is indispensable for forensic facial identification, approximation, reconstruction, depiction, and image comparison (Schmidlin *et al.*, 2018). The related anatomical and visual manifestations thereof are extremely relevant to this current research. Facial comparisons in this study as well as in many real life scenarios utilising face data for identification or recognition purposes, are often between faces of different ages, if no recent ante-mortem photograph is available (Taister and Holliday, 2000; White *et al.*, 2017). Most studies on facial ageing focus on age-related changes over several years or even decades, but a thorough search of the relevant literature yielded only one result that explores diurnal variations within the face. Those are changes that occur within the course of just one day. Although limited by small sample size (n = 38; Japanese population group) and acknowledging that this topic requires further research, Tsukahara *et al.* (2004) found that noticeable changes occurred even in such a seemingly short time frame. Capturing facial images of the same individuals in the morning and then again in the evening, increased skin elasticity and a reduction of skin thickness were observed later in the day. During the night, whilst in supine position, gravitational forces are not causing a downward pull on facial features. In the morning, some fluid retention may cause the face to look rounder and more filled, and the lack of facial muscle movement may even out some facial lines. Throughout the day, lines and wrinkles deepened and became more visible. If, as was shown in the aforementioned study, changes do occur and are detectable in the course of a day, how much does the face change within a week, a month, or a year? There are no studies as of yet that have analysed these rather short-term facial changes and it will be incredibly difficult to do so in a quantifiable manner. The closest self-directed longitudinal studies visualising such changes, can be found in individuals taking facial images or selfies every day over a long period of time, as are available in abundance on YouTube, e.g. Kalina (2020).

But what are the effects of age-related facial changes on face matching, facial recognition, and identification? In a 1-to-10 face matching experiment, using high quality facial photographs that were obtained on the same day, Bruce *et al.* (1999) observed an error rate of approximately 30%. A two-phase study by Megreya, Sandford and Burton (2013) compared 1-in-10 and pair-wise face matching accuracies for images taken on the same day and in a second step, with images taken on average 17 months apart. The results showed a drop in accuracy of around 20% between same-day and months-apart image comparisons (79% and 90% vs. 58% and 70% respectively). This suggests that there is "*strong evidence, that age-related changes in facial appearance could increase the mistaken identification of unfamiliar faces*" (Megreya, Sandford and Burton, 2013; p. 705). The authors criticise that most studies, and the databases used to conduct them, rely on same-day facial photographs or video footage, which constitutes a discrepancy between laboratory and real-life scenarios, where same-day data is hardly ever used. With identification documents (e.g., passports, ID cards, drivers' licence) often valid for several years, even the most basic face matching at airports or borders occurs between age-different faces. In this regard, it can be assumed, that faces change more significantly over those years, resulting in a correlating decrease of accuracy for face matching, recognition and identification than was found in the 17 months difference by Megreya, Sandford and Burton (2013).

Ageing is one of the components of person-based factors than can influence and impact facial recognition and identification. It has been found that recognition rates for unfamiliar faces are higher and more accurate, the more similar the face appears in compared images (White, R.I. Kemp, Jenkins, Matheson, *et al.*, 2014). As a logical consequence, unfamiliar face matching, recognition or identification becomes increasingly challenging, the more someone's facial appearance is altered, e.g. by weight gain/loss, ageing, or variations in facial expression (White *et al.*, 2017). In other words, as the time span between capture or observation of face pool, ante-mortem and target, or post-mortem image or face increases, recognition becomes more and more challenging, an issue also known as *temporal performance degradation* (Kaur *et al.*, 2015).

4.2 Facial Aesthetics and Body Modifications

Body modifications are defined as "*intentional alteration[s] to [the] human biological phenotype*" (Black and Thompson, 2007; p. 379). They range from temporary or short-term (e.g., make-up, eyebrow shaping, temporary tattoos) to permanent (e.g., tattoos, piercings, earlobe elongation) and extreme modifications (e.g., eyeball-tattoos, tongue splitting, doughnut or other distorting implants, scarring, cutting, branding, neck stretching, head molding) (Featherstone, 1999; Kosut, 2015). Some are associated with culture and/or religious beliefs and rituals (e.g., scarring and tattoos of the Mãori in New Zealand), or with a sub-culture or organisational affiliation (e.g. Punk, Gothic, and Japanese Yakuza) (Roberts, 2015). Such modifications are often viewed as a means to define one's identity or to have one's identity perceived by others in a certain way. Whilst some are aimed at making an individual stand out from the crowd, others are desired and performed to help with blending in (e.g., maxillofacial surgery to correct deformities). The same body modification can fit into either category, depending on the social, situational, or regional context (Gowland and Thompson, 2013a). Especially in Western or developed countries, body modifications are often linked to either negative self-mutilation or positive self-expression (Weiler *et al.*, 2021).

Whilst most of the aforementioned body modifications would classify as voluntary, there are of course also involuntary changes that can occur and affect facial appearance. Accidents (e.g., burns, cuts, lacerations), attacks (e.g., blunt and sharp force trauma, acid), or diseases that leave lasting marks on the skin (e.g., severe acne, skin cancer, chickenpox in adults) can alter facial features to a greater or lesser extent and therefore potentially impact facial identification (Costa *et al.*, 2014; Choi, Giu and Kang, 2016; Burg, 2019). Wanted criminals often seek to change their facial features by means of plastic surgery for this exact reason (Maliniak, 1935; Furlong, 1984; Atlas, 1994; Girardin and Helmer, 1994; Willsher, 2011; De Marsico *et al.*, 2015; Department of Justice, 2016; Remnick, 2016; Encyclopedia Britannica, 2020). This section will focus on facial body modifications only, as post-cranial changes are not relevant to this current research.

Make-up is one of the oldest and most universal forms of bodily adornment (Russell *et al.*, 2019). Today, the cosmetics and beauty industry has a global market value in excess of $532 billion (Danziger, 2019), and several studies have shown that faces are generally rated as more attractive when the individual is wearing make-up (e.g., Cox and Glick, 1986; Cash *et al.*, 1989; Etcoff *et al.*, 2011; Batres *et al.*, 2019; Fardouly and Rapee, 2019; Comfort *et al.*, 2021). There is however very little research on how make-up affects human perception and identification (Ueda and Koyama, 2010; Tagai, Ohtaka and Nittono, 2016). Russell *et al.* (2019) investigated the impact of make-up on perceived age and found that younger women (age ~20) were perceived to look older when wearing make-up, older women (aged 40+) appeared younger and for middle aged women (aged 30-40), there was no difference in perception. For men, several studies have found that facial hair enhances perceptions of masculinity, confidence, age, and social maturity (e.g., Roll and Verinis, 1971; Kenny and Fletcher, 1973; Wogalter and Hosie, 1991; Neave and Shields, 2008). Studies on attractiveness relative to facial hair however have produced mixed results (please refer to Dixson and Vasey, 2012 for a comprehensive review of facial hair perception research). Still very little to no research discusses hairstyle changes (colour, length, shape), dermal fillers, Botox, eyebrow shaping, facial hair (length, style, pattern), and other modifications in relation to facial comparison effects. In maxillofacial surgery, the extraction of upper 3rd molars results in a more narrow facial contour (Taister and Holliday, 2000), but how this affects human or machine based face matching is not understood and a large research gap remains. However, as surgical interventions and aesthetic treatments are on the rise, their impact has to be further investigated and considered in facial identification, comparison, and verification efforts.

4.2.1 Socio-Cultural Phenomena and Current Trends in Facial Aesthetics

The *Selfie* was declared word of the year 2013 by the Oxford Dictionary, and CNN journalist Ben Brumfield even went on to call it "*debatably the most embarrassing phenomenon of the digital age*" (Brumfield, 2013; online). However, surgeons have warned that selfies, social media, and advertising cannot be trusted in this regard and are potentially a rather concerning influence contributing to a rise in body image disorders (Kremer, 2017). Therefore, the British Association of Aesthetic Plastic Surgeons has requested a ban of such advertisements (Ellson, 2020). Nevertheless, there has been a significant increase in aesthetic treatments and surgeries over the past decade, specifically rhinoplasties, which many attribute to selfie-culture (Pendolino and Ottaviano, 2019; Lee *et al.*, 2022). More precisely, the often ignored fact that the mid-face region – primarily the nose – is distorted and appears larger, bigger, and wider than the actual features when an image is taken at arm's length (i.e., object-to-camera distance) and with a mobile phone camera (Ward *et al.*, 2018). Despite cosmetic surgeons explicitly pointing this effect out to their patients, many still choose to undergo the procedure to have a selfie-perfect face, irrespective of potential real-life implications (Pressler *et al.*, 2022).

Another new term that has emerged recently is the *Kardashian effect* (Brooks, 2017; Kremer, 2017; Kisyova *et al.*, 2019). It refers to a rise in filtered and edited photographs, and increasing demand for both non-surgical and invasive aesthetic treatments, as young people aspire to follow beauty standards portrayed and promoted by the Kardashian-Jenner clan through their respective social media channels and the TV show *Keeping up with the Kardashians* (Brooks, 2017; Sood, Quintal and Phau, 2017; Werner, 2018).

There is a known demographic shift in the age structure of industrialised countries. Not only is the population as a whole getting older, but there is also growing interest, pressure, and progressively easier, more affordable access to counteracting the biological process (Venkatesh *et al.*, 2019). In more liberal societies today, *"we are seen as being responsible for what happens to our bodies, especially with regard to health and beautification ([…] described as the visual evidence of well-being)"* (Kosut, 2015; p.32). It appears that – opportunity and funds aside – there is arguably an expectation towards individuals to engage in body modifications. In July 2020, L'Oréal Paris started to run an advertising campaign in the UK for their *Revitalift Laser Renew SPF 20* range of skincare products with the slogan "*postpone the procedure*". The advert featured on podcast apps (e.g., Podcast Addict), as well as magazines and newspapers, such as *Elle online* (Munson, 2020b)*, the Daily Mail* (MailOnline, 2020), and *Good Housekeeping* (Munson, 2020a). Advertising the same product range, other magazines promoted the products under somewhat questionable headlines, for instance "*5 Women on the Clever Ways They're Future-Proofing Their Skin*" by Lawrence (2020) for *Women's Health* magazine, and "*6 effective skincare alternatives to booking a treatment*" (Vince, 2020) in *Harpers Bazaar*. The casual way in which *the procedure* is mentioned as an inevitable step that has to be taken sooner or later - by women - contributes to the current narrative of and societal pressure towards normalisation and expectance of self-optimisation by any means possible. The researcher has submitted a complaint about this particular advert to the Advertising Standards Authority (ASA) in September 2020. After having received multiple similar complaints about the same advertising campaign, the ASA contacted L'Oréal, who in turn provided reassurance that the slogan would be taken down and further advertising campaigns would not imply the need for cosmetic procedures.

4.2.2 Impact of Body Modifications on Face Recognition and Identification

There appears to be a distinct lack of published research regarding body modifications of the head and face and its relevance to and impact on facial identification. Studies on adding simple lines (Ellis, Davies and Shepherd, 1978) and make-up (Ueda and Koyama, 2010) on faces have shown a decline in recognition performance. An even bigger impact was found when adding spiral-shaped facial tattoos like the Maori would traditionally wear: Buttle and East (2010) found that recognition performance for upright, non-tattooed faces was at 84%, which is in agreement with other face recognition studies on human recognition ability (see section 6.3 Human Face Recognition below for further details). When one side of the face was tattooed, performance dropped to 65%, and to almost chance-like levels of 55% if the entire face was displayed with spiral tattoo patterns. A rather dramatic effect and comparable to performance on inverted faces (e.g., Young, Hellawell and Hay, 1987; Tanaka and Farah, 1993; Megreya and Burton, 2006).

There is a lack of more recent research on this subject, as studies increasingly appear to focus predominantly on automated face recognition systems in this regard (e.g., Singh *et al.*, 2010; Nappi, Ricciardi and Tistarelli, 2016; Zuo, Saun and Forrest, 2019; Bouguila and Khochtali, 2020; Rathgeb *et al.*, 2020). Quite recently, Urbanová, Eliášová and Dostálová (2022) have evaluated the effects of facial corrective surgery in the context of face recognition on 3D before and after facial models, using semi-automatic landmark-based systems. Results showed a significant impact on facial identification accuracy, prompting their demand for further testing in this context of body modifications, not only using automated systems, but also for manual facial comparison approaches.

Plastic surgery – both cosmetic and reconstructive – is now understood to be another factor impacting automated systems, alongside more well researched factors, such as pose, illumination, expression, ageing, and disguise (Zuo *et al.*, 2019). However, the exact impact is not yet well understood. A decade ago, Singh *et al.* (2010) created a plastic surgery face database and found that automated systems were unable to overcome this hurdle. Fairly recently, Rathgeb *et al.* (2020) created a newer plastic surgery face database, after others had criticised the former database by stating issues with low quality images. Rathgeb *et al.* (2020) further entailed automated face recognition using the new database and found that for deep-learning-based algorithms, the impact was not quite as significant as for previously researched approaches. The issue with using the new database, and hence high-quality images only, is that methods may not be transferrable to real-world scenarios, where images are often taken under uncontrolled conditions and this requirement is not viable.

To the best of the researcher's knowledge, there is currently no published research on the identification of the dead from facial features that were altered, by either temporary or permanent intentional body/face modifications, between the time an ante-mortem photograph ("*before*") was taken, and the deceased ("*after*") was recovered.

The effect of body modifications (e.g., facial tattoos, piercings, plastic surgery, and cosmetic treatments such as dermal fillers) on visual personal identification and recognition based on facial data is not at all well researched and understood at this time. Anecdotally, a recent case made the headlines as three women, who had travelled abroad for cosmetic surgery, failed identity verification at passport control, as their new (still swollen) faces no longer matched the passport photographs (Hurst, 2017).

Major ethical questions and health concerns resulting from societal pressure and social media trends aside (for further reading on this subject, please refer to Fardouly and Rapee, 2019; Hogue and Mills, 2019; Kisyova *et al.*, 2019; Eggerstedt *et al.*, 2020), it becomes more and more apparent that photographs no longer necessarily portray the actual face of an individual, and can therefore not be blindly trusted and relied upon as comparative data without keeping the aforementioned socio-cultural developments in mind. In addition to real face modifications, virtual modifications – through photo editing software and filters – can often blur, distort, or fully erase certain morphological features and characteristics that might otherwise be unique to an individual and would therefore be particularly helpful for facial comparison.

However bad the image may be, selfies available from social media platforms can still be incredibly useful for the identification of the missing and unidentified (Miranda *et al.*, 2016; Nuzzolese *et al.*, 2018). Most people do not have severe body modifications, and oftentimes there are photos available that show the face after such changes have been implemented. A smiling facial expression may expose teeth, which can be compared to dentition of unidentified remains (University of Dundee, 2019). But orthodontic treatments are also on the rise, as they become more affordable for the wider population through services like SmileDirectClub™ and Invisalign® (Kravitz and Bowman, 2016), and need to be considered in comparison tasks.

4.3 Stability of Facial Features and Facial Expression

FISWG identified and compiled a list of factors that influence the physical stability of facial features, excluding variations that are introduced through different capturing devices (camera, scanner, etc.) or imaging conditions (lighting, resolution, etc.) (FISWG, 2021). These factors can change the appearance of a face quite considerably, which in turn has an effect on face recognition and identification efforts. Most of the factors are already considered in other sections (under sections 4.1 Facial Ageing and 4.2 Facial Aesthetics and Body Modifications above) (e.g., changes over time, weight loss/gain, impact of health and illness, and body modifications), but facial expression also needs to be considered.

Facial expression is defined as "*any deviation from a relaxed face*" (FISWG, 2021; p.2). A relaxed or neutral expression is common for facial photographs used for identification documents (e.g., passport, driver's licence), but non-standardised, natural images of a person often do not depict the face in a neutral expression or facial angle. This is important as ante-mortem images used for identification are often not standardised but the deceased face in the early post-mortem period likely reflects more of a relaxed and slack expression. This gap of differences is a challenge that human experts and machine algorithms used for face recognition and identification must acknowledge, take into consideration, and aim to overcome. The dead face and how its appearance deviates from the entirely relaxed living face, has not yet been quantified.

5 Human Taphonomy

The Russian palaeontologist and geologist Ivan Antonovich Efremov was the first to introduce the term *taphonomy* (Greek: *taphos* = burial/grave; *nomos* = law) into the scientific community in 1940 (Efremov, 1940). By its more recent definition, it describes the study of post-mortem processes which "*affect (1) the preservation, observation, or recovery of dead organisms, (2) the reconstruction of their biology or ecology, or (3) the reconstruction of the circumstances of their death.*" (Haglund and Sorg, 2006b; p. 13). Taphonomy is a multidisciplinary field that requires expertise and research input from several disciplines, such as biology, microbiology, soil science (pedology and edaphology), anthropology, chemistry, and entomology. It is a relatively new field that has significantly developed over the past three decades, and entails all matters of the burial itself as well as the peri- and post-mortem processes relating thereto (Randolph-Quinney, Haines and Kruger, 2018). Schotsmans, Márquez-Grant, and Forbes (2017), three current experts in the field, argue that not only peri- and post-mortem events should be included under this umbrella term, but also the related storage and subsequent analyses.

Forensic taphonomy, considered a part of forensic anthropology, includes the collection of samples and evidence of the depositional context and human remains, which enable a reconstruction of peri- and post-mortem events (Haglund and Sorg, 2006b). As opposed to taphonomy in an archaeological, paleontological and/or paleoanthropological context, which encompasses considerably wider time frames, forensic taphonomy is concerned with the events immediately surrounding death. The post-mortem period of interest in forensics is often merely days, weeks, or months, and very rarely years.

5.1 Stages of Decomposition

Within the scientific literature, the number of stages of decomposition is predominantly listed as the following five: fresh, bloat, active decay, advanced decay, and skeletonisation (Goff, 2010). However, human decomposition is somewhat predictable but also highly variable, and stages are not distinctly separated, rather several stages may be present simultaneously in one deceased individual and are commonly overlapping (Haglund and Sorg, 2006b). Although the general overarching pattern of macro- and micro-changes occurring is somewhat similar, the onsets and durations can vary, as can other variables or phenomena that may be present or developing in some but not others (e.g., mummification, adipocere).

Environmental conditions and individual body composition can significantly influence the stages of decomposition and timing thereof (Tibbett, 2008; Wilkinson, 2014). The interplay of environment, entomology, scavenger activity, and biological processes, results in the necessity for specialist analysis of both, the aforementioned, as well as the human remains themselves. This underlines the requirement for an interdisciplinary approach to research and casework that involves taphonomy and human decomposition, at all stages a cadaver will undergo (Forbes, 2008).

Internal (autolysis, putrefaction) and external (decay) deterioration of the human body usually advances in the following order (Gill-King, 2006):

1. Large and small intestines, pancreas, liver, gall bladder, stomach, blood, heart
2. Lungs and airways
3. Bladder and kidneys
4. Nervous tissue and brain
5. Skeletal muscle tissues
6. Hair, connective tissues, skin

As all PM data used in the current study was collected from individuals in the earlier stages of decomposition (i.e., fresh, slightly bloated), the focus of this section will primarily be on changes occurring in the earlier post-mortem interval, although other stages are also briefly described. The changes in relation to facial appearance and surface tissue shall be discussed separately in the next section, and details of underlying chemical processes and later stages of decomposition are kept somewhat brief in this review. Methods for post-mortem interval (PMI), time since death (TSD), and/or accumulated degree days (ADD) are not discussed, as they bear no relevance for the author's research presented in this thesis. The same applies to bog bodies, glacial environments, and other more unusual deposition scenarios. For a more in-depth description and further reading on the above mentioned areas of human taphonomy, please refer to Haglund and Sorg (2006a), Gunn (2011), and Hayman and Oxenham (2016).

The rate of decay, degree of severity of individual post-mortem changes, and manifestations of deviations from the textbook processes (e.g. mummification, adipocere) are all heavily influenced by a number of intrinsic and extrinsic factors (Blau and Forbes, 2016; Knobel *et al.*, 2019):

- Intrinsic factors
    - Body build
    - Age
    - Bone type / size
    - Disease / traumatic injury
    - Enteric microorganisms
- Extrinsic factors
    - Time between death and burial/disposal
    - Environment
        - Topography
        - Type of soil/sediment
        - Water/humidity
        - Temperature
        - Flora & fauna
            - Insects
            - Macro fauna
            - Vegetation
    - Fire
    - Body treatment
    - Clothing/coverage
    - Mode of disposal
    - Type of coffin
    - Burial depth

5.1.1 Fresh

The fresh stage begins at time of death and is characterised by a number of external and internal signs of death. In fair skinned individuals, the first external observation may be a distinct paleness (or pallor), caused by cessation of circulatory functions (Simmons and Cross, 2013)s. Another distinction is made between uncertain and certain signs of death. The former are unreliable for determining the death of an individual, as some or all can potentially be misleading and/or possibly reversed. Those include cessation of cardiac activity, areflexia, wide and fixed pupils, respiratory arrest, and decreased body temperature (hypothermia). Certain signs of death however are commonly used to determine and certify death, and consist of livor mortis, algor mortis, rigor mortis, as well as the lack or decrease of supravital reactions through electronic or manual stimulation, e.g. subsiding electrical excitability of mimetic muscles (from ipsilateral facial half 2.5-3.5 hours PM to limitation of reactions to immediate area of stimulation 8.5-13.5 hours PM)

(Madea *et al.*, 2014). Autolysis, the internal disintegration and - in essence - self-digestion of bodily structures through loss of cell integrity caused by the body's own enzymes occurs throughout the body.

With the cessation of blood circulation, oxygen supply to the tissues is no longer sustained leading to anoxia. Depending on the type of cell, these will die within either minutes or hours after death (e.g., brain cells irreversibly die after 3-7 minutes, skin cells can survive for up to 24 hours) (Gunn, 2011). Cell death and other autolytic alterations can first be observed under the microscope, whereas macroscopic changes manifest most prominently in the following three phenomena: livor mortis, algor mortis, and rigor mortis (Marks and Tersigni-Tarrant, 2016). Autolysis is influenced by both ambient and body temperature, with an accelerated progression at higher temperatures (e.g. ante-mortem fever, hot climate) and a slower development in colder climates and/or due to refrigeration (Clark, Worrell and Pless, 2006).

5.1.1.1 Livor Mortis

Once the heart stops beating, blood is no longer circulating around the body and - due to gravitational forces - settles in areas closest to the ground. This pooling forms distinct reddish-purple discolourations (see Figure 9 below), referred to as livor mortis (Latin *livor* = blue-ish discolouration/bruise; *mortis* = death), cadaveric/post-mortem lividity or hypostasis. The appearance thereof presents initially with smaller blotches which extend and spread, as well as deepen in colour over time. Eventually, haemolysis will occur, causing for the blood cells to disintegrate and haemoglobin pigment to be spread to neighbouring tissues. There the pigment will be transformed into sulphaemoglobin, which is greenish in colour and exudes the characteristic rotten egg smell, due to its hydrogen sulphite components (Gunn, 2011). The rate in which livor mortis forms and how well it is visible is dependent on several factors, such as certain medical conditions, body size, cause of death, blood volume within the body, and skin colour (Tsokos and Byard, 2016; Byard, 2020). It can be observed as early as 1 hour PM, and will become 'fixed' at around 10 hours PM, due to adipose cells in the dermis solidifying and sealing the blood-filled capillaries (Dix and Graham, 2000).

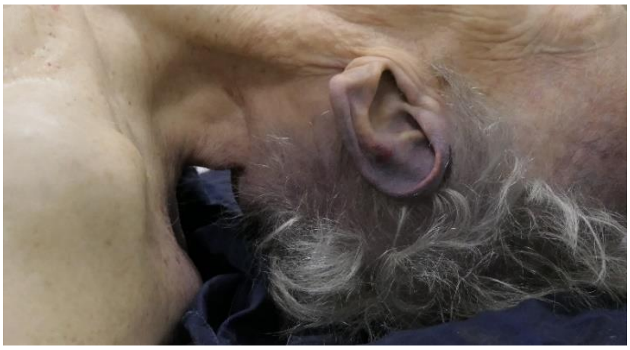

Figure 9: Livor mortis visible as purple discolouration of the ear and back of the head (occiput) in the early post-mortem period (image: author's own; taken at FARF, TXST; used with permission).

5.1.1.2 Algor Mortis

Metabolism halts and with it thermoregulation, the latter regulated during life by the hypothalamus (Rattenbury, 2018). In some cases, body temperature might remain at a plateau for a few hours, before cooling sets in and will progress until it eventually adjusts to the ambient temperature (Smart and Kaliszan, 2012). This process is also referred to as *algor mortis* (Latin *algor* = coldness; *mortis* = death). As well as *livor mortis* mentioned above, it is a sign of decomposition that is non-reversible, therefore represents a *certain sign of death* (or Latin *signum mortis*). Other such signs would be physical traumas (e.g., decapitation, complete severing of the torso) that are also incompatible with life. The normal average body core temperature in living humans is approximately 36.9°C, but evaporation, convection, conduction, and radiation will cause for the deceased to become cold. The rate of cooling in the early post-mortem period is dependent on several factors, as shown below (Smart and Kaliszan, 2012; Byard, 2020):

Acceleration:

- Low Body Mass Index (BMI)
- Wind, moisture
- Hypothermia
- Hypothyroidism
- Greater difference between body core and ambient temperature
- Naked body, limbs extended
- Children, infants, petite adults

Delay:

- High Body Mass Index (BMI), as adipose tissue retains heat
- High levels of physical activity prior to death
- Fever
- Clothing, blankets, etc.
- Curled-up body position
- Hot and/or windless climate

Sex is also an important factor, since males tend to have less subcutaneous fat deposits, hence cooling in males tends to progress faster compared to females of a similar height and stature.

5.1.1.3 Rigor Mortis

Once death has occurred, muscles and joints relax, which – depending on the position of the person – may cause injuries due to the body collapsing or falling, and/or defecation, enuresis, and regurgitation of stomach contents due to flaccidity of the sphincter muscles (Gunn, 2011). Post-mortem, an increase in calcium ion concentration inside the muscle cells leads to the muscle filaments actin and myosin interlocking. Function (contraction and relaxation) of muscle fibres in the living is reliant on the interplay of those filaments, a process also known as *sliding filament theory* (Powers *et al.*, 2021). This is enabled by the adenosine triphosphate (ATP) and adenosine di-phosphate (ADP) cycle. Since ATP is no longer formed in the deceased, the cross-bridge cannot be released, causing a sustained contraction of muscle fibres along with the loss of extensibility leading to stiffness of muscles and joints. This process or condition is referred to as *rigor mortis* (Latin *rigor* = stiffness, rigidity; *mortis* = death). For an in-depth explanation of the underlying chemical processes, please see Gunn (2011) or Tsokos and Byard (2016).

According to Nysten's law (aka. Nysten's rule), rigidity begins in the face and neck, moving down on the body, affecting small muscles or muscle groups prior to larger ones (Tsokos and Byard, 2016). The French physician and paediatrician Pierre-Hubert Nysten (1771-1818) also observed that – in the event of physical exercise or a struggle/fight - any muscles activated during the ante-mortem period immediately preceding death are affected first (Nysten, 1811). The timeline can vary, with an approximate onset of rigor mortis 3-4 hours after death (often at eyelids and jaw), and a peak of full body rigidity at 9.5 hours (Madea *et al.*, 2014). Influencing factors are, inter alia, stature, level of activity prior to death, and temperature (earlier onset and shorter duration in higher ambient temperatures). With progressing autolysis, the muscles and joints will eventually become flexible again after several hours or days, due to protein degradation.

5.1.2 Bloat

Bacteria are a natural part of the intestinal flora, but during life, those are kept under control in terms of location and propagation. Following death, the acid-alkaline balance and oxygenation are no longer sustained, rendering the environment acidic and anaerobic (Rattenbury, 2018). Due to autolysis, gut bacteria are able to break down those chemical and biological barriers and are now able to spread to other parts of the body and multiply. This process is known as putrefaction, and its manifestations in the cadaver mark the transition from fresh to bloat stage (Gunn, 2011). An initial greenish-blue discolouration is commonly visible on the outer abdominal wall above the caecum (Tsokos and Byard, 2016). This will soon spread through blood vessels, rendering them darker in purple, brown-black, and/or blue-green hues, which is referred to as *marbling* or *intravascular haemolysis* (see Figure 10 below), less commonly also known as suggillation (Simmons and Cross, 2013; Marks and Tersigni-Tarrant, 2016).

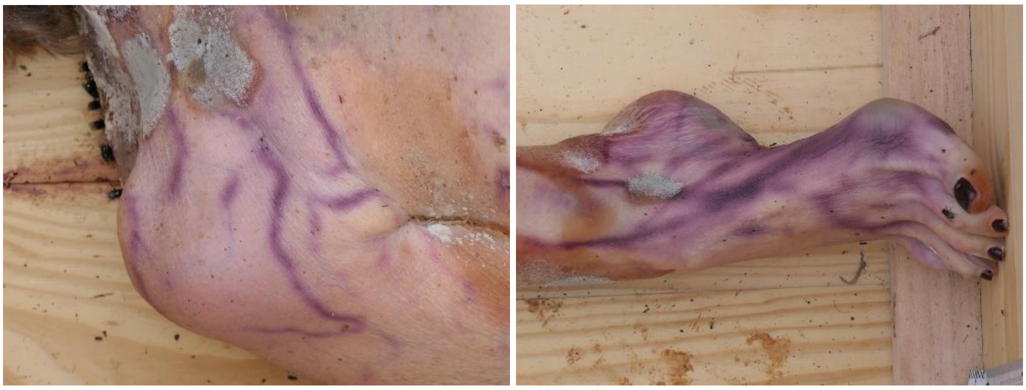

Figure 10: Right shoulder and right foot showing distinct discolourations along superficial vein paths, visualising the decomposition phenomenon of marbling (image: author's own; taken at FARF, TXST; used with permission).

The formerly fair or lighter skin in is now appearing yellowish-green and will turn progressively darker towards dark brown to black. The changes in colour are less visible in darker skinned individuals. Blisters may form on the skin surface and skin slippage due to separation of dermis and epidermis can occur (Tsokos and Byard, 2016)t. The face and abdomen (in males also the scrotum) are increasingly bloated.

5.1.3 Active Decay - Putrefaction

The active decay stage is overlapping with both the previous bloating and following advanced decay stages. This stage begins often with the leaking of putrefactive fluids from bodily orifices, as soft tissues – beginning on the inside – liquify as a result of continuous autolysis and putrefaction, as well as microbial, invertebrate, and vertebrate activity (Gunn, 2011). If bloating is so advanced that the abdominal skin can no longer contain the gasses, or the skin is damaged (e.g., ante- or peri-mortem injury, surgical or autopsy sutures), the body will eventually release internal gasses and collapse, and with this invite more entomological species as well as other opportunistic animals to feed on the remains. The majority of biomass is reduced during this stage and insect species present on the body can be useful in PMI estimation, as well as DNA analysis of remains, by means of determining the growth stage or analysing the invertebrates' gut content (Campobasso *et al.*, 2005; Goff, 2016; Mukherjee *et al.*, 2019).

5.1.4 Advanced Decay - Liquefaction

Putrefaction may almost be complete, and the skin is visibly changing colour to darker hues, such as green-brown and brown to black. The body starts to dry and entomological activity is significantly reduced (e.g. maggot migration away from the remains, or to the underside of the body into the shade for pupation) (Arnott and Turner, 2008; Gunn, 2011). Fungal colonisation may be present, especially in humid environments. Decomposition liquid has saturated the surrounding grounds/soil and in case of bodies deposited above ground, a Cadaver Decomposition Island (CDI) will form, defined by vegetation death surrounding the body (Janaway, Percival and Wilson, 2009; Fancher *et al.*, 2017). Later on, this soil - saturated with what can be described as natural fertiliser from waste products of decomposition - will likely show a more pronounced and flourishing vegetation, which can indicate shallow graves and/or a (former) body deposition site. During advanced decay, tissues such as hair, cartilage, and bones may remain, and any soft tissue present is most likely discoloured to darker brown to black tones. If the skin does not mummify, most of it will be lost at this stage and the transition to skeletonisation is somewhat fluid. It is important to note that skin and bone lesions at this stage can both mimic and conceal trauma on the body, and what is often mistaken for a gunshot or stab wound, may simply be soft tissue damage caused by insect or scavenger activity (Gunn, 2011).

5.1.5 Putrid Dry Remains - Skeletonisation

The rate of skeletonisation is influenced by a multitude of factors, such as temperature, humidity, insect and scavenger access to remains, and the type of burial or body deposition (Janaway *et al.*, 2009). This stage is defined by the loss of soft tissues, however not all bodies will skeletonise, or at least not fully. What can occur instead is mummification and/or adipocere formation, two forms of natural preservation. Mummification is the result of dehydration of skin or other soft tissue, which is more likely to occur in arid environments and faster in hotter climates, when bodies are deposited above ground or remain indoors. This will inhibit further decompositional processes to the affected tissue, by means of interfering with bacterial growth (Sledzik and Micozzi, 1997; Clark *et al.*, 2006; Blau and Forbes, 2016). As the skin desiccates, it appears leathery and dark brown to black. Depending on the environment, it may only be the skin that is affected by mummification. However, especially in hotter and drier climates, it is possible for the internal organs to mummify as well (Simmons and Cross, 2013). Eventually only dry bones will remain underneath the skin layer in most cases.

Adipocere (Latin *adeps* = fat, and *cera* = wax), also referred to as corpse/grave wax, forms from adipose tissue during a process known as saponification, which is essentially the hydrolysis of fatty acids (Gunn, 2011). For this to occur, the environmental factors have to be aligned favourably in what Tyler O'Brien (1994) termed the *Goldilocks Phenomenon*[1]. This is given if the ambient conditions are just right, meaning a interplay of a certain temperature, plus humidity or submergence in water, bacterial activity or lack thereof,

[1] *Goldilocks and the Three Bears* is a 19th century fairy-tale, said to have originated in Britain. In the story a girl intrudes the home of a bear family, trying out their chairs, bowls of porridge, and beds. Amongst all sets of three she finds the perfect one (in height, temperature, and size respectively) to be that of the bear cub/child. O'Brien's analogy refers to different elements in the taphonomic environment which need to be exactly right to facilitate adipocere formation.

and a certain amount of adipose tissue on the remains. If only one of the influencing factors is not quite right, soft tissue will desiccate (too cold and/or dry), or liquefy (too warm and/or wet) (O'Brien and Kuehner, 2007).

The formation process of adipocere can be simplified as follows: The membrane of an adipocyte (= fat cell) consists of fatty acids, glycerols, proteins, and phospholipids. In an environment where those are submerged, the cell membranes will absorb any excess of water until the cell ruptures and releases its fatty contents. Chemical decomposition, caused by aerobe bacteria, leads to the fatty contents being hydrogenated and hydrolysed. The neutral fat transforms into hydroxy fatty acids, as the bacteria are secreting toxins containing proteases and phospholipases that subsequently destroy the cell membrane. As the anaerobic organisms create ammonium, amongst other waste materials, the environment becomes alkaline, which subsequently prevents further bacterial activity and growth, therefore halting further decay of the remains (O'Brien, 2006). In most cases, the taphonomic degeneration of the body will come to a halt, and adipocere preserves the remains, either partially or wholly, depending on the localisation thereof (Rodriguez, 2006). The grey-white substance, also known as grave or corpse wax, will form more likely on and within individuals that possessed a higher fat content during life, as compared to cachectic or more athletic builds (Tsokos and Byard, 2016). The overall process of adipocere formation is usually slow and can sometimes take only 1-6 months but mostly years (Hochmeister, Grassberger and Stimpfl, 2007), depending on the environment and amount of adipose tissue on the corpse. O'Brien and Kuehner (2007) pointed out a lack of longitudinal research into the transformational processes and timelines regarding adipocere, which still persists today.

Skeletonisation is defined as "*the removal of soft tissue from bone*" (Clark, Worrell and Pless, 2006; p. 159). Once the majority of soft tissues have vanished, what remains are ligaments, tendons, hair, finger- and toenails, as well as bone, as those are more difficult to disintegrate. After death, bacteria spreading predominantly from the abdominal area are able to infiltrate the bones through pores. Hence most notable changes in and natural damage to skeletal tissues during the earlier post-mortem period is likely to be found within the torso and pelvic region (Gunn, 2011). Exposed skeletal remains can become white through bleaching by direct sunlight, become porous and brittle in acidic soil, and/or show other signs of discolouration caused by items of clothing and/or the surrounding environment (Simmons and Cross, 2013). The integument, soft and connective tissue, as well as muscles will usually disintegrate to allow for disarticulation of the bones (Tsokos and Byard, 2016). Eventually, skeletal remains will also break down, due to microbial activity and exposure of the remains to environmental influences, but there is no distinct end point to this stage (Goff, 2016).

5.2 Decomposition in Water Environments

Taphonomic processes can vary in water environments compared to decomposition on land. In the UK, decomposition progresses at approximately half the speed in water compared to land. This is mainly due to water usually being cooler than the outside temperature, and the absence – or significant reduction of – necrophagous insect activity (Ellingham *et al.*, 2017). This interval can be stretched further in flowing bodies of water, and shortened when water is stagnant and/or heavily polluted (Lawler, 1992). If water temperature remains constant at a temperature of below 5°C, the onset of any decompositional changes can be deferred by several weeks, whereas in warm and hot climates, it can rapidly progress within 24 hours. Decomposition in aquatic environments is influenced by a multitude of intrinsic and extrinsic factors, such as saline content of the water, temperature, currents, oxygen saturation, clothing, body composition, and whether the person was still alive or already deceased when submerged (please refer to Sorg *et al.,* 2006 for a comprehensive list of variables).

The sink or float pattern can vary, as not all bodies will sink. Some bodies sink, float, and then sink again, whilst others might get trapped under water. When water is inhaled into the lungs or swallowed and hence weighing down internal organs, the body will likely sink once the person has died (Mateus *et al.*, 2013 in Ellingham, Perich and Tidball-Binz, 2017). Saline content of the water and body composition can also affect buoyancy (Lawler, 1992; Doberentz and Madea, 2010; Simmons and Heaton, 2013). Fresh water is more readily absorbed into vessels and tissues, causing swelling and eventually rupture of soft tissues. Whereas in water with higher saline content, the body's water content is drawn outwards, which in turn can slow down bacterial activity (Simmons and Heaton, 2013). A denser physique will favour sinking. Hydrostatic pressure will cause for gasses and air within the body to be compressed more, the further down it is sinking. Once decomposition reaches the bloating stage, bodies tend to resurface in a supine position until putrefactive gasses have been released, after which the remains may sink again (Reh, Haarhoff and Vogt, 1977; Simmons and Heaton, 2013).

Rigor mortis is mainly affected by water temperature, with colder water postponing onset and prolonging duration (Lawler, 1992). Algor mortis will occur faster in water than on land (approx. twice as fast in UK) and is further expedited when water is not stagnant. Livor mortis is often more pronounced around the neck and face, anterior chest, hands and forearms, lower legs and feet, as most bodies tend to initially float in a prone position, with limbs extended downwards (Lawler, 1992; Sorg *et al.*, 2006; Keil, 2021). In more rapidly moving water, livor mortis may not develop at all, as the body will not remain in one position for long enough. When remains are submerged for prolonged periods of time, adipocere formation is likely to occur (Reh *et al.*, 1977; Lawler, 1992; O'Brien, 2006). Decompositional changes in water usually follow a certain pattern, but like the previously mentioned stages of decomposition on land, overlap is likely to occur (e.g., skull skeletonised, whilst body protected by clothing is largely preserved but bloated) (Reh *et al.*, 1977; Simmons and Heaton, 2013).

Maceration is the most typical post-mortem sign associated with water immersion. At first, the skin appears increasingly pale to white, then soggy, and wrinkled (aka. *washer woman* skin). The epidermal layer eventually detaches, as will hair and nails (Lawler, 1992). Other possible taphonomy related phenomena in water include the formation of a carpet of algae on the skin (Keil, 2021), and adipocere (O'Brien, 1994, 2006; Forbes, Wilson and Stuart, 2011; Simmons and Heaton, 2013), the latter potentially preserving (partial) facial features. Injuries are common and likely to occur during horizontal and vertical displacement of the body. As most bodies will initially float in a prone position, skin abrasions on the face (primarily forehead and nose), hands, knees and feet can be caused by contact with the bottom of a body of water (e.g., sand, rocks, debris) (Lawler, 1992; Keil, 2021). Water currents may bring a body in contact with other hard structures, such as bridge pillars and driftwood, and severe damage can also occur by marine propellers. Depending on individual circumstances, injuries may have occurred prior to submergence, e.g., as a result of fall from great height and/or hard surface tension of the water, or from contact with projecting rock ledges or piers. Exposed body parts, such as face and hands that are not covered by clothing, will often show scavenger damage from aquatic wildlife (Ellingham *et al.*, 2017).

No studies do date have explicitly investigated facial changes during decomposition in water, neither with nor without a link to certain timelines or contexts. At this time, it is therefore impossible to predict for how long into the post-mortem interval visual methods of identification would be feasible. For many migrant deaths in the Mediterranean, comparison of visual characteristics (e.g., tattoos, face morphology, teeth if visible in photographs) may be the only way to identify someone, because AM data for primary methods of identification are generally not available in these cases (Alemán Aguilera, 2022). Although the use of non-primary identifiers is not accepted as legal basis for identification in some countries or jurisdictions (Ellingham *et al.*, 2017), they can still be extremely useful if AM photographs exist (Caplova, Obertova, *et al.*, 2018). These methods may allow to find a possible AM-PM match, potentially trace back the deceased's migration journey, and ultimately connect with their family to possibly obtain other AM data samples and/or provide answers and closure.

An estimated one in ten people die during crossing attempts in the Mediterranean (Cattaneo *et al.*, 2022), and only approximately 13% of people reported missing or dead are recovered (Baraybar, 2022). The identification rate of the recovered bodies remains unspecified at this time, as there is neither a centralised database nor sufficient, organised data sharing that would allow for such a statistic (Olivieri *et al.*, 2018; Wilkinson and Castaneyra-Ruiz, 2021; Alemán Aguilera, 2022). About 60% of recovered maritime DVI victims (e.g., ship accident) are identified by means of odontological comparison (AM data availability permitting), and facial comparison would only be feasible if the damage and changes to the soft tissue is minimal. On the whole, research on taphonomy in water is still rather sparse and most reports and findings are derived from forensic casework (O'Brien, 2006; Sorg *et al.*, 2006; Doberentz and Madea, 2010; Ellingham *et al.*, 2017). There is however a dire need for further investigation, as this scenario is quite common in DVI (e.g., migrant crisis, natural disasters).

5.3 Post-Mortem Changes to the Face

There are only a handful of studies in the published literature to date that have been undertaken to specifically investigate changes to the human face following death, particularly the timeline thereof (Wilkinson and Lofthouse, 2015). Hence the common scientific consensus at this time being that more research on this topic is urgently needed (Caplova, Obertova, *et al.*, 2018). However, some general patterns have been observed that seem to be similar in their progression, although the rate of onset and timings, as well as possible overlap of decomposition related changes to facial features seem to be significantly influenced by environmental conditions, as is to be expected (O'Brien and Kuehner, 2007; Blau and Forbes, 2016).

5.3.1 Physiological Post-Mortem Changes to the Face

In their exhibition and eponymous book *Nochmal Leben vor dem Tod* (German, meaning *life before death*), Beate Lakotta and Walter Schels portray the life (text) and face (photograph) of terminally ill individuals in hospice care, shortly before death as well as within the early post-mortem period (Lakotta, 2003). Examples of their work are shown in Figure 11 below – with the photographer's permission – as they capture the very essence of what humans naturally perceive as the living face and the visible changes within self-same once death has occurred.

A number of alterations to the facial features may appear either in the days or hours leading up to death (ante-mortem), during death (peri-mortem), and/or following death (post-mortem), depending on the circumstances of the individual person. Therefore, prior to discussing post-mortem changes of the face, it should be acknowledged that there is potentially some overlap and co-dependency, requiring a level of consideration in terms of correct identification and understanding, which prevents possible misinterpretations and false conclusions.

Figure 11: Three terminally ill individual's faces captured shortly before and after death (Lakotta and Schels, 2004) – (Images used with kind permission from Walter Schels and Beate Lakotta).

Visible changes of the facial appearance occurring in the last days and hours leading up to a natural death can follow a somewhat predictable pattern, however, those may not develop or show in everyone. One example would be Kirchhofrosen (German: *Kirchhof* = churchyard / cemetery; *Rosen* = roses) or intravital lividity, which is characterised by irregular pinkish-red areas around the lower cheeks and behind the ears, and are a result of diminishing cardiac function which leads to circulatory slackening and local stasis (Prokop and Göhler, 1975; Madea, 2016). This skin discolouration can be mistaken for livor mortis and sometimes also affects the lower legs (aka. mottling/mottled skin). A second example is the *face of the dying* or Hippocratic facies (aka. *facies hippocratica*; Latin *faciēs* = face, *hippocratica* after the Greek physician Hippocrates of Cos c. 460-370 BC), which is often a sign of impending death and well known in palliative medicine

(Tinsley, 1918; Hach and Hach-Wunderle, 2016). Related facial changes present in a pinched nose, slack but hardened, pale, and wax-like skin, hollow cheeks, as well as sunken temporal and orbital regions (Kanathur, Sarvajnyamurthy and Somaiah, 2013).

Once death has occurred, the first apparent post-mortem changes in the face are those relating to facial expression. Facial muscles relax fully - to an extent that cannot be mimicked wilfully by a living individual - and the overall expression is blank (Cornett *et al.*, 2019). Initially, all muscles become flaccid but with the onset of rigor mortis the resulting changes can temporarily 'freeze', often beginning at the eyelids and shortly thereafter also including the jaw. The position in which someone has died or is found dead in is rather significant to their post-mortem facial changes and respective appearance. In general, the lack of muscle tone and effects of gravity are distorting the features (Taylor, 2001). Being in a supine position, lateral (on their side), or sitting, potentially with the head tilted backwards, the muscles holding the mandible in place (e.g., masseter, temporalis) are now relaxed, causing a slack jaw. When in a lateral, but especially in a prone position (on their front and face down), gravity will cause blood to pool in the soft tissues of the face, giving it a dark red, purple, blue and/or even black-ish appearance, where the facial surface is not compressed (Wilkinson and Lofthouse, 2015; Catanese, 2016). Lividity can also be observed in the sclerae, depending on body position. If the individual is lying on their side, most likely only one sclera will be affected (Catanese, Levy and Catanese, 2021).

During the early post-mortem period (i.e., within hours after death), the orbital region is void of any tension, often causing the eyelids to remain partially open (Clark *et al.*, 2006). The cornea will begin to turn cloudy approximately two hours after death due to progressive lack of moisture, leading to an imbalance in hydration of the eye tissues. Endothelial cells of the innermost corneal layer deteriorate and thus fail to inhibit the aqueous humour – located in the anterior and posterior chamber of the eye in the living – from entering the cornea. Additionally, degradation of collagen fibres, changes in proteoglycan hydration and ion concentrations can contribute to and intensify opacity (DelMonte and Kim, 2011; Trokielewicz, Czajka and Maciejewicz, 2020). A reddish-dark brown to black line can form transversely across the partially exposed sclera of the eye, where those remained open, which is known as *tache noire* (French *tache* = stain/spot/mark; *noire* = black) (Trokielewicz *et al.*, 2020). These changes are associated with the fact that the eyes are slowly drying out, which further leads to "*wrinkling and creasing across the iris as well as the [eventual] collapse of the cornea*" (Sauerwein *et al.*, 2017; p. 1601). The pupils usually remain in a mid-dilated state, which is then referred to as *cadaveric position*, but the size can somewhat differ between both eyes of the same individual due to rigor mortis potentially influencing the iris sphincter muscles non-equilaterally (Trokielewicz *et al.*, 2020). After several days to weeks, dehydration and autolysis will cause a decrease in intra-ocular pressure, the ocular globe will decompress and appear indented, flaccid, and the entire orbital area sunken (Belsey and Flanagan, 2016). Sauerwein *et al.* (2017) pointed out that the desiccation and collapse of the eye and with it the iris, will render this feature unusable for identification purposes.

At later stages of decomposition, the internal build-up of gasses (methane, hydrogen sulphide, mercaptans, etc.) and subsequent bloating will lead to a swelling of the facial soft tissues (eversion of lips, bulging eyes and cheeks, tongue protrusion) and neck, as a result of increasing internal pressure (Gunn, 2011; Simmons and Cross, 2013; Wilkinson, 2014; Tsokos and Byard, 2016). According to Cornett *et al.* (2019), bloating of the soft tissues predominantly affects the lower face around cheeks and jawline. Discolouration of the skin advances simultaneously, blisters can form on the skin surface, and separation of the dermal from epidermal layers may occur, possibly leading to skin slippage (Marks, Love and Dadour, 2009; Bolme *et al.*, 2016). Bloating or swelling of the facial soft tissues is arguably one of the biggest factors in natural post-mortem distortion of features, rendering the face unrecognisable – even by individuals familiar with the face – after only a short period of time (Taylor, 2001; Gunn, 2011). This cannot be underestimated and in recent DVI events has been recognised as an important issue, hence – wherever feasible - the importance of cooling or refrigerated storage of the deceased victims as soon as possible following recovery (Tsokos *et al.*, 2006; Wright *et al.*, 2015).

In arid conditions, mummification will render the skin grey or dark brown to black, with a leathery appearance (Cornett *et al.*, 2019). Hair often detaches in one piece and eventually slides off the head, which is related to the process of skin slippage during the bloat and decay stages (Janaway *et al.*, 2009; Simmons and Cross, 2013). If adipocere forms around the head and neck area, it is possible for facial features to remain preserved and visible for months or years after death has occurred, as the substance may act like a natural death mask (Kahana *et al.*, 1999; Tsokos and Byard, 2016). Numerous cases have been reported whereby the preservation of facial features was sufficient to allow for visual identification of the body and face (Dix and Graham, 2000; Fiedler and Graw, 2003; Stuart, 2013). In mummification, facial hair and/or skin may remain intact, but with possible further discolourations and distortions (shrinkage), due to loss of water content within the soft tissues (Sledzik and Rodriguez, 2002). The latter is often most noticeable around the mouth/lips and nose (Taylor, 2001; Clark *et al.*, 2006).

Insect and scavenger activity also contributes to severe facial changes during human decomposition, when entomological and vertebrate species have access to the remains. Flies will deposit eggs in the facial orifices and their larvae and subsequent maggots feed on the soft tissues of the face (Bisker and Ralebitso-Senior, 2018). Colonisation by a variety of entomological species may severely obstruct individual features and hinder data capture, as well as significantly advance the decomposition process (see Figure 12 below) (Gunn, 2011; Cornett *et al.*, 2019). Some insects are likely to deposit their eggs under the eyelids and in between eyelashes, as well as inside the nostrils and/or mouth, which can distort both shape and size of the respective features (Taylor, 2001; Anderson and Cervenka, 2002).

Figure 12: Decomposition related changes to facial features within two weeks at FARF, TXST (May 2019). From severe maggot colonisation of the face and orifices (active decay), further fly egg deposits and maggots at different stages of development (advanced decay), to partially skeletonised and mummified facial features with ceasing of insect activity (skeletonisation). Remains deposited above ground in shallow ditch (image: author's own; taken at FARF, TXST; used with permission).

Post-mortem insect and scavenging activity may be mistaken for facial wounds or trauma (Schulz *et al.*, 2006). Scavengers are more likely to access and gnaw on exposed body parts, such as hands and face (eyelids, nose, mouth), the latter potentially causing difficulties for identification (Cornett *et al.*, 2019). Minimal to no bleeding suggests the soft tissue damage was inflicted post-mortem, and certain wound patterns and margins, abrasions and/or gnaw marks can provide clues as to which animal is responsible (Tsokos *et al.*, 1999; Schulz *et al.*, 2006; Gapert and Tsokos, 2013). Please see Tsokos *et al.* (1999) for case studies and photographic illustrations of related facial soft tissue damage.

In summary, the general pattern of post-mortem external facial feature changes – which can deviate, depending on the environment and other influencing circumstances – has been described as follows (Perper, 1980; Utsuno *et al.*, 2005; Clark *et al.*, 2006; Wilkinson and Tillotson, 2012; Wilkinson and Lofthouse, 2015; Tsokos and Byard, 2016; Caplova, Gibelli, *et al.*, 2018):

1. Cessation of blood circulation → pallor, algor, livor, and rigor mortis
2. Loss of muscle tone → flaccidity, slack jaw
3. Dehydration and desiccation of the eyes → tache noir, corneal opacity, sunken orbits
4. Internal autolytic and putrefactive changes → skin discolouration; bloating of the lateral face regions (cheeks, jaw, temples) – very little change in volume to upper face regions (except for possibly protruding eyeballs at bloating stage)
5. Insect and scavenger activity → orifices and/or entire face colonised; soft tissue destruction
6. Desiccation of remains → sunken cheeks, lip shrinkage, further loss of soft tissues (skin)

5.3.2 Pathological Post-Mortem Changes to the Face

There are several pathological ante- and peri-mortem conditions that affect facial appearance, especially in relation to the skin, that should be taken into consideration when interpreting post-mortem changes. Some of those conditions might mimic natural post-mortem changes and can give the impression of advanced decomposition. Most such ante- and peri-mortem changes are caused by chronic illness, substance abuse, and/or poisoning, and have the potential to interfere with identification efforts that rely on facial features.

The *purple head sign* is indicative of sudden death, especially when caused in relation to cardiac failures (Dix and Graham, 2000). Although not distinctively researched as of yet, it is assumed that the discolouration in this case stems from an uncontrolled terminal discharge by the sympathetic nervous system, causing an opening of the capillary filters so that capillaries become severely filled with blood, which gives the dark red to purple appearance of the face (Catanese and Rapkiewicz, 2016). This should not be confused with carbon monoxide poisoning, which can cause bright, raspberry red discolouration of the blood and hence result in lighter coloured patches (livor mortis), reddish soft tissue and viscera (Downs, 2016). Other toxins, such as cyanide and fluoroacetate, can alter skin appearance in a similar manner (Tsokos and Byard, 2016). Jaundice or Icterus, visible as jaundiced skin and/or sclerae, can be caused, inter alia, by haemolysis associated with sickle cell crisis, hepatic cirrhosis due to chronic alcoholism, a hepatitis C infection, or hepatitis caused by chronic intravenous drug abuse (Catanese and Labay, 2016; Catanese and Rapkiewicz, 2016), Gilbert Syndrome, and/or a disruption in the gall flow from liver to the intestines (Preuss *et al.*, 2001). The reason behind the yellow skin discolouration is a disturbance in the bilirubin metabolism, causing a build-up of bilirubin in the blood; bilirubin being a degradation product of haemoglobin. Yellow to green discolouration of the skin is also associated with the onset of putrefaction (Clark *et al.*, 2006; Hochmeister *et al.*, 2007).

Petechiae are small, pin-point shaped haemorrhages, which can often be found in the face and sclerae as a result of hypoxia in combination with a sudden increase in the vascular pressure within the head region, e.g., through strangulation or hanging (Sundal *et al.*, 2020). Fatal heat stroke (possible scenario: children left inside a car in high temperatures) may also result in petechiae, due to severely increased intracranial pressure (Ikeda *et al.*, 2019). It is unlikely, that the aforementioned conditions would significantly affect facial identification, as facial features are still intact and minor discolourations would probably not interfere with familiar or unfamiliar face recognition. However, to the best of the researcher’s knowledge, no studies have been conducted on this particular subject as of yet.

Facial trauma (e.g., lacerations, puncture wounds, scrapes, bruising, avulsions, fragmentation of the skeletal structures, compression injuries) on the other hand has the potential to distort and/or obstruct facial features to an extent that complicates identification efforts. Forensic artist Karen Taylor (2001) describes three levels of facial injuries: slight, moderate, and extensive. In the first category, the features are still easily recognisable. Moderate trauma also still allows for features to be determined with relative ease, whereas in extensive facial damage, distortion and obstruction often make a visual assessment difficult and would likely cause distress for relatives or the public. This is why identification efforts of unknown deceased individuals in this category are often supported by post-mortem depictions, which are either drawings or digitally edited post-mortem photographs that hide injuries or post-mortem changes, so someone who knew that person in life would be able to recognise them, without getting distracted and distressed by signs of trauma and post-mortem changes (Wilkinson and Tillotson, 2012; Wilkinson, 2014). In its entirety and during life, the face and facial surface have great potential for identification, as they are a unique makeup of individual features and shapes. In opposition therefore, post-mortem changes and/or facial trauma can have major effects on identification efforts (Gowland and Thompson, 2013b). The interplay of all factors creating the taphonomic environment, and the methodological approaches chosen to recover and analyse the remains, can all impact on the ability and success of establishing a biological profile and/or an identity (Connor, Baigent and Hansen, 2017).

5.4 Taphonomic Research Facilities

Research and subsequent acquisition of knowledge relating to taphonomic processes in humans has been, and still, is rather complicated for a number of reasons. Haglund and Sorg (2006a) argue that the shortage of human cadavers is the main issue, but also mention a lack of collaboration between occupational fields involved in death investigations. One might argue that there is either an apparent lack of understanding in the medical or law enforcement community regarding the value of other disciplines that would be beneficial in this interdisciplinary matter. Or, if such an awareness is present, insufficient resources often prevent effective collaborations not only for research purposes but also in active investigations. To address this issue and create a regulated framework for such incredibly valuable and important research and training opportunities, forensic anthropologist Dr William M. Bass founded the first Anthropological Research Facility in the late 1970s and with that pioneered Taphonomic Research Facilities (TRFs) (Rattenbury, 2018; Oostra *et al.*, 2020). Also known as Human Taphonomy Facilities (HTFs) or Human Decomposition Facilities (HDFs), these are often colloquially referred to as *body farms*, the name originating from Patricia Cornwell's eponymous novel *Body Farm*, first published in 1994 (Cornwell, 2011). As taphonomy is considered a relatively new field, there is undoubtedly still a lot to be discovered and learned with regards to human decomposition. These facilities provide an invaluable opportunity to conduct experimental research, with the potential to apply the outcomes thereof to real life scenarios or, in reverse, to re-create real-life scenarios to gain a better understanding for influencing factors and the timeline of events.

Currently there are eleven such TRFs in operation, that utilise human body donations rather than porcine analogues: Eight are located in the United States, one in Australia, one in Canada, and one in The Netherlands. The first TRF, the University of Tennessee Anthropological Research Facility (ARF), opened in Knoxville at the University of Tennessee in 1981, and is part of the University's Forensic Anthropology Center (FAC). This facility was founded in the late 1970s by Dr William M. Bass who, through his role as consultant forensic anthropologist in police investigations, saw the urgent need for facilitating the controlled study of human decomposition in a variety of different environments (Zejdlik, Passalacqua and Williams, 2018). The ARF is situated in a semi-urban partially wooded area next to the Tennessee River and encompasses 2 acres of land after a recent expansion from originally 1.3 acres. The Forensic Osteology Research Station (FOREST), part of Western Carolina Human Identification Laboratory at Western Carolina University in Cullowhee, North Carolina was the second TRF and opened in 2007. Even though it is merely a two-hour drive away from ARF, the physiographic zones are very different from each other, with FOREST being located in a more rural mountainous region (Zejdlik *et al.*, 2018).

Recognising the importance of establishing TRFs not only in different terrains (forest, mountains, grassland, desert, etc.), but also in different climate zones, and exploring the observed processes with regards to a variety in vertebrate and invertebrate species, makes the findings much more transferrable to real-life scenarios. For example, high acidity and poor drainage of the soil, together with high wind velocity at a low elevation at the Complex for Forensic Anthropology Research (CFAR) at Southern Illinois University, Carbondale IL are important influencing factors for decomposition rate and patterns. In contrast, the Forensic Investigation Research Station (FIRS) at Colorado Mesa University in Grand Junction, is located at a high altitude (1447,8m AMSL) in a very dry climate (8" of rain/year). The recently established Forensic Research Outdoor Station (FROST) at Northern Michigan University in Marquette, Michigan USA and The Secured Research Facility for Thanatology Studies / Site Sécurisé de Recherche en Thanatologie (SSRT) at L' Université du Québec à Trois-Rivières in Canada, are two of three facilities on more northern latitudes. The third cold climate TRF is the Amsterdam Research Initiative for Sub-Surface Taphonomy and Anthropology (ARISTA) in the Netherlands, the first of its kind in Europe (Oostra *et al.*, 2020).

Data for the current research study was collected at the Forensic Anthropology Research Facility (FARF) in San Marcos, Texas. It is part of the Forensic Anthropology Center at Texas State (FACTS), at Texas State University. With its 26 acres, the FARF comprises the largest surface area of any such facility in the world (Texas State University, 2018) and is not only a valuable resource for students and researchers, but also offers training for law enforcement agencies, medico-legal personnel, external research students, and cadaver dog units in search and recovery of human remains. Since opening in 2008, approximately 650 donations have been included in research into time since death, decomposition processes, and the PMI. During the researcher's study visit to this facility in 2019, the director of FACTS Dr Daniel Wescott confirmed that more than 1000 living individuals are currently registered to donate their body to FARF after death. Following decomposition and/or mummification of soft tissues, the skeletal remains of all donors are processed at the Osteology Research and Processing Lab (ORPL) and are subsequently integrated into the permanent Texas State Donated Skeletal Collection.

Unfortunately, in most countries the legal requirements and ethical restrictions still do not allow for human remains to be used for taphonomic research (Knobel *et al.*, 2019). Current efforts are underway to establish such a facility in the UK (Williams, Rogers and Cassella, 2019). There are several TRFs that use pigs and other animals as human analogues, but this has recently been deemed somewhat unreliable, and results from decomposition studies have shown to often be non-transferrable to human cadavers

(Connor *et al.*, 2017; Keough, Myburgh and Steyn, 2017; Knobel *et al.*, 2019). It can be safely assumed that body donations to a TRF in the UK would follow the same ethical and legal guidelines that currently apply to body donations to medical schools, namely informed consent, the Human Tissue Act regulations, and a right to veto a donor's decision by their next of kin (Williams *et al.*, 2019). Until such a facility has been established in the UK, there will be certain deficits in the training and education of students and professionals in the field of forensic science, as very few of them are able to see, practice on, and/or learn from real human cadavers, before being confronted with a real death scene in their careers. This brings implications for both the competitiveness on the global job marked for lack of experience compared to others (e.g., from the U.S.), but also unreliable or non-transferrable decomposition studies and other evidence required for expert witness work in UK Courts of Justice (Cross and Williams, 2017; Williams *et al.*, 2019).

As taphonomic research has shown that natural human variation, differences in diet, lifestyle, medications, etc. and varying climate conditions have a huge impact on decomposition, more such TRFs are needed globally to understand these processes better and develop more reliable models that may aid forensic case work and ultimately benefit the living. It is crucial to mention those facilities and recognise their importance, as they provide an immeasurable amount of valuable research opportunities, and therefore advances and progress in many scientific fields, such as forensic anthropology, biology, taphonomy, entomology, funerary archaeology, cadaver dog/canine unit training, search and rescue team training, etc. Their uniqueness being facilitated by the generous donation of human bodies and in most cases also the provision of ante-mortem data of the deceased donors, all with the intent to enable research studies such as the current project discussed here, that aim to provide answers to important but yet unsolved issues.

6 Face Recognition

6.1 Biometrics

Derived from ancient Greek, biometrics (*bios/βιός* = life; *metron/μέτρον* = measure) is in essence the measurement and surveying of living beings (Burghardt, 2012). Applied to humans, by means of capturing and comparing physiological and behaviouristic features or characteristics, this enables a differentiation between individuals. In the wider literature, a distinction is often made between either soft and hard, or physiological and behavioural biometrics, the latter not corresponding to the former, as illustrated in Figure 13 below (Zewail *et al.*, 2004; Jain *et al.*, 2016). A biometric trait is interested in the "*who am I?*" question, as opposed to "*what can I remember?*" (e.g., password, pin code) or "*what do I have on me?*" (e.g. passport, ID card), ultimately asking what makes someone unique by means of physiological characteristics or behaviour and not via something that could potentially be misplaced, stolen, forgotten, copied, or transferred from one person to another (Ulrich, 2017; Zuo *et al.*, 2019; Bolme *et al.*, 2020). The term *biometric recognition* is defined as "*the automated recognition of individuals based on their biological and behavioural characteristics*" (Jain, Nadnakumar and Ross, 2016;p. 80). According to Jain (2004), the definition of the ideal biometric identifier incorporates the following factors:

<u>Physiological</u>

Uniqueness	The trait should be different in every individual
Permanence	The trait should be stable over time
Universality	Everyone should possess the trait

<u>Technological</u>

Collectability	It must be possible to measure and compare the traits
Performance	The trait must be assessed by means of a scientifically sound method, with the right balance between speed and accuracy

<u>Interactive</u>

Acceptance	Data collection and use must be accepted by society, therefore non-invasive, passive methods of data collection and analysis are preferrable
Circumvention	The technology should be difficult to manipulate or deceive, and robust to spoofing attacks

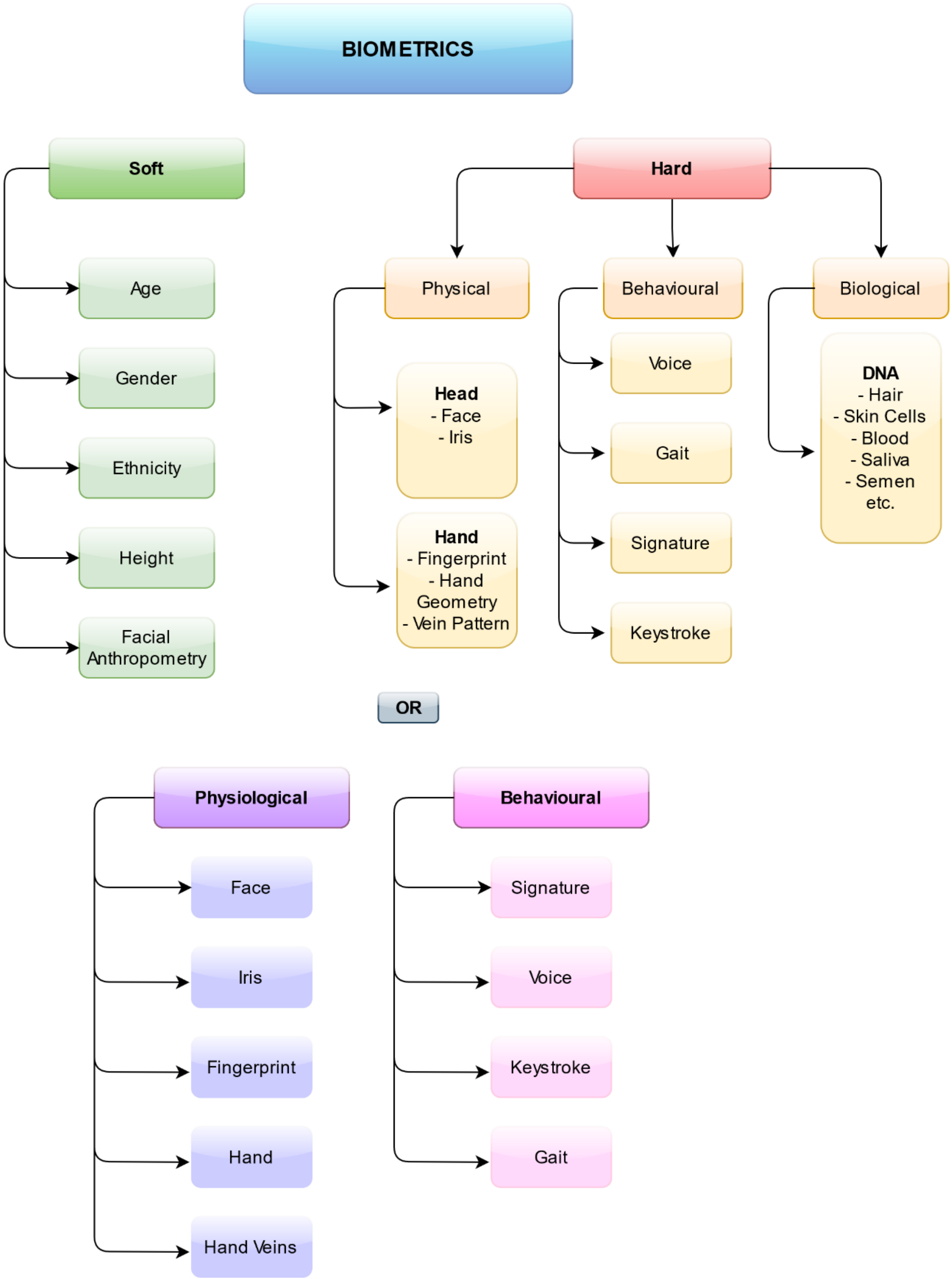

Figure 13: Two categorisation models for biometrics in humans, as suggested in the wider scientific literature (adapted from Stephen and Reddy, 2011; Kodituwakku, 2015).

However, it is important to note that none of the currently employed biometric traits meet all those requirements (Wilson, 2010; Burghardt, 2012; Oloyede *et al.*, 2020). A number of different features have been proposed as biometric identifiers for person recognition/identification (e.g. fingerprints, face, iris, DNA, palmprints, hand geometry, hand vein patterns, signature, voice, gait, keystroke dynamics, ears, scars, skin marks, tattoos), or a combination of different traits, but the most widely used are still fingerprints, iris, and face (Hurley *et al.*, 2008; Bhanu and Govindaraju, 2011; Jain *et al.*, 2016). Since 2005, all European and UK Passports have integrated biometric data in the shape of 2D facial images and some also contain fingerprint information (Bundesamt für Sicherheit in der Informationstechnik and Frauenhofer Institut, 2006). Biometric systems, irrespective of the trait they are designed to focus on, are fundamentally pattern recognition systems that obtain the relevant data (i.e., biometric entity) from an individual, extract a set of features or properties from this entity, and compare it against a template stored in the existing database (Kaur *et al.*, 2020; Oloyede *et al.*, 2020).

As two samples from the same individual are never 100% the same, due to variations in lighting, pose, position, physiology (e.g., bruises, cuts, weight gain/loss) etc., the response or result of the system will be given as a similarity score. This score reflects the evaluation of the input and stored template in comparison, and in most systems, the higher the score, the more similar the two are and the more certain the system is of it being the same individual as opposed to an impostor or unknown person (Jain, 2004; Ulrich, 2017). In other words, the intra-subject samples (from same individual) should be more similar and therefore generate a higher score than inter-subject (from different individuals) data (Jain *et al.*, 2016). Biometric systems are generally used for two types of recognition: verification/authentication or identification (Andress, 2014). As a matter of relevance to the current study, this review will focus predominantly on the entirety of the face as a biometric in face recognition by both humans and computer algorithms.

6.1.1 The Face as a Biometric

The human face can be utilised as a biometric identifier, as it is a unique feature which can be non-invasively, passively or actively observed, and the acceptance of the general public for compliance in this regard is rather high (Tolba *et al.*, 2006; Huang *et al.*, 2011; Kaur *et al.*, 2020). It is also a non-contact biometric identification method, and does not require as much active participation by the user than contact biometric systems, e.g., fingerprints (Guo *et al.*, 2003). The human face possesses a wide range of features that, by themselves, and – especially in combination and relation to one another – are very distinctive. However, the defining statistical model which quantifies this human facial uniqueness, and the inter- and intra-subject variations that come into play in automated face recognition, has yet to be developed (Jain *et al.*, 2016). With rapidly improving systems, one has to bear in mind, that "*the revolution in biometrics marks a new era of surveillance in which the body is both the target and the instrument of control*" (Maguire, 2009; p.9). Therefore, ongoing debates and research need to consider – if not prioritise – data security and privacy rights, instead of solely focusing on technological advances.

6.2 Automated Face Recognition Systems

Automated face recognition (AFR) is a biometric-based method to establish, verify, and/or authenticate someone's identity based on individually unique facial features (Oloyede et al, 2020). All AFR algorithms and systems are currently developed using, and trained on, face databases of living individuals, as their intended application evolved from the need to verify and identify the living. The larger the database, the better and more robust the training, as the algorithm can learn to distinguish between more nuanced differences and learn more diverse features (Hassan *et al.*, 2015). However, this is only true, if there is also diversity in the database with regards to illumination, pose, expression, etc. With state-of-the art systems performing with near perfect accuracy in controlled/constrained environments, the challenge to achieve acceptable results in the wild under unconstrained conditions still remains the largest hurdle of all (Hassan *et al.*, 2015; Kaur *et al.*, 2020; Oloyede *et al.*, 2020).

Recent scientific publications and newspaper articles have highlighted the issue of bias in face recognition, because most training databases only contain Caucasian data that is predominantly male, therefore introducing not only a gender and race bias, but sometimes even making the system so blind to darker skin tones, that the face is not recognised as a face at all (Buolamwini and Gebru, 2018; Castelvecchi, 2020). The idea and ambition to utilise ARF algorithms for the identification of the dead in both humanitarian and forensic scenarios, has only very tentatively emerged over the past decade (Jain, Klare and Park, 2012; Bolme *et al.*, 2016; Gómez *et al.*, 2018; Cornett *et al.*, 2019; Dobay *et al.*, 2020). This introduces yet another challenge for AFR systems that should not be underestimated, namely post-mortem changes to the face (see section 5.3 Post-Mortem Changes to the Face).

6.2.1 Automated Face Recognition Workflow

The basic stages of the automated face recognition workflow are very similar to those of any other biometric system and are designed around the general principles of pattern recognition (Wilber, Shmatikov and Belongie, 2016; Kaur *et al.*, 2020):

1. Data acquisition / Image capture — One or more probe or query images are captured as the required biometric modality via still or CCTV camera in digital format. Most AFR systems are designed for 2D data, but some have experimented with 3D models as query image, which is easier to align to the stored 2D image in the database.

2. Face detection — The face is detected within the overall image and isolated from the background and other potential distractions. Not all the information within the image is required for authentication or identification, which means the data is reduced in size and complexity, allowing for a faster processing time.

3. Pre-processing — Enhancement of image quality as well as face image size normalisation to re-scale all face images to the same size/format and angle. Face alignment, where main features points of the face (eyes, nose, mouth) are

		located, which blends into the next step of feature extraction. This last step however is still a challenge in unconstrained environments (Jin and Tan, 2016).
4.	Feature extraction	A template of the face is generated by extracting a pre-defined set of features or feature properties; this simplifies the large amount of data present and removes the noise and other distractions from the image.
5.	Feature storage / enrolment	In most FR algorithms, the information extracted will be converted into and stored as a feature vector; however, the decision on which features are extracted (i.e., classifiers) and what is emphasised remains a challenge regarding optimal performance of face recognition in the wild (Ding and Tao, 2017). Examples of classifiers used are nearest neighbour, minimum distance, or convolutional neural networks (CNN) (Oloyede *et al.*, 2020).
6.	Feature classification and database matching	The extracted classifiers are compared against those from stored templates within the existing database. Face recognition algorithms operate either holistically or feature based.
7.	Authentication / Verification or Identification	For authentication or verification, the 1:1 comparison will tell whether someone is who they claim to be. In identification, the ID is initially unknown and the aim is to establish it through a 1:n comparison of known individuals (Bowyer *et al.*, 2004). Successful or positive authentication or identification is achieved, if the facial features match to an extent that generates a similarity score which meets the required threshold.

The relevant terminology for ARF systems is defined as follows (ISO/IEC 2382-37:2017 Information technology – vocabulary – Part 37: Biometrics; available under https://www.iso.org/standard/66693.html [accessed 27.05.2021]):

- Type II error: two samples of one person do not match (i.e., false negative)
- Type I error: samples from two different persons match (i.e., false positive)
- Match: defined by the biometric system, when the score assigned to the comparison meets a particular pre-defined threshold
- Similarity scores: the higher the value, the more likely the samples belong to the same person
- Dissimilarity scores: the higher the value, the more likely the samples are from different persons (in face recognition applications, systems traditionally generate similarity scores)
- Genuine, authentic = mate; impostor = non-mate (in verification applications)
- False non-match rate (FNMR) vs. false match rate (FMR)
- Laboratory conditions for image capture = constrained, controlled
- Real-world image capture = uncontrolled, unconstrained, in-the-wild

6.2.2 Historical and Developmental Background

In the early 1960's, with support from the U.S. Department of Defence (DOD), a small research group formed around Woodrow W. Bledsoe, that quickly became an established company called the Panoramic Research Incorporated in Palo Alto, California (Boyer and Boyer, 1991; Raviv, 2020). Their research focused on pattern recognition and is now referred to in the literature as pioneering work on the ability of face recognition by computers (Jain *et al.*, 2016). Their earliest set-ups were "*man-machine*" systems (Bledsoe, 1965, 1966), that required initial manual input from human experts on facial landmark placement on 2D photographs. Data entry was achieved by using a Grafacon/Rand Tablet (see Figure 14 below), that enabled digitisation of the analogue features at approximately 40 images per hour. An automated comparison step performed by the computer resulted in a narrowing down of the face pool to closest matches, presented to a human expert for the final identification decision (Bledsoe, 1964; Boyer and Boyer, 1991).

Figure 14: Photograph of the Rand Tablet (aka. Grafacon) from the RAND Corporation Archives, which is said to be the predecessor for modern iPads and tablets (Rand Corporation, 2021).

The "*first fully automated face recognition system*" (Jain *et al.*, 2016, p.89) was proposed nearly a decade later by Kanade (1973), as a result of his PhD research at Kyoto University. His computer system design performed the first successful human face recognition task, by means of extraction of facial features by the software, via a double computer connection, and feature-by-feature comparison or template-matching against an image database. Between then and now, the advances in computer vision and image processing have made major developments in the field of AFR possible, some of which are listed in rough chronological order in Table 1, with reference links for further reading. The differentiations lie predominantly in the type of feature extraction method utilised within the algorithm (see Oloyede *et al.*, 2020 for a more detailed review).

Table 1: A non-comprehensive list of different ARF techniques and systems specifications, with reference links to key literature and further reading.

ARF Techniques	Key References and Further Reading
Geometrical Feature Matching	(Kanade, 1973; Tamura, Kawai and Mitsumoto, 1996); performs well on small, low-resolution images (8x6 pixels); struggles to cope with larger datasets
Principal Component Analysis (PCA)	(Sirovich and Kirby, 1987); reduces data size and facial characteristics within the images
Artificial Neural Networks (ANN)	(Stonham, 1986; Tolba *et al.*, 2006); issues due to black-box nature and unknown specifications
Eigenfaces	(Turk and Pentland, 1991); aka. Karhunen-Loève expansion; performs well on standardised data; recognition accuracy of 90% on FERET database (Tolba *et al.*, 2006)
Template Matching	(Bichsel, 1991; Brunelli and Poggio, 1993); identification of faces by using only limited regions; simple application but not robust to pose and illumination changes (Kaur *et al.*, 2020)
Fisherface Method	(Belhumeur, Hespanha and Kriegman, 1997); recognition by nearest neighbour; similar to but outperforms Eigenfaces
Linear Discriminant Analysis (LDA)	(Thakur *et al.*, 2009; Kaur *et al.*, 2020); feature selection method; based on Euclidian distance; can be applied to face recognition, tracking and face detection; expanded by Discriminant Common Vectors (DCV) (Abate *et al.*, 2007)
Elastic Bunch Graph Matching	(Lades *et al.*, 1993); extension to ANN (Tolba *et al.*, 2006); model-based face recognition, e.g. Hidden Markov Models (HMM) (Wiskott *et al.*, 1997); based on dynamic link structures; cannot perform well on small images

ARF Techniques	Key References and Further Reading
Hausdorff Distance / Edge Map	(Huttenlocher, Klanderman and Rucklidge, 1993; Guo *et al.*, 2003); doubly modified Hausdorff (M2HD) distance (Takács, 1998); more robust to illumination variations
3D Morphable Model	(Vetter and Poggio, 1997; Blanz and Vetter, 2003)
Line Edge Map (LEM)	(Gao and Leung, 2002); performance equal to Eigenfaces, superior to Edge Map (Tolba *et al.*, 2006)
Support Vector Machine (SMV)	(Phillips, 2001; Elmahmudi and Ugail, 2019); feature extraction with PCA; improved performance over older systems
Viola Jones Face Detection	(Jones and Viola, 2003; Viola, Jones and Snow, 2003; Hefenbrock *et al.*, 2010; Gupta, 2019)
Sparse Representation Coding	(Wright *et al.*, 2008; Wright and Ma, 2010; Yang *et al.*, 2011)
Face Analysis for Commercial Entities (FACE)	(De Marsico *et al.*, 2013); outperformed PCA and others on several databases, incl. Faces in the Wild
Deep Learning / Deep Face Recognition	(Grm *et al.*, 2018; Masi *et al.*, 2019; Wang and Deng, 2021); machine learning techniques; uses convolutional neural networks (CNNs)
MATLAB®	(Nagi, Ahmed and Nagi, 2008); low computational requirements; high processing speed;

Earlier AFR systems have focused on a classification of faces "*based on normalized distances and ratios among feature points*", and the algorithm was designed to create a face model representation by detecting and extracting features individually and analysing the link between them (Turk and Pentland, 1991; p.586). Many older ARF technologies rely solely on the correct programming of the respective algorithm, but the system is not adaptable to change when confronted with more diverse data (Dobay *et al.*, 2020). These models struggled to cope with even slight variations in pose, illumination, and facial expression. Starting with *Eigenfaces* in the early 1990s, face recognition became slowly but increasingly more robust to such influencing factors (Zhao, Han and Shao, 2018).

6.2.3 State of the Art and Current AFR Systems

Current AFR systems, or more specifically their machine learning algorithms, are developed using a training face database, which can vary greatly in number of individuals, images, and conditions under which the latter were captured. Algorithms are now able to adjust and learn from the information that is presented to them. This relatively new approach is referred to as deep learning, a type of machine learning, which uses artificial neural networks (ANN) to process and analyse the data. Many such training datasets are openly available to researchers and as a rule, the bigger and more varied (e.g. illumination, pose, facial expression, race, etc.) the training database, the better the algorithm will be able to perform in terms of accuracy, as it learns more subtle variations between individuals (Tolba *et al.*, 2006). However, those databases are still very small compared to the vast amount of data big tech companies such as Meta Platforms (formerly known as Facebook) and Google Inc. can access to train their respective algorithms (refer to 6.2.7.2 Training Databases and Machine Bias below). Google's FaceNet facial recognition system is currently the best performing in terms of accuracy (99.63%), when applied to the Labelled Faces in the Wild (LFW) dataset, which contains unconstrained images (Schroff *et al.*, 2015). This is closely followed by Facebook's DeepFace, with a recognition accuracy on the same dataset of 97.35% (Taigman *et al.*, 2014). Both systems are based on deep learning and distance metric learning, incorporating CNNs. Zhao, Han and Shao (2018; p.2679) sounded a note of caution with regards to these excellent and promising results, stating that datasets such as the LFW are "*still far from reality*".

It would go beyond the scope of this thesis to include and elaborate on all AFR techniques. It is also not entirely relevant to the research at hand, as the objective here was not to find or even build the ideal AFR algorithm, but to test currently available systems on deceased individuals, which are not cloud based (due to ethical and image rights concerns) and are freely available. However, there are some very comprehensive literature reviews detailing the different approaches and outlining both advantages and disadvantages thereof. For further and more in-depth reading on this subject, please see: Grother *et al.*, 2017; Kaur *et al.*, 2020; Oloyede, Hancke and Myburgh, 2020; Grother, Ngan and Hanaoka, 2021.

6.2.4 2D vs. 3D Face Recognition

In real world applications of face recognition, a differentiation is commonly made between three types of data: Static images, video footage, and 3D models (Ulrich, 2017). For static images, only one or a small number of individual 2D images are used, whereas for video footage, a series of images (screenshots) are considered. A clear advantage of 3D images or facial models is the additional depth and spatial information, which is not visible and available from frontal or lateral view 2D images. The 3D models are also much more adaptable to variations in pose between the query data and database template information, as was recognised by the BKA in their report on using photographs as a search tool in criminal investigations (Bundeskriminalamt, 2007). They conceived that 3D data might be collected from the accused upon booking them and pose variations could be overcome, making the 3D faces more adaptable to 2D images in comparisons. This has yet to be implemented.

Early 3D AFR systems were highly inefficient regarding the required computational power, could not overcome differences in facial expression, and hence were deemed unsuitable for real-life application (Phillip *et al.*, 2003). Gupta *et al.* (2007) developed a new method based on complex-wavelet structural similarity metrics (CW-SSIM), which achieved a 98.6% recognition rate on a database of 360 3D facial images (12 individuals; 30 images per person) and required considerably less processing time. However, no specific information on their dataset regarding possible facial expression or illumination variations was provided by the authors. In 2006, the EU Project *3D-Face* started (funded by the European Commission), which analysed biometric border control systems. In a preliminary report presented at the 10th German IT-Security Congress, Busch and Nouak (2007) stated that the primary objectives were to increase security through 3D systems, as they are expected to outperform 2D systems and are also more resistant against spoof attacks, and achieve a higher recognition accuracy. Their results revealed a <2% false recognition rate (FRR), and a false acceptance rate (FAR) of 0.25%, for a 3D-2D comparison model, piloted at 2 airports in Germany and Austria (Busch *et al.*, 2012). The low error rates are very promising; however, every misidentified person can be a potential security threat.

De Angelis *et al.* (2009) proposed a manual 3D-2D superimposition method for comparisons of a suspect with CCTV footage or still images, based on facial landmarks and a final identification evaluation by an expert. They showed that even partial face scans can be matched extremely well to the target or serve as basis for exclusion. Although recognising that their approach is more qualitative in nature, it incorporates expert human judgement and input, which was entirely disregarded by fully automated, earlier systems (Yoshino *et al.*, 2000; Goos, Alberink and Ruifrok, 2006). A pilot study by Lynnerup *et al.* (2009) – previously mentioned under 3.2 Craniofacial Superimposition above – also managed to overcome head pose variations with ease, by using 3D face models acquired with a handheld laser scanner for identification through superimposition on 2D images. Roughly a decade ago, major difficulties with using 3D data in face recognition lay in often encountering missing parts in the data (incomplete mesh), quality issues with the facial model, or other scan issues that would either obscure or not capture features correctly (e.g., blurring of features, missing parts due to laser absorption in darker areas or hair) (Berretti *et al.*, 2013). The cost of 3D laser scanners for 3D data acquisition was also extremely high (De Angelis *et al.*, 2009).

Since then, the scanning and post-processing technology for creating 3D facial models has significantly improved. The hardware has not only become more user-friendly (i.e., portable, faster, more precise), but the cost has also decreased quite substantially (Guo *et al.*, 2016; Pala *et al.*, 2019). A definite advantage of 3D data is that the acquisition is less dependent on environmental conditions (such as lighting), and the model can be more easily positioned to mimic the head pose in the 2D image template within the database (please refer to section 3.2 Craniofacial Superimposition above). Nevertheless, most current ARF systems still operate on 2D data alone, despite the known advantages of 3D data. Cui *et al.* (2018) presented a face depth estimation method from 2D RGB images, using fully convolutional networks (FCN) and CNNs, which achieved higher accuracy rates than using generic RGB 2D data. 3D face recognition systems are generally divided into either holistic or feature based approaches, the latter being more robust against variations in occlusion, facial expression, and missing data (Guo *et al.*, 2016). Feature-based methods are further sub-divided into landmark-based, patch-based, and curve-based algorithms. With their expression-invariant 3D-3D face recognition method (EI3D), (Guo *et al.*, 2016) measured similarities between two face models and achieved a false acceptance rate of 0.001, with identification rates at 97%, despite variations in facial expression. This approach incorporated local feature matching as well as 3D point cloud registration.

6.2.5 Applications

Nearly 60 years after the first publication on automated recognition based on finger ridge patterns in 1963 (Trauring, 1963), automated fingerprint matching, iris scans, and facial recognition are now routinely applied in a variety of identification and verification scenarios (Kaur *et al.*, 2020; Oloyede *et al.*, 2020; Trokielewicz *et al.*, 2020). If the objective is to authenticate or verify an individual, the AFR system will perform a 1:1 comparison of the input query image to a stored template of the person they are claiming to be, and either accept or reject the individual based on a similarity score either within or outside of a certain pre-defined threshold (Ulrich, 2017). As it is virtually impossible to take the same photograph twice, and through factors such as environmental conditions and biological changes (e.g., illumination, ageing, change of hairstyle, pose), the query image will never match the stored template 100% (Jain, Ross and Prabhakar, 2004; Birngruber *et al.*, 2010). The threshold defines how much leeway the system allows for such intra-subject variations before the difference between images of the same person exceed those from different individuals. Verification is employed, inter alia, for security restricted buildings or devices, visa applications, and e-border control, to grant access and protect from potential impostors. The user is aware of and collaborating with the process (e.g., controlled head pose, facial expression, very little movement).

Identification, on the other hand, is a 1:n (one-to-many) comparison of an unknown individual to a database of known persons (Jain *et al.*, 2016). However, a 1:n comparison can be useful in both authentication and identification, if more than one template image is available per individual in the database, or if more query images of the same person are obtainable. Identification is mostly applied in law enforcement, where a perpetrator is matched to images of a mug shot or driver's licence database, or in the context of child sexual exploitation and human trafficking (Grother and Ngan, 2015; Liu, 2018). AFR has also been applied to counter terrorism operations, when video surveillance is utilised at borders or mass gatherings, to filter through the crowd and find potential suspects (Ganor, 2019). The latter has caused a heated debate in the literature, and algorithms are said to be unreliable in this context, as too many false positive matches are generated, leading to endangering innocent people of being falsely targeted (Bisgaard Munk, 2017; Verhelst, Stannat and Mecacci, 2020). In the identification applications of AFR, the subject is not always aware of or not compliant with the process, which gives rise to some ethical and privacy concerns (Huang *et al.*, 2011). For a more comprehensive list on possible applications of AFR systems, please see Table 2.

Table 2: Areas of application for automated face recognition systems (adapted from (Jafri and Arabnia, 2009; Kaur et al., 2020))

Area of Application	Examples
Law enforcement	Static matching of 2D data with mug shots or driver's license photographs; false identity claims; post-event analysis
Security	Unlocking personal devices; access control points to secure facilities; network or software security
Surveillance	CCTV video surveillance, either day-to-day, in counter terrorism, or the tracking of known suspects; it can also be applied in cases of missing persons to track their movements prior to the disappearance, or for post-event analysis by forensic experts (e.g., robbery)
Identity verification	Checking that the image in an identity document matches the document holder (e.g., border control; voter registration; immigration; driver's licences)
Image database investigations	Missing persons; human trafficking; police criminal checks and bookings
Miscellaneous	Computer graphics; psychology (e.g., facial expression and mood research); neural networks; multidisciplinary approaches in image indexing; face-based searches; advertising; interactive gaming; social media image tagging

Interpol's Face Recognition system (IFRS) has collated images from over 179 countries and is utilised as a global criminal database. Their strong recommendation is to only use photographs with high image quality as the query data (e.g., standardised passport image with neutral background), otherwise the IFRS might fail to deliver results and the accuracy thereof can be compromised (Interpol, 2021).

The tentative application of automated face recognition systems to the identification of the dead has only started to emerge over the past decade. Sauerwein *et al.* (2017) analysed three different biometric traits (iris, fingerprint, face) in manual AM-PM comparisons at the Forensic Anthropology Center (FAC) at the University of Tennessee (UOT). Face data remained recognisable for 3 days during spring/summer and on average 40.75 days during winter. The authors recognised the important value of potentially including more post-mortem biometric analyses in forensic cases in the future. Cornett *et al.* (2019) pointed out the ease of PM data collection (facial images) and were the first to apply and evaluate deep neural network algorithms to the identification of the deceased through PM vs. PM comparisons. Their research is linked to the previously mentioned study by Sauerwein *et al.* (2017), and was conducted using parts of the same dataset (N=52: 544 images) from UOT's FAC. Six categories according to levels of decomposition and possible insect related occlusion were defined. Intake (recently deceased) images were used as their database, and images from the later stages of decomposition as query/test images. Although results are promising, they were unable to provide further information on the algorithms used, due to licencing agreements, and acknowledge that training datasets containing deceased individuals are necessary to achieve better and more reliable results, through sufficient AFR system training.

6.2.6 Independent Evaluation of Existing AFR Systems

Starting in 1995, the DOD initiated the government funded Face Recognition Technology program (FERET), which was a large-scale independent evaluation, to test and review existing face recognition technology (Tolba *et al.*, 2006). After 1996, the program underwent an update, extension, and name change to Facial Recognition Vendor Test (FRVT) (Beveridge *et al.*, 2010), which is regularly conducted and updated by the National

Institute for Standardization (NIST). The German *Bundesamt für Sicherheit in der Informationstechnik* (BSI; = Federal Office for Security in Information Technology) has been analysing face recognition algorithms in various independent projects, such as BioFace, and BioP I and II from 2002 until 2004 (Busch, Daum and Graf, 2003; Bundeskriminalamt, 2006). Between 2004 and 2005, NIST ran the Facial Recognition Grand Challenge (FRGC), with a dataset of 50,000 2D images, which were captured under both controlled and uncontrolled conditions, as well as 3D data (Phillips *et al.*, 2005; Survey, 2007). Results suggested that head pose/position had the biggest impact on recognition accuracy and that most algorithms could not identify individuals from profile view images (Grother *et al.*, 2017).

According to Jain *et al.* (2016; p.92) "*NIST evaluations serve as an excellent resource to benchmark the current recognition performance of various biometric systems*". The most recent FRVT is still ongoing, as of May 2021, and is split into verification and identification testing, with separate reports on each (Grother *et al.*, 2021, 2022). Since the last published report in 2019, two datasets of unconstrained images and further algorithms were added to the FRVT and are being reviewed. The analysis entailed a 1:n identification set-up using 2D facial images. Unfortunately, there is no operational information on individual algorithms available within the report, as prototypes are submitted to the NIST in a black-box format. In line with results from previous reports, accuracy appears to be considerably improving for each evaluation period (2010-2013; 2013-2018; 2018-2020), suggesting that significant developments arise equally or even more fast-paced. On high quality frontal view images, Rank One miss rates are nearing 0.1% (Grother *et al.*, 2021). But real-life factors, such as ageing and facial expressions, remain a challenge.

6.2.7 Factors Influencing and Affecting AFR Systems

AFR systems are now able to achieve near perfect accuracy rates, when operating under controlled conditions and influencing factors, such as lighting, pose, facial expression, and age variations are minimised (Hassan *et al.*, 2015; Oloyede *et al.*, 2020). Same age also refers to capturing all data of the same individual either on the very same day or with very little temporal distance, so that challenges such as ageing or body modifications (e.g., change in hair style, colour, facial hair) are not present (Castelvecchi, 2020). However, in real-life application scenarios, facial query images are usually acquired through CCTV footage, digital cameras, or mobile phones and are often captured under unconstrained and uncontrolled conditions (Jain *et al.*, 2012). But if the AFR systems are trained using idealised databases, it is predictable that they will not perform well when confronted with more realistic data/image variations.

6.2.7.1 Challenges for AFR

Large discrepancies between different images of the same individual can result from changes in facial expression, pose variations, different lighting conditions, camera settings, head/face angles, obstruction of features (e.g., sunglasses, facemasks), ageing, and body modifications (Lin *et al.*, 2017). Changes in facial expression also apply to AM-PM data comparisons, due to the loss of muscle tone and gravity-impacted distortion of facial features. All of those factors affect AFR systems quite significantly, as intra-subject variations can overtake inter-subject differences (Ramanathan *et al.*, 2004; Oloyede *et al.*, 2020). In their review on biometric research of the past 50 years, Jain *et al.* (2016) state five sources for intra-subject variations:

1. Sensor limitations
2. Intrinsic ageing of the biometric trait
3. Variations in user interaction
4. Changes in the acquisition environment
5. All other factors affecting the biometric trait

Body modifications include anything from wearing make-up, to shaving, accidental injuries/scars, and/or undergoing plastic surgery (see section 4.2 Facial Aesthetics and Body Modifications for details). The latter is a relatively new hurdle for ARF, as the number of procedures steadily increases (Nappi *et al.*, 2016). This can have a very minor impact if facial images on social media are no longer being tagged or recognised automatically, or more serious issues unlocking the smartphone via FR, to major complications when getting rejected at border control (Bouguila and Khochtali, 2020).

Face modifications can be used by criminals to intentionally avoid automatic detection but can also cause unexpected issues, if the passport image no longer resembles the live face. Recent studies have found that the performance of AFR in this regard is still highly unsatisfactory, as larger training databases including individuals who have undergone facial procedures need to be expanded and need to be more demographically balanced (De Marsico *et al.*, 2015; Bouguila and Khochtali, 2020; Rathgeb *et al.*, 2020).

Computer vision would classify ageing in the human face as a "*function of face shape and texture in time*" (Osman and Viriri, 2018). The same authors state in their review, that current systems are not suitable to overcoming the ageing issue, and – unsurprisingly – more diverse datasets for training are urgently needed, which incorporate the age parameter in a realistic fashion. A real-life issue in this context is the fact, that ID document images (driver's licences, visa face images, passport images) are not updated regularly, and the face can change considerably during even one decade (as previously outlined in 4.1 Facial Ageing above). The 3D FR approaches mentioned previously have been motivated by the aim to overcome facial expression, pose, and illumination issues (Blanz and Vetter, 2003; Drira *et al.*, 2013). But until 3D systems can be considered for utilisation in the real world, more 3D databases are required for training and their applicability needs further evaluation in the wild. Another possible solution to overcome pose and age variations, would be the implementation of numerous and more varied template images per individual in a database, so face variations for the same person can be learnt better by the algorithm and would not automatically get rejected as a non-match. This is not always practicable, as some current algorithms won't allow for the analysis of multiple images per individual in the face pool.

Issues with scale, resolution, blur, and/or distortion within the image can partially be overcome by image enhancement algorithms, which does not apply to noise (Hassan *et al.*, 2015). As with regards to pose or head angle/orientation, the previously mentioned issue of being unable to align two different 2D objects, or extract the same information from both, applies. Occlusion adds more difficulty for the AFR systems, as important information on the features is lost. No proposed solution has yet managed to overcome this issue. The above mentioned factors, especially in combination (as would be expected in real-life ARF applications), still remain problematic for current ARF systems (see Oloyede *et al.*, 2020 for a comprehensive review on current databases, ARF algorithms, and challenges). Why they are significantly less impactful on the average human (manual) face recognition ability – hence the superior recognition accuracy over machines – is further explained in section 6.3 Human Face Recognition below.

6.2.7.2 Training Databases and Machine Bias

Most training databases will either contain ideal, controlled images, or feature only one or very few constraints. To date, there are only a small number of publicly available databases that feature a true variety of inter- and intra-subject variations in their images, such as Labelled Faces in the Wild (LFW) (Huang *et al.*, 2007), and Public Figures (PubFig) (Kamar *et al.*, 2009). Since more and more people are (mostly) willingly sharing their personal data (incl. facial images) online, big tech companies, such as Google and Facebook, have access to billions of images as training data for their respective algorithms FaceNet (Google) (Schroff *et al.*, 2015) and DeepFace (Facebook) (Taigman *et al.*, 2014). Neither their image pool nor algorithm specifications are publicly available, and if – as previously mentioned – the algorithm's performance is strongly tied to the size and variety of information contained in a training database, the ordinary researcher will not be able to develop an AFR system that meets or even surpasses the accuracy rate of those systems. On a smaller scale, Crosswhite *et al.* (2018) found that the performance of their template adaptation approach plummeted by 19% if the template databases merely contain a single image per person, allowing only 1:1 matching, rather than 1:n. This was reversed when more template images were present, as performance then matched state-of-the-art alternatives.

With regards to quality rather than quantity of training databases, an ongoing debate has emerged in recent years around the matter of racial and gender bias within training databases, that directly translates to a bias in the AFR systems. This issue has recently been brought to a much wider audience through the extremely popular and thought provoking documentary *Coded Bias©* (Kantayya, 2020). Featured amongst other female researchers and scientists is founder of the Algorithmic Justice League (AJL), Joy Buolamwini (AJL, 2021). She refers to algorithmic bias as *coded gaze*, after having discovered that several commercial FR algorithms were unable to *see* (detect) her face as a person of colour, unless she put on a white mask (Ings, 2019). Such *face blindness* can be explained by the lack of diversity in the training dataset, which results in the algorithm being unable to recognise – for example – non-white faces. Further research revealed just that: three major commercial face recognition systems (IBM, Microsoft, Face++) were tested on a new dataset which was balanced by skin type and gender, and all performed best for gender classification on lighter skinned males, worst for darker skinned females. (Buolamwini and Gebru, 2018). In the past, the emphasis was on the algorithms overall performance, but very little attention has been paid to the training database composition. The solution seems reasonably straight forward: identifying any bias, repeatedly auditing existing software, creating more inclusive training datasets, and being aware of the social and economic impact of the biometric system that is created.

After having reviewed a substantial amount of the existing literature on facial recognition, it became evident that most studies only focus on either one or a select few of the challenges mentioned above under 6.2.7 Factors Influencing and Affecting AFR Systems (De Marsico *et al.*, 2013; Drira *et al.*, 2013; Jourabloo and Liu, 2015; Guo *et al.*, 2016; Nappi *et al.*, 2016; Keshtgar *et al.*, 2019; Rathgeb *et al.*, 2020). The same applies for a number of literature reviews on the subject (Zhang and Gao, 2009; Ghiass *et al.*, 2014; Benta and Vaida, 2015). While it could be seen as tackling one problem at a time, by either utilising an existing algorithm, adapting self-same or developing an entirely new approach, initially, this is a plausible and logical strategy. However, research should then be continued by exposing the same system to a different challenge, after the first problem has been solved. If multiple algorithms are created that individually succeed in tackling only one or a select few different challenges, there will not be a satisfactory overall practical outcome or solution in the near future. In addition, each research group – whether academic or industry-based – has slightly different agendas and priorities with regards to the challenges at hand, and with the aforementioned in mind, it is very difficult to make out the overall best performing system. Currently, most of the high-ranking algorithms appear to perform well on a limited number of challenges, but there is no coherence or reliable transferability into a real-world setting (Dodge and Karam, 2016).

6.3 Human Face Recognition

Since the 1950's, there has been an increased interest and research effort in face perception, processing, and recognition (Bruner and Tagiuri, 1954; Gilbert, 1998). The field has therefore become incredibly vast and includes the determination of holistic vs. feature-based face processing (Tanaka and Simonyi, 2016; Towler *et al.*, 2017), the analysis of facial uniqueness and its value in recognition and identification (Lucas and Henneberg, 2015; Balazia *et al.*, 2021), facial expression (Karayanidis *et al.*, 2009; Drira *et al.*, 2013; Benta and Vaida, 2015), face memory (Frowd, 2012; Croydon *et al.*, 2014), right vs. left hemisphere dominance in recognition tasks (Levine, Banich and Koch-Weser, 1988; Schweinberger *et al.*, 2003; Lopatina *et al.*, 2018; Harrison and Strother, 2020), conditions affecting natural face recognition ability, and super-recognisers (Palermo *et al.*, 2017; Phillips *et al.*, 2018; Dunn *et al.*, 2021; Rossion, 2021). Intertwined disciplines include psychology, neuroscience, biology, along with forensic and computer sciences.

The large variability in the unique morphology and geometric constellation of facial features in humans allows for individuals to recognise familiar faces and match unfamiliar faces in scenarios such as everyday interpersonal interactions, for security, verification and/or forensic purposes, essentially both as a form of identification and basis for human social interaction (Bindemann, Attard and Johnston, 2014; Lee *et al.*, 2019). The difficulty lies in the fact that the same individual can look very different in two images, depending on lighting, pose, age, body modification and other variables, whereas different individuals may look similar, even though their identities do not match (Abudarham, Shkiller and Yovel, 2019).

As per the current state of research, humans remain superior to automated systems at face recognition, as we are generally able to overcome variations between the same face in two images or other face representations, especially for familiar faces (Blauch, Behrmann and Plaut, 2021). However, whilst humans appear to be rather good at familiar face recognition, unfamiliar faces can cause significant matching accuracy deterioration, depending on the scenario (Bindemann *et al.*, 2013; Megreya *et al.*, 2013; Towler *et al.*, 2019). The research in this thesis is restricted to unfamiliar faces only, hence the focus of this section will predominantly lie on the human ability to match unfamiliar faces, the *what* and *why* of matching errors, and a brief background on the cognitive neuropsychology involved in face recognition and processing.

6.3.1 Cognitive Neuropsychology Behind Human Face Recognition

The human visual system within the central nervous system, which is tasked with perception, begins to develop its face processing abilities by showing a distinct preference of face-like stimuli, which are first received as separate parts before a holistic reception mechanism is formed (Chellappa, Sinha and Phillips, 2010). Face detection is defined as a process by which humans detect faces within their visual surroundings (e.g., Simpson *et al.*, 2019; Barton, Davies-Thompson and Corrow, 2021). This automatic and incredibly fast mechanism precedes all other face related tasks, such as face recognition, face matching, facial identification, gaze perception, and emotion recognition (Pongakkasira and Bindemann, 2015). Both hemispheres are active during an attempt to match a current representation of a face with one that has already been stored in memory, a process commonly referred to as facial recognition (Carey, Diamond and Woods, 1980). Face recognition takes place predominantly within the inferotemporal cortex, which implies that the process is learned through repeated exposure to facial stimuli during early development (Nelson, 2001). Babies recognise faces as a "*separate class of objects*" during the first 6 months of their life and the utilisation of this particular stimulus precedes the onset of language as a form of non-verbal communication (Nelson, 2001; p.3). New-born babies are already able to distinguish between familiar (e.g., parent) and unfamiliar faces (e.g., stranger) (Nelson and Ludemann, 1989; Johnson and Morton, 1991; Pascalis and de Schonen, 1994).

There is a documented improvement in face memory skills from age five to twelve, possibly due to an overall advancement in processing performance, memory, concentration, and attentiveness (Crookes and McKone, 2009; McKone, Crookes and Kanwisher, 2009). Adult-like recognition performance and accuracy is generally reached in adolescence, with an observed decline around 65 years of age (Chung and Thomson, 1995), attesting to the widely accepted theory that "*cognition improves with development and declines with ageing*" (Megreya and Bindemann, 2015; p.5). However, there appears to be a difference between familiar and unfamiliar face recognition here. The former developing and reaching peak performance earlier in life, with simultaneous maturation of name memory and recognition for inverted faces in the early 20's. Unfamiliar face memory and recognition abilities generally reach their highpoint in the early 30's (Germine, Duchaine and Nakayama, 2011). Early face identity recognition ability arguably becomes expertise in adults, following a lengthy and extensive developmental progress, with deviations from this trajectory in individuals affected by autism and/or prosopagnosia (aka. face blindness) (Pascalis and de Schonen, 1994; Croydon *et al.*, 2014).

The human face provides an abundance of information which we rely on as important social cues, such as personal identity, emotional and health state, as well as approximate age and gender (Karayanidis *et al.*, 2009). The competency to extract this information from visual cues relies on cognitive abilities, which begin to form in early childhood and appear to further develop until adolescence (Bruce and Young, 1986; Hancock *et al.*, 2000). This applies to facial identity as well as facial emotion processing skills (Chung and Thomson, 1995).

Besson *et al.* (2017) compared three different face processing levels: (1) superordinate face categorisation (i.e., detecting human faces amongst other stimuli, such as animal faces), (2) face familiarity (i.e., recognising famous faces amongst unfamiliar faces), and (3) face verification (i.e., recognising a target face). Their findings are in strong agreement with the *superordinate advantage hypothesis*, which suggests that human face categorisation always precedes any other subordinate level, such as familiar or specific target face recognition. Recognising faces and being able to interpret facial expressions is vital for human interaction and therefore assumed to be selected for through evolutionary processes (Nelson, 2001). There are certain conditions, such as autism, prosopagnosia, and later stages of dementia, that appear to negatively affect or altogether inhibit face recognition and face memory abilities (Wang and Olson, 2018; Young and Burton, 2018).

6.3.2 Face Recognition and Matching Ability

Unfamiliar face matching appears to be rather error prone, even under controlled laboratory conditions, with mock eyewitness scenarios producing error rates between 20% and 40% (Shapiro and Penrod, 1986; Bruce *et al.*, 1999; Megreya and Burton, 2008). In the wild, eyewitness identifications from police line-ups were found to be erroneous in 41% of cases (Havard and Memon, 2014). In face matching tasks for unfamiliar faces, error rates appear to be slightly lower at around 10-20%, but still worryingly high, considering the relating experiments were yet again conducted under ideal, controlled conditions (Burton *et al.*, 2010; Bindemann, Avetisyan, *et al.*, 2012). Upon adding imaging variations, such as degradation and/or age differences – which would be expected to be common when checking ID documents in a real life scenario – matching accuracy is further decreased (Kemp *et al.*, 1997; Bindemann *et al.*, 2013; Megreya *et al.*, 2013). It has been claimed that currently used forms of photographic identification are unsuitable and inadequate, given that correct identification in face matching tasks from driver's license and passport images could only be established in 75-79% of cases (Kramer, Mohamed and Hardy, 2019). In another study, when mimicking a border control scenario, accuracy levels reached 85% (Wirth and Carbon, 2017).

6.3.2.1 Super-Recognisers vs. Prosopagnosia

It has been demonstrated and is widely accepted that there is considerable inter-person variability with regards to face recognition abilities (e.g. Balsdon *et al.*, 2018; Phillips *et al.*, 2018; Dunn *et al.*, 2020). This hereditary cognitive skill has been shown to be stable within individuals across different tasks (Dunn *et al.*, 2021). Face recognition ability must be envisioned as a continuous scale, with individuals affected by either developmental or acquired prosopagnosia, who are unable to perform face identity recognition altogether, at the bottom end of the scale (Rossion, 2021). In stark contrast to this are so-called super-recognisers, who consistently excel at face recognition and perception tasks, but for whom there is no formal definition (de Haas, 2021; Ramon, 2021). Super-recognisers tend to present a significant deviation from the mean with regards to their matching and recognition abilities. This invaluable skill has not only been recognised by the scientific community, but is also utilised for real-life forensic case work in the form of specialist task forces, such as the Central Image Investigation Unit (CIIU) at the Metropolitan Police – also known until 2017 as the Super Recognisers Unit (Balsdon *et al.*, 2018; de Haas, 2021). However, super-recognisers make up only 2-3 % of the general population (Dunn *et al.*, 2020).

For the more average human, face recognition varies greatly depending on whether a face is familiar or unfamiliar (Young and Burton, 2017). There is commonly a lack of insight into one's own face matching and recognition abilities, but it has been found that people generally tend to overestimate themselves in this regard (Bindemann, Avetisyan, *et al.*, 2012; Bindemann *et al.*, 2014; Ritchie *et al.*, 2015; Palermo *et al.*, 2017; Towler *et al.*, 2019). This baseline talent or lack thereof appears to be very difficult to improve through training, repeated exposure, face learning or other methods (Young and Burton, 2017). In 1907, Galton published his observations that the median derived from cumulative estimates by a group of people outperformed the best performing single person's guess on the weight of an ox. Over the past century, this crowd-over-individual advantage has been found in many other areas of decision-making, predictive tasks, and judgements (Surowiecki, 2004). The so-called *wisdom of the crowd effect* in facial recognition and identification tasks has been explored by many researchers and is well known in the literature (e.g. Dowsett and Burton, 2015; White, Phillips, *et al.*, 2015; Jeckeln *et al.*, 2018). As an example, Towler, White and Kemp (2017) found that combining responses from seven untrained individuals produced near-perfect accuracy, and collating seven trained examiner's ratings, perfect accuracy was achieved. Further supporting this theory is a study by Phillips *et al.* (2018), which concluded that the rather large difference in face matching ability and hence individual accuracy can be overcome by fusing human judgements, as collaborative efforts of four specialist facial examiners or three super recognisers resulted in no matching errors or 100% accuracy. This is especially relevant for unfamiliar face matching in forensic settings, where these scientific insights should be considered and whenever possible, necessary adjustments to the current procedures implemented to accommodate a joint effort, predicting better accuracy and results (Jeckeln *et al.*, 2018).

6.3.2.2 Human vs. Machine Face Recognition Ability

Humans are said to have a very good recognition ability, which still outperforms most ARF systems, especially under unconstrained conditions in the wild (Hassan *et al.*, 2015). However, it is arguably not quite that clear-cut, as for humans, it depends largely on whether a face is familiar or unfamiliar, on inter-image variations, and on the level of facial examination expertise (Young and Burton, 2017; Heyer, Semmler and Hendrickson, 2018; Devue, Wride and Grimshaw, 2019). For machines, the development and training of the algorithm will determine which hurdles the AFR system is able to overcome or likely to fail at (Ings, 2019; Kortli *et al.*, 2020). It has been suggested that one of the superior human traits is the ability to "*make multiple fixations and* [utilise those to] *perform detailed featural comparisons between image pairs*", an ability that is not available to AFR systems (Blauch, Behrmann and Plaut, 2021; p. 12).

Automatic Border Crossing (ABC) systems (aka. Electronic Passport Gates, e-Gates), in use in the European Union and the United Kingdom, are not actually fully automatic – as many may think – and involve a human controller, who can overturn decisions made by the systems if they are perceived as incorrect or confirm perceived correct decisions (Fysh and Bindemann, 2018). *Perceived*, because it has been shown that humans are not exactly infallible in face matching tasks (Kemp *et al.*, 1997; Burton *et al.*, 2010; Fysh and Bindemann, 2017b). In addition, the human decision or validation process in this set-up appears to be influenced by the machine's preceding identification decision, and is therefore not – as was believed in the past – independent from the system's results (Bindemann, Burton and Jenkins, 2005; McCaffery and Burton, 2016; Fysh and Bindemann, 2018). Heyer and Semmler (2013) voiced concerns over a possible confirmation bias when an initial hypothesis by the system is displayed and needs to be accepted or rejected by a human operator.

According to most recent FRONTEX (2016) guidelines, one human operator can monitor between three to ten e-Gates simultaneously, with the average being five in the field and a time limit of approx. 30 minutes (depending on traveller volume and demand), after which the operator has to take a break, as the task requires enormous concentration and attention to detail. Laboratory tests however have shown that human judgement deteriorates considerably when a person has to make more than one identity decision simultaneously (Megreya and Burton, 2006; Bindemann, Sandford, *et al.*, 2012). In summary, the joint efforts of humans and machines in the field are error prone from both sides, and in combination, meaning that the outcome is often unreliable. Furthermore, there are currently no controls in place that can evaluate accuracy levels in the field, and false acceptance decisions at border controls or other official identity checks can have far reaching consequences.

Heyer, Semmler and Hendrickson (2018) investigated a human-machine collaborative set up, wherein the AFR system would display different length candidate lists in an unfamiliar face matching task. Their results showed a correlation between increasingly inaccurate responses and the increased length of candidate lists for both novices and experienced facial reviewers. Furthermore, the daily amount of comparisons is also linked to accuracy, as the more exposure and therefore training in perceptual learning tasks, the better the results (Bruce and Burton, 2002; Heyer *et al.*, 2018). This contradicts previous findings by White, Kemp, Jenkins, Matheson, *et al.* (2014), who found no difference in performance between trained passport officers and students. Previous research by Heyer and Semmler (2013) found that using ARF systems alongside human examiners can change the decision making process, and to some extent inhibit the training effect for aspiring experts.

There can however be an advantage in the collaborative efforts of humans and machines. For instance, Phillips *et al.* (2018) explored different set-ups with the aim to maximise facial identification accuracy, with a mimicked real-life face matching task using living individual's 2D frontal view photographs. Their findings showed that professional facial examiners (93% accuracy), facial reviewers (87% accuracy), and super recognisers (83% accuracy) outperform control groups (~72% accuracy), and that their median human face matching accuracy is comparable to that of a high performing AFR system (85% accuracy). Best results were achieved when facial examiners and super recognisers worked collaboratively with the AFR system, suggesting that certain weaknesses or shortcomings in humans and machines can be counterbalanced by the other's strengths respectively. Limitations of this study are Caucasian subjects and frontal viewpoint only, and rather high-quality images, therefore future research is required that involves more image variation and a more diverse demographic. An in-depth analysis of the differences, challenges, advantages and disadvantages of humans vs. machines in relation to face recognition and possible collaborative efforts is a task that needs to be addressed by further research (Blauch *et al.*, 2021).

6.3.3 Face Familiarity

Familiar faces are defined as those of people that are known to us, such as family, friends, colleagues, and certain celebrities (depending on individual exposure). Unfamiliar faces on the other hand are those that we may encounter *en masse* on a daily basis, but which essentially belong to strangers. Our own face is considered to be a "*more than highly familiar face*" (Alzueta *et al.*, 2019; p. 100). Humans are generally faster and more accurate at recognising their own face vs. other faces, which is known as the *self-face advantage* (Bortolon and Raffard, 2018). It could be argued that the level and intensity of social bonding is a sliding scale, as most will be more familiar with their mother's face than they are with a colleagues, in other words, making a distinction between the informal and formal relationships and hence closeness or intensity of any given bond. The same applies for celebrity faces. If one has followed an actor's career path or seen multiple movies in which the actor is of a different age and playing a variety of roles, the level of familiarity with the face is considerably higher than having seen one picture in a magazine or having seen a single movie and hence only having internalised the face at a particular age and in a particular state, without much diversity (Kramer *et al.*, 2018). Whilst most of the scientific literature in the past has made a clear binary distinction between familiar and unfamiliar face recognition and face matching (for reviews see Johnston and Edmonds, 2009; and Young and Burton, 2017), it has recently been acknowledged, that it is not quite as simple, and that face familiarity should also be understood as a sliding scale (Kramer *et al.*, 2018).

Different categorisations have been proposed for face recognition over the past few decades (Fysh and Bindemann, 2017b). According to Bruce and Young (1986), human face recognition can be explained by seven different codes, namely structural, pictorial, identity-specific semantic, expression, name, visually derived sematic, and facial speech. However, this applies predominantly to familiar face recognition. When a face is familiar to us, we have the code for an individual's unique combination of features and expressions stored in our memory. This is even robust against certain changes in pose, lighting, facial expression, ageing, and other modifications (Bruce and Young, 1986; Burton *et al.*, 2005; Davis *et al.*, 2012). With repeated exposure to familiar faces and subsequent perceptual learning, we are able to extract identifying, unique characteristics from faces, suggesting that features used in familiar face recognition may vary slightly depending on the individual feature set for a certain person, which is stored in our face memory (Abudarham *et al.*, 2019; Alzueta *et al.*, 2019). Faces change significantly throughout life. These changes are dependent on a number of variables such as age, facial expression, body modifications such as colouring the hair, plucking eyebrows, wearing make-up, etc. It has been suggested that humans are able to cope with changes in style, facial expression, distortion, or lighting to a certain extent when it comes to recognising a familiar face (Bruce and Young, 1986; Burton *et al.*, 2005; Zheng, Mondloch and Segalowitz, 2012; Davies, 2016; Kramer *et al.*, 2018).

Unfamiliar faces remain faces we have not previously been exposed to, but factors, such as age, race, and overall experience and level of exposure to faces over a lifetime seem to play an important role in relation to our ability to identify individuals, even when the face is unfamiliar. Recognition or matching of unfamiliar faces is nevertheless considered challenging and often unreliable, even under optimal conditions (Kemp *et al.*, 1997; Davis and Thasen, 2000; Young and Burton, 2017; Bindemann and Burton, 2021). Bruce and Young (1986) indicated that unfamiliar faces are saved in our memory as a picture, as opposed to a combination of several codes and unique features for familiar faces. This suggests that familiar and unfamiliar faces are processed very differently (Bruce *et al.*, 1999; Towler *et al.*, 2017; Young and Burton, 2017). There are many important real-life applications involving unfamiliar face matching, but the information that can be used in comparisons is limited to what is present in a photograph or other face representation, which is unlikely to be ideal data, such as most images used in face recognition research (Jenkins *et al.*, 2011; Ali, Spreeuwers and Veldhuis, 2012).

In eyewitness or face matching tasks, the *source* of human errors is believed to lie in the missing information or incomplete representation of a face, due to lack of familiarity. When additional image/face variations are introduced (lighting, age difference, expression, etc.), the incomplete mental image becomes increasingly harder to match or associate with the target (Megreya *et al.*, 2013; Bindemann *et al.*, 2014). The *cause* for such errors however may potentially lie in a commonly overestimated, self-centred face recognition ability, probably in parts resulting from the prevailing lack of feedback provided for such tasks (Bindemann *et al.*, 2014). Underlining this theory are findings from two studies, that show clear performance enhancement in cases were feedback was given (Alenezi and Bindemann, 2013; White, R.I. Kemp, Jenkins and Burton, 2014).

With familiar faces feedback is usually instantaneous, as, in social settings, we rely on being able to recognise friends, family, and colleagues instantly and it would be problematic if we were unable to do so. Even with famous faces, which are considered to lie on the familiarity continuum between familiar and unfamiliar, but – depending on exposure – closer to familiar faces, a picture or other media coverage is usually accompanied by a name and context, which helps us to match the face (and name) to a mental image (Bindemann *et al.*, 2014).

There appears to be an overriding terminological issue with regards to *unfamiliar face recognition* in the scientific literature. The dictionary's definition of *recognition*, as in the act of *recognising*, postulates "*the fact of knowing someone or something because you have seen or heard them or experienced it before*" (Cambridge Dictionary, 2021; online). The term *unfamiliar* however contradicts this, as it is commonly used as referring to unknown individuals that have not been encountered before and whose faces are therefore unknown or new to the analyst (Hancock *et al.*, 2000; Young and Burton, 2017; Rossion, 2018). One cannot recognise what is unknown and has not been encountered previously. Hence, u*nfamiliar face recognition* is a contradiction in itself (oxymoron) and should therefore be referred to as unfamiliar face matching or facial comparison instead, with a clear distinction from familiar face recognition.

6.3.4 Face Processing

6.3.4.1 Holistic vs. Feature-Based Face Processing

Holistic face processing involves the innate, automatic human ability or process to recognise faces as a whole rather than a cluster of individual features, that is somewhat variable between different individuals and dependent on factors, such as face familiarity (Galton, 1896; Tanaka and Farah, 1993; Tanaka and Simonyi, 2016; Megreya, 2018). Since it is more of a natural, automatic process, holistic face processing is not considered a method of facial comparison *per se*, but has to be taken into account for all comparative efforts, as it provides an almost unavoidable, subjective human baseline for other methods, and individual holistic matching ability strengths or weaknesses are likely an unknown (because untested) factor in most cases (OSAC, 2021). Richler, Cheung and Gauthier (2011) were the first to empirically study the link between holistic face processing and face recognition abilities, and found that "*differences in both face matching and face identification were related to holistic processing*" (p.470) for the average individual, but it remains unclear as to why individuals with clearly impaired face recognition abilities (e.g. prosopagnosia, schizophrenia) can have normal holistic face processing abilities, when the two are supposedly so closely linked (Maurer *et al.*, 2002; Barton *et al.*, 2021).

Following a long debate over whether faces are processed in parts or as wholes, and which is preferential or more dominant in face recognition and/or comparison tasks (for comprehensive reviews please see Richler and Gauthier (2014), Tanaka and Gordon (2011), and Tanaka and Simonyi (2016)), it was established that unfamiliar faces tend to be processed and compared on a featural level (Hancock *et al.*, 2000; Megreya and Burton, 2006; Megreya *et al.*, 2013), whereas with increasing familiarity, a holistic, more gestalt-like processing mechanism becomes dominant (Megreya and Burton, 2006; Lobmaier and Mast, 2007). Therefore, unfamiliar face matching is similar to other tasks that require a feature-based processing approach, such as non-face object recognition, and individual facial components (e.g., nose, mouth, eyes, ears) become important on their own more so than as a composition. Results from studies investigating the benefits of feature-by-feature comparison training in unfamiliar face matching tasks are somewhat inconsistent (e.g. (Woodhead *et al.*, 1979; Berman and Cutler, 1998; Towler *et al.*, 2014, 2017). Older studies came to the conclusion that a feature-based approach to face matching did not constitute an advantage (Woodhead *et al.*, 1979) and was even a hindrance (Berman and Cutler, 1998), whereas more recent findings by White, Phillips, *et al.* (2015) showed that individuals trained in the feature-by-feature approach outperformed control groups on unfamiliar facial comparisons.

Megreya (2018) found that verbal instructions to concentrate on a feature-by-feature comparison did not improve unfamiliar face matching, and whilst verbal instructions for a holistic focus accelerated the matching speed, they caused a decrease in accuracy. In two different studies, feature instructions for certain features (e.g., eyebrow region) were found to be beneficial to enhancing matching accuracy, even when photographs had been taken months apart (Towler *et al.*, 2017; Megreya and Bindemann, 2018). Research findings from the last two decades also suggests that what constitutes a distinguishing feature in one face or stimulus may be different to the next face, and it may not only vary between different individuals, but also between different images of the same person (Jenkins *et al.*, 2011; Hills and Pake, 2013; Hills, Cooper and Pake, 2013; Megreya and Bindemann, 2018). Furthermore, there appears to be an element of the own-race bias involved, as faces of different races may present different features that are key to defining identity, which need to be considered in order to increase accuracy levels (Hills and Pake, 2013; Hills *et al.*, 2013; Towler *et al.*, 2017; Megreya and Bindemann, 2018).

Abudarham and Yovel (2016) found that different perceptual sensitivity (PS) features are critical in unfamiliar face recognition and identification of Caucasian faces to varying degrees. PS is therein considered a measure for discriminative power. A further distinction was made between high and low PS features, meaning that some features hold more value; fewer high value features are more effective than more low value features in combination (e.g., 4-6 high PS vs. 16 low PS). High PS features are therefore considered *critical features* for identification and changing those would alter the identity of a face for human perception. Their findings showed that the *type* of feature that was changed had more impact than the *level or degree* to which certain features were changed (Abudarham and Yovel, 2016). High PS features tend to be those that remain relatively stable even under varying imaging conditions or facial expressions (e.g., lip thickness, hair colour, eye colour, eye shape, eyebrow thickness, ear protrusion, forehead height, hair length, eye size, skin texture, jaw width, eyebrow texture). Low PS features appear to be more variable and less stable under different imaging or expression conditions (e.g., mouth size, eye distance, chin shape, cheek shape, face proportion, skin colour, nose shape and size). This should be taken into account in any facial comparison tasks, as high PS features would thus hold more potential to be dependable characteristics that could improve accuracy levels.

6.3.4.2 External vs. Internal Features

Whether internal or external features are processed preferentially or dominantly, appears to be strongly linked to face familiarity. In familiar face processing, the focus lies more on internal facial features, such as the eyes, nose, mouth, and eyebrows (Ellis, Shepherd and Davies, 1979; Clutterbuck and Johnston, 2002; Longmore, Liu and Young, 2015), whereas unfamiliar face processing is more affected by changes in external features (e.g., ears, overall face shape, hair, hairline, jawline) (Bruce *et al.*, 1999). With regards to internal facial features, facial identity judgements appear to depend heavily on the upper face (e.g., eyes, forehead), whereas expression or emotion judgements rely predominantly on the mouth and therefore the lower face region (Cunningham and Odom, 1986 in Karayanidis *et al.*, 2009). Faces that are familiar to us are stored in our memory with enough information that we are able to overcome variations in appearance (e.g., facial hair, make-up changes, age and weight differences), or as Young and Burton (2017; p. 215) phrased it *"[m]ultiple exposures to a familiar face allow our perceptual systems to separate transient within-person differences from the stable characteristics of the face*". This largely contributes to making familiar face recognition accuracy much more effective and reliable compared to unfamiliar face matching tasks. Human visual perception mechanisms appear to be somewhat confused when confronted with versatility in unfamiliar faces, which seems strange given our fast exposure to so many different faces on a daily basis (Young and Burton, 2017). If only encountered once and briefly, it appears that a face cannot be sufficiently processed and learned.

The exact definition of inner vs. outer facial features has been somewhat neglected in previous studies, as it was simply implied that the facial contour marks the border, but further investigations into facial periphery and its influence were lacking (Young, Hellawell and Hay, 1987; Frowd *et al.*, 2007). However, studies on unfamiliar face matching that forced participants to focus on internal, rather than external facial features, by obstructing or masking the former, have found that facial periphery is processed holistically and important for matching accuracy, which is lost when only focusing on internal features (García-Zurdo, Frowd and Manzenero, 2018). Additionally, matching accuracy can be improved through what is believed to be linked to face learning from internal features, when external features are not available (Longmore *et al.*, 2015; Kemp *et al.*, 2016), but there appears to be no benefit in providing more viewing time with regards to higher discrimination abilities (Fletcher, Butavicius and Lee, 2008). The fact that internal facial features tend to be more stable over time and external features are more prone to alterations, the natural processing cues for human unfamiliar face matching (e.g., hair) are rather unreliable (Young *et al.*, 1985).

Unfamiliar face matching is an extremely important task in forensic cases. One of the major issues is that eyewitnesses are unable to remember or recall the facial features that are required for identification efforts, since the witness will more likely remember external features but in order to prompt familiar face recognition from composite images, internal features are more important (Ellis *et al.*, 1979; Frowd *et al.*, 2007). In other real-world face comparison tasks, it can be assumed that changes from a living to a deceased face pose even more complications, as internal as well as external features become more varied, and add to other differences (e.g., age, body modifications, weight gain/loss, or image modalities) which in itself appear to be rather difficult to overcome (Bindemann *et al.*, 2014; Young and Burton, 2017).

7 Summary of the Literature

After reviewing the subject of facial identification and related fields of study to date it becomes evident that this is not an isolated topic or discipline and that an interdisciplinary approach is both inevitable and crucial. In summary, there are a few key points that stand out with regards to gaps in knowledge, the need for further validation, and limitations of existing approaches.

Unfamiliar facial identification of the living appears to be highly error prone, even under controlled, assumingly ideal conditions. Most studies tend to focus on one particular method of facial identification, but methods are generally not used in conjunction or hierarchical order, despite FISWG not recommending certain approaches at all (anthropometry) or as a stand-alone method (superimposition) (FISWG, 2019a). There is also a notable imbalance in terms of research focus, as some methods (e.g., facial composites, craniofacial superimposition, craniofacial reconstruction/approximation) are studied more in depth and volume compared to others (e.g., facial comparison/matching).

Overall, there is a lack of standardisation, legislation, and uniform approaches, hence most facial comparison, recognition, and identification work is practitioner led and largely dependent on personal training and experience, as well as on available resources and equipment. There are issues with regards to terminology and feature descriptions, which are generally not used homogenously. Current feature lists are non-exhaustive, although recently published guidelines aim to address this problem (FISWG, 2018a).

It is not defined how many morphological features or landmarks have to align or match to justify either a positive identification or exclusion decision. Validation studies remain extremely scarce in all areas of facial identification, but in particular in relation to using facial data for post-mortem identification. Another issue observed is that error rates are often not stated in the publications. There is also no uniform system for reporting on matching accuracy and results. Those range from similarity scores, levels of support, to rating scales, or an exclusion / inclusion decision without further justification. Again, this appears to be very much practitioner led, and more standardisation, consistency, and coherence are urgently needed to make studies and subsequent results more comparable. The current study ultimately aims to provide an additional piece of the puzzle, as well as highlight the blind spots of research to date that may be addressed in the future.

Facial ageing research has shown that the overall linear process of ageing does not occur at equal pace and intensity in all individuals, as it is linked to natural human variation as well as both intrinsic and extrinsic factors, which can vary widely. Not at all well understood are methods for age estimation in the living based on facial appearance. Certainly, evaluating age from visual cues alone is rather subjective and there are considerable limitations to quantitatively assess highly variable clinical signs of facial ageing. This is nonetheless a factor in facial comparison cases, and age estimation from the deceased face has not been discussed at all at this time, to the best of the researcher's knowledge. This is linked closely to post-mortem resilience of facial features, another topic that has only been discussed in very few studies to date (Hadi and Wilkinson, 2014; Labati *et al.*, 2021). Furthermore, it is yet undetermined, to which extent aesthetic treatments, plastic surgery or merely filters (e.g., social media) affect facial comparison and matching accuracy.

Automated face recognition research has shown that AFR systems are currently able to reach near perfect accuracy rates under controlled conditions (Grother *et al.*, 2017, 2022).

However, to date, humans remain mostly superior to machines in face matching and face recognition scenarios, particularly under unconstrained conditions in the wild. A human-machine collaborative effort and approach may be promising, as certain shortcomings on either side could possibly be overcome and strengths may complement each other. However, this particular area needs to be explored by further research. So far, most ARF studies rely on standardised images, and/or introduce only a limited, controlled number of challenges (e.g., illumination, blur, obstruction, age). Although it appears sensible for systems to initially be exposed to and trained on individual challenges separately, those systems are then not exposed to additional challenges in a consecutive manner. In other words, separate studies focus on different issues, but their findings cannot be amalgamated or collated, as applied algorithm designs vary considerably. Hence, those lab based AFR approaches are not transferable to the wild at this stage, as in reality one is confronted with a complex combination of several challenges and facial image variations. Researchers are therefore calling for more diverse training databases. Merely a handful of studies have explored the potential of applying ARF to the identification and recognition of the deceased (Bolme *et al.*, 2016; Sauerwein *et al.*, 2017; Cornett *et al.*, 2019). The reported recognisability of faces varied widely depending on environmental conditions, and overall findings were limited by data availability issues.

Published research on the applicability of facial identification and recognition methods, both manual and automated, to the identification of the dead to date is very scarce (Alonso *et al.*, 2005; Wilkinson and Lofthouse, 2015; Cornett *et al.*, 2019; Khoo and Mahmood, 2020). This is despite the recognised potential of the unique configuration and morphology of facial features for identification (Kreutz and Verhoff, 2004; Stephan *et al.*, 2008; Alt, 2012; Wilkinson, 2015) and several case studies showing that currently defined primary and even secondary methods of identification (ICRC, 2008; Interpol, 2018) cannot always be applied, for lack of comparative AM data, decompositional changes, limited access and resources, etc. (e.g., Dror, Charlton and Péron, 2006; Hinchliffe, 2007; ICRC, 2008; Hartman *et al.*, 2011; Mulawka, 2014; Wright *et al.*, 2015; Kobus, Kirkbride and Raymond, 2016; Cornett *et al.*, 2019; Khoo and Mahmood, 2020).

Overall, data quantity and quality appears to be one of the largest obstacles in relation to facial identification of the dead, both in laboratory environments and in the field. AM data can often not be controlled for, is mostly circumstantial by nature, and availability thereof can be problematic (Alonso *et al.*, 2005; Cornett *et al.*, 2019). Applicability of visual identification methods is dependent on available resources and state of preservation of remains (Ranson, 2016). Limitations regarding primary identifiers may warrant the use of secondary identifiers, which – in combination – can reach or even surpass the threshold required for positive identification or exclusion. It is however problematic that this threshold is not clearly defined for facial identification. Unlike DNA evidence, where a set number of matching loci is defined as a match (Crown Prosecution Service, 2022), approaches and identification decisions are very much practitioner led, although some guidelines do exist and there are efforts to standardise the relevant methods.

IV. Methodology, Methods and Materials

This chapter describes the background and considerations regarding the current research approach. It also provides an overview of the methods of analysis and data collection, the materials used, as well as ethical considerations relevant to the project.

1 Research Methodology

1.1 Background

The motivation for this research project arose from the ongoing humanitarian identification efforts, in response to manmade and natural disasters occurring on a global scale (e.g., Migrant deaths in the Mediterranean, Africa, and around the US-Mexico border; earthquakes in Nepal 2015, and the Indian Ocean 2012). Primary and secondary methods of identification can often not be applied, due to lack of required AM data (Black and Bikker, 2016). As facial photographs tend to be more readily available, facial identification has and should increasingly be considered in these scenarios (Wilkinson and Tillotson, 2012; Yoshino, 2012; Ranson, 2016; Broach *et al.*, 2017; Caplova, Obertova, *et al.*, 2017, 2018; Nuzzolese *et al.*, 2018; Labati *et al.*, 2021).

In order to find suitable methods that are transferable to the aforementioned context, it appears plausible to focus on the PM face and related identification methods, as well as validation studies thereof. However, most facial comparison research to date has been conducted on living faces and/or under highly controlled, lab-based conditions. Visual identification of human remains has proven both incredibly useful and also prone to error in the past; e.g. following the Estonia ferry disaster 1994, the Bali bombings 2002, the Asian Tsunami 2004, and during the ongoing migrant crisis in the Mediterranean (Soomer *et al.*, 2001; Morgan *et al.*, 2006; Tsokos *et al.*, 2006; Caplova, Obertova, *et al.*, 2017). Hence, there is an urgent need to validate and/or improve existing manual human face comparison approaches, and further to explore the application of manual and automated methods for the identification of the dead. Only a few published (case) studies analysing the PM face exist (Wilkinson and Lofthouse, 2015; Bolme *et al.*, 2016; Čaplová, 2017; Caplova, Obertova, *et al.*, 2018; Cornett *et al.*, 2019), and there is a general lack of validation studies of proposed standards and guidelines (ENFSI, 2018; FISWG, 2018a; OSAC, 2021), which also test transferability to real world conditions (i.e., use of non-standardised data, portability and practicability of equipment). It is obvious that the application of both manual and automated analyses to the identification of the deceased is an extremely neglected area of research.

One explanation for the scarcity of such research is certainly the difficulty of obtaining the required data. As an example, the FAST-ID Project (2010-2013) failed to pursue one of its main objective, namely the "*[c]omparison of AM and early-PM facial images*", stating that "*[n]o suitable research databases of AM and PM images were identified for this research as the use of face images is subject to legal and ethical restrictions with respect to privacy protection and missing or dead persons are not able to give their consent to use their face images*" (Crabbe *et al.*, 2013; p.4) - despite a budget of almost 3 Million Euros and a research collaboration between several industry and academic experts and organisations from four different countries. Current events (e.g., Migrant crisis in the Mediterranean) require facial identification methods to be applied, as primary AM comparative data is not available. However, these approaches have not yet been scientifically tested and evaluated scientifically. Stakeholders are currently awaiting published results on related research to refine their respective approaches and make those more effective in the future.

A number of important ethical, moral, and legal considerations arise, and as an umbrella question, one might ask: who owns the face (Newitz, 2019; Wong, 2021)? And what constitutes the underlying ethical and legal requirements for facial data, and for identifiable data of recently deceased individuals? Facial data is defined as any data related to the human face, including facial maps, meshes, and models, facial landmark and coordinates data, facial biometrics, facial images, and data extracted from any such data that can be used for facial analysis, verification, and/or identification (Law Insider, 2022). Additionally, although there is some reglementation in place (e.g., UNESCO Universal Declaration on Bioethics and Human Rights; UK Data Protection Act 2018; General Data Protection Regulation 2016; Texas revised Uniform Anatomical Gift Act 2009), facial recognition research and post-mortem data in particular appear to fall into somewhat of an ethical and legal limbo. According to the GDPR, a person loses their right to personal data after death. In other words, "information relating to a deceased person does not constitute personal data and therefore is not subject to the UK GDPR" (ICO, 2021; online).

1.2 Research Approach

The exact differences between living and deceased faces have not yet been sufficiently established in the context of both manual and automated face comparison and recognition. It can be tentatively assumed, that methods developed for and tested on living individuals are not always applicable to the PM face, as the latter presents with its own unique challenges (see section III.5.3 Post-Mortem Changes to the Face for details) and that manual methods may require certain adaptations to account for PM related facial changes. One of the aims of the current research is therefore to test applicability of existing manual and automated facial comparison and recognition methods to the deceased.

The PM face generally undergoes a somewhat predictable sequence of changes soon after death (see section III.5.3 Post-Mortem Changes to the Face for details), rendering both manual and automated identification efforts increasingly unreliable, the more those changes progress (Tsokos *et al.*, 2006; Cornett *et al.*, 2019). Hence, only recently deceased individuals were considered in this current project. The element of progressing post-mortem changes, diminishing identifiability and/or stability of facial features following death has been explored in previous studies (Tillotson, 2011; Wilkinson and Tillotson, 2012; Hadi and Wilkinson, 2014; Bolme *et al.*, 2016; Sauerwein *et al.*, 2017; Cornett *et al.*, 2019) and such longitudinal elements were not considered in the current research.

It is known that error rates for unfamiliar face matching in the living are quite high but also largely dependent on the experimental setup and data used (Shapiro and Penrod, 1986; Bruce *et al.*, 1999; Bindemann *et al.*, 2013; Havard and Memon, 2014; White, R.I. Kemp, Jenkins, Matheson, *et al.*, 2014; Megreya, 2018; Kramer *et al.*, 2019). Recognition ability and accuracy for deceased faces in human face matching and recognition tasks remains unknown. The researcher's own face matching ability has not been tested or otherwise evaluated prior to this research (e.g., through Glasgow face matching test), although it is known that this ability varies between individuals (Balsdon *et al.*, 2018; Dunn *et al.*, 2020). There is a known lack of research and especially validation studies regarding manual facial comparison methods in general (Bacci, Houlton, *et al.*, 2021). Furthermore, individual visual observations and manual comparisons are generally very difficult to quantify, even when following guidelines, best practice recommendations, and feature lists.

Although manual and automated approaches tested here may not proof sufficient as a stand-alone method, they may contribute to identification efforts in narrowing down face pools and streamlining efforts, therefore potentially saving valuable resources. It is understood that in casework scenarios, a match or potential decision should always be validated through other methods, whenever possible, prior to definitively confirming an identification.

1.2.1 Manual Comparison Approach

Facial comparison methods – both manual and automated – are predominantly developed on and applied in the context of identification of the living. Only a small number of publications discuss the application of related methods in the context of identifying the dead (Jayaprakash *et al.*, 2001; Lee *et al.*, 2011; Čaplová, 2017; Caplova, Gibelli, *et al.*, 2018; Olivieri *et al.*, 2018; Davis *et al.*, 2019b; Potente *et al.*, 2021). Each method has its own advantages and limitations (please refer to section III.3.1.1 Facial Identification Methods above for details). The FISWG (2019a) recommend morphological comparison as the gold standard. However, in-depth, manual feature-by-feature comparisons are extremely time consuming and not practical for large datasets. (Cranio-) facial superimposition is only recommended in conjunction with other face identification methods (FISWG, 2012, 2019a), and considered more accurate when comparing 3D to 2D data, rather than two 2D objects, as the facial angle cannot reliably be aligned with the latter approach (Damas *et al.*, 2020). Given the aforementioned, it may be advantageous to utilise more than one manual method of facial identification, as this could potentially balance out individual shortcomings of each method. This prompted the application of a hierarchical, combined-methods, 3- or 2-step process (depending on data format) for the manual comparison in the current research.

With the intent to filter down the AM face pool in a fairly rapid and straightforward manner as early on in the process as possible, and therefore streamline subsequent analyses, a preliminary, feature-based comparison was implemented as the first method. By using a feature-based approach, a mere holistic comparison was avoided, as this had been shown to be prone to error and bias in previous research (Wiese, Komes and Schweinberger, 2013; Tanaka and Simonyi, 2016; Megreya, 2018). The chosen preliminary approach is somewhat comparable to a slightly more in-depth passport photograph check in a border control scenario, following brief feature lists and assessing inter-feature relationships.

The second step implemented in the current hierarchical approach (for 3D vs. 2D data only), namely facial superimposition, is not recommended as a stand-alone method and not at all for use on 2D-2D data (ENFSI, 2018; FISWG, 2019a; OSAC, 2021), whereas 3D-2D, or even 3D-3D comparisons allow for a much better and more reliable alignment (Damas, Cordón and Ibáñez, 2020). Differences in shape, feature size, and spatial relationships can be assessed rather well with this method, although limitations remain with regards to differences in facial expression, age, weight, and image distortion. Geomagic® Freeform® Modelling Plus software will be used for this step of the analysis. Limitations associated with this software are that it is only possible to display shape/geometry and any texture information on the 3D model is unavailable. Furthermore, landmark placement is not feasible or reliable in this approach and any deviations between the 2D image and 3D model cannot be quantified. Nevertheless, this approach was recommended by the researcher's supervisor, who had implemented self-same in the past, *inter alia* for an authentication analysis on the death mask of John Dillinger for the FBI (University of Dundee Press Office, 2010). The MEPROCS project has acknowledged that applicability of recommended standards and best practice guidelines with regards to superimposition is always limited by the software and hardware available to practitioners (Damas, Cordón and Ibáñez, 2020). Most of the previous research on (cranio-)facial superimposition to date has focused on skull-to-face comparisons, and to the best of the researcher's knowledge, none have explored the aspect of PM vs. AM face-to-face comparison through superimposition.

A detailed morphological comparison, as recommended by FISWG (2018a, 2019a), was chosen as the final step of the combined methods, hierarchical approach for manual comparison. This is deemed to be very thorough and reliable, but also extremely time consuming and therefore not suitable for large datasets and associated high numbers of comparisons, given it requires the analysis of over 600+ individual features and facial characteristics. As a final step in the hierarchical approach however, this was intended and expected to provide more evidentiary weight to the final decision making. Finer details, such as facial line and wrinkle patterns, hair growth patterns, or detailed ear morphology have been shown to be extremely valuable in distinguishing between faces (Iannarelli, 1989; Stavrianos *et al.*, 2012; Schüler and Obertová, 2020), that in previous stages may have appeared very similar, as those details were likely missed or not considered heretofore. Past research has shown that even shorter and incomplete feature lists and instructions considerably improve matching performance (Towler *et al.*, 2017; Megreya and Bindemann, 2018), therefore using a slight adaptation of the particularly extensive feature list provided by FISWG (2018a) was deemed rather promising. Limitations of this approach are, that research on detailed morphological comparisons to date is severely lacking and it is understood that individualising facial features remain extremely difficult to quantify (Gibelli *et al.*, 2016; Bacci, Houlton, *et al.*, 2021).

Using the aforementioned manual comparison methods in a hierarchical, combined-methods approach allows for the element of a safety net and is intended to act as an effective filtering tool in ultimately reaching the correct identification decision. Previous research has shown that response aggregation, collated results, and combined efforts can considerably improve the accuracy levels in unfamiliar face matching (White et al., 2013, 2015; Dowsett and Burton, 2015; Balsdon et al., 2018; Jeckeln et al., 2018). Terminology and descriptions of facial features, and their respective size, shape, expression, and position or relation to one another, will be based on and derived from previous literature (İşcan, 1993; Vanezis *et al.*, 1996; Dunn and Harrison, 1997; Wilkinson, 2004) and the *Facial Image Comparison Feature List for Morphological Analysis* (FISWG, 2018a). However, it is a known issue, that such descriptions are not universally standardised nor uniformly applied (Caple and Stephan, 2016; Steyn *et al.*, 2018; Bacci, Davimes, *et al.*, 2021). A certain level of subjectivity cannot be avoided when describing size and shape (e.g., small, large, slight, strong), for lack of standardised, uniform, and more delimiting adjectives. However, manual facial comparisons are practitioner led and therefore inherently somewhat subjective, even when following best practice guidelines and standards (ENFSI, 2018).

To evaluate matching decisions in the manual comparisons, a five-point rating scale is used in this research, as opposed to the widely known Bromby scale, which entails six levels of support (Bromby, 2006). The latter has been criticised in the past for not providing transparent guidance on the differentiation and decision making (Steyn et al., 2018). However, any decision making in manual facial comparison has been and still is primarily practitioner-led, and therefore has a somewhat subjective component that cannot be definitively quantified at this stage. It has previously been labelled as "*justified subjectivism*" (Biedermann et al., 2017; p.477) and the lack of validation studies and quantifiable alternatives have resulted in the decision to use somewhat subjective probabilities and more qualitative methods (Schüler and Obertová, 2020). ENFSI (2018) also stated that it is currently impossible to reliably calculate the likelihood of a given face match, as the underlying population data required for such efforts is not yet available. The five-point rating scale used in this research provides a small quantification element to the results and is rather self-explanatory and easy to use (1 = non-match; 2 = probable non-match; 3 = inconclusive; 4 = possible match; 5 = match). A similar five-point scale is commonly used in South Africa for facial comparison casework (Steyn et al., 2018).

The rationale for the order in which the three or two (depending on data format) manual methods were applied, is that the methods increase in the details that are taken into consideration and as a result, also increase in their respective temporal requirements for each 1-to-1 analysis conducted.

1.2.2 Automated Comparison Approach

Manual face comparison methods can be rather time consuming, hence it would be highly beneficial to have an automatic tool that could reliably aid in narrowing down large face pools and potentially even achieve true, positive recognition or identification decisions in a much shorter timeframe. To test the applicability of existing methods to the deceased face, two automated face recognition procedures will be included in this research and results compared to those of the manual comparison outcomes. Past research and recent NIST surveys have shown that most FR algorithms are developed using, and hence only perform well on standardised, frontal view data of living individuals (e.g., Burton et al., 2001; Phillips et al., 2005; Jain, Nadnakumar and Ross, 2016; Grother et al., 2017; Kaur et al., 2020). As the AM data used in this project are purposely not standardised, with the aim to mimic real-life scenarios, there is an expectation that automated systems may struggle to produce accurate results, but it is unknown to which extent. It is also not understood, which impact the element of the dead face (e.g., flaccid muscles and skin, discolouration, bloating, feature obstruction or clouding, decomposition related changes, destroyed or missing tissue, etc.) will have on recognition results, as AFR systems are generally developed for and tested on living faces only.

Although there are several automated face recognition algorithms available, a number of important restrictions in relation to the sensitive data used in this current project apply. Most AFR systems are online, cloud-based applications, where data security and protection of image rights are often not guaranteed (Decker and Ford, 2017; Ashley, 2020; Verhelst *et al.*, 2020; ICO, 2021). As for PM data, informed consent by the participants and donors was only given for research purposes and not for public distribution and third-party utilisation. The only options for AFR systems to be considered in this current study are therefore applications that can operate entirely offline, are user friendly (time constraints), freely available (limited funding), and will not permanently store or distribute data to third parties. The options for freely accessible, offline face recognition software, which do not require a computer science or similar background, are extremely limited.

One of the latter is the MATLAB® *simple face recognition* algorithm by MathWorks, which has achieved good results on standardised, frontal view data (Mathworks UK, 2014, 2021). The code for this *simple face recognition algorithm* (Mathworks UK, 2014, 2021) is freely obtainable online (see Appendix E – *Automated Comparison – MATLAB® Scripts* for details), and the MATLAB®*R2020b* offline desktop application software required is available to the researcher at her institution. No extensive training or coding knowledge is required to operate this AFR option, and following data upload and input, results are generated within mere seconds. Performance is expected to be less accurate for non-standardised data and the added challenge of the PM face.

However, the ultimately extremely poor results obtained from using MATLAB® simple face recognition in this context, led to the implementation of a second (semi-) automated method, for which similar baseline conditions apply. The photo managing software Google Picasa 3 was chosen, as it is available as freeware and easy to use without extensive training. It entails a face recognition function that operates automatically on all images that are imported into the software by the user. Google discontinued supporting the PICASA desktop application as of May 2016 (Betters, 2016; Mediati, 2016) and users have since been encouraged to use Google Photos instead, which operates online and can therefore not be considered for the current project as an alternative option. Past research has found that Google's face recognition function performed better compared to other similar applications (Miller et al., 2015; Schroff, Kalenichenko and Philbin, 2015) and Picasa has successfully been used in research on face recognition and age progression (Liu, 2018). As this software package has been developed on much larger training datasets (Google) and is intended for use on everyday photographs, the researcher expects better recognition results for the non-standardised data compared to those obtained from MATLAB®. The PM element is again an unknown variable with regards to predicted outputs, as training data and intended application here is also in the context of the living face only.

To the best of the researcher's knowledge, no published research is available on applying either the MATLAB® simple face recognition or PICASA 3 approach to the identification of the dead, and what potential impact this additional challenge would pose in terms of recognition accuracy, nor how both algorithms perform on non-standardised data. Both MATLAB®'s *simple face recognition* algorithm and Google PICASA 3 are black box systems, meaning that the detailed functionality and transparency of the algorithm is largely hidden from and therefore unknown to the user (Oscar, 2020).

The researcher is aware that cloud-based services from big tech companies, such as Google AI or Microsoft Azure, are likely more accurate and efficient for data analysis and may provide better results, as they are trained on extremely large and more varied datasets. However, most cloud-based services cannot be entrusted with sensitive data, such as images of deceased individuals, or even just facial data that is not meant to be shared uncontrollably or unintentionally. Input data is often retained, used, and shared without this being explicitly declared by the manufacturer or provider when using such systems (Decker and Ford, 2017; Ashley, 2020; Verhelst *et al.*, 2020; ICO, 2021). This was an extremely restraining factor, as the range of free, offline services in this context is very limited.

1.3 Research Ethics

Potential research collaborations with forensic institutes in the UK and Germany were precluded by the fact that most deceased individuals undergoing post-mortem examinations have been subject to sudden or unexpected deaths. This scenario allows neither for advance, ante-mortem consent by the subject, nor would a next-of-kin in this phase of sudden loss and acute grief be able to give their objective permission on behalf of the deceased, without the mere mention of a research study likely only adding to their distress. Hence, a collaboration with Texas State University was sought, as body donors at their Forensic Anthropology Research Facility either provide informed consent whilst still alive, or a next-of-kin does so on their behalf after the individual has died.

The University Research Ethics Committee (REC) at Liverpool John Moores University approved the application for this project involving deceased individuals on 26th February 2019. The REC reference code is 19/LSA/002. For more detailed information, please refer to Appendix F – *Ethics*. The labels attributed to PM and AM subjects within this thesis have been assigned by the researcher and guarantee anonymity, as well as data protection and privacy rights. Images of the deceased, as well as AM data and foils of donors have intentionally been kept out of the main section of this thesis (with the exception of two photogrammetry models) and can be found in Appendix G – *Dataset images*, as well as the *Supplemental Materials Folder* submitted in conjunction with this thesis (see Appendix H – *Supplemental Material – Guide to Digital files* for details).

2 Research Methods

2.1 Methods of Analysis

2.1.1 Manual Comparisons

In this first section of the main study, three or two (depending on data format) manual, partially computer-based facial identification methods were combined in a hierarchical approach to establish either positive identification or exclusion:

3D PM vs. 2D AM	**2D** PM vs. 2D AM
- Preliminary, feature-based analysis	- Preliminary, feature-based analysis
- Facial superimposition	
- Detailed morphological comparison	- Detailed morphological comparison

Matching decisions were evaluated based on a 5-point rating scale, as detailed in Table 3 below.

Table 3: Rating score for manual comparison on 1-5 scale with explanations for each scoring category.

Scale	Definition
1 = non-match	Definite excluding differences observed, that cannot be explained by natural changes, e.g., age, weight, health or body modifications within reason, no distinct similarities.
2 = probable non-match	Definite excluding differences observed, some features may have been obstructed, some may seem slightly similar.
3 = inconclusive	No or very little excluding differences observed, as well as similarities that warrant further analysis, e.g., similar face shape and feature relationship/size.
4 = possible match	No excluding differences observed, some similarities found that correspond with PM subject, some features may be obstructed and/or age/weight/health/body modification changes can explain differences
5 = match	No excluding differences found, similarities observed correspond with PM subject in identifying features, e.g., line/wrinkle pattern, ear shape, dental features, moles or scar pattern

After each stage, non-matches (1) and probable non-matches (2) were excluded from further analysis, in an attempt to filter down the AM face pool, to ultimately obtain a match (5) or possible match (4) for the PM subject after the last step of the comparison. Inconclusive (3) identification decisions were retained and moved along to the next step of the comparison. The analysis was conducted as a blind study approach, as AM-PM matches were not known to the researcher at this point. The final identification decision was verified by the researcher's director of studies, who held the key to the true matches in the face pools. Figure 15 below shows a flowchart of the algorithm followed for the manual comparison process.

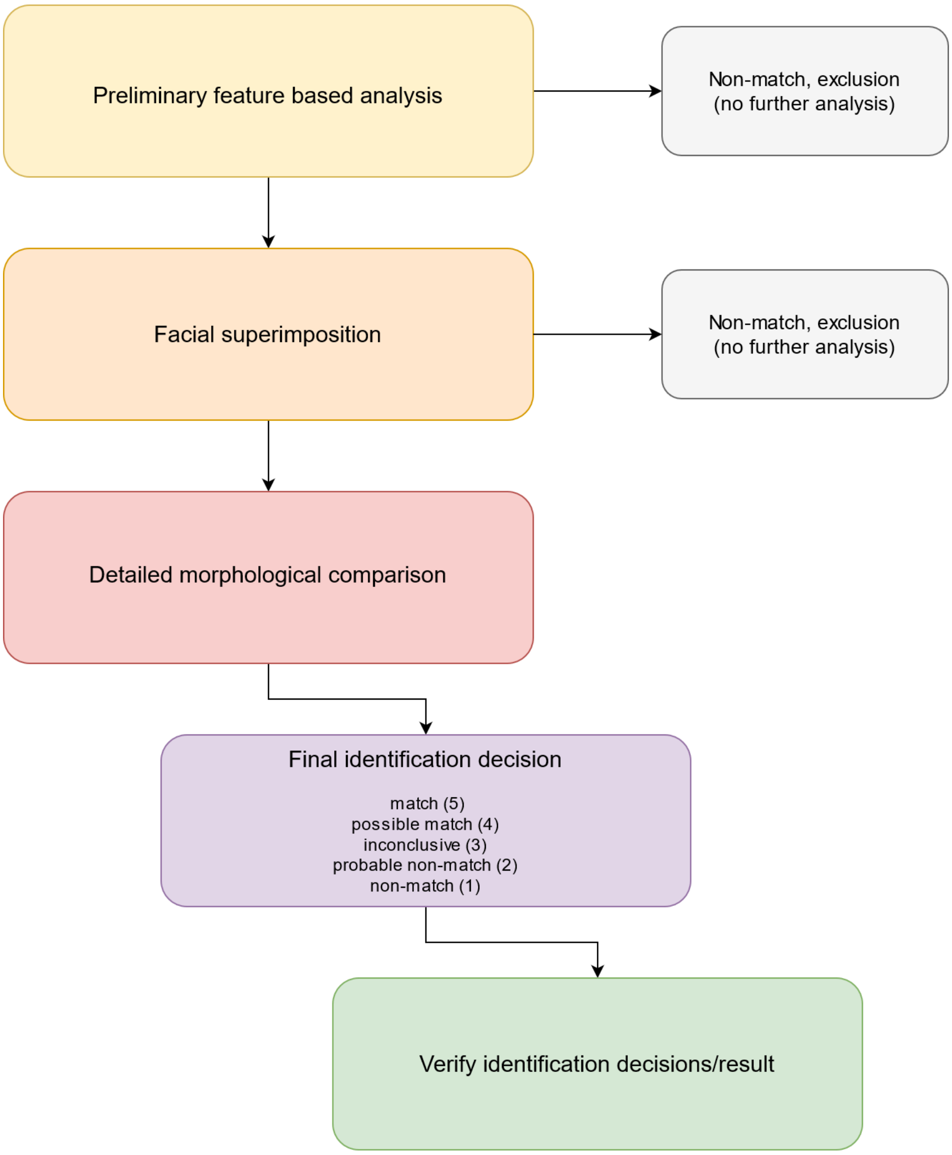

Figure 15: Flowchart showing the algorithm followed for the manual comparison process of 3D PM facial surface scan models & 3D photogrammetry model vs. 2D AM facial photographs.

2.1.1.1 Preliminary Feature-Based Analysis

During the first step of the manual comparison process, the preliminary, feature-based analysis, a 3D PM face model (N=3) or 2D PM facial photographs (N=6) of the PM subject were compared 1-to1 against an AM face pool of 30 2D facial images (N=60; 30 males, 30 females; separate face pools used). Features considered during the comparison process of the preliminary analysis are summarised in Table 4 below. Each feature was described in accordance with FISWG guidelines and other published literature (as mentioned in the methodology section above), with a clear focus on differences that would warrant exclusion from the face pool (e.g., ear lobes attached vs. detached, hairline straight vs. widow's peak, and/or differences in facial line and wrinkle pattern). The focus was on overall face shape, relative dimensions, and inter-feature relationships within the face, without using exact anthropometric measurements (e.g., forehead height is approximately ½ of forehead width, nasolabial height is approximately 1/3 of labio-mental height). Shape and size in relation to other morphological structures were described and compared, age- or weight-related changes, and other body modifications were also noted (see also Figure 16 below). In general, outer/external features were considered first (e.g., neck, hair, ears), then internal features (e.g., eyebrows, eyes, nose, mouth) in a top to bottom manner, before focusing on finer details such as lines, wrinkles and skin marks.

Where individual features and/or overall appearance of the AM face did not potentially match the PM face, and noted differences could not be explained by ageing processes, weight changes, or body modifications alone, the individuals were eliminated from the AM face pool and not included in the next stages of the analysis. If no excluding differences were visible, or the analysis was inconclusive, the AM photograph was kept in the face pool for further investigation in the next step of the comparison.

Table 4: Features considered during the preliminary, feature based analysis and reasons for inclusion/exclusion with corresponding rating scores.

Preliminary Analysis		
Features Considered	**Exclusion**	**Inclusion**
– Overall face shape – Upper face shape & rough dimensions – Lower face shape & rough dimensions – Brow ridge – Eyebrow length / shape – Eye shape / details – Nasal tip shape / details – Nasal root / bridge – Alae – Ear shape / detail / protrusion / size – Jaw line – Chin shape / details – Facial marks / wrinkles / lines – Neck shape / details – Hairline	Individuals from AM face pool eliminated based on incompatible feature differences with PM subject. *Rating score: 1 or 2* *1 = non-match* *2 = probable non-match*	If no excluding features were found, or feature changes could be explained due to changes in age, weight, body modification (within reason), health issues, differences in facial expression, AM individuals remained in face pool and moved on to the next stage of the analysis. *Rating score: 3, 4, or 5* *3 = inconclusive* *4 = possible match* *5 = match*

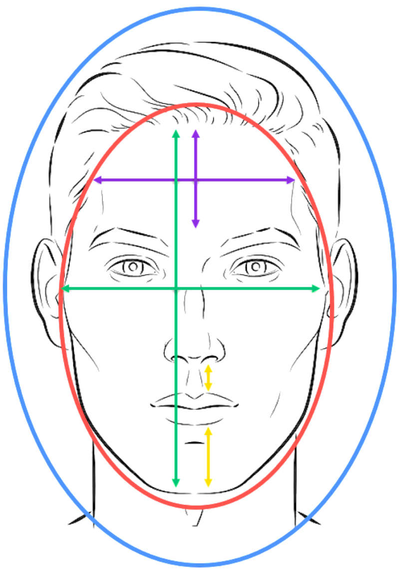

Figure 16: Facial features considered during preliminary analysis: external features (blue), internal features (red), inter-feature relationships (e.g., forehead height vs. width ratio (purple)).

Facial comparisons in this first step were performed with a two-screen set up, viewing 3D PM face models in Artec Studio Professional v.14 software (3D laser scan models) and Agisoft Metashape Standard Edition 1.7.2. (photogrammetry model), and 2D images in Windows Picture Viewer side-by-side. Each comparison was achieved in a timeframe of approximately 10 minutes (3D-2D), with more lenient identification decisions and assigned rating scores, aiming to prevent incorrect exclusion of the matching AM targets prematurely.

For 2D PM vs. 2D AM data comparisons, the preliminary analysis was conducted slightly more in depth, to compensate for the missing second step of facial superimposition, and approximately double the time (15-20 mins) was spent on comparing each set of faces. The main goal of this step was to filter down the face pool quite considerably and relatively swiftly, as the second and third steps comprise of methods that are significantly more time consuming. Therefore, having less AM images in the pool, hence less comparisons to perform, was intended to make the overall process more efficient.

2.1.1.2 Facial Superimposition (for 3D vs. 2D only)

The second step of the manual comparison process for 3D-2D data analysis involved a computer based manual facial superimposition of the 3D PM face models (subject) and 2D AM photographs from the face pool (possible target; foils), using Geomagic® Freeform® Modelling Plus software and Geomagic® Touch™ Haptic Device. Only AM photographs and corresponding individuals that were assigned a rating score of 3, 4, or 5 in the preceding preliminary feature-based analysis, and therefore remained in the face pool, were considered in this stage for facial superimposition. The exact step-by-step workflow for the facial superimposition can be found in Appendix D – *Manual Comparison – Superimposition Workflow and Morphological Feature List*. Texture is lost, when importing 3D data into the software, which only shows shape/geometry. Furthermore, as it was not possible to import the photogrammetry model, superimpositions of the PM photogrammetry model and AM photographs were conducted using the Glass2k freeware, which allows for windows to be made transparent on screen, so two different software applications can be overlayed. The photogrammetry model was displayed, manually rotated, and scaled in Agisoft Metashape Standard Edition 1.7.2. and AM photographs in Windows Picture Viewer respectively.

In each case, superimpositions were evaluated by comparing alignment and overall shape correspondence and relationships of individual features within the faces. The focus was on morphological differences that would warrant exclusion. The last comparative statement would evaluate whether there were any distinct morphological differences between the 3D model and 2D photograph and based on the previous observations for the same superimposition with one AM individual and the respective PM subject. A short conclusion was drawn, summarising key similarities and/or differences between the two faces/objects. Geomagic® Freeform® Modelling Plus software does not allow for marking landmarks on either the model or the face and no measurements between landmarks can be obtained. The alignment was therefore assessed on a subjective, practitioner-led basis (see Figure 17 below for a superimposition example). The facial superimposition was concluded by assigning a score on the previously mentioned 1-5 ranking scale. Again, only AM images that scored a 3, 4, or 5 were included in the next and final step of the manual comparison process, all others were eliminated at this stage.

Figure 17: Screenshot of a manual superimposition example (2D facial photograph, 3D face model) produced using Geomagic® Freeform® Modelling Plus software and Geomagic® Touch™ Haptic Device (pretend-dead, living individual) (image: author’s own; used with permission).

2.1.1.3 Morphological Comparison

The final step of the manual comparison section of the analysis – irrespective of data format – involved a detailed morphological comparison, which was based on a slightly adapted version of the *Facial Image Comparison Feature List for Morphological Analysis* compiled and recommended by (FISWG, 2018a). Overall, more than 600 individual features and characteristics (= parameters) were considered and compared for each subject. Please refer to Appendix D – *Manual Comparison – Superimposition Workflow and Morphological Feature List* for the complete features list used in this current research. The remaining 3D PM or 2D PM face data was compared to the remaining AM images in the face pool. 3D PM face models were viewed in Artec Studio Professional v.14, 2D PM and AM facial images in Windows Photo Viewer respectively, in a two-screen, side-by-side set up. The 3D PM photogrammetry model was displayed using Agisoft Metashape Standard Edition 1.7.2. A morphological assessment was created for each PM subject, followed by separate assessments for the remaining individuals in the AM face pool. Those were then compared 1-to-1 and evaluated as to whether differences could be explained by natural changes (e.g., age, weight, body modification) or were not compatible with a (possible) identification match, and leading to an exclusion.

2.1.2 Automated Comparisons

The automated analysis section of this research was divided into two parts, as detailed below:

2.1 Automated Comparison – MATLAB®
2.2 Semi-Automated Comparison – Picasa

2.1.2.1 MATLAB® - MathWorks' Simple Face Recognition

The MATLAB®*R2020b* offline desktop application software, and a slightly adapted version of the *simple face recognition algorithm* (Mathworks UK, 2014, 2021) were used in the first automated comparison approach in this current research (please refer to Appendix E – *Automated Comparison – MATLAB® Scripts* for all MATLAB® scripts and codes used). Some minor adjustments to the code were required to adapt it to the PM-AM database and formats used, which were implemented with support from David Burton MEng (Hons) CEng MIET ASEP MINCOSE CPRE, Systems Engineering Consultant. The *simple face recognition algorithm* (Mathworks UK, 2014, 2021) in MATLAB® matches the face of an input image (= PM subject) to a person from the gallery (= AM face pool). The MATLAB® script/algorithm is merely a template, and the researcher's own data (PM vs. AM) was imported and used here; as opposed to a generic, readily available face database. Face verification, a subset of face recognition, then determines if the two images presented belong to the same person. The result output is automatically provided as a similarity score on a scale from -1 to 0 (incl. 4 place values after the decimal point; e.g., "-0.0561"), with 0 being a perfect match – as deemed by the algorithm. The final stage required for the observer/researcher to either verify or reject the face recognition match decisions.

Figure 18 below shows an example of the output format in MATLAB® for one subject and related comparisons made, with the highest similarity score assigned to the best match according to the algorithm (here: false positive / incorrect match).

person = 6 **PM subject ID**

scores = 1×62 single row vector **number of AM subjects in the pool / number of comparisons made**

Columns 1 through 18

-0.0097 -0.0119 -0.0195 -0.0108 -0.0158 -0.0216 -0.0150 -0.0175 -0.0129 -0.0147 -0.0193 -0.0185 -0.0192 -0.0217 -0.0187 -0.0207 -0.0102 -0.0198

Columns 19 through 36

-0.0121 -0.0169 -0.0191 -0.0168 -0.0133 -0.0159 -0.0236 -0.0233 -0.0194 -0.0114 -0.0126 -0.0083 -0.0196 -0.0175 -0.0069 -0.0157 -0.0199 -0.0220

Columns 37 through 54 **similarity score for best match (according to MATLAB)**

-0.0138 -0.0192 -0.0148 -0.0139 -0.0209 -0.0195 -0.0141 -0.0202 -0.0181 -0.0186 -0.0098 -0.0161 -0.0156 -0.0148 -0.0104 -0.0190 -0.0101 -0.0070

Columns 55 through 62

-0.0174 -0.0180 -0.0226 -0.0180 -0.0090 -0.0128 -0.0180 -0.0165

integerIndex = 33 **location of match within AM face pool (AM subject no. 33/62)**

Figure 18 Example of the raw output in MATLAB® with similarity score ratings from -1 to 0 (incl. 4 place values after the decimal point).

2.1.2.2 Google Picasa 3

Google Picasa 3's (semi-) automated face recognition was tested on the same baseline data as before. In this software, photos can be imported, tagged, edited, and sorted into groups or albums. To use the face recognition function, a person/face in an image has to be manually tagged, after the software detected a face within the image, and Picasa will automatically run a comparison with all uploaded images to find the same person/face in other photos. Recognised faces are then grouped together, and matching results need to be verified and manually counted by the human operator.

The following entails only a brief description of the workflow that was followed in the experiments using Google Picasa 3 for semi-automatic face recognition. A considerably more extensive description of the Picasa workflow can be found in Appendix I – *Automated Comparison - Picasa Workflow*.

- Folder manager settings were set to "remove from Picasa" for all but one folder, the latter containing data relevant to the respective experiment was set to "scan always"
- To avoid bias in the face recognition process, all files for one experiment were imported at the same time, hence organised in a single folder (one PM subject's PM data files + combined AM face pool images with augmented AM images)
- Import -> Import from [e.g., removable hard drive] -> import to [create temporary folder on laptop/PC] -> Import all
- Picasa automatically scans all imported images for faces (i.e., face detection; approximately 5-10mins)
- None of the files had yet been tagged or named; face photographs (as detected by Picasa) were not displayed under "People" -> "unnamed"
- Picasa creates groups for face images; groups are faces of individuals that the software perceives to belong to the same person/individual
- "Expand groups" to see all files individually. If PM files are not displayed under "unnamed", Picasa did not recognise the files as face images, and need to be manually tagged; Can find undetected PM face images under the temporary folder in Picasa (here: 'Picasa trail')

- To tag PM subjects (i.e., labelling the images), select single PM image -> double click after “expanding groups” -> in new window, select blue bust symbol on bottom right; If face was detected in image, sidebar on the right will display “People – who is in the photos?” with option “add a name” -> PM subject was manually tagged/labelled by entering their unique identifier (e.g., DM3) as a name -> enter -> new window opens -> “create a new person” -> select “new person” -> “OK”.
- If Picasa did not recognise a file as containing a face, find image under “Picasa trail” [= temporary folder created earlier in Picasa] -> double click on image -> in new window, select blue bust symbol on bottom right -> select face region manually by adjusting the frame -> tag/label as before
- Picasa creates an album for all labelled PM images (ideally, containing all imported PM images for a single subject). Generally, one PM image is tagged/named and Picasa will automatically recognise all or most of the other PM images for that person from the imported folder. Those “recognised” PM files will be displayed in the folder and need to be manually approved or rejected. If not all PM data files were found by the software, the missing files may need to be tagged manually, as described before.
- After tagging/labelling is complete, return to “library” [top left]
- Picasa now automatically scans all files in the imported folder (i.e., face pool) for matches or similar faces, which are then either displayed as matches or suggestions in the PM subject’s folder
- Difference between match and suggestion: Match images are automatically added to the PM folder; suggestions are also added but appear with two options underneath the image, so the user can either “accept” [green tick] or “reject” [red cross] the suggestion
- After manual acceptance and/or rejection, Picasa has *learned* more information about the face and will scan all imported files again
- The final recognition results are manually counted as numbers of correctly recognised AM images, after no further suggestions are presented by the software
- After each experiment -> delete “person” folder (remove tags) created by Picasa when tagging PM images -> delete subject album -> delete all files from ‘Picasa

trail' folder on the laptop/PC -> remove/replace old with new PM subject folder (incl. PM and AM data) for next experiment

Picasa provides options for a "suggestion threshold" and a "cluster threshold", both can be manually set anywhere between a 50% and 95% range. A low suggestion threshold results in more "suggestions", and potentially less accurate true "matches"; a high suggestion threshold will only suggest or match face data that Picasa is very confident about. A low cluster threshold groups more faces together on the naming page but may get some/more wrong; a high cluster threshold means that only very similar photographs/faces (as perceived by the software) are grouped together. In the current study, the cluster threshold was less important, as the PM subject images were tagged manually and the initial automatic grouping following data import was largely ignored. Each experiment analysed recognition rates of Picasa for a single PM subject against a combined male and female AM face pool. Each experiment was run three times, at three different threshold settings.

- Under "Tools" -> "Options" -> "Name tags" -> "Threshold" -> set threshold to 70%, 60% and 50% respectively for each experiment [3 separate analyses per PM subject].
- Between testing face recognition in Picasa at different thresholds, all data had to be deleted and re-imported, as described above to avoid bias.

Unlike MATLAB® simple face recognition algorithm, Google Picasa 3 does not provide an output in the format of a similarity score, likelihood ratio, or other numerical value that is comparable. Results are displayed in the shape of facial images only (either matched or suggested) in a subject folder, after tagging/naming the PM subject images. Those images have to be manually counted, which was done three times per PM subject. The first count occurs after all PM images of one subject were tagged or labelled and Picasa has automatically scanned the entire folder, aiming to find matching data. The second count was done after initial suggestions were either rejected (false positive) or accepted (true positive). For count one and two, the researcher was also interested in how many correct and incorrect matches were found/assigned by Picasa. After rejecting suggestions, those are then no longer displayed in the folder. The third and final count occurs at the end, when no further suggestions are displayed. When counting displayed results or matches in this manner, it can be determined how many, if any, of the matching 12 AM face images Picasa recognised either incorrectly (false positive) or correctly (true positive).

2.1 Methods of Data Collection

All data for this research (3D PM, 2D PM, 2D AM) were collected by the researcher at the Forensic Anthropology Research Facility (FARF), and the Osteology Research and Processing Lab (ORPL) at Texas State (TXST) University, San Marcos USA during a ten-week research visit between 11th March and 24th May 2019. The research collaboration with Texas State University was officially confirmed by Dr Daniel Wescott on 22nd January 2019 and is ongoing. All subjects are 18+ years of age and representative of the current Texan population (World Population Review, 2022). However, PM subjects are predominantly elderly Caucasians. The latter is entirely coincidental, as data collection was only performed over the period of 2.5 months, and the deceased are a random sample of registered donor deaths which occurred during this period, and donors that were already present at the research facility at that time. However, neither the exact age of the deceased (other than 18+), nor the donor's age at the time the AM photographs had been taken is known to the researcher. Aspects of biological sex, age, and ethnic origin in the dataset were purely circumstantial and not pre-selected for. Only recently deceased individuals were included.

3D PM facial laser surface scans (N=12), and 2D facial photographs (N=12; 12 individuals, 2 images per person) were captured during *intake* (includes taking hair, nail, and blood samples, photography of the donor body, mesh-bagging of hands and feet to prevent scavengers displacing bones, etc.) at ORPL, when donations arrived at the facility and were processed prior to being placed outside at FARF by undergraduate and graduate student volunteers from TXST University. Donors were lying in a supine position when facial data (scans, photographs) were collected. Only recently deceased donors were considered for this study, so only the recently deceased and donors that had been stored in freezers (e.g., to be used later on in workshops at FARF) were used. No names, age, date of birth, or other personal information (e.g., medical history, cause of death) was collected from the donors, and the donor numbers assigned during intake at ORPL at TXST were further anonymised and adapted for the purpose of this study by the researcher (e.g., DM3 = 'Deceased Male 3').

All 2D AM photographs (N=60) had previously been provided by families and next-of-kin when body donation forms were signed and submitted to FACTS. Therefore, AM data had been captured under uncontrolled conditions, hence non-standardised, showing great variation in terms of image quality, resolution, pose, size, illumination, distortion, facial expressions, weight gain/loss, age differences, and occlusion (e.g., glasses, hair). These circumstances closely mimic real-life identification scenarios, in which relatives, friends, or law enforcement agencies often provide similar non-standardised AM photographs for comparison.

The number of available images per person varied (N=60; 1-3 images per person; mean = 1.15). All AM data was compiled by Laney Faser (final year MSc Forensic Anthropology student at TSXT in 2019) at FACTS (100+ AM images), stored on a passport protected external drive and given to the researcher's director of studies, who compiled the AM face pools for this research, enabling a blind-study approach. There was no set limit to the number of AM images per individual, but for most, only a single image existed in the database. AM images were collected from the existing donor database at TXST and chosen on the basis of the face pool including respective AM target images and foils matching the approximate age range and ethnic background present in the PM sample. Final AM face pools for this study were created from all AM data provided by TXST by the researcher's director of studies Prof Wilkinson at LJMU, ensuring a blind-study approach and the presence of target AM images within the respective pools.

The 3D PM data was obtained using an Artec Spider handheld 3D scanner and corresponding Artec Studio v.14 Professional software for capture and post-processing of the face models. The Artec Spider handheld laser surface scanner has a remarkably high scan accuracy with a resolution of 0.1mm and precision of up to 0.05mm (GoMeasure3D, 2017). The scanner has 3x3D cameras, 6 LED flashing lights surrounding the texture camera in the middle, and a regular flash for 3D capture. The Artec Spider allows for real-time capture and creation of 3D models and is able to record up to 7.5 pictures per second. The measuring field of the Artec Spider ranges from 9x7cm to 18x35cm, with the option to combine several fields or scans into one model in the post processing stage. An advantage of this scanner model is the texture tracking function, which allows the user to also obtain texture information for the model, rather than geometry/shape only, as would be the case with lower end laser scanners (e.g., Artec Eva) or CT scans. The Artec Spider handheld laser scanner was one of the best portable models available at the time of data collection (pre-2020) and has since been replaced by the follow-up model, the Artec Space Spider (Kersten, Przybilla and Lindstaedt, 2016; Tokkari et al., 2017). All scanning equipment used in this research is rather expensive but was readily available at Face Lab, LJMU and the researcher had familiarised herself with the scanner and software during other projects prior to conducting this research (e.g., Beaumont et al., 2018; Ord, 2018).

Due to technical problems with this particular scanner model and issues with matching AM data availability for the 3D facial surface scan models obtained, one photogrammetry model was also included in the 3D vs. 2D manual comparison section of this research. The respective data for which had also been collected in 2019 by a fellow visiting researcher at FARF, Dr Clara Alfsdotter, who kindly shared their data with the researcher after obtaining permission from Dr Daniel Wescott at TXST.

The 2D PM frontal and lateral facial photographs were taken using a Panasonic Lumix Bridge Camera DMC-FZ2000EB, following FISWG's *Standard Guide for Postmortem Facial Image Capture* (FISWG, 2018b) whenever possible.

3 Materials

A total of 12 PM 3D facial laser surface scans were collected at ORPL, TXST in 2019. However, due to equipment malfunction and sensitivity, only six scans were usable. Issues resulted from a combination of prolonged warm-up times and subsequent failure to calibrate within the timeframe of notification of arriving donors and intake procedures, as well as the Artec Spider scanner being unable to capture a complete mesh surface from frozen/defrosting remains, which displayed a more reflective skin surface (condensation) compared to donations stored in coolers only. It had been communicated previously that all donors have matching AM data available, which was not the case. From the six complete scans, unfortunately only two had matching AM photographs in the database. Table xyz below gives an overview of scanned donations, usable scan quality, and matching AM data availability (see Table 5 below for details). On the basis of this, one PM 3D photogrammetry model was then included in the materials, to have a total of 3 PM subjects available for the 3D PM vs. 2D AM comparison section of this research.

Table 5: Table providing an overview of all 12 donations scanned, successful scans obtained (6) and corresponding AM data availability (6); No. 13 refers to the photogrammetry model.

Donors		**Available data**		
No.	Assigned Code	3D PM	2D PM	2D AM
1	DF1	-	X	X
2	-	-	X	-
3	-	-	X	-
4	-	X	X	-
5	DM1	-	X	X
6	DM2	-	X	X
7	DF2	-	X	X
8	-	X	X	-
9	DF3	X	X	X
10	-	X	X	-
11	DM3	X	X	X
12	-	X	X	-
(13)	DF4	X	-	X

2D PM facial photographs were also collected from each donor (one frontal, one lateral), but only 6 PM subjects could be used in the 2D PM vs. 2D AM comparison section, due to the lack of matching AM data as described above.

The AM 3D facial photographs in the face pool consisted of 30 males and 30 females respectively. Separate face pools were used in the comparison, depending on the PM subject's sex in 1:1 comparisons for all 30 AM images (1 target, 29 foils; N=30).

3.1 Manual Comparison

The following materials were included in the manual comparison sections of this research, depending on PM data format (3D/2D) explored:

<u>**3D** PM vs. 2D AM</u>:

PM data:

- 3 3D PM subjects (1 male, 2 females)
    - 2 PM facial laser surface scans (male, female)
    - 1 photogrammetry model (female)

AM data:

- A total of 60 AM subjects (30 males, 30 females)
    - Mostly only 1 image per person in the AM pool (mean = 1.15 per person)
    - 1 target AM image per PM subject; 29 foils in the respective AM face pool

<u>**2D** PM vs. 2D AM:</u>

PM data:

- 6 2D PM subjects (3 males, 3 females)
    - 1 frontal, 1 lateral PM facial photograph per PM subject

AM data:

- A total of 60 AM subjects (30 males, 30 females)
    - Mostly only 1 image per person in the AM pool (mean .15 per person)
    - 1 target AM image per PM subject; 29 foils in the respective AM face pool

Figure 19 below represents examples of 3D PM, 2D PM and 2D AM data of a living individual pretending to be deceased and respective AM images, as all original dataset images (with the exception of two photogrammetry models) were intentionally kept out of the main thesis to adhere to data protection and permission protocols. Those can be found in Appendix G – *Dataset images*.

Figure 19: Examples of a frontal screenshot of 3D (pretend) PM, 2D (pretend) PM and three AM images of the same individual (images used with permission by Josie Ide).

3.2 Automated Comparison

The same baseline dataset as in the Manual Comparison section was used for the Automated Comparisons, to ensure comparability with regards to matching accuracy and recognition ability. An overview of the data used in both automated comparisons (MATLAB® and Picasa) is provided below:

Mimic **3D** PM vs. 2D AM

PM data

- 3D PM subjects (1 male, 2 females) (N=3)
    - PM facial laser surface scans (1 male, 1 female)
    - photogrammetry model (1 female)
        - 30 2D screenshots for each 3D model/PM subject

AM data

- A total of 60 AM subjects (30 males, 30 females) (N=30)
    - 1 image per person in the AM pool
        - Artificially augmented to 8 per person
    - 1 target AM image per PM subject; 29 foils in the respective AM face pool

2D PM vs. 2D AM

PM data

- 6 2D PM subjects (3 males, 3 females) (N=6)
    - 1 frontal, 1 lateral PM facial photograph per PM subject

AM data

- A total of 60 AM subjects (30 males, 30 females) (N=30)
    - 1 image per person in the AM pool
        - Artificially augmented to 8 per person
    - 1 target AM image per PM subject; 29 foils in the respective AM face pool

Different data combinations were explored, hence each results section in V Results. entails a brief description of materials used at the beginning, which led to the respective results.

Due to the software (both MATLAB® and Picasa) only allowing for 2D data input, 30 screenshots per 3D PM subject were taken and used as ‘mock 3D’ data. Each original AM image per person was artificially augmented to eight images, using the images processing toolbox in MATLAB®. This created a larger face pool, since a single image per person does not allow the algorithm to *learn* a face well. Augmentation was achieved by rotating, mirroring, adding noise, scaling, blur, reflection, and grey scale (see Figure 20 below). This process allowed for the AM face pool to be artificially enhanced from 1 image per person to 8 images per person, or from 61 to 488 images in the AM face pool respectively. As number 9 was missing from the original male face pool but there was a number 31, the researcher included her husband as placeholder for number 9 and hence a total of 61 subjects are part of the face pool in this section. MATLAB® codes used for resizing and image augmentation can be found in Appendix E – *Automated Comparison – MATLAB® Scripts*. Figure 20 below gives an example of a single AM image artificially augmented to 12, representing and introducing different image variations (e.g., blur, zoom, rotation, noise, illumination changes, grain), as used in the automated comparisons. All original dataset images were intentionally kept out of the main thesis (with the exception of two photogrammetry models) to adhere to data protection and permission protocols. Those can be found in Appendix G – *Dataset images*.

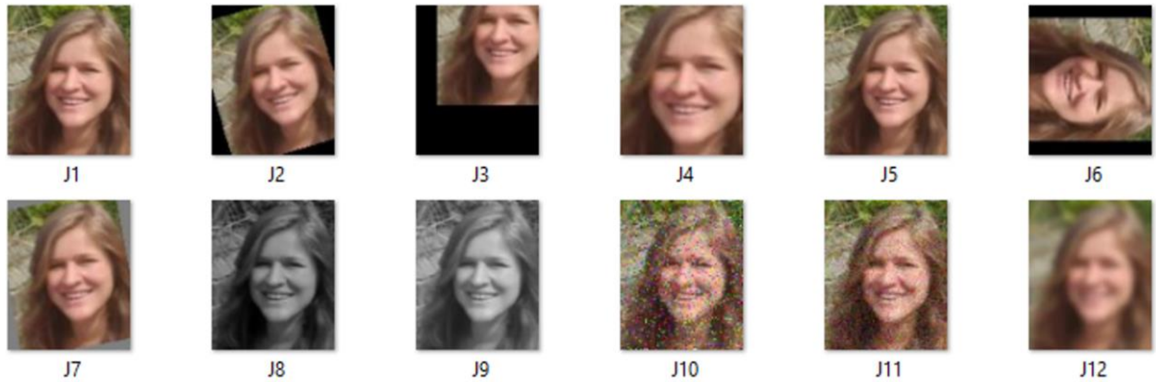

Figure 20: J1 shows the original image, J2-J12 represent examples of image augmentation options used to enhance the number of images per individual in the face pool for automated comparison.

3.2.1 MATLAB®

All PM and AM images were manually cropped to a size that is devisable by 112x92 to avoid distortion in the next step, then automatically resized in MATLAB® v.R2020b to fit the 112x92 pixel requirement of the face recognition algorithm. If the manual cropping step is not included, automatic resizing will cause distortion of the image content (i.e., face). In the current approach, PM images are generally used as *query image* in the comparison. Only one image per PM subject can be considered by the algorithm, but multiple PM subjects can be compared against the face pool at any one time. In a real-life scenario, the deceased would be the unknown and the aim is to match the unknown subject with a known AM image/identity from a pool of AM data/individuals (e.g., missing persons database). To re-create the 3D-2D approach from the manual analysis (section 1A), multiple 2D screenshots of the 3D PM facial surface scan models needed to be included. For the algorithm to consider multiple PM images per individual, the pools had to be switched from multiple AM to single AM and single PM to multiple PM images per person respectively.

However, this means that for the mimic 3D to 2D comparisons, the AM face pool (now marked as "*targets*" in the code) only contained one image per individual (no augmented data), whereas the PM subjects ("*pool*") contained 30 images from different angles per PM subject. With the simple face recognition code used, there was no option to consider multiple AM and multiple PM images at the same time. For the 2D-2D comparisons, AM images always served as the *pool* and contained eight AM images per person. Artificially augmented images of the same photograph were used, rather than a variety of different images from the same individual. The differences in the datasets and modes of comparison are listed in *Table 6* below.

Table 6: Summary of different data combinations considered in the automated analysis, the respective comparison set-ups and description explaining main differences and reasoning for inclusion.

MATLAB® – Automated Analysis			
Dataset		**Comparison Modes**	**Description**
Main Study Data	3D PM face models; 2D AM facial images	1. Multiple PM vs. single AM 2. Single PM vs. multiple AM	Multiple 2D screenshots of 3D models (30pp); single AM images (artificially augmented to 8pp); actual deceased individuals; only 1 AM image per person available; trying to mimic 3D data, but also consider potential impact of augmented AM images
	2D PM facial images; 2D AM facial images	3. Multiple PM vs. single AM 4. Single PM vs. multiple AM	2D PM facial images (2pp; frontal, lateral); single AM images (artificially augmented to 8pp); actual deceased individuals; only 1 AM image per person available; can only consider 1 PM image pp. if using augmented AM data (multiple), when testing more than one PM image, only 1 AM image pp can be considered

3.2.2 Google Picasa 3

The same baseline data was used as in the previous manual and automated (MATLAB®) comparison sections. The experiments on semi-automated face recognition using Google Picasa 3 is sub-divided into two sections:

2.2.1 – 3D PM vs. 2D AM semi-automated comparison - Picasa

2.2.2 – 2D PM vs. 2D AM semi-automated comparison – Picasa

Separate folders for each PM subject were created, containing the respective PM data plus all AM face pool images in the augmented format. Augmented images have the original AM pictures as a baseline, from which image variations are created (e.g., tilt, blur, black and white, noise, etc.). One original AM image per individual in the face pool was used with an additional 11 artificially created augmented images (see Appendix E – *Automated Comparison – MATLAB® Scripts* for details), so 732 images in total for a combined face pool of male and female AM data. Below is a summary of the data used, depending on data format considered:

3D PM vs. 2D AM

PM data

- 30 screenshots of each 3D PM face model (N=3; 90 images)

AM data

- 732 AM images for the face pool (= 31 individuals x 12; incl. 11 artificially augmented original images)

2D PM vs. 2D AM

PM data

- four 2D PM photographs per subject (N=6; 24 images)

AM data

- 732 AM images in the face pool (= 31 individuals x 12; incl. 11 artificially augmented original images)

V. Results

The results section is divided into two main parts, both are further subdivided as follows:

- 1. Manual Comparisons
    - 1.1 3D PM vs. 2D AM
    - 1.2 2D PM vs. 2D AM

- 2. Automated Comparisons
    - 2.1 MATLAB®
    - 2.2 Google Picasa 3

The first part aimed to establish whether using different manual facial comparison methods in combination would provide accurate and reliable unfamiliar face matching results, and whether implementing multiple steps would be a reliable tool to filter down a larger face pool with the possibility of applying this to real-life identification scenarios as a pre-screening tool. In actual casework, this could then be followed by primary identification methods to confirm identification decisions.

In the second part, face recognition ability of one automated and one semi-automated system were tested under real-life conditions with unconstrained data, providing most of the known challenges to the respective system's algorithm. The results aim to determine whether those or similar systems could be applied in cases involving non-standardised AM data and recent PM data for identification purposes, and whether the recognition accuracy would be superior, similar, or inferior to that of a human observer conducting manual facial comparisons.

Subject images have intentionally been kept out of the main text of this thesis (with the exception of two photogrammetry models), in order to comply with image usage permission and to protect privacy rights of both donors and participants. The dataset images can be found in Appendix G – *Dataset images*, and also in the Supplemental Materials folder submitted in conjunction with this thesis (see Appendix H –

Supplemental Material – Guide to Digital files for details); the latter also containing all screenshots, augmented, and resized images and their respective face pools from the analyses conducted.

1 Part 1 – Manual Comparisons

The manual comparison results section is divided into two sub-sections as follows, depending on PM data format:

- 1. Manual Comparisons
  - 1.1 3D PM vs. 2D AM
  - 1.2 2D PM vs. 2D AM

1.1 Manual Comparison – 3D PM Facial Surface Scans & Photogrammetry Model vs. 2D AM Photographs

The results section for *2A Manual comparison – 3D PM facial surface scan models & 3D photogrammetry model vs. 2D AM photographs* is organised per PM subject (N=3) and hence in three main sections, each further subdivided into three sections by type of analysis conducted (preliminary analysis, facial superimposition, detailed morphological comparison). Each of the three PM subjects were compared against an AM face pool of 30 individuals (separate face pools for males and females; 1 target AM image per PM subject). Figure 21 below provides an overview of the manual 3D vs. 2D results, with AM face pool numbers included at each step, and final identification decisions respectively. All assigned rating scores and results are also summarised per PM subject in a table at the end of the respective sections. More detailed notes on all stages of the manual analyses can be found in Appendix C –*Manual Comparison – Notes* and the supplemental material. Please see Appendix H – *Supplemental Material – Guide to Digital files* for details.

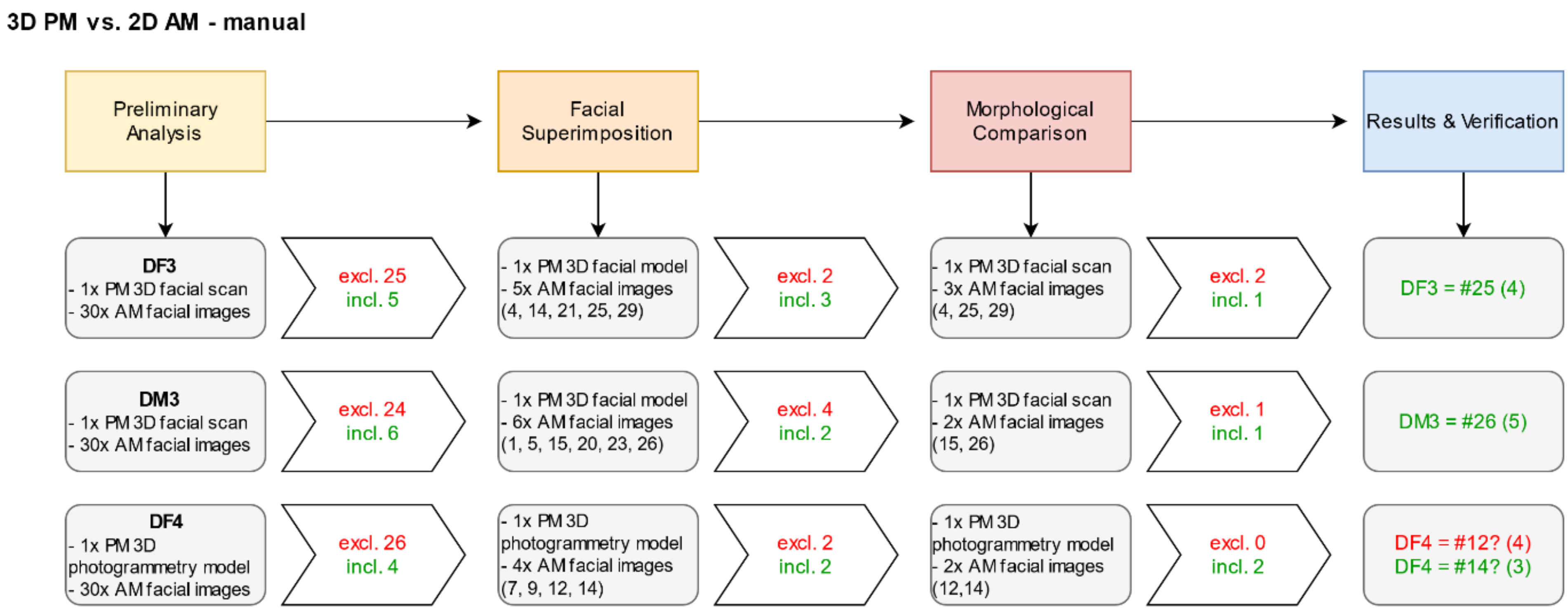

Figure 21: Flowchart showing the different stages of the 3D-2D manual comparison workflow with face pool numbers for included individuals, and results with final ranking score and confirmed matches highlighted in green. DF4 was wrongly matched with AM face no. 12 (false positive), the actual match was no. 14 (false positive).

1.1.1 1A.4.1 Subject DM3 – PM Face Scan Model vs. 2D AM Photographs

1.1.1.1 Preliminary, Feature Based Analysis: 3D PM vs. 2D AM

The 3D PM face scan model of DM3 was assessed, and then compared against all males in the face pool (N=30) and a rating score from 1-5 assigned to each comparison by the researcher.

During the preliminary feature-based analysis of PM subject DM3, 24 individuals were excluded from the AM face pool with rating scores of either 1 or 2. Six individuals (1, 5, 15, 20, 23, 26) remained in the pool due to rating scores of 3 or higher (potential matches) and were moved along to the next step of the comparison process.

1.1.1.2 Facial Superimposition: 3D PM vs. 2D AM

After the facial superimposition of PM subject DM3, four individuals were excluded from the face pool (1, 5, 20, 23) with rating scores of either 1 or 2. Two individuals (15, 26) remained in the pool due to rating scores of 3 or higher (potential matches) and were moved along to the next step of the comparison process.

1.1.1.3 Detailed Morphological Comparison: 3D PM vs. 2D AM

Following a detailed morphological comparison of PM subject DM3 vs. no. 15 and no. 26 from the AM face pool, no. 15 was eventually excluded (score: 1) as the differences observed were not compatible with DM3, whereas no. 26 was given a rating score of 5 and deemed as an AM match for the PM subject. The differences observed between no. 26 and DM3 could all be explained due to PM changes to the face, distorted features due to lack of muscle tone or autopsy damage to soft and hard tissues, and differences between the scan image and a photograph (e.g., moles on forehead not clearly visible in scan, distortion, illumination, and differences).

The match was confirmed as correct. Table 7 below shows a summary of each analysis in the manual comparison process for 3D vs. 2D with PM subject DM3, and respective rating scores assigned per individual in the AM face pool at each step.

Table 7: Summary of manual 3D-2D comparison results for male subject DM3, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

3D-2D MALE		AM Pool Number Males																													
		1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
DM3	*Preliminary Comparison*	3	1	2	2	3	1	1	1	1	1	1	2	1	4	1	1	1	1	3	1	1	3	1	1	4	1	1	1	1	1
3D Facial Surface Scan	*Facial Superimposition*	1	.	.	.	2	.	.	.	.	.	.	.	.	4	.	.	.	.	2	.	.	2	.	.	4	.	.	.	.	.
	Morphological Analysis	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	5	.	.	.	.	.
	Final Score	**1**	**1**	**2**	**2**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**5**	**1**	**1**	**1**	**1**	**1**

1.1.2 Subject DF3 – PM Face Scan Model vs. 2D AM Photographs

1.1.2.1 Preliminary, Feature Based Analysis: 3D PM vs. 2D AM

The 3D PM face scan model of DF3 was compared against all females in the face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

During the preliminary, feature based analysis of PM subject DF3, 25 individuals were excluded from the face pool with rating scores of either 1 or 2 and five individuals (4, 14, 21, 25, 29) remained in the pool due to rating scores of 3 or higher (potential matches) and were moved along to the next step of the comparison process.

1.1.2.2 Facial Superimposition: 3D PM vs. 2D AM

During the facial superimposition of PM subject DF3, two individuals were excluded from the face pool (14, 21) with rating scores of either 1 or 2 and three individuals (4, 25, 29) remained in the pool due to rating scores of 3 or higher (potential matches) and were moved along to the next and final step of the comparison process, a detailed morphological comparison.

1.1.2.3 Detailed Morphological Comparison: 3D PM vs. 2D AM

Following a detailed morphological comparison of PM subject DF3 vs. no. 4, 25, and 29 from the AM face pool, no. 4 and 29 were eventually excluded (score: 2) as the differences observed were not compatible with subject DF3. Especially differences in facial line and crease patterns and facial dimensions were deciding factors for the respective rating scores. No. 25 was given a rating score of 4 and deemed as possible AM match for the PM subject. The differences observed between no. 25 and DF3 can all be explained by PM changes to the face, distorted features due to lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight.

No. 25 was confirmed as true positive and correct match for DF3. Table 8 below shows a summary of each analysis in the manual comparison process for 3D-2D with PM subject DF3, and respective rating scores assigned per individual in the AM face pool at each step.

Table 8: Summary of manual 3D-2D comparison results for female subject DF3, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

3D-2D FEMALE		**AM pool number Females**																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
DF3	*Preliminary Comparison*	1	1	1	3	1	1	1	1	2	1	1	1	1	3	1	1	2	1	1	1	3	1	1	1	3	1	1	1	3	1
3D Facial Surface Scan	*Facial Superimposition*	.	.	.	3	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	1	.	.	.	4	.	.	.	3	.
	Morpho-logical Analysis	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4	.	.	.	2	.
	Final Score	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**4**	**1**	**1**	**1**	**2**	**1**

1.1.3 Subject DF4 – PM Photogrammetry Model vs. 2D AM Photographs

1.1.3.1 Preliminary, Feature Based Analysis: 3D PM vs. 2D AM

The cropped facial section from the full-body 3D PM photogrammetry model of DF4 was compared against all females in the face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

During the preliminary, feature based analysis of PM subject DF4, 26 individuals were excluded from the face pool with rating scores of either 1 or 2 and four individuals (7, 9, 12, 14) remained in the pool due to rating scores of 3 or higher (potential matches), and were moved along to the next step of the comparison process.

1.1.3.2 Facial Superimposition: 3D PM vs. 2D AM

During the facial superimposition of PM subject DF4, two individuals were excluded from the face pool with rating scores of either 1 or 2 and two individuals (12, 14) remained in the pool due to rating scores of 3 or higher (potential matches) and were moved along to the next step of the comparison process.

1.1.3.3 Detailed Morphological Comparison: 3D PM vs. 2D AM

Following a detailed morphological comparison of PM subject DF4 vs. no. 12 and 14 from the AM face pool, no. 14 was rated as inconclusive. Similarities and differences did not warrant neither exclusion nor a possible match decision, therefore no. 14 was kept in the final pool. In a real-life scenario, further analysis and testing with other ID methods could be undertaken on both no. 12 and no. 14 at this stage, to help establish accurate and reliable positive identification. No. 12 was considered a possible match following the analysis, based on the similarities found and the fact that no definite excluding features could be found during the comparison. However, the quality and resolution of the photogrammetry model of DF4 is rather low and the angle unfavourable, all potentially impacting the analysis and comparison in a negative way.

Upon confirming the results, no. 12 was revealed as false positive, and no. 14 as false negative match for female PM subject DF4. Table 9 below shows a summary of each analysis in the manual comparison process for 3D-2D with subject DF4, and respective rating scores assigned per individual in the AM face pool at each step.

Table 9: Summary of manual 3D-2D comparison results for female subject DF4, with respective rating scores assigned at each step of the analysis and final ID decision (false positive) highlighted in red and false negative (actual match) in yellow.

3D-2D FEMALE		**AM Pool Number Females**																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
DF4	*Preliminary Comparison*	1	1	1	2	3	1	3	1	3	1	1	3	2	3	1	1	2	2	1	2	2	1	1	1	2	2	1	2	1	1
3D Photo-grammetry Model	*Facial Superimposition*	.	.	.	.	2	.	1	.	2	.	.	3	.	3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	Morphological Analysis	.	.	.	.	.	.	.	.	.	.	.	4	.	3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**2**	**2**	**1**	**1**	**1**	**2**	**1**	**1**	**4**	**2**	**3**	**1**	**1**	**2**	**2**	**1**	**2**	**2**	**1**	**1**	**1**	**2**	**2**	**1**	**2**	**1**	**1**

1.1.4 Summary of Results and Findings – 3D PM vs. 2D AM manual

Results of the 3D PM vs. 2D AM manual comparisons (Chapter V, section 1.1) show that two out of three 3D PM face models were matched correctly with their corresponding AM data. In case of the third 3D PM subject DF4, the AM face pool was narrowed down to two candidates following the third and final step of the comparison process, one of which was the true match. However, the final identification decision was ultimately wrong. The issue of misidentification for the third PM 3D face model was most likely caused by the low-quality photogrammetry head model. The latter had been collected by another researcher as a full-body and surroundings model (i.e., cadaver in coffin) for their project and was only included in this study due to the extremely small sample size and lack of data availability. It was found that the image quality and polygon count was insufficient for creating an adequate 3D face model, and during the comparison, not enough facial detail could be extracted for facial comparison.

3D PM face data was collected using an Artec Spider handheld laser surface scanner and Artec Studio v.14 Professional software. The photogrammetry model was added later, due to scan issues and low PM subject count. It needs to be stated that this particular model was not collected for purpose and lacked in detail (e.g., low polygon count), making it not necessarily comparable with or equal to the other 3D face models in this current research, and therefore potentially distorting the interpretative value of results.

Nevertheless, despite the issue of using a low-resolution photogrammetry model and although the correct match was not identified in the final step, the true match was not excluded from the face pool with an assigned, final rating score of 3 (inconclusive). In casework scenarios, visual methods may be followed up by other methods of identification. Therefore, narrowing the initial AM face pool down – here from 30 individuals to just two – would still have been extremely useful, in order to save precious time and resources in this regard. The other two 3D PM subjects and their respective target 2D AM images were matched correctly, hence the approach of combining three different facial comparison approaches in hierarchical order seems to ultimately have been successful, although rather time consuming.

No statistical analysis of the results was conducted, as rating scores and associated results are rather self-explanatory and – after consulting with a statistician – it was decided to not over-complicate simple results. Furthermore, standard statistical methods would not have been applicable, due to the data not being independent in this hierarchical filtering approach. Given that all PM subjects and respective AM target matches were either correctly identified (DM3, DF3) or remained in the face pool until the end (DF4), it suggests that the preliminary, feature-based analysis in particular has a strong discriminating power and was an effective and useful tool for quickly excluding the majority of non-matches from respective face pools at an early stage in the process.

1.2 Manual Comparison – 2D PM vs. 2D AM Facial Photographs

The results section for '*1B: Manual comparison – 2D PM vs. 2D AM photographs*' is organised per PM subject (N=6) and hence in six main sections, each subdivided into two sub-sections by type of analysis conducted (preliminary analysis, detailed morphological comparison). Figure 22 provides an overview of the different stages, with numbers of included and excluded AM face pool numbers for each PM subject per stage, as well as final identification results with rating scores, and confirmed matches. All assigned rating scores and results are also summarised per PM subject in a table at the end of the respective sections. More detailed notes on all stages of the manual analyses can be found in Appendix C –*Manual Comparison – Notes* and the supplemental material. Please see Appendix H – *Supplemental Material – Guide to Digital files* for details.

Figure 22: Flowchart showing the different stages of the 2D-2D manual comparison workflow with face pool numbers for included individuals, and results with confirmed AM matches and final ranking scores highlighted in green.

1.2.1 Subject DM1 – 2D PM Photographs vs. 2D AM Photographs

1.2.1.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DM1 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

After the more detailed preliminary, feature based analysis of PM subject DM1, 29 individuals were excluded from the face pool, with rating scores of either 1 or 2. Only one individual (10) remained in the pool due to a rating score of 3 or higher (potential match), and was moved along to the next step of the comparison process.

1.2.1.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

Following a detailed morphological analysis of PM subject DM1 vs. no. 10 from the AM face pool, no significant differences between the two faces were observed that would warrant an exclusion. No. 10 was assigned a rating score of 4 and deemed as possible AM match for the PM subject, given the similarities particularly in eyebrow growth pattern, overall facial dimensions and feature relationships, facial line and crease pattern, as well as lip shape. The differences between no. 10 and DM1 can all be explained by PM changes to the face, distorted features due to lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight. No. 10 was confirmed as true positive and correct match for DM1. Table 10 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DM1, and respective rating scores assigned per individual in the AM face pool at each step.

Table 10: Summary of manual 2D-2D comparison results for male subject DM1, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

2D-2D MALE		**AM Pool Number Males**																														
		1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
DM1	*Preliminary Comparison*	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	.	.	.	.	4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
	Final Score	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	

1.2.2 Subject DM2 – 2D PM Photographs vs. 2D AM Photographs

1.2.2.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DM2 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

Following the more detailed preliminary, feature based analysis of PM subject DM2, 27 individuals were excluded from the face pool, with rating scores of either 1 or 2. The remaining three individuals (8, 12, 22) remained in the pool due to a rating score of 3 or higher, and were moved along to the next step of the comparison process.

1.2.2.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

After a detailed morphological analysis of PM subject DM2 vs. no. 8, 12, and 22 from the AM face pool, no. 8 and 12 were eventually excluded (score: 2) as the differences observed were not compatible with subject DM2. Especially differences in facial line and crease pattern, relative facial dimensions, and feature relationships, as well as hair growth pattern and distribution were deciding factors for the respective rating scores. No. 22 was given a rating score of 4 and deemed as possible AM match for the PM subject. The differences observed between no. 22 and DM2 can all be explained by PM changes to the face, distorted features due to expression, lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight, as well as image variation. No. 22 was confirmed as true positive and correct match for DM2. Table 11 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DM2, and respective rating scores assigned per individual in the AM face pool at each step.

Table 11: Summary of manual 2D-2D comparison results for male subject DM2, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

2D-2D MALE		AM Pool Number Males																													
		1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
DM2	*Preliminary Comparison*	1	1	1	1	1	2	1	3	1	1	3	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	.	.	.	2	.	.	2	.	.	.	.	.	.	.	.	.	4	.	.	.	.	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**4**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**

1.2.3 Subject DM3 – 2D PM Photographs vs. 2D AM Photographs

1.2.3.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DM3 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

Following a more detailed preliminary, feature based analysis of PM subject DM3 against the AM face pool , 29 individuals were excluded from the face pool, with rating scores of either 1 or 2. Individual 26 remained in the pool due to a rating score of 4 (potential match), and was moved along to the next step of the comparison process.

1.2.3.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

After a detailed morphological analysis of PM subject DM3 vs. no. 26 from the AM face pool, no. 26 was given a rating score of 5 and deemed as AM match for the PM subject. The minor differences observed between no. 26 and DM3 can all be explained by PM changes to the face, distorted features due to expression, lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight, as well as image variation (e.g., nose bent in DM3 likely due to body bag pressing on feature). Similarities between no. 26 and DM3 are very distinct (e.g., lower lip crease, skin marks on lower neck, hair growth pattern, detailed ear shape, overall facial dimensions, and relative feature size) and lead to the confident ID score of 5. No. 26 was confirmed as true positive and correct match for DM3. Table 12 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DM3, and respective rating scores assigned per individual in the AM face pool at each step.

Table 12: Summary of manual 2D-2D comparison results for male subject DM3, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

2D-2D MALE		AM Pool Number Males																														
		1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
DM3	*Preliminary Comparison*	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**5**	**1**	**1**	**1**	**1**	**1**

1.2.4 Subject DF1 – 2D PM Photographs vs. 2D AM Photographs

1.2.4.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DF1 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

Following a more detailed preliminary, feature based analysis of PM subject DF1, 29 individuals were excluded from the face pool, with rating scores of either 1 or 2. Individual 5 remained in the pool due to a rating score of 4 (potential match), and was moved along to the next step of the comparison process.

1.2.4.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

In summary, after a detailed morphological analysis of PM subject DF1 vs. no. 5 from the AM face pool, no significant differences between the two faces were observed that would warrant an exclusion. No. 5 was given a rating score of 4 and deemed as possible AM match for the PM subject, given the similarities particularly in hairline, eyebrow shape and growth pattern, overall facial dimensions, and feature relationships, as well as the facial line and crease pattern (e.g., glabella lines, creases below lower lip, forehead lines). The differences between no. 5 and DF1 can all be explained by PM changes to the face, distorted features due to lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight. No. 5 was confirmed as true positive and correct match for DF1. Table 13 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DF1, and respective rating scores assigned per individual in the AM face pool at each step.

Table 13: Summary of manual 2D-2D comparison results for female subject DF1, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green

2D-2D FEMALE		**AM pool number Females**																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
DF1	*Preliminary Comparison*	1	1	1	2	4	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	1
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**2**	**4**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**2**	**1**	**1**

1.2.5 Subject DF2 – 2D PM Photographs vs. 2D AM Photographs

1.2.5.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DF2 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

Following a more detailed preliminary, feature based analysis of PM subject DF2, 29 individuals were excluded from the face pool, with rating scores of either 1 or 2. Individual

8 remained in the pool due to a rating score of 4 (potential match), and was moved along to the next step of the comparison process.

1.2.5.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

In summary, after conducting a detailed morphological analysis of PM subject DF2 vs. no. 8 from the AM face pool, no significant differences between the two faces were observed that would warrant an exclusion. No. 8 was given a rating score of 4 and deemed as possible AM match for the PM subject, given the similarities particularly in dentition, prognathism, eyebrow shape and growth pattern, overall facial dimensions, and feature relationships, as well as nasal shape and philtrum width. The differences between no. 8 and DF2 can all be explained by PM changes to the face, distorted features due to lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight. No. 8 was confirmed as true positive and correct match for DF2. Table 14 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DF2, and respective rating scores assigned per individual in the AM face pool at each step.

Table 14: Summary of manual 2D-2D comparison results for female subject DF2, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

2D-2D FEMALE		**AM Pool Number Females**																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
DF2	*Preliminary Comparison*	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	:	.	.	4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**4**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**

1.2.6 Subject DF3 – 2D PM Photographs vs. 2D AM Photographs

1.2.6.1 Preliminary, Feature Based Analysis: 2D PM vs. 2D AM

The 2D PM facial photographs (frontal, lateral) of subject DF3 were assessed, then compared against all males in the AM face pool (N=30), and a rating score from 1-5 assigned to each comparison by the researcher.

Following a more detailed preliminary, feature based analysis of PM subject DF3, 28 individuals were excluded from the face pool, with rating scores of either 1 or 2. The remaining two individuals (19, 25) remained in the pool due to a rating score of 3 or higher (potential matches), and were moved along to the next step of the comparison process.

1.2.6.2 Detailed Morphological Comparison: 2D PM vs. 2D AM

In summary, after conducting a detailed morphological analysis of PM subject DF3 vs. no. 19, and 25 from the AM face pool, no. 19 was eventually excluded (score: 1) as the differences observed were not compatible with subject DF3. Especially differences in facial line and crease pattern, relative facial dimensions, and feature relationships, as well as nasal tip angle, overall nasal and mouth width were deciding factors for the final rating score. No. 25 was given a rating score of 4 and deemed as possible AM match for the PM subject, based *inter alia* on consistent facial line and wrinkle pattern, matching nose and mouth shape, finer ear feature shape matching and diagonal fold in right ear lobe. The differences observed between no. 25 and DM3 can all be explained by PM changes to the face, distorted features due to expression, lack of muscle tone, gravitational pull, and other positional variations, as well as naturally occurring differences in age and weight, and image variation. No. 25 was confirmed as true positive and correct match for DF3. Table 15 below shows a summary of each analysis in the manual comparison process for 2D-2D with PM subject DF3, and respective rating scores assigned per individual in the AM face pool at each step.

Table 15: Summary of manual 2D-2D comparison results for female subject DF2, with respective rating scores assigned at each step of the analysis and final ID decision (correct) highlighted in green.

2D-2D FEMALE		**AM Pool Number Females**																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
DF3	*Preliminary Comparison*	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	3	1	1	2	1	1	4	1	2	2	1	1
2D Facial Photograph	*Morphological Analysis*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	4	.	.	.	.	.
	Final Score	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**2**	**1**	**1**	**4**	**1**	**2**	**2**	**1**	**1**

1.2.7 Summary of Results and Findings – 2D PM vs. 2D AM manual

Results from the 2D PM vs. 2D AM manual facial comparison (section 1.2) showed that all PM subjects were matched correctly with their corresponding AM data. This shows that the applied methods and filtering process were successful. It was particularly useful to have both a frontal and lateral view image for all PM subjects available, providing more information about the facial features than a single image. The AM facial images were non-standardised and presented with a variety of different head angles, poses, and in some cases occlusions. The ear is also believed to be unique in its morphology and more resistant to age-, weight-, and post-mortem-related changes. This feature was also found extremely useful as a distinguishing element. It is not ideal to compress a 3D object into a 2D image, as a considerable amount of spatial and depth information gets lost in translation. The additional lateral view does somewhat compensate for this and was deemed very useful in the current study.

Utilising a slightly more detailed preliminary, feature based analysis as the first step to compensate for the missing superimposition appears to be a very useful tool in narrowing down AM face pools very effectively, whilst all correct matches were successfully retained and moved along to the detailed morphological comparison step. Due to the FISWG morphological comparison feature list (FISWG, 2018a) being based on the living face, some aspects thereof were found not to be applicable to the deceased. Certain features of the deceased face were not available or not deemed reliable for detailed comparison (particularly orbital and oral regions), due to PM changes, distortion, or obstruction.

2 Part 2 – Automated Comparisons

The automated comparison results section is divided into two sections, which are further sub-divided as follows, depending on software, input modality, and PM data format:

- 2. Automated Comparisons
    - 2.1 MATLAB®
        - 3D PM vs. 2D AM Automated Comparisons
            - Multiple augmented AM vs. Single PM images
            - Single AM vs. Multiple PM images
        - 2D AM vs. 2D PM Automated Comparisons
            - Multiple augmented AM vs. Single PM images
            - Single AM vs. Multiple PM images
    - 2.2 Google Picasa 3
        - 3D PM vs. 2D AM Semi-Automated Comparisons
        - 2D AM vs. 2D PM Semi-Automated Comparisons

2.1 Automated Comparison – MATLAB®

The automatic face recognition analysis section for MATLAB® is divided into separate results sections for the different data combinations and formats used, followed by a brief discussion and conclusion section.

The simple face recognition algorithm in MATLAB® matches the face of an input image (query) to a person from the face gallery (pool). Only one query image per person can be used against multiple images in the pool, and MATLAB® only allows for 2D data input. To be able to mimic 3D data, screenshots of the face scans and photogrammetry model were taken from various facial angles. In the *multiple augmented AM vs. single PM* approach, only one 2D PM screenshot image (frontal view) was used (query) against multiple, augmented AM images (8pp) (pool). For the *single AM vs. multiple PM* approach, 30 2D PM screenshots (pool) were compared against a single AM image per person (query).

Whilst rather confusing to the reader, this was the only way to consider all factors, if not simultaneously, then consecutively in different combinations. The first factor is the 3D data element, or in this context, the algorithm potentially *learning* a PM face from multiple angles. Secondly, using more than one AM image per person in the face pool in the shape of artificially augmented images, as this may be advantageous to only using one AM image.

For 2D PM vs. 2D AM data comparisons in MATLAB®, a similar principle was followed, as two PM images per subject were available (frontal, lateral), and both had to be included to utilise the same data as in the manual comparison. Hence, this results section is divided into two parts, which are further subdivided as follows:

- 3D PM vs. 2D AM
    - Multiple augmented AM vs. single PM images
    - Single AM vs. multiple PM images
- 2D PM vs. 2D AM
    - Multiple augmented AM vs. single PM images
    - Single AM vs. multiple PM images

The raw output data from MATLAB® simple face recognition algorithm for all analyses below can be found in the supplemental material (see guide under Appendix H – *Supplemental Material – Guide to Digital files*).

2.1.1 3D PM vs. 2D PM Automated Comparison – MATLAB®

2.1.1.1 Multiple Augmented AM vs. Single PM Images for 3D-2D Analysis

The first automated test compared 488 images from the combined male and female AM face pools (resized and augmented, 61 individuals, 8 images per person) against single PM images of the subjects (frontal view scan screenshots). The targets in this scenario were deceased individuals from which 3D facial surface scans or a photogrammetry model had been obtained (DF3, DM3, DF4). Frontal view screenshots of the scans and model respectively were used in MATLAB®. By using only one PM image per individual, the 3D element of the analysis is lost, and the comparison became one of 2D vs. 2D.

The similarity score in the MATLAB® results output is displayed on a numerical scale ranging from -1 to 0, with 4 place values after the decimal point (e.g., -0.0054). The closer to 0, the more likely it is the same individual, with 0 being a perfect match, according to the algorithm. However, none of the AM target images was recognised correctly and matched to the PM subject images by the MATLAB® simple face recognition algorithm used. Results are shown in Table 16 below, which provides a summary of the matching decisions made by the simple face recognition algorithm. The highest score/best match according to MATLAB® is labelled here as either “correct” (true match) or “false positive” (false match), and the score assigned to the correct AM target image under “false negative” (false rejection), if MATLAB failed to identify it. The table only summarises the results for the target AM image (correct or false negative) and the wrong AM match (false positive). Multiple scores were assigned during the analysis, depending on number of comparisons made, and the raw MATLAB® outputs can be found in the supplemental materials folder, submitted in conjunction with this thesis.

Table 16: MATLAB® simple face recognition result outputs for 3D PM subjects (multiple AM vs. single PM images) with respective similarity score (-1 to 0; 0 being a perfect match), showing that none of the targets were recognised correctly.

PM Subject	Correct	False Positive	False Negative
DF3	-	-0.0054	-0.0207
DM3	-	-0.0072	-0.0098
DF4	-	-0.0072	-0.0222

2.1.1.2 Single AM vs. Multiple PM Images for 3D-2D Analysis

In order to re-create the 3D element of the manual comparison (section 2A) in the automated approach, the data input in MATLAB® had to be switched from multiple AM images to multiple PM images (pool) and single PM images to single AM images (query subjects). This is the only way that the simple face recognition algorithm will consider more than one PM image per person in the analysis. It also implies however, that the AM data for comparison has been reduced from 488 (= 8 per person) to 61 (= 1 per person). For both PM facial scan and photogrammetry screenshots, 30 images per deceased individual were used from varying facial angles. As mentioned under section IV.3 Materials above, only 2 deceased individuals who's facial surface scans had been collected at TXST had matching AM data available. A single photogrammetry model and matching AM data for one other person was later acquired and included in the manual and automated comparisons respectively. Results are displayed in Table 17 below. The two PM subjects whose facial scans were obtained using the handheld Artec Spider were not matched correctly against their AM data. Multiple screenshots of the PM photogrammetry model, however, were correctly matched to a single AM photograph of subject DM4, with a perfect similarity score of 0.

Table 17: MATLAB® simple face recognition result outputs for 3D PM subjects (single AM vs. multiple PM images) with respective similarity score (-1 to 0; = being a perfect match), showing that none of the targets were recognised correctly

PM Subject	Correct	False Positive	False Negative
DF3	-	-0.0125	-0.1751
DM3	-	-0.0717	-0.1440
DF4	0	-	-

Running the same experiment again but excluding the low-quality photogrammetry model (DF4) PM and matching AM data thereof from the analysis, the result outputs were different for the two remaining PM subjects, but again, neither was matched correctly, as shown in Table 18 below.

Table 18: MATLAB® simple face recognition result outputs for 3D PM subjects (single AM vs. multiple PM images) excluding photogrammetry model data, with respective similarity score (-1 to 0; 0 being a perfect match), showing that none of the targets were recognised correctly.

PM Subject	Correct	False Positive	False Negative
DF3	-	-0.1303	-0.1675
DM3	-	-0.1372	-0.1355

2.1.2 2D PM vs. 2D AM Automated Comparison - MATLAB®

2.1.2.1 Multiple Augmented AM vs. Single PM Images for 2D-2D Analysis

For this part, 488 images from the combined male and female AM face pools (resized and augmented, 61 individuals, 8 images per person) were compared against single, frontal view PM images of six subjects (DF1, DF2, DF3, DM1, DM2, DM3). The 2D PM photographs used here are the same as in the manual comparison section under 1B above. Although two 2D PM photographs per subject were available, only the frontal view image was used here, as only one *query* image per subject can be compared against the AM *pool* in MATLAB®. The results are shown in Table 19 below. The similarity score in the MATLAB® results output is displayed on a numerical scale ranging from -1 to 0. The closer to 0, the more likely it is the same individual, with 0 being a perfect match. Only one out of the six PM subjects/queries (DM1) was recognised correctly by the MATLAB® simple face recognition algorithm used.

Table 19: MATLAB® simple face recognition result outputs for 2D PM subjects (multiple AM vs. single PM images), with respective similarity score (-1 to 0; 0 being a perfect match), showing that only DM1 was recognised correctly.

PM Subject	Correct	False Positive	False Negative
DF1	-	-0.0077	-0.0188
DF2	-	-0.0095	-0.0211
DF3	-	-0.0041	-0.0151
DM1	-0.0088	-	-
DM2	-	-0.0087	-0.0194
DM3	-	-0.0084	-0.0181

2.1.2.2 Single AM vs. Multiple PM Images for 2D-2D Analysis

In this section, both PM images for each subject were used as *pool* (6 subjects, 2pp, 12 in total) and only one AM image per person (61 in total) was used as *query*. Multiple queries (1pp) can be run at once and compared to the same pool of images. The augmented, multiple images element for the AM data could not be considered here. Results are shown in Table 20 below. None of the PM subjects were matched correctly.

Table 20: MATLAB® simple face recognition result outputs for 2D PM subjects (single AM vs. multiple PM images), with respective similarity score (-1 to 0; 0 being a perfect match), showing that none of the PM subjects were recognised correctly.

PM Subjects	Correct	False Positive	False Negative
DF1	-	-0.1136	-0.2325
DF2	-	-0.1330	-0.1656
DF3	-	-0.0701	-0.2093
DM1	-	-0.0699	-0.0968
DM2	-	-0.1091	-0.1289
DM3	-	-0.1201	-0.2329

2.1.3 Summary of Results and Findings – MATLAB®

Although collating the data in the correct format (i.e., resizing, augmentation) for MATLAB® was somewhat time-consuming, the automated face recognition process itself ran and produced similarity score outputs within merely seconds. Data used in this current research was mostly non-standardised, and PM face changes are also potential hurdles, therefore results were expected to show lower accuracy rates. The issues of using non-standardised and PM data in MATLAB® had not been explored previously. MATLAB® does not allow for 3D data inputs and comparisons. Furthermore, only one query image per person can be considered in the automatic comparison against the pool. In the pool, multiple images per person can be run. In the first test, multiple augmented 2D AM images (pool) were compared against single PM screenshots from the 3D face models (query/input) of subjects DM3, DF3, and DF4 (3D PM vs. 2D AM; section 2.1.1.1 Multiple Augmented AM vs. Single PM Images for 3D-2D Analysis). Results showed that ultimately none were matched correctly. In this first test, the 3D element of the data is lost, as only single frontal view screenshots of the PM models could be used. In the second test, the query and pool were reversed, so 30, 2D screenshots of the 3D models per PM subject were used here. The second test of single AM vs. multiple PM images (3D PM vs. 2D AM; section 2.1.1.2 Single AM vs. Multiple PM Images for 3D-2D Analysis) included screenshot images from the photogrammetry model (DF4) and the two facial surface scans (DF3; DM3).

Results suggested a potential bias towards the low-resolution photogrammetry model screenshots, as almost every AM image was matched to the PM data of DF4. A possible explanation is that with the blurry, low-resolution face model, distinct features were not clearly recognisable to the algorithm, and therefore it likely resembled multiple images. This would be in line with the phenomenon, in which the same individual can appear very different in two images, depending on image factors (e.g., pose, illumination, age, body modifications), whereas different individuals may look similar to each other, even though their identities do not actually match (Abudarham *et al.*, 2019). This prompted a re-run of the same analysis with PM facial surface scan model screenshots only, excluding the photogrammetry data, in an attempt to eliminate this potential bias. Results varied with regards to numerical similarity score outputs. However, still none of the AM and PM data was matched correctly. The photogrammetry model (DF4) likely introduced further issues, as the quality thereof was so low that extracted features may not have been sufficiently unique.

The third test explored 2D PM vs. 2D AM automated comparisons (section 2.1.2.1 Multiple Augmented AM vs. Single PM Images for 2D-2D Analysis), with multiple augmented AM images as the pool, and a single PM photograph (frontal view) as the query/input. The AM and PM data for subject DM1 was matched correctly, with a similarity score of -0.0088. The other 5 query images were not correctly recognised by the algorithm. As a final test using MATLAB® simple face recognition algorithm, a single AM image per individual from the pool was used as query/input and compared against multiple PM images of the deceased subjects (pool) (section 2.1.2.2 Single AM vs. Multiple PM Images for 2D-2D Analysis). None were matched correctly. This reverse use of pool and query images again allowed for multiple PM images to be considered, similar to the second test mentioned above. Given the poor performance and results obtained in this experiment, MATLAB®'s simple face recognition algorithm is not deemed suitable as a face recognition tool for this, or similar non-standardised PM data and respective AM face pools.

2.2 Semi-Automated Comparison - Picasa

The experiments on semi-automated face recognition using Google Picasa 3 are divided into two sections:

2.2.1 – 3D PM vs. 2D AM Semi-automated comparison – Picasa

2.2.2 – 2D PM vs. 2D AM Semi-automated comparison – Picasa

2.2.1 3D PM vs. 2D AM Semi-Automated Comparison – Picasa

In this experiment of 3D PM vs. 2D AM semi-automated comparison, 30 2D screenshots of PM subjects DM3 and DF3 were used (N=2) to mimic the 3D aspect of the data, in conjunction with the combined (males + females) 2D AM face pool (N=61) which included artificially augmented images (11 per AM subject; 732 images total).

Results are summarised in Table 21 below, and show that no correct targets were recognised, except for 9 out of 12 AM images for PM subject DM3 at a threshold of 50%. Some false positive suggestions were made by Picasa, which for the female PM subject DF3 all belonged to male AM faces. The experiments were repeated using separate AM face pools (female AM pool for DF3; male AM pool for DM3). Results did not differ, except that no false positives were suggested for DF3 at the 50% threshold.

Table 21: Results for the semi-automated face recognition/comparison using Google Picasa 3 for mimic 3D PM vs. 2D AM face data at varying threshold settings, using the combined male and female AM face pool and screenshots of the 3D PM face models for subjects DF3 and DM3.

PM Subject	PM Data	Matching AM Data	AM Pool Data Total	Threshold	False Positives	True Matches
DF3	30	12	732	70	0	**0**
DF3	30	12	732	60	0	**0**
DF3	30	12	732	50	5	**0**
DM3	30	12	732	70	0	**0**
DM3	30	12	732	60	0	**0**
DM3	30	12	732	50	5	**9**

2.2.2 2D PM vs. 2D AM Semi-Automated Comparison – Picasa

In this experiment of 2D PM vs. 2D AM semi-automated comparison, 4 2D PM facial images per PM subject (N=6) [males: DM1, DM2, DM3; females: DF1, DF2, DF3] were used in conjunction with the combined (males + females) 2D AM face pool (N=61) which included artificially augmented images (11 per AM subject; 732 images total).

Results are summarised in Table 22 below, and show that no correct targets were recognised, and only some false positive suggestions were provided for DM3, DF1, DF2, and DF3 at the 50% and for DM3 also at the 60% threshold level. Experiments were repeated using separate AM face pools (female AM pool for DF3; male AM pool for DM3). Results did not differ, except for some minor deviations in the false positive suggestions provided by Picasa.

Table 22: Results for the semi-automated face recognition/comparison using Google Picasa 3 for 2D PM vs. 2D AM face data at varying threshold settings, using the combined male and female AM face pool and 4 2D PM images for each PM subject (DF1, DF2, DF3, DM1, DM2, DM3).

PM Subject	PM Data	Matching AM Data	AM Pool Data Total	Threshold	False Positives	True Matches
DM1	4	12	732	70	0	0
DM1	4	12	732	60	0	**0**
DM1	4	12	732	50	0	**0**
DM2	4	12	732	70	0	**0**
DM2	4	12	732	60	0	**0**
DM2	4	12	732	50	0	**0**
DM3	4	12	732	70	0	**0**
DM3	4	12	732	60	1	**0**
DM3	4	12	732	50	2	**0**
DF1	4	12	732	70	0	**0**
DF1	4	12	732	60	0	**0**
DF1	4	12	732	50	4	**0**

PM Subject	PM Data	Matching AM Data	AM Pool Data Total	Threshold	False Positives	True Matches
DF2	4	12	732	70	0	**0**
DF2	4	12	732	60	0	**0**
DF2	4	12	732	50	2	**0**
DF3	4	12	732	70	0	**0**
DF3	4	12	732	60	0	**0**
DF3	4	12	732	50	1	**0**

2.2.3 Summary of Results and Findings – Picasa

It was noted during the analysis, that Picasa does not always automatically upload all files within a folder. There have been similar reports of compatibility issues from other users in the past (Pohran, 2009; Castellini, 2012; Microsoft, 2015), especially regarding incomplete file imports/uploads in Microsoft Windows 10, the same operating system the researcher used. This might be due to the software being developed for previous operating systems (e.g., Windows 7, Vista, XP, and the Apple MAC OS X), and support from Google – including any further development and improvement on Picasa – was discontinued in March 2016 (Betters, 2016; Mediati, 2016). In this current study, all files were copied into the temporary folder (“Picasa trail”) on the laptop prior to importing it into Picasa, but not all files were then also displayed in Picasa itself, after the entire folder was uploaded. A major issue encountered in relation to this was that PM subject DF4 could not be analysed or included in this semi-automated section, as none of the photogrammetry model screenshots for this subject were displayed in Picasa. It is unknown as to why exactly this occurred. It was therefore paramount, to always ensure that at least all PM and AM data for the subject in question had been imported successfully prior to running the experiment.

Effects on result outcomes when using either combined (males + females) as opposed to separate AM face pools were essentially non-existent, with the exception of some false positive suggestions by Picasa at the lowest threshold settings. Hence, the above sections

(2.2.1, 2.2.2) show results for combined face pools only. Overall, the face recognition results were very disappointing, as only one PM subject was recognised and matched to its correct AM data in the 3D PM vs. 2D AM approach (section 2.2.1 above). Nine out of 12 AM images were correctly assigned to the 30 tagged PM screenshots of subject DM3. A possible explanation for this result and Picasa's inability to match other subject's PM and AM data, is that for DM3, the AM and PM facial images presented very little variation or differences in terms of age, weight, or body modification changes, as the AM image was likely taken shortly before death occurred. For all other PM subjects, their respective AM data was generally showing a much younger face, and in many cases, more varied facial expressions, and weight fluctuations, which might have been hurdles that the algorithm could not overcome.

It was noted by the researcher that Picasa had issues with recognising some of the faces in PM photographs or screenshots, predominantly the ones with larger facial volume, and when shown in profile view or from a superior/inferior angle, as was the case with some 3D PM screenshots and half the 2D PM images. Faces in frontal view were generally detected automatically. However, two PM subjects in the 2D PM vs. 2D AM comparisons (DM1, DM2) were not recognised as faces at all and their faces had to be manually defined and tagged in each PM file. Potentially worth noting here, is that other photograph organising applications which also include face detection and recognition functions (e.g., Tag That Photo), state, that manually selected faces will automatically not be considered by the program for face recognition (Tag That Photo, 2021). Unfortunately, there is no information available on whether this also applies to manually selected/defined faces in Picasa. Zero suggestions or matches were found by Picasa for the two PM subjects, whose faces had to be entirely manually defined, possibly suggesting a similar issue with this algorithm.

Given the poor performance and results obtained in this experiment, Google Inc.'s Picasa 3 is not deemed suitable as a face recognition tool for this, or similar non-standardised PM data and respective AM face pools.

VI. Summary

The summary chapter of this thesis is comprised of three main sections: the overall discussion, overall critical review and limitations, and an overall conclusion. Those are further subdivided, which is elaborated accordingly at the beginning of each section. Whilst the overall discussion aims to position the current research within the wider field of facial identification and related disciplines, past studies, as well as current developments, the critical review and limitations section will predominantly focus on matters specifically related to and encountered as part of this current project.

1 Overall Discussion

This section discusses topics related to ethics and data of the current research, such as the ethical and legal framework, data availability and formats, the equipment used for data collection, as well as the different methods of analysis that were applied. It also includes a summary of the overall findings and proposes future considerations.

The current research focused on the identification of recently deceased individuals by means of manual and (semi-)automated facial comparison and recognition, in a 3D vs. 2D and 2D vs. 2D data format approach. The main aims were to use manual methods in a hierarchical, combined approach, and apply both manual and automated techniques, that are generally developed and primarily used on the living, to the deceased. The initial concept for this study arose from the known and increasingly prevalent issue of post-mortem identification (e.g., Migrant Disaster Victim Identification), in scenarios where standard primary and secondary methods cannot be applied and facial identification has been and should increasingly be considered (Wilkinson and Tillotson, 2012; Black and Bikker, 2016; Ranson, 2016; Caplova, Obertova, et al., 2018).

Most facial identification methods are generally used in conjunction with other approaches and can be valuable in helping to direct investigations, narrow down face pools, and in some cases also facilitate an identification. Testing methods that can be implemented offline and remotely (i.e., portable, affordable equipment, non-cloud-based software, freeware), and the aspect of not requiring extensive training in the use of neither respective equipment nor software applications, would make techniques more transferrable and more practical for use in casework scenarios.

Although facial comparison, face matching, and face recognition of the living appear to be rather well represented within the existing literature and ongoing research (some methods more so than others), facial identification of the dead is severely lacking in both published research and case reports, not to mention validation studies. It is therefore difficult to compare and contrast current findings to previous research, as the latter is so incredibly scarce. To the best of the researcher's knowledge, the only studies to date, which aimed to validate the FISWG morphological comparison approach, were published after the practical part of this current project had ended. Those studies discuss facial image comparison of the living (Bacci, Briers, *et al.*, 2021; Bacci, Houlton, *et al.*, 2021) and no published research involving PM data exist in this context.

On the whole, there appears to have been a shift in challenges over time with regards to facial identification. Whereas in the past, technical and image quality problems were seemingly at the forefront of issues that needed to be addressed, those are now considered mostly basic and have been overcome to some extent with rapid technological advances, continuously improving equipment and capturing devices, as well as related post-processing software. Current challenges involve *inter alia* image manipulation (e.g., filters, photo editing), body modifications, ethical and legal challenges, and involuntary data sharing resulting from a lack of control over cloud-based services (Nappi *et al.*, 2016; Rajanala, Maymone and Vashi, 2018; Phillips and Knoppers, 2019; Stark and Hoffman, 2019; Oloyede *et al.*, 2020; Rathgeb *et al.*, 2020). In the current research, the added challenge of the deceased face and its effect on human face matching ability and (semi-)automated face recognition applications has been explored.

Unfamiliar face recognition has shown to be error prone (Chapter II, section 6.3.2 Face Recognition and Matching Ability), although recently, Bacci, Briers, *et al.* (2021) and Bacci, Houlton, *et al.* (2021) reported good results. Is nevertheless an important, indispensable tool for personal identification that is applied in a wide variety of contexts, such as border control, security measures (e.g., building access), and forensic casework (Robertson *et al.*, 2015; García-Zurdo *et al.*, 2018; Megreya and Bindemann, 2018). Facial appearance in an image can be altered rather drastically by a multitude of factors, such as lighting, facial expression, pose, differences in age and weight, or certain diseases (İşcan, 1993). The vast majority of manual and automated face recognition and comparison research does however focus on the living, whilst using highly standardised data with little image variation, and recognition rates for unfamiliar deceased faces and transferability of findings to forensic casework remains largely unexplored. The deceased face introduces its own unique obstacles and variations, and an additional understanding of taphonomic processes is essential. The interdisciplinary nature of this research has to be emphasised here, which also accounts for the rather extensive literature review.

1.1 Ethical Considerations

Ethical considerations for the current research were rather complex, as data protection and informed consent had to be ensured at all times. In this context, the researcher has completed a mandatory *Research Ethics Training* at LJMU in 2017, and the NHS National Institute for Health Research's *Good Clinical Practice Online Course* in 2018. The main study of this research required 2D and 3D facial data of recently deceased individuals, matching 2D ante-mortem data, and random 2D ante-mortem images for the face pool. Given the current study prerequisites of PM facial data, the researcher found that clear guidelines and binding regulations are lacking, as to what is possible/allowed/ethical and what is not. Facial images are classed as a biometric and therefore represent sensitive data under the GDPR, that have to be protected, but anonymisation of facial data has limitations, as its biometric information remains identifiable. Most of the current frameworks are based on and referring to living or AM data, and within the grey zone of PM data, it is unclear who owns this data and should decide over it. Informed consent should of course be the gold standard for all research, whenever practicable.

Unsurprisingly, the first major hurdle encountered in this current research were the ethical and moral considerations related to the study of recently deceased individuals and the use of inherently identifiable facial data. The main aspect that needed to be carefully explored and satisfactorily covered was the issue of informed consent, as the deceased are unable to give their consent at the time of post-mortem data collection. However, when consent for research and/or the wish to donate one's remains has not been decided upon and appropriately recorded in life, it can become extremely difficult and complicated to either make those decisions after death on behalf of a deceased next of kin, or for the researcher to obtain consent from next of kin of the deceased. What the individual him-/herself, family, friends, and the respective society regard as appropriate or desired when it comes to dealing with a human body once it is no longer alive, varies widely depending on cultural, religious, and societal beliefs and background.

Identification efforts often require comparisons of large datasets. This prompts the aspiration to develop or adapt a method that would allow one to do so automatically, since manual methods are extremely time consuming and difficult to adapt to large scale operations. Existing AFR algorithms are developed for and hence trained on living faces only, their applicability to the deceased face has not yet been sufficiently explored, as such data are extremely difficult to acquire and ethical considerations as well as data protection regulations limit these aspirations to a large extent (ICO, 2021). The particularly sensitive nature of face data – especially PM data – therefore require the use of offline, non-cloud-based systems that will not share or store data and are less vulnerable to hacking or other forms of data theft. This was a considerable limitation for the current study, as only very limited options exist for secure, offline AFR applications.

Bearing in mind the controversies surrounding tech giant Clearview AI and the current use of their automated face recognition algorithm in the context of identifying the war dead in Ukraine (Clayton, 2022; Ramage, 2022), there appears to be a large divide in automated face matching of the dead, between ethical principles and being able to provide answers to bereaved relatives (Van Noorden, 2020; Clayton, 2022; Ramage, 2022; Romero Moreno, 2022). However, is does not equate to a the end justifies the means approach, and although at first glance a potentially promising approach, the company has repeatedly drawn strong criticism for unlawfully obtaining face data for their database from social media and other online sources without consent, blatantly ignoring privacy rights and ethical consent issues. The Information Commissioner's Office (ICO) in the UK has filed a lawsuit against Clearview AI Inc. in excess of 7.5 million GBP in May 2022, on the aforementioned basis (ICO, 2022). This current study therefore aimed to test two offline applications that respect the ethical requirements, but results show that the two algorithms tested are not reliable, and approaches therefore not transferrable to the real world. Future AFR systems need to be trained on larger datasets of PM and unconstrained

AM faces, if an implementation of the algorithm in a PM-AM scenario is aspired. Systems should be used offline and adhere to privacy rights and informed consent regulations.

For this current research, data from mortuary settings could not be considered, and related potential collaborations with forensic institutes in the UK and Europe eventually had to be abandoned, as it was not possible to obtain informed consent prior to death. It was deemed unethical and too distressing to ask relatives and next-of-kin for consent, in the aftermath of sudden, unexpected loss. Even though the institutes in Leicester and Germany both had permission for data to be used for research, no agreement could be reached between the University's Research Ethics Committee and external ethics boards, likely caused by the grey zone of post-mortem identifiable data and the question of who should decide over the use of such data. Overall, this caused a 2.5-year delay to the project's timeline.

Unable to obtain data from within the UK or EU, a collaboration with Dr Daniel Wescott and Texas State University was ultimately sought, as donors or their next of kin had already given informed consent for their own or their loved ones' bodies to be used for research purposes following their death (please refer to Appendix F – *Ethics* for details). When collecting PM data in other circumstances (e.g., at the morgue), the deceased are generally not able to provide informed consent in advance. Obtaining self-same from relatives at a time of sudden loss and often traumatic bereavement can potentially cause further distress, which could have conceivably outweighed the importance of this study.

Whilst a solid ethical framework is extremely important and its necessity understandable, being overly cautious and restrictive can potentially lead to considerable limitations that ultimately have a negative impact on the progress of research internationally. It can be – and to a certain extent already is – also problematic in terms of research possibilities and competitiveness between different global regions. If research is to be conducted internationally, collaboratively, or the results thereof aimed to be applied in casework on a global scale – an aim of the current research – the ethical frameworks become incredibly restrictive. Data availability and collection for research involving deceased individuals is rather limited, highly restricted, and quite challenging, as was the case with this current project.

1.2 AM Data Availability

The reasoning behind choosing a 2D and 3D PM facial comparison approach with uncontrolled, circumstantial AM data for this current research, was to mimic real-life scenarios of data availability and how this could potentially be addressed using different methods of facial comparison and matching, on a cost effective and ethically sound basis. Unlike the many face recognition databases in existence – most of which consisting of standardised, highly controlled frontal view images – there is no open PM database of facial images or other PM face data available to researchers at this time. This of course is due to the ethical, legal, and moral implications this would entail.

It is understood, that AM comparative data for primary or even secondary identification methods are not always available (see chapter II, section 1.2.1 Primary and Secondary Methods of Identification above for details), especially in disaster victim identification or mass fatality incident scenarios (ICRC, 2008; Bennett, 2020; Khoo and Mahmood, 2020; Cappella *et al.*, 2022). The most likely and oftentimes only available data, are AM photographs (Yoshino, 2012; Broach *et al.*, 2017; Caplova, Obertova, *et al.*, 2017; Nuzzolese *et al.*, 2018; Olivieri *et al.*, 2018; Labati *et al.*, 2021). Therefore, it seems only plausible to focus on the potential of the post-mortem face as a biometric and related comparative methods, as well as validation studies thereof. Although yielding promising results in lab-based conditions, most such approaches thus far have only been researched and in parts applied in the context of living faces (see chapter II, section 6 Face Recognition) and have yet to be explored in field conditions and/or on non-standardised AM and PM data (Caplova, Obertova, *et al.*, 2018).

In their very recent publication, Cappella *et al.* (2022) provide a case study of Migrant deaths and identification efforts related to the Lampedusa shipwreck in 2013. The authors give information about related AM data acquisition processes and difficulties, and an overview of what type of AM images are generally made available to the forensic practitioner in such circumstances. Ideally, comparative AM images would be of high resolution, depict great detail, the time interval between AM image acquisition and PM event and data capture should be as short as possible, and image capture modalities similar (e.g., angle, subject-to-camera distance, lighting, facial expression) (Stephan *et al.*, 2019). Additionally, multiple AM images per person from different angles would be advantageous (Glaister and Brash, 1937; Webster, 1955; Al-Amad *et al.*, 2006; Wilkinson and Lofthouse, 2015; Gordon and Steyn, 2016).

In reality however, those criteria are often not met. AM images are often of low resolution, or photos derived from identity documents. The latter have been shown to be less than ideal for facial identification, verification, and recognition in the past (White, Burton, Jenkins, *et al.*, 2014; White, R.I. Kemp, Jenkins, Matheson, *et al.*, 2014; Gaudio *et al.*, 2016; Papesh, 2018). For the Lampedusa identifications (Cappella *et al.*, 2022), the quality and quantity of AM images per unidentified deceased individual varied greatly from 1-33 (average 6pp). Previous studies have found that a memory component has a positive influence on the matching accuracy of unfamiliar faces, meaning that if multiple ante-mortem photographs are available, the face can be learned and more reliable results achieved (Sandford and Ritchie, 2021). However, just as in most real-life scenarios, obtaining more than one AM photograph was not possible for most individuals included in this current study. Therefore, whilst the theoretical background for an improved approach is well known, it is not always possible to implement this in actual case work or research aiming to mimic real-life scenarios.

AM data for the main study of the current research was entirely random and hence comparable to casework conditions. Images had previously been provided to TXST University by either donors in life, or their next of kin. Image quantity ranged from 1-3pp, with having only one image available for the majority of subjects in the face pools. The non-standardised AM data used in the current research could be regarded as a considerable challenge. However, the quantity and quality of AM images does reflect circumstances that would likely be encountered in a casework scenario, it makes both the methods and results more transferrable to the wild. Past research has also shown that low-quality photographs can still be extremely useful for exclusions (Schüler and Obertová, 2020), which would also narrow down an AM face pool and hence aid in overall identification efforts.

Additional issues with AM images that are increasingly gaining ground, are the manifold ways in which images can be manipulated (e.g., photo editing, deep fakes) or image quality can be lost (e.g., generation loss in electronic file formats like .jpg, manual scanning, and photocopying). In the current dataset, the majority of AM images were scanned-in or photocopied analogue photographs, rather than digital image files. Most donors had only provided a single AM image and it was deemed morally unjustifiable by the researcher to approach relatives or next of kin retrospectively to request additional data.

However, some caution has to be exercised regarding photographs provided by families and friends, as they may want to present their loved ones in a good light, which may not be the most accurate or suitable AM data. Posed, edited pictures or those that do not show distinctive and perhaps physically detracting features (e.g., facial asymmetry or birthmarks) are less useful. Also, AM images should ideally be as recent as possible. However, any demands designed to acquire anatomically correct pictures must be tempered with sympathy. Whilst a photograph of a broadly smiling individual may help the forensic odontologist by revealing details about dentition, it may not be as helpful when aiming to compare other features. However, the rule for real-life scenarios should be to not reject any picture or, indeed, any information, no matter how irrelevant it may seem initially, as it may prove to be more helpful than expected at a later stage, or could be the only data available. Nevertheless, the data quality needs to be carefully considered and reliability of findings assessed accordingly.

The majority of AM images in the current study were of low resolution, by default non-standardised, and did not show finer details such as skin texture and facial hair growth pattern, which would have aided greatly in the comparison process. To provide a range of intra-person variability, previous research has shown that using multiple images for comparison in face matching tasks of unfamiliar faces is favourable, as it allows for a face learning component that positively influences accuracy (Sandford and Ritchie, 2021). However, concluding from the above that having less comparative data available would by synonymous with less evidentiary value, is not necessarily true, as even a single AM image may hold unique, distinguishing morphological details that can lead to a positive identification or confident exclusion (Black and Thompson, 2007; Black *et al.*, 2014; Caplova, Obertova, *et al.*, 2018; Cappella *et al.*, 2022). For the manual comparison process, the quantity and quality of the AM images proved to be sufficient to establish correct identification.

1.3 3D vs. 2D Data Formats

Several studies have shown an advantage of and hence recommend using 3D over 2D face data for comparison against 2D images in facial identification (e.g., Nickerson *et al.*, 1991; Thomas, 1998; Lynnerup *et al.*, 2009; Cattaneo and Gibelli, 2014; Caple and Stephan, 2016; Miranda *et al.*, 2016; Caplova *et al.*, 2017; Caplova, Gibelli, *et al.*, 2018; Caplova, Obertova, *et al.*, 2018; Lee *et al.*, 2019) (see also chapter II, section 3.2.1.4 3D-2D Data Comparison Preferred Over 2D-2D). The dataset used in the current study is admittedly too small and not enough comparisons were made for a conclusion to be drawn based on final results, as to whether 3D or 2D data is favourable or advantageous over the other in AM vs. PM comparisons, or whether there is simply no difference for the outcome, irrespective of data format. Final identification decisions for the 3D vs. 2D and 2D vs. 2D manual comparisons were accurate in all cases, except for the 3D photogrammetry PM face model of DF4; the low quality of which was ultimately deemed unsuitable for facial comparison, as it did not relay sufficient identifiable features. Therefore, models of similar low polygon count and resolution should not be included in identification scenarios going forward. It is important to state here that photogrammetry as a mode of 3D data collection should not be overlooked or dismissed, but face models need to be collected for purpose, with sufficient resolution and ultimately a much higher polygon count / finer mesh surface. The data used in the current study was originally collected as a full body photogrammetry model and the cropped facial section did not display sufficient detail.

Obtaining 3D PM face models allowed for an additional method of comparison to be considered and included in the current research (facial superimposition), and for the PM face to be aligned with whatever angle the head/face was portrayed in within the AM images. Cappella *et al.* (2022) stated that only half of the real-life AM images they obtained showed faces in frontal view. Therefore, only obtaining standard, frontal and profile view PM photographs (ICRC, 2016; Caplova, Gibelli, *et al.*, 2018) may not be particularly useful in the comparison process, but a 3D model would easily overcome this hurdle and offer a much wider range of angles, as shown in the current study. Lynnerup *et al.* (2009) achieved very good results by implementing a 3D laser scanner and comparing a 3D face model to a 2D image taken from living subjects under uncontrolled conditions. Since the AM data in a real-life scenario can also be assumed to come from uncontrolled conditions in most cases, the researcher felt a similar approach would be suitable and potentially transferrable to the AM-PM comparisons in this current study, with the additional element of utilising PM rather than AM 3D data.

Nevertheless, research and case reports entailing 3D PM face data are still extremely scarce (Tillotson, 2011; Miranda *et al.*, 2016; Caplova, Gibelli, *et al.*, 2018), hence comparisons to other research findings are very limited at this stage. Fairly new technologies aim to create 3D models of a person's face based on a single 2D digital photograph, in order to generate more nodal points and a wider range of perspectives for comparison (Afzal *et al.*, 2020; Zhao *et al.*, 2022). However, such technology is inherently susceptible to error given that the computer is extrapolating a three-dimensional model from a two-dimensional photograph and more research is needed.

1.4 Methods of Analysis

For the identification of the deceased by means of facial imaging approaches, there is only a small but significant window of opportunity to collect suitable visual representations of facial features (e.g., photographs, facial scans, death masks) during the early post-mortem interval, when the soft tissues are still mostly present and intact (Tillotson, 2011; Alt, 2012; Wilkinson and Tillotson, 2012; Caplova, Obertova, *et al.*, 2017, 2018). Afterwards, features generally become unrecognisable to human observers and those methods are no longer applicable (see chapter II, section 5.3 Post-Mortem Changes to the Face).

The research described in this thesis involved the use of PM subject to AM face pool comparisons in a blind-study approach, utilising a hierarchical combination of manual face matching methods (preliminary feature-based analysis, facial superimposition, and detailed morphological analysis), and comparing findings with an automated (MATLAB® simple face recognition) and a semi-automated (Google Picasa 3) method of face recognition. Facial comparison methods – both manual and automated – are predominantly developed on and applied to the identification of the living. Only a small number of publications to date have discussed their application in the context of identifying the deceased (Jayaprakash *et al.*, 2001; Lee *et al.*, 2011; Čaplová, 2017; Caplova, Gibelli, *et al.*, 2018; Olivieri *et al.*, 2018; Cornett *et al.*, 2019; Davis *et al.*, 2019a; Potente *et al.*, 2021). Very few studies have been published on automated methods tested in research settings on PM facial biometrics (Bolme *et al.*, 2016; Cornett *et al.*, 2019).

Although identification of the dead by means of visual comparison has the potential to be incredibly error prone (Tsokos *et al.*, 2006; ICRC, 2008; Caplova, Obertova, *et al.*, 2017), this may simply result from the *how* and not the *what* that is taken into account. When disaster strikes and families are grief stricken, and likely in a state of shock and emotional turmoil, they are not able to make objective decisions. In any circumstance, where a relative, partner, or friend has to identify someone by means of face and/or body recognition, a scientific approach is often lacking or not feasible due to the aforementioned elements, and errors seem almost predestined and unavoidable.

Sometimes desperately claiming a body, fuelled by the religious or cultural prerequisite of a timely burial (e.g., burial before sunset the following day (Schwikart, 2010)), or the longing for closure can render individuals consciously or subconsciously blind to the necessity of correct identification and repatriation. Errors in identification have a domino effect and consequences are severe and often traumatic. In some countries (e.g., USA) and in certain DVI scenarios, it is still common practice for relatives to be called to the morgue to identify their loved ones by means of visual recognition (Anderson, 2008; Čaplová, 2017), which is a somewhat questionable approach, bearing the aforementioned in mind. In DVI or other forensic casework, when the visual identification is being conducted by a practitioner rather than family members, reliable, accurate, and validated methods are lacking to make the overall process more efficient and standardised; from data capture, over comparison approaches, down to interpretation of results.

Human identification of the dead from facial features is an important but not particularly well researched area of forensic casework, and stakeholders involved in the identification of deceased migrants with no comparative AM data for primary identification methods are awaiting publications of validated approaches (e.g., LABANOF Italy, . Extensive knowledge of facial anatomy, identification methods, taphonomic processes, and technological aspects involved are required. This calls for an interdisciplinary approach to the subject matter. Reliable, more precise approaches are still lacking, which are needed to make identification methods more objective, quantifiable, and standardised, so they provide a higher evidentiary value.

1.4.1 Manual Comparison

Face matching, or facial comparison and imaging methods reported in the literature are predominantly based on data derived from living individuals (see e.g., Frowd *et al.*, 2007; Alenezi and Bindemann, 2013; White *et al.*, 2013; White, R.I. Kemp, Jenkins and Burton, 2014; Papesh, 2018; Phillips *et al.*, 2018; Bindemann and Burton, 2021). Most such studies are conducted under lab-based conditions, and influencing factors and variables are highly controlled for within the datasets. Whilst exploring individual factors under the umbrella of disciplines such as psychology and facial anthropology, the applicability of approaches and research findings to real-life scenarios is often highly limited (Stacchi *et al.*, 2020). Research into applying manual face comparison to the deceased is extremely scarce, and most observations to date are derived from case reports (Caplova, Obertova, *et al.*, 2018).

Each manual facial comparison method has its own advantages and limitations (please refer to chapter II, section 3.1.1 Facial Identification Methods for details). The FISWG (2019) has recommended morphological comparison as the gold standard. However, in-depth, manual feature-by-feature comparisons are extremely time consuming and not practical for a large face pool of AM data. Superimposition is only recommended in conjunction with other face identification methods (FISWG, 2019a), and considered more accurate when comparing 3D to 2D data, rather than two 2D objects, as the facial angle cannot reliably be aligned with the latter approach (Damas *et al.*, 2020). Research by Gordon and Steyn (2012) on skull-to-face comparisons has stated to have used a combined methods approach to cranio-facial superimposition. However, in their study, two different methods were used in isolation (morphological superimposition and landmark-based superimposition) and results later combined. This does not constitute a combined methods approach, but rather the merging of results which may likely obscure limitations for each individual method.

Atsuchi *et al.* (2013) suggested a combined methods approach for 3D vs. 2D data analysis, concluding findings from their study on superimposition of age and expression variant and twin data of living subjects. Given the aforementioned, it was deemed advantageous to use more than one approach for the manual comparisons, as this could potentially balance out individual shortcomings of individual methods. This prompted the application of a hierarchical, three- and two-step process respectively in the main study of this current thesis, depending on data format.

Using manual comparison methods in a hierarchical approach provided a safety net and was intended to act as an effective filtering tool in ultimately reaching the correct identification decision. Previous research has shown that response aggregation, collated results, and combined efforts can considerably improve the accuracy levels in unfamiliar face matching (White *et al.*, 2013, 2015; Dowsett and Burton, 2015; Balsdon *et al.*, 2018; Jeckeln *et al.*, 2018). However, there are currently no standard procedures, and practitioners tend to use and adapt methods for each individual case or scenario, and approaches are often difficult to quantify. This introduces personal bias, further elements of subjectivity, and implies that results and findings cannot be compared and validation studies are severely lacking (Bacci, Houlton, *et al.*, 2021; Martos Fernández, 2021).

1.4.1.1 Preliminary, Feature-Based Analysis

To significantly filter down the AM face pool in a fairly rapid and straightforward manner, a preliminary, feature based comparison was implemented in the main study. This comprised of a considerably less in-depth feature-based approach compared to the morphological analysis but aimed to avoid a mere holistic comparison. The latter has been shown to be prone to errors and bias (Wiese, Komes and Schweinberger, 2013; Tanaka and Simonyi, 2016; Megreya, 2018). It does not allow for the analysis of finer details or subtle shape differences and is somewhat comparable to a slightly more thorough passport photograph check in a border control scenario. Utilising feature instructions, however, was found to be advantageous for face matching (Megreya, 2018; Megreya and Bindemann, 2018). The PM subject was evaluated first and a profile based on overall feature spatial relationships and face morphology established, which was then compared to the AM face pool one-by-one. Anthropometry was not applied in this first step of the analysis, as this is not recommended (İşcan, 1993; Campomanes-Álvarez *et al.*, 2015; FISWG, 2019a; Lee *et al.*, 2019; OSAC, 2021), whereas more general, broader spatial relationships between features were found to be rather useful here (e.g., nasolabial height to labio-mental height ratio, forehead height to width ratio, inter-canthal distance to eye-fissure width ratio). Not all such spatial relationships are available for each comparison, as they are dependent on facial angle and feature visibility in the AM and PM data. Nevertheless, this was found to be a good tool to quickly establish a basic overall outline of the respective PM face and whether those same or similar details were also present in the AM image.

For the 3D PM vs. 2D AM comparisons, the preliminary analysis was conducted rather rapidly, within approximately 10 minutes per comparison, whereas for the 2D PM vs. 2D section of the main study, a slightly more in-depth preliminary analysis was aimed to compensate for the missing superimposition step. Still, no more than a maximum of 20 minutes per comparison was required. The most useful features and those yielding the highest evidentiary value, are those which remain stable over time and are not or less affected by factors such as age, weight, and facial expression. Given that none of the correct AM matches were excluded during this initial stage of the process, but the AM face pools were narrowed down quite drastically, meaning all following steps were more streamlined to a very limited number of comparisons, shows that this approach was a highly effective and reliable filtering tool.

1.4.1.2 Facial Superimposition

Facial superimposition (FS) – the second step of this current study's hierarchical approach for 3D PM vs. 2D AM comparisons – is debated with some controversy in the literature. Most research to date has focused on the cranio-facial overlay, rather than a face-to-face comparison, and FS is currently not recommended as a stand-alone method and not at all for use on 2D-2D data (ENFSI, 2018; FISWG, 2019a; OSAC, 2021). 3D-2D, or even 3D-3D comparisons allow for a much better and more reliable alignment and are increasingly being explored (Damas *et al.*, 2020; Martos *et al.*, 2022). Differences in shape, feature size, and spatial relationships can be assessed rather well with this method, although limitations remain with regards to differences in facial expression, age, weight, and image distortion.

A potential disadvantage in using Geomagic® Freeform® Modelling Plus software for this step of the analysis was that landmark placement is not feasible or reliable in this approach, and any deviations between the 2D image and 3D model cannot be quantified. The extensive differences between AM and PM data in the current research due to facial expression, age, weight, and other factors were quite significant. This would have likely resulted in large inter-landmark discrepancies, which would not have been an accurate account of compatibility. Geomagic® Freeform® Modelling Plus software is only able to display shape/geometry on the 3D model and texture information is unavailable. Even though one might argue that superimposition is based on geometry, results from the pilot study (Appendix A – *Pilot Study*) have shown that it is advantageous for matching accuracy to have texture information available as well. Nevertheless, texture vs. geometry requires further investigation, not only in relation to (cranio-)facial superimposition, but for the purpose of facial comparison in the broader sense. Ultimately, this determines which data (CT, photogrammetry, laser) should be obtained or considered in identification scenarios.

The manual best fit or alignment approach is rather subjective but allows for morphological details to be carefully observed and compared during the overlay process and was utilised for this study. Previous studies have concluded that a perfect superimposition is never possible, due to factors such as lens distortion, degree of head tilt, and variations in facial volume or facial expression (Stavrianos *et al.*, 2012; Stephan, 2017; Lee *et al.*, 2019). Decision making was based on feature alignment and shapes, their relative size, and both presence or absence of morphological details observed in the 3D PM face model and 2D AM image. According to Cornett *et al.* (2019), superimposition is essentially a 1:1 biometric comparison in which the practitioner's purpose is to confirm a suspected identity rather than comparing the unknown remains against a database of likely individuals. The researcher would argue that it is still a useful tool to narrow down a face pool and is just as valid for exclusion decisions.

Despite its limitations, this approach had been implemented in the past, *inter alia* for an authentication analysis on the death mask of John Dillinger for the FBI (University of

Dundee Press Office, 2010). The MEPROCS project has acknowledged that applicability of recommended standards and best practice guidelines is always limited by the software and hardware available to practitioners (Damas *et al.*, 2020). Most of the previous research on (cranio-)facial superimposition to date has focused on skull-to-face comparisons, and to the best of the researcher's knowledge, none have explored the aspect of PM vs. AM faces through superimposition. This is likely due to the limited availability of the necessary comparative data to researchers.

There are arguably more suitable methods for facial superimposition available, that would provide more quantifiable results (Lynnerup *et al.*, 2009; Atsuchi *et al.*, 2013; Stephan *et al.*, 2019). The approaches reported in the literature, however, have only been applied to living face data or in a skull-to-face comparison, and would have to be tested and validated on PM facial data, prior to any consideration for their application in casework. The FISWG currently do not recommend superimposition as a stand-alone method, and it should only be used in combination with other methods to provide additional information (FISWG, 2012, 2019a). Although it was used with other methods in the current context, the manual combined methods approach was hierarchical rather than complimentary/parallel. Even though it was effective in the filtering approach for this study, the approach remains controversial and requires further evaluation, in particular with face-to-face, rather than skull-to-face, and PM vs. AM data.

1.4.1.3 Detailed Morphological Comparison

The detailed morphological comparison as the third and final step for both 3D vs. 2D and 2D vs. 2D analyses, was based on current FISWG guidelines and feature lists (FISWG, 2018a, 2019a). An excel template was created by the researcher, which listed over 600 possible features, morphological details, and spatial relationships, as well as body modifications. Respective data points were first analysed in and collected for the PM subject, then again for all AM subjects still remaining the face pool after previous steps. A comparison between the PM and AM data points was undertaken and evaluated, leading to an identification decision on the five-point rating scale. This approach was incredibly time consuming (on average 6 hours per individual), but also very detailed and provided significantly more detail than the previous two manual comparison methods.

Not all features were found to be applicable for the deceased face. Especially the oral and orbital regions, which were distorted by early PM changes and hence were deemed unreliable. Some features were not visible in the face model or image, due to facial angle, obstruction, or data quality issues. It is understood that this morphological feature list was primarily designed for the application to living individuals and describes a maximum number of features and components that may be analysed. It is understood that this list would have to be adapted to each individual case, according to feature availability within the comparative data. The added element of the deceased face and associated factors and challenges, further narrow down the data points which can be compared between the PM and AM face.

Outer facial features, such as ears and hair, were often not available in the 3D PM face models, as hair appeared blurry, and ears were either obstructed by hair or did not scan well. Ears were identified as very useful in facial comparison efforts in the past (see also section 1.4.2 Identifying Features below) (Iannarelli, 1989; Abaza *et al.*, 2013; Krishan, Kanchan and Thakur, 2019). Feature details, such as hair colour or eyebrow shape can easily and quickly be changed through body modification and are therefore less reliable for the morphological comparison, whereas hair growth patterns (e.g., facial hair), or set line and wrinkle patterns remain stable over time.

Face matching accuracy has shown to increase when utilising a morphological feature-based approach, even when comparisons are conducted by amateurs (Towler *et al.*, 2017). It should be noted here that the FISWG guidelines do not provide definitions or instructions for the comparison process itself, but rather an extensive feature list with schematic representations as a visualisation aid, and some descriptions of individual characteristics. The detailed morphological comparison, as recommended by FISWG (2018, 2019), was deemed to be very thorough and reliable, but also extremely time consuming and therefore not suitable for large datasets and associated high numbers of comparisons.

As a final step in the hierarchical approach however, this approach was intended and expected to provide more evidentiary weight to the final decision making. Finer details, such as facial line and wrinkle patterns, hair growth patterns, or detailed ear morphology have shown to be extremely valuable in distinguishing between faces (Iannarelli, 1989; Stavrianos *et al.*, 2012; Schüler and Obertová, 2020), which may have appeared very similar in previous stages, as such details were likely missed or not considered heretofore. Past research has shown that even shorter and incomplete feature lists and instructions considerably improve matching performance (Towler *et al.*, 2017; Megreya, 2018; Megreya and Bindemann, 2018), therefore using a slight adaptation of the particularly extensive feature list provided by FISWG (2018a) was deemed favourable.

Validation studies are slowly emerging but are still too scarce to paint a comprehensive picture. One such study using the FISWG morphological feature list deemed it *"highly accurate and reliable"* for comparing standardised to non-standardised facial photographs (Bacci, Davimes, *et al.*, 2021; p.6). Limitation here are the male-only, AM data used. The PM element is not represented, and to the best of the researcher's knowledge, no other study to date has explored the FIWSG morphological comparison feature list in this context either. Research on detailed morphological comparisons to date is still severely lacking and it is understood that individualising facial features remain extremely difficult to quantify (Lucas and Henneberg, 2015; Gibelli *et al.*, 2016; Bacci, Houlton, *et al.*, 2021).

1.4.2 Identifying Features

It has been stated that the morphological diversity in human faces is more varied compared to other species (Sheehan and Nachman, 2014) and that indeed "*no two faces are alike*" (Wilkinson, 2004; p.5). It is not clear, however, how many similarities or differences an identification or exclusion decision should be based on. To recount and define differences, feature descriptions and a clear taxonomy, as well as universal guidelines are required. Nevertheless, the present state of affairs sees incomplete feature descriptions lacking set parameters. Although there are a number of atlases, guidelines, and feature description lists available (İşcan, 1993; Vanezis *et al.*, 1996; Dunn and Harrison, 1997; Ohlrogge *et al.*, 2008, 2009; Allanson *et al.*, 2009; FISWG, 2018a), they are not used uniformly and consistently in research and practice, as there are no universal standards.

During the manual facial comparison in this study, neither of the aforementioned feature lists or atlases were found to be entirely comprehensive. Although following the FISWG guidelines for morphological comparison for the final step of the hierarchical approach, the researcher resorted to compiling feature descriptions and terminology from the aforementioned and other sources (Knußmann, 1992; Dunn and Harrison, 1997; Wilkinson, 2004; Hadi, 2014; Hadi and Wilkinson, 2017). This is a known issue in manual facial comparison, that inevitably reinforces the subjective element and lack of comparability of this approach (Čaplová, 2017; Stephan *et al.*, 2019). As all feature lists and descriptions are based on living faces, their applicability to the deceased remains largely untested. It is understood, that when utilising an extensive morphological feature list (such as FISWG, 2018), not all features or details may be found in the images or other comparative face data (e.g., hair obstructing ear). Any such list can therefore only be seen as a template, which has to be adapted somewhat for each individual case.

The same was observed in this current study, as even the recently deceased face renders certain features unrecognisable or unreliable. Areas particularly affected were the orbital and oral regions. The orbital region in all PM subjects showed early signs of PM changes, on the one hand related to chemical alterations and dehydration of the eyes (corneal clouding, tache noir, loss of volume within the vitreous body), and on the other hand feature distortion due to loss of muscle tone and/or rigor mortis, and positional distortion (e.g., supine body position and related gravitational pull). Similar factors also affected the oral region. Lips appeared mostly dehydrated and flat, lip creases were not always visible, a slack jaw caused severe feature distortion.

Past research found hair to be an important feature in human face matching tasks and facial recall (Wright and Sladden, 2003). Whilst familiar face recognition is more dependent on internal facial features, unfamiliar face processing and matching appears to be more affected by changes in external features (e.g., ears, hair, hairline, jawline, overall face shape), external features appear to play a more important role (Bruce *et al.*, 1999) (see chapter II, section External vs. Internal Features). Hair in the 3D laser scan and photogrammetry models used in PM vs. AM manual comparisons, often lacked detail. As the scanner struggled to capture hair well, this sometimes resulted in gaps within the scan mesh, that are than displayed as blurry and distorted sections. The low quality and resolution of the photogrammetry model portrayed the hair as almost molten wax-like, devoid of useful detail, such as texture, fullness, and hairline shape.

An external feature that was found to be particularly useful in the current study – whenever visible in both AM and PM data – are ears. Although Iannarelli's ear classification system was published in 1964 and has now been in use for more than half a century (Iannarelli, 1964, 1989), research on this particular feature remains scarce. Atsuchi *et al.* (2013) found that ear morphology, even in monozygotic twins, shows great variation and is very useful for morphological comparison. This is in agreement with previous studies that found ears to be suitable as distinguishing features (Abdel-Mottaleb and Zhou, 2006). During the manual comparison of the current study, it was observed that ears of the same individual showed bilateral asymmetry and variation, an aspect that has been reported previously (Krishan *et al.*, 2019). This feature appears to be relatively stable in the living (see chapter II, 4.1.4.2.3 Ears), as it remains largely unaffected by facial expression (Hurley *et al.*, 2008). From what was observed during the current study on recently deceased individuals, ears should get more attention in PM identification, as they appear to also be unaffected by early PM changes but hold great differentiation potential.

It has been criticised in the past that age-related features such as wrinkles, lines, and folds are often not considered in studies concerned with facial identification and/or recognition, which mostly rely and focus on broader facial features of the mid-face, such as eyes, nose, and mouth. Only more recent studies have taken facial lines and wrinkles into account when designing and testing facial recognition software (Elbashir and Yap M.H., 2020; Tarasova and Mintii, 2021). A limiting factor for facial lines and creases can be the effect of facial expression on their location, shape, depth, and visibility. However, the shape and location of facial lines and wrinkles will not change once they have formed and set, making them a stable feature. Facial line and crease patterns were found to be particularly useful during both the preliminary, feature based analysis and detailed morphological comparison, as those are very unique in their manifestation and stable even in the early PM period, which endorses previous research findings by Hadi and Wilkinson (2014). However, image quality, especially within the AM face pool, was often insufficient to assess such detail.

Differences have been observed between male and female individuals in the current research. Especially during the preliminary feature-based comparison and detailed morphological comparison, males oftentimes presented with more comparable features due to considerably less feature obstruction or modification compared to females. In the latter, the hairline, forehead, and/or ears were often not visible in either or both AM and PM data. Even temporary body modifications, such as makeup caused for some features to be unreliable, hence having to be disregarded (e.g., eyelashes, eyebrows, lips). This finding could be seen as an extension to findings from Tagai, Ohtaka and Nittono (2016), stating that faces with only light make-up are more easily identifiable compared to faces with heavy make-up in the living.

1.4.3 Automated Methods

The past two decades have shown an exponential growth and fast-paced developmental advances in automated face recognition (AFR) technologies. Two (semi-) automated face recognition procedures were included in this research and results compared to those of the manual comparison outcomes. Past research and the recent NIST surveys have shown that most algorithms are developed using, and hence only perform well on standardised, frontal view data (e.g., Burton *et al.*, 2001; Phillips *et al.*, 2005; Jain, Nadnakumar and Ross, 2016; Grother *et al.*, 2017; Kaur *et al.*, 2020). As the AM data used in this project was non-standardised, there was an expectation that automated systems may struggle to produce accurate results, but it was unknown to what extent. It was also not understood, what impact – if any – the element of the *dead face* (e.g., flaccid muscles and skin, discolouration, bloating, feature obstruction or clouding, decomposition related changes, destroyed or missing tissue, etc.) would have on recognition results, as AFR systems are generally trained on, developed for, and tested on living faces only.

Although there are several automated face recognition algorithms available, a number of important restrictions in relation to the sensitive data used in this current project applied here. Most AFR systems are online, cloud-based applications, where data security and protection of image rights are often not guaranteed (Decker and Ford, 2017; Ashley, 2020; Verhelst *et al.*, 2020; ICO, 2021). As for PM data, informed consent by the participants and donors was only given for research purposes and not for public distribution and third-party utilisation (see Appendix F – *Ethics*, section B.1.2 for details). Therefore, AFR systems could only be considered for this current study, if they can operate entirely offline, are user friendly (e.g., time constraints, limited computer science experience), freely available (limited funding), and do not permanently store or distribute data to third parties. The options for free to use, offline face recognition software are extremely limited. One such application is the MATLAB® simple face recognition algorithm by MathWorks, which has achieved good results on standardised, frontal view data (Mathworks UK, 2014, 2021). No extensive training or coding knowledge was required, and following data upload and input, results are generated within mere seconds, making this a very time efficient approach.

Performance was expected to be less accurate for non-standardised data and the added challenge of using a PM face. However, the ultimately extremely poor results led to the implementation of a second (semi-)automated method, for which similar baseline conditions applied. Google Picasa 3 was chosen, as it is available as freeware and easy to use without extensive training. It entails a face recognition function that operates automatically on all images that are imported into the software by the user. Google discontinued supporting the Picasa desktop application as of May 2016 (Betters, 2016; Mediati, 2016) and users have since been encouraged to use Google Photos instead, which operates entirely online and could therefore not be considered for the current research as an alternative.

Past research has found that Google's face recognition function performed better compared to other similar applications (Miller *et al.*, 2015; Schroff *et al.*, 2015) and Picasa has successfully been used in research on face recognition and age progression (Liu, 2018). As this software package has been developed on much larger training datasets (Google!) and is intended for use on everyday photographs, the researcher expected better recognition results for the non-standardised data. The PM element was again an unknown variable with regards to predicted outputs, as training data and intended application here also lies in the context of living faces only.

To the best of the researcher's knowledge, no published research is available on applying either the MATLAB® simple face recognition or Picasa 3 approach to the identification of the dead, and what potential impact this additional challenge would pose for recognition accuracy, nor how both algorithms perform on non-standardised data. Both MATLAB®'s simple face recognition algorithm and Google Picasa 3 are black box systems, meaning that the detailed functionality and transparency of the algorithm is largely hidden from and therefore unknown to the user (Oscar, 2020).

1.4.3.1 MATLAB®

The operation of the MATLAB® *simple face recognition* algorithm was rather straight forward once the script had been adapted to the current dataset. This was somewhat challenging, with only very limited previous computer science training and coding experience, and the initial adaptation of the script would not have been possible without support from an engineering consultant with multiple years of coding experience. The algorithm used in this current study was built around and trained on a highly standardised, frontal-view only dataset (Mathworks UK, 2014, 2021), therefore performance was expected to be negatively impacted by the non-standardised nature of the AM data and the element of the PM face. Applying this particular system to the identification of *wild* data and in particular to the dead, had not been explored previously.

As for most subjects in the face pool only a single AM image was available, each was artificially augmented to eight per person, using a separate script/code (see Appendix E – *Automated Comparison – MATLAB® Scripts*). However, results suggest that this did not help to improve recognition, as the face within the image remains unchanged by augmentation, even when variations such as blur, black and white, and image rotation are introduced. Therefore, no variety is given, and face or machine learning cannot occur. It became evident rather quickly during this current study, that the software requires considerably more training data in order to function optimally. Something the scarce AM data (generally 1 image per person, even when artificially augmented to 8) was unable to provide.

Although the face recognition function itself was performed by the software within seconds, it was rather time consuming to compile and adapt all respective data to the requirements of this code/script. Therefore, the initially envisaged element of saving time by using automated as opposed to manual methods overall was much less profound than expected. MATLAB® does not allow for 3D data input or comparisons, and the element of 3D data had to be mimicked by using 2D screenshots of the 3D PM face models. Furthermore, the algorithm only enables a 1:N image comparison, wherein the '*n*' can contain multiple images per person. As an example, a single PM image can be compared against the AM face pool, and within the face pool, multiple images per person can be considered. It is also possible to create multiple folders for different PM subjects containing only a single image each, which are then compared to the aforementioned AM face pool at the same time. Whilst this was adequate for 2D PM vs. 2D AM data comparisons, for the 3D element to be represented, the comparison had to be reversed, as now multiple images per person had to be considered for the PM subject and only a single image per person for the AM subjects could be used.

It is unknown, whether the facial recognition algorithm can adequately process non-frontal images at all, given it has been developed for and trained on frontal view images only. If lateral views cannot be considered, the mimicked 3D element is lost, and the non-standardised AM data poses a huge hurdle. The dataset used in the current research is particularly small and in the realm of automated face recognition not sufficient as a training database or to provide enough information about a face (variation, multiple images) to enable for a sufficient *face print* to be extracted and ultimately matched. It is also apparent that this particular algorithm in MATLAB® was unable to cope with the multitude of *wild*, uncontrolled factors introduced by the current dataset (i.e., variations in age, weight, pose, illumination, and the deceased face), given it was initially designed using and applied on highly standardised data. As this is a black box system, the exact mechanisms are unknown to the researcher and the issues cannot be further specified at this time.

It was also explored, whether MATLAB® simple face recognition algorithm, although seemingly unreliable for instant face recognition in the current setting, may potentially still be suitable as a filtering tool for face pools. If the correct matching image from the AM face pool would fall within the top 20% of similarity scores, this approach would still be useful. Even if correct target matches would reliably (!) lie within the top 30% or even 50%, it would essentially half the face pool and hence still potentially streamline identification efforts. However, with the exception of two correct matches under the top 20%, all other correct AM data would have erroneously been excluded by this approach from the face pool and distribution of correct matches in the result output's similarity scores appeared completely random, and mostly not even within the top 50%. This shows that this algorithm is neither appropriate for face recognition using PM facial data vs. non-standardised AM images, nor as a filtering tool for large datasets containing similar variation in its current configuration.

1.4.3.2 Picasa

Google Picasa 3 was initially designed for organising digital images and creating albums, but does "*rely on a semi-automatic training procedure*" (De Marsico *et al.*, 2013; p.162). This requires the user to manually label important query images, which are then automatically compared against all other images by Picasa in an attempt to find matches of the same object or person, and for the user to then provide feedback by distinguishing correct from incorrect matches or suggestions (see chapter V, section 2.2 Semi-Automated Comparison - Picasa and Appendix I – *Automated Comparison - Picasa Workflow* for details).

Picasa has been used in a previous study for the purpose of face recognition on age progression data, and the face recognition function found to be reliable (Liu, 2018). Although face detection and recognition functions in Picasa are operating within a black box system and intricacies thereof remain unknown to the user, it was considered potentially more suitable for the purpose of this project in comparison to other offline applications, such as Eigenfaces (Turk and Pentland, 1991; Belhumeur *et al.*, 1997) and Viola-Jones (Jones and Viola, 2003; Gupta, 2019). Eigenfaces, now considered a dated method, only performs well on standardised data, which the current set was not. The Viola-Jones algorithm and software on the other hand is more suitable for object detection rather than face recognition, and in a brief trial with the current data failed to detect the face in most of the PM images.

Interestingly, this current study showed that Picasa – just like Viola Jones – failed at the face detection stage for some of the deceased subjects, as their face was not recognised as such within the image. AFR systems are generally designed to train on and operate with living data, therefore the PM factor appears to be severely neglected in the past and current research landscape. There are only a few exceptions, which have explored iris recognition and other facial biometrics using AFR and PM data (Bolme *et al.*, 2016; Sauerwein *et al.*, 2017; Cornett *et al.*, 2019). Datasets in the aforementioned studies were rather small and therefore findings yielded somewhat limited statistical power and general informative value (see chapter II, section 6.2.5 Applications). Nevertheless, some of the observations reported were also made in the current study. Bolme *et al.* (2016) identified three types of failure during their analyses: the failure to collect data, for the face or relevant biometric to be detected, and failure to verify or find correct matches. Issues encountered during data collection in the current research are further elaborated in section 2.1 Data, Equipment & Ethics below. Picasa failed to detect some of the faces within PM images or from non-frontal view images, and neither MATLAB® not Picasa managed to produce reliable and accurate matching results, with only a few exceptions.

As Picasa is a black box algorithm (Becker and Ortiz, 2013), it can only be speculated as to why the recognition performance was so poor. Possible reasons may be the small dataset and insufficient training data for effective *face learning*, large variations between the appearance of the faces in AM vs. PM images (e.g., weight, age, facial expression), and/or the additional hurdle of the deceased face – which in two instances was not recognised as a face at all. In addition to the aforementioned factors, there are known compatibility issues with Picasa 3 and Windows 10. Although those were reported as minor, the exact effect of a new operating system on the discontinued desktop application remains unknown.

Ultimately, the non-standardised AM data in the current research, the added factor of the PM face and related variables, as well as only having a single AM image available for most AM subjects in the pool, were seemingly too high of a hurdle for Google Picasa 3 to provide accurate face recognition decisions. Therefore, the researcher would not propose to use Picasa 3 in a similar context and does not see a potential transferability into real-life identification scenarios for this particular approach.

1.4.4 Decision Making and Outputs

A five-point rating scale was used for all identification decisions in the manual comparison section of this research (see Table 23 below), as opposed to the widely known Bromby scale, which entails six levels of support (Bromby, 2006). The latter has been criticised in the past for not providing transparent guidance on the differentiation and decision making (Steyn *et al.*, 2018). However, any decision making in manual facial comparison is still primarily practitioner-led, and therefore has a somewhat subjective component that cannot be definitively quantified at this stage. This has previously been labelled as "*justified subjectivism*" (Biedermann *et al.*, 2017; p.477) and a lack of validation studies and quantifiable alternatives have resulted in the current use of somewhat subjective probabilities and more qualitative methods (Schüler and Obertová, 2020). ENFSI (2018) also stated that it is currently impossible to reliably calculate the likelihood of a given face match, as the underlying population data required for such efforts is not yet available. The five-point rating scale used in this research provides a small quantification element to the results and is rather self-explanatory and easy to use. A similar five-point scale is commonly used in South Africa for facial comparison casework (Steyn *et al.*, 2018).

Table 23: Six-point Bromby scale vs. five-point scale used in current research for facial comparison assessment, with explanatory descriptions (Bromby, 2006).

Bromby Scale	Current Research
1 = lends no support	1 = no match
2 = lends limited support	2 = probable non-match
3 = lends moderate support	3 = inconclusive
4 = lends support	4 = possible match
5 = lends strong support	5 = match
6 = lends powerful support	

In the automated and semi-automated comparison sections of the current research, other result formats were obtained by default. MATLAB® provides a similarity score on a scale from -1 to 0 by default; the closer to 0, the more similar the match is considered to be by the algorithm. In Picasa, results have to be manually counted based on displayed images of correct matches as suggested by the software. Those then need to be categorised into correct matches, false positives, and false negatives by the researcher, based on known true match pairs. Limitations here are that all three approaches (manual, automated, semi-automated) provide different result formats that cannot be translated into a common denominator, hence cannot be compared in a statistical manner.

1.4.5 Temporal Requirements

Current guidelines recommend a detailed morphological comparison as the most reliable manual face matching approach (FISWG, 2018a, 2019a). Whilst this is an incredibly thorough way to compare facial data, it is also extremely time consuming and hence not suitable for larger datasets or face pools. The initial dataset for 3D PM vs. 2D AM required 90 comparisons (3 PM subjects vs. 30 AM face pool images in the respective face pools), and for the 2D PM vs. 2D AM comparisons, a total of 180 initial comparisons were required (6 PM subjects vs. 30 AM face pool images). This equates to a total of 270 initial manual comparisons.

The first step of the manual comparison process in this current research, a preliminary feature-based analysis, required on average 15 minutes per comparison, the facial superimposition approximately 3 hours for each case, whereas the morphological analysis and comparison of two faces took on average 6 hours. That equates to an average of 67.5 hours for the preliminary comparisons (270 comparisons x 0.25hrs); whereas conducting a detailed morphological analysis on the same amount of data would have required 1620 hours. To give a true comparison, it has to be considered that if only morphological comparison were conducted for all comparisons (not a 2- or 3-step approach), the 1620 hours have to be seen in relation to how long the overall hierarchical, combined-methods approach took.

Calculating the latter with 0.25 hrs per preliminary analysis (0.25 x 270 = 67.5), plus 3 hrs for each facial superimposition (3 x 15 = 45), plus 6 hrs for each detailed morphological comparison (6 x 16 = 96), this equates to 208.5 hours in total, which is 12.87% of 1620 hrs for 270 morphological comparison at 6 hrs each. This demonstrates that in the current study, applying a combined methods approach in hierarchical order, saved a total 1411.5 hours compared to using the detailed morphological comparison as a stand-alone method on the same dataset (see Figure 23 below). The aforementioned calculations take all manual 3D-2D and 2D-2D comparisons into account that were conducted in the current study.

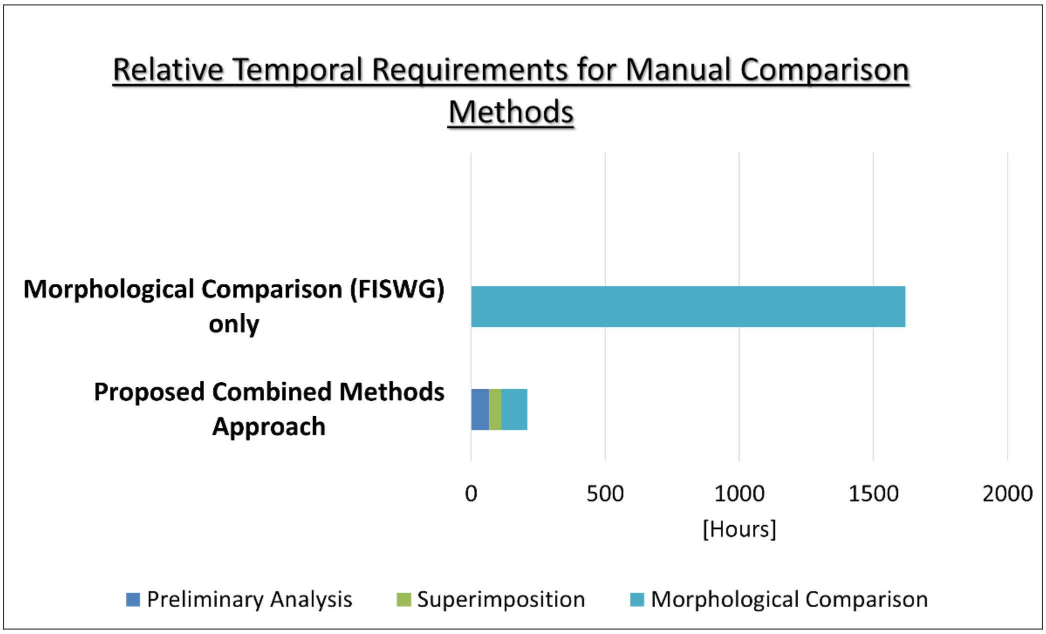

Figure 23: Comparing combined methods timescales from the current research to likely timescales for FISWG recommended morphological comparison as a stand-alone method.

Although the aforementioned calculations for the manual, combined-methods hierarchical approach of this study are somewhat subjective and based on only one observer, temporal requirements have to be understood and reported for potential applicability of methods to casework. Especially in DVI, the three initial questions about the methodological approach are generally '*how accurate is the method*?', '*how expensive will it be to implement*?', and '*how long does it take?*' (Julie Roberts, Company Scientific Adviser at Alecto Forensics & Forensic Anthropology national point of contact for UK DVI; personal correspondence 14th May 2023).

In casework scenarios, it is likely that the data either equates to a very limited number of required comparisons (e.g., 1-2 suspects vs. CCTV footage) or large face pools (DVI; missing person databases). In the first scenario, a detailed morphological comparison only approach would be feasible, but not in the second with a large number of required comparisons. Therefore, a relatively quick filtering tool to narrow down a face pool is not only useful but also necessary. The preliminary analysis proposed in this current research is not simply holistic, it is based on a feature list, hence any inclusion or exclusion is still evidence based and not merely a subjective and/or biased ruling.

The application of automated facial recognition systems would solve the time requirements issue, as comparisons are made within mere seconds for large amounts of data. Although datasets may require some preparation to fit the software requirements (e.g., formatting, augmentation), the actual face detection and recognition function following data upload/input is computed almost instantaneously. However, options are severely limited due to data safety and ethical concerns, and results from the current study showed that the two (semi-) automated methods tested here (MATLAB® and Picasa 3) are not suitable for this purpose.

1.5 Overall Findings

A review of the research hypotheses (as defined under chapter II, section 4 Hypotheses) in relation to the overall findings from this current project are summarised in Table 24 below.

Table 24: Brief summary of research hypotheses in the context of current research findings, with corresponding accept/reject decisions of hypotheses.

Research Hypotheses	Research Findings	Accept / Reject
Manual face matching accuracy will outperform (semi-) automated systems in this current study, given the PM element.	Results show that the hierarchical manual comparison approach led to correct identification decisions (1 exception), whereas the success rate was almost reversed for automated methods and AFR approaches. MATLAB® and Picasa are therefore deemed inaccurate and unreliable.	Accept
Using several manual comparison methods in a combined, hierarchical approach will lead to reliable identification decisions.	Results showed that the hierarchical, combined methods approach leads to accurate identification decisions. The multi-step filtering appears to provide a reliable safety net, resulting in correct exclusions and requiring multiple comparisons on the same data prior to a positive identification decision.	Accept

Research Hypotheses	Research Findings	Accept / Reject
Manual facial comparison methods developed for and tested on the living are expected to be easily transferrable and adaptable to PM data, without losing their potential success rate.	The PM element of the face was not found to be a big hurdle for manual facial comparisons. Some features proved to be rather unreliable due to PM changes (e.g., oral, orbital regions), but results were very promising.	Accept
Automated methods will struggle with PM and non-standardised AM data, likely resulting in a negative impact on accuracy rates.	Picasa failed to detect a number of PM faces within the images and accuracy rates were extremely low for both (semi-)automated methods used, likely due to small datasets and non-standardised AM images. A blanket statement for all AFR systems cannot be made, as only two options were explored here (MATLAB®, Picasa).	Accept partially
3D PM facial data will be more beneficial compared to 2D PM data, as various facial angles and detail are available and alignment with head pose in 2D AM images can be achieved.	The 3D element allowed for use of an additional comparison method (superimposition) that is not recommended for 2D-2D data. Varying facial angles from 3D face models and alignment to AM head angles was found to be very useful here. Limitations are connected to equipment and software issues (e.g., blurry hair and distorted lateral features).	Accept partially
Facial laser surface scans collected with the Artec Spider handheld laser scanner will be a fast and efficient way to capture high detail and create accurate 3D PM face models.	The Artec Spider was more a hindrance than help in data acquisition, given the many issues encountered. Face models are very useful and show great morphological detail, but some scans failed entirely or in part. Therefore, this particular scanner model was deemed unfit for use in field conditions, under time pressures, and on frozen/defrosting remains. Other scanner models may perform better. Therefore, the overall approach cannot be ruled out. Photogrammetry may have been the better option under the circumstances given in this study.	Reject

Overall, the proposed combined methods, hierarchical filtering approach to manual comparisons appears to be a reliable tool for PM vs. AM unfamiliar face matching, with 3D and 2D PM data vs. a 2D AM face pool. However, the 3D PM subject numbers were particularly limited, hence findings should be interpreted with caution at this stage. Although the low number of PM subjects was somewhat compensated for by the large number of comparisons conducted (i.e., larger AM face pool), no definitive statement can be made as to whether 3D or 2D PM data is preferrable in this process, other than what has already been described in previous research (e.g., ability to adjust 3D model to facial angle in AM photograph, potential of more possible points of comparison). The research objectives have therefore been met, as the manual comparison approach resulted in verified true matching and hence correct identification results.

It was expected that automated methods – especially given the limited availability of suitable options of self-same (i.e., ethical, and legal / data rights concerns) – would likely not perform as well as the human observer in manual comparisons. As most systems are designed for and trained on large sets of living, standardised facial data, it was not surprising, that the non-standardised AM data and the element of the deceased face have apparently introduced some image variations the systems were not equipped to process and overcome sufficiently.

MATLAB®'s simple face recognition algorithm resulted in only two correct recognitions: for PM subject DF4 in the mimicked 3D vs. 2D, *single AM images vs. multiple PM screenshots* scenario (chapter V, section 2.1.1.2 Single AM vs. Multiple PM Images for 3D-2D Analysis) and for DM1 in the *multiple, augmented AM images vs. single PM images* scenario (chapter V, section 2.1.2.1 Multiple Augmented AM vs. Single PM Images for 2D-2D Analysis). For DF4, aforementioned issues with the photogrammetry model may have introduced a bias. Matching results in MATLAB® appeared rather random and unreliable.

To explore, whether this approach would be at least suitable as a filtering tool for face pools, rather than actually produce accurate matching decisions, the top 20% closest matches (as defined by the similarity score output) were analysed to see if the correct AM or PM match would fall within this threshold. In casework scenarios, narrowing down the face pool could be incredibly useful to streamline identification efforts, and the filtering approach could then be followed up by other methods of identification on a much smaller set of data. However, it was found that this approach would have excluded the majority of correct corresponding AM/PM data from the overall pool and is therefore not suitable. Overall, the researcher does not recommend MATLAB® simple face recognition algorithm for comparably small datasets (not enough training for the algorithm), or for face recognition using PM and non-standardised AM data.

Google Picasa 3 was chosen as a second (semi-)automated method, which also complied with the necessary software requirements regarding data security. Unfortunately, results were equally disappointing, as only one correct recognition was made for PM subject DM3 in the mimicked *3D PM vs. 2D AM* scenario (section V.2.2.1 3D PM vs. 2D AM Semi-

Automated Comparison – Picasa, section 2.2.1 3D PM vs. 2D AM Semi-Automated Comparison). An additional observation here was that Picasa already failed at the face detection stage on some PM subject's facial images, as the face within the image was not found and had to be manually selected and tagged. Reviews for a comparable software application, Tag That Photo, stated that whenever face detection fails, the respective image/data is then automatically not considered for face recognition by the program, even when it has been selected and tagged manually. This may in part explain the poor performance of this system on PM data. It was assumed by the researcher, considering it is a Google product, that the software would have been trained on large and varied facial datasets during developmental stages. However, not only were some PM faces not detected as such, but faces showing in lateral, superior, or inferior angles within the images were also not detected as faces. Given the poor results outcome using Picasa, this application is not recommended for use on non-standardised AM vs. PM data at this time. The automated comparison approach using both MATLAB and Google Picasa 3 have been deemed unsuccessful, as defined in the research objectives, as the algorithm's face recognition did not result in correct matching and true identification decisions.

Considerable difficulties were encountered in obtaining ethical approval and securing suitable data for the current research. The aforementioned, as well as the fact that not all PM subjects that were scanned and photographed had matching AM data available, in combination with issues encountered with the equipment (see section 2.1 Data, Equipment & Ethics below), resulted in a rather small PM subject cohort for this research study. Findings are therefore somewhat limited in their universality and transferability.

Several statistical concepts and methods were considered and explored in depth for the analysis of the results and outputs of this research. External advisors and LJMU's statistical support were approached for support, resulting in contradicting and ultimately unsuitable propositions, overall causing a significant delay to this project's timeline. After finally consulting with a statistician (Dr Pietà Schofield, University of Liverpool; personal communication 27th October 2021), it was decided not to include any statistical analyses in this research. The rating scale output from the hierarchical, manual comparisons is self-explanatory and results are very clear. It was considered best not to over-complicate this by introducing unnecessarily complex statistical outputs, but rather to use plain language to explain and interpret results instead.

Moreover, the hierarchical manual comparison output values are not independent. Independence in this context would be defined as the value of one observation not affecting the value of another, whereas in the current study almost the opposite is the case. Observations from preceding hierarchical steps have an effect on subsequent analyses, as results from step one (and two) determine who is excluded from or included in further analysis. This is likely to affect statistical analyses by producing false positive results. In addition, it is not possible to compare the manual and automated comparison numerical value results against each other, as each output is in a different format and cannot be translated into a common denominator. Nevertheless, rating scores and ultimate results from the manual comparison show that the preliminary, feature-based analysis in particular was a very powerful, efficient, and accurate filtering tool. The AM face pool was narrowed down quite drastically by this approach, whilst the correct targets remained in the pool for the next steps. The first, drastic filtering is extremely important, as it can streamline resources as well as further analyses, but also has the potential to render the entire manual comparison methodology nil and void if an incorrect initial decision is made.

In summary, manual methods in this current study yielded far superior identification and matching results compared to the two (semi-)automated AFR algorithms employed here. The human factor in facial comparison tasks, especially for non-standardised data appears essential for a detailed examination, analysis, and interpretation of results, as well as final identification decisions. The human observer appears to be better able to overcome challenges, such as pose, illumination, and obstruction variations, as well as feature distortion. Facial changes associated with the early PM period apparently pose a significant obstacle for automated systems but were not found to be an issue in manual comparisons. Both MATLAB® and Picasa are black box systems, and the exact points of failure in the recognition process are unknown. It can be assumed that all of the following factors contributed to poor overall AFR performance:

- very small dataset
- (mostly) single AM image per person not providing variation of the same face as learning opportunity
- non-standardised, non-frontal view data
- PM face data and related facial changes

However, this does not equate to ruling out AFR systems and approaches in general, as there may be better, more suitable approaches available or yet to be developed, which remain to be tested and results verified with suitable datasets for PM applications.

1.6 Future Considerations

The fact that AM images are almost always obtainable in forensic casework, more so than other comparative AM data, needs to be recognised and utilised more within the global scientific community but also by practitioners in the field. There is a pressing need for methods utilising the overall face as a means of identification to be explored and validated further. Findings from the current research suggest that a combined, hierarchical methods approach to manual facial comparisons is a valid tool for both 3D-2D and 2D-2D PM vs. AM data analyses. Although still time-consuming, compared to a detailed morphological comparison as the stand-alone method, the current technique is more time efficient and includes the element of a safety net. Hierarchical combined methods approaches are almost completely lacking in research and case reports, and – given the promising results obtained here – should be further investigated and validated, using much larger datasets and including inter-observer studies in the future.

Existing manual facial comparison and identification methods of the living should be further tested and validated on PM vs. AM datasets, to establish the transferability of methods and their respective reliability, accuracy, and applicability on non-standardised, wild data. Only then can further standards and guidelines be formulated, that would cover visual, facial comparison methods in the context of PM identifications in forensic casework, ideally based on solid evidential findings and research than is the case today. As the feature list for morphological comparison (FISWG, 2018a) was developed for the identification of the living, further scrutiny and evaluation of which features might be challenging or even obsolete in comparisons involving deceased faces is necessary (e.g., iris colour due to corneal clouding and tache noir, closed/sunken eyes, slack jaw, distorted orbital and oral regions, skin discolourations, stability of features).

The two automated approaches explored here did not perform well on non-standardised AM data and with the added challenge of the PM face. Options for offline systems that conform with ethical and data security requirements remain extremely limited at this time. It is likely, that cloud based AFR systems would perform much better, as they are trained on larger, more varied datasets, but are not an option given the sensitive nature of the data. Current development and research of such systems appears to be focused primarily on standardised data, adding only a very limited, controlled amount of challenges. The first step would have to make systems resilient to and still efficient on wild data. In a second step, the PM element can potentially be introduced, as this appears to entail its very own challenges.

Human unfamiliar face matching ability and accuracy rates on PM data have not been explicitly explored in past research, whereas numerous studies discuss self-same in the context of the living. To the best of the researcher's knowledge, there are no published studies to date that evaluated potential differences between matching or identifying deceased individuals from their physical body/face in person, compared to viewing PM photographs or 3D models. The aforementioned gives an incentive for future scientific investigations, as it would make underlying causes for potential misidentifications by family members or lay people more transparent.

Skeleton-ID™'s semi-automated superimposition feature (Panacea Cooperative Research, 2021) seems to be a promising tool for potential future facial comparison research. Their recently published validation studies were focused on craniofacial (i.e., skull-to-face) superimpositions (Martos *et al.*, 2022). Developers expect a facial (i.e., face-to-face) superimposition to be easier to achieve (i.e., comparison of the same and not two different objects) and should yield even more promising results (Andrea Valsecchi, personal communication 7th October 2022). The definitive advantage of this approach would be the element of quantification in the analysis that is provided by an automated numerical output of deviations between landmarks. Therefore, it would be a step towards standardisation of methods and a more uniform approach. This is not given in the manual superimposition efforts employed in the current study using Geomagic® Freeform® Modelling Plus software and Geomagic® Touch™ Haptic Device. In addition, it would be important to investigate, if facial expression changes associated with the PM face (e.g., gravitational pull, slack jaw, sunken eyes) pose too much of an obstacle for Skeleton-ID™'s algorithm, as numerical outputs might highlight significant discrepancies that may not necessarily be an accurate account of actual compatibility or correct identification decisions.

Only very few studies to date have attempted to determine and evaluate facial features that remain stable over time, (Sforza, Grandi, Catti, *et al.*, 2009; Hadi and Wilkinson, 2014; Caplova, Compassi, *et al.*, 2017; Ubelaker *et al.*, 2019), which could represent a very valuable asset to facial comparisons from age-different images or footage, a challenge within the current dataset. A very valuable expansion of research in this area is the study of features which show resilience to early PM changes and associated timelines in different environments. Only very few studies have aimed to explore this to date (Tillotson, 2011; Wilkinson and Tillotson, 2012; Hadi and Wilkinson, 2014; Wilkinson and Lofthouse, 2015; Bolme *et al.*, 2016; Sauerwein *et al.*, 2017; Cornett *et al.*, 2019), all presenting with considerable limitations (e.g., no AM data available, very small datasets, difficulty with quantification of results or repeatability of approaches). It has to be established and understood, which – if any – approach is suitable, in which context (e.g., DVI) and when (e.g., severity of PM changes and/or facial trauma).

In essence, to productively address remaining challenges in facial identification, as well as develop and validate novel approaches, interdisciplinary and cross-institutional collaborative efforts are indispensable.

2 Overall Critical Review and Limitations

2.1 Data, Equipment & Ethics

Considerable difficulties and hurdles associated with obtaining ethical approval for this study resulted in initially envisaged research collaborations with institutes in the UK (2016-2019), Germany (2017-2018), and the Netherlands (2018) to be unsuccessful. This caused a considerable delay to the overall timeline of this project. Ultimately, a new collaboration was sought with Texas State University and ethical approval granted in February of 2019, two and a half years after this project started in 2016 with a three-year funding plan. Data collection for the current project commenced in Texas between March and May of 2019. Time and funding constraints meant that a longer research visit and gathering more data respectively was not feasible. Unfortunately, it transpired only after 2D and 3D PM data collection had already been completed and the AM face pool was being collated by Laney Feeser, that AM data was not available for all donors, resulting in some 2D and 3D PM images and face models having to be excluded from this research for lack of comparative data. Below mentioned problems with the scanning equipment resulted in some 3D data not being usable, therefore the 3D PM subject number in this current study was particularly low.

AM data was captured under uncontrolled conditions by unknown third parties, and the quantity and quality thereof was purely circumstantial. A firm component within the body donation forms for TXST is the provision of AM images (see Appendix F – *Ethics*). Unfortunately, this had obviously not been complied with by many donors in life or their next of kin. Recent recommendations by Cappella *et al.* (2022) state that the most recent AM images should be used whenever possible, as older photographs would imply limitations for the comparison process. However, in the current study, most data showed quite considerable differences in age, weight, and appearance between the AM and PM face of the PM subjects, and for most individuals, only a single AM photograph was available. Past research has already shown that a single AM image does often not provide enough important and characteristic detail, strongly advising to use more than one AM image for comparison (Wilkinson and Lofthouse, 2015; Caplova, Obertova, *et al.*, 2018; Kramer *et al.*, 2018).

The Artec Spider handheld laser surface scanner was chosen for 3D data capture, as it has a remarkably high scan accuracy and great precision (see chapter IV, section 2.1 Methods of Data Collection for details). This particular scanner was one of the best portable models available at the time of data collection (pre-2020) and has since been replaced by the follow-up model, the Artec Space Spider (Kersten *et al.*, 2016; Tokkari *et al.*, 2017; Leipner *et al.*, 2019). All scanning equipment used in this research is rather costly but was readily available at Face Lab, LJMU and the researcher had familiarised herself with the scanner and software during other projects prior to conducting this research (e.g., Beaumont *et al.*, 2018; Ord, 2018). Unfortunately, a number of issues were encountered regarding the equipment used for PM data collection and AM data availability.

The Artec Spider handheld laser surface scanner proved to be a rather delicate piece of equipment. It was not possible to set up a permanent scanning station at TXST, as the premises were used daily for a multitude of different activities (e.g., teaching, workshops, forensic casework), hence the scanner and laptop had to be set up anew each time when donors were taken in at the facility. There were issues with this particular scanner, as the expected warm-up time of 20-30 minutes was closer to 1hr45mins, and re-calibration was required every time the scanner was unpacked and prepared for use, which is not usually the case. Technical support was contacted about long warm-up times and the scanner overheating quickly, hence not always capturing surfaces accurately. This could not be resolved remotely.

Scanning had to be conducted in a rather short time span during *intake*, whilst other preparations were undertaken by the TXST anthropology student volunteers (e.g., taking hair and blood samples, nail clippings, wrapping hand and feet in mesh bags, taking full body and close up photographs, completing paperwork, etc.). The researcher had very short notice (as had the volunteers), when a body was ready for collection and brought to ORPL for processing/*intake* prior to placing it outdoors at FARF or storing it in the cooling facilities at ORPL for future use (i.e., workshop reserve). The timeframe for data collection was therefore limited and was not supposed to hold up or disrupt the intake or subsequent body placement process. Hence, any prolonged warm-up and re-calibration delays posed serious obstacles.

Rigor mortis or the frozen state of some remains made it difficult to position the head or get a complete scan and/or photograph of both lateral aspects of the face. Donors that had previously been stored in coolers or freezers had a moist to wet skin surface from condensation. Additionally, illumination at the facility came from a motion sensor activated fluorescent lighting, which could not be dimmed or adjusted. Both of the aforementioned factors contributed to the skin/face presenting as a rather reflective surface, which the scanner had considerable issues with and did not always manage to overcome. This resulted in misaligned sweeps and/or holes in the mesh surface, that were ultimately not usable as no coherent 3D model could be produced during post-processing. In fact, all scans collected from donors in a frozen or defrosting state were not usable. Measures to minimise this effect were taken, as the skin was dabbed with clean paper towels, and the white body bags covered with dark coloured cloth behind the donor's head to keep reflection to a minimum.

Facial scans collected from donors brought in via transport from a morgue or funeral home on the same day and those stored in the fridge did not pose the same issue as bodies which had been stored in a freezer. A clear pattern emerged here: From a total of 12 donors that were scanned during the researcher's stay at TXST, five were either frozen or defrosting and the facial scans were unusable, as scans were incomplete and misaligned, hence postprocessing failed. From the remaining seven scanned donations, only two had matching AM data available. All twelve donations were also photographed (frontal, lateral view), to have 2D PM data available as backup for comparisons. Out of the twelve, AM data was only available for six donors.

In a real-life scenario, it is also unlikely that the ambient light can be controlled. Environmental conditions in general can be rather unpredictable, also regarding elements such as dirt and moisture, quickly rendering such delicate tools unsuitable. Even though the scanner comes with an external battery/power bank that would in theory allow use in the field without a current source, the powerful laptop required to run the Artec software to capture sweeps, the time-consuming nature of setting it up – particularly as encountered here – as well as the delicate nature of the equipment, would not have allowed for outdoor/field use. The portability factor is therefore somewhat limited.

Since collecting data in 2019, Artec have launched a new model of the scanner, the Artec Space Spider, which does not have the same issues that in part resulted from a lack of temperature stability (Matthew McMillion, Senior Editor & Writer at Artec 3D; personal correspondence October 2022). The new software package is also a considerable improvement, hence the issues encountered by the researcher may no longer be applicable when using updated and currently available scanner models and software. However, this has not yet been tested under similar conditions, therefore conclusions here have to be drawn based on the previous scanner model, the Artec Spider. For that reason, the researcher would deem the Artec Spider overall not suitable for this or a similar purpose, as too many obstacles were encountered.

Therefore, the recommendation would be to revert to photogrammetry rather than laser surface scans. Capturing multiple photographs from different angles to create a dome-like perspective and creating a 3D photogrammetry model with software such as Agisoft Metashape seems a more practical and inexpensive approach. Acquisition costs for the scanner were approximately 15.000 GBP in 2015, including the software package (Kersten *et al.*, 2016). Photogrammetry data (2D images) can be captured with any digital camera or even a mobile phone, commonly entailing considerably lower acquisition costs and better portability due to much lighter and less delicate equipment (Clara Alfsdotter, fellow visiting researcher at TXST; personal correspondence, April 2019). It would make this approach more suitable to field applications, as the camera or phone is generally readily available and more robust against environmental conditions and electricity requirements. Additionally, research has shown that the difference in accuracy between laser surface scan and photogrammetry models is negligible (Urbanová, Hejna and Jurda, 2015; Leipner *et al.*, 2019). For a comprehensive review on the aforementioned and other geomatic techniques, which are currently or could potentially be applied in the field, please see Berezowski, Mallett and Moffat (2020).

In addition to the 3D facial surface scans and based on the limited number of PM subjects with matching AM data available from the 3D laser scanning process (N=3), it was decided to include a 3D photogrammetry model of a recently deceased individual within the study. This data was also collected at FARF by fellow visiting researcher Clara Alfsdotter (PhD researcher at Linnaeus University Sweden in May 2019) for her own PhD research project in early 2019. Permission was granted from Texas State University for this model and matching AM data to be used in the current study.

It needs to be emphasised that this photogrammetry model was captured as a full body model and not for the purposes of this particular study, hence some limitations in utilising it for facial comparison research were to be expected. It was found that the model was ultimately not of sufficient quality and resolution (i.e., low polygon count/mesh density), blurring most identifying features and making the comparison extremely difficult and unreliable. However, if photogrammetry data is collected for purpose (i.e., head only, high resolution images, high polygon/mesh count), it would be a viable and more cost-effective alternative to laser scanners. An example of two photogrammetry models is given in Figure 24 below. This illustrates the difference between data collected with the intended purpose of utilising it for facial comparison, and a cropped model that was not initially collected for the same purpose. The contrast in overall quality and morphological detail gives a clear example of which model would ultimately be desirable and more suitable as comparative data.

Figure 24: Cropped full-body photogrammetry model of DF4 (left) and face-only photogrammetry model of donor 12 (right), showing differences in overall quality and morphological detail (images: author's own; taken at ORPL/FARF, TXST; used with permission).

Given the small sample size for the main study, an extension of the pilot study (Appendix A – *Pilot Study*) was considered in late 2019 and early 2020. However, the Covid19 pandemic and lockdown resulted in no or only restricted access to equipment, as well as social distancing guidelines that prevented this endeavour and any further data collection on living individuals pretending to be deceased. The researcher is also acutely aware that pretend-dead living faces are not an adequate substitute for PM data. This was further confirmed by the inability of Picasa to detect some of the faces within PM images, and the difference in face detection between actual PM (main study) and pretend-PM (ancillary studies) data.

The overall difficulty in and ultimate impossibility of obtaining AM and PM face data for this research in the UK and Europe – whilst plausible to a degree from the informed consent perspective – is an issue. The immense delays caused to the project timeline in pursuing this matter aside, it undoubtedly hinders important research in the UK and Europe in general, if the necessary data can seemingly only be obtained from much further afield. It is not feasible for all researchers to conduct research visits to existing taphonomic research facilities (TRFs) overseas, and as already mentioned in chapter II section 5.4 Taphonomic Research Facilities of the literature review, the geographical location of the TRF dictates both the donors' demographic background and the taphonomic environment (Williams *et al.*, 2019). Therefore, research findings and results are not always transferable to casework in the UK or Europe.

Furthermore, whilst the researcher does acknowledge the sensitivity of biometric data and yet again importance of consent, it appears almost paradoxical to have cadavers undergo severely invasive procedures for research or teaching in medical schools on a daily basis, whilst this current study was looking to obtain facial images and conduct analyses in an entirely non-invasive fashion yet failed at accessing donor data in the UK and Europe.

Although there is a TRF in the Netherlands, its size and donor count is extremely small (currently one?) and can therefore only be viewed as a pilot facility in Europe. It also currently only allows for research to be conducted on interred remains. The endeavours to establish larger TRFs in the UK and/or Europe appear rather dormant at this point, so maybe an option to access similar data to what was used in this current study closer to home, could be minor amendments to already existing body donation forms. Those would ask donors in life or their next of kin to submit AM images on a voluntary basis to have comparative facial data on record, and possibly open up access to more fields of research. Nevertheless, the researcher strongly advocates for one or more TRFs in the UK and Europe with human rather than porcine cadavers, which would hugely benefit not only research but also teaching and training for a large variety of fields, as is already the case in the US, Canada, and Australia.

Difficulties in obtaining suitable, required data for similar research have previously been acknowledged by Cornett *et al.* (2019). Their study compared early to later stage PM facial images using commercial and custom facial recognition systems. Comparative AM data was not available, but its acquisition aimed to be pursued in the future for further research. The arguably well-funded (~3 million Euros), 3 year-long (2010-2013), interdisciplinary and collaborative (INTERPOL, BKA, Pklass Data Software A/S, Frauenhofer Gesellschaft, Crabbe Consulting Ltd., and University of Dundee) FastID project (Crabbe *et al.*, 2013) had to ultimately abandon one of their main objectives of obtaining and comparing PM and AM facial images, and instead opt for pretend-dead facial data. This does indeed explain the lack of research if the necessary data is so incredibly difficult to come by.

2.2 Methods of Analysis

2.2.1 Manual Comparison

As the preliminary, feature-based comparison was created by the researcher as an initial filtering tool in the hierarchical process, this particular method had not been applied previously. It was intended to be slightly more thorough than a passport control-type comparison, but not as extensive and time-consuming as the detailed morphological analysis. The main limitations for the facial superimposition comparisons using Geomagic® Freeform® Modelling Plus software, are that texture on the 3D model cannot be displayed and facial landmark placement and inter-placement distances are not included. The latter is recommended for quantification and determination of spatial differences and the approach applied here does not conform with the MEPROCS recommended guidelines (Damas *et al.*, 2020; Martos *et al.*, 2022), which were published after this current research had already begun.

The FS method utilised in the current study is more suited and was intended for exclusion rather than identification decisions. There is still considerable doubt in the scientific community as to the validity and reliability of (cranio-) facial superimposition (see e.g., Gordon and Steyn, 2016; Strathie and McNeill, 2016; FISWG, 2019a; Damas *et al.*, 2020; Martos *et al.*, 2022). Research and published casework involving PM and AM facial superimpositions are lacking, hence a comparison with current findings is not possible. The researcher would have likely reached the same identification decisions if superimposition had not been included in the combined methods approach at all. However, the temporal requirements for the overall process (see section 1.4.5 Temporal Requirements above) would have substantially increased.

Due to timeline and funding constraints, the sensitive nature of the data, and personnel availability, no inter- or intra-observer studies were included in this project. There is an almost unavoidable element of potential bias for any observer in face comparison research and casework. The researcher’s ability to apply consistent analytical judgement may not have been specifically tested previously, hence it could be argued that the consistency in these inter-researcher judgments and descriptions of the same facial data was not evaluated here. However, results show that the applied methods led to accurate outcomes in the more qualitative, manual assessment of the data. Ultimately, it was decided that the use of terminology and exact approach for both the qualitative facial assessments, descriptions and evaluations and automated comparison methods was not required to be replicated by others, and transferability therefore not specifically tested, as neither approach was entirely new, only its application to the deceased.

2.2.2 Automated Comparison

As both offline (semi-) automated comparison methods applied in the current research (MATLAB® and Picasa) resulted in highly erroneous matching decisions, and further failed to detect most non-frontal view faces and deceased faces within the images, both applications are unsuitable for data variation as given in the current research and likely to be encountered in casework scenarios. The aforementioned also implies that it was not possible to mimic the 3D data element, by using multiple 2D screenshots of the 3D face models from varying perspectives. Both MATLAB® and Picasa are black box systems, it is not known how the systems operate. The small sample size appeared to be insufficient in training the MATLAB® simple face algorithm effectively and providing enough distinguishing features for comparison. The commercial desktop application Google Picasa 3 was expected to have been trained using Google's own image database, which is rather extensive and varied. Recognition results were therefore expected to be better compared to MATLAB®, which was not the case. Training these algorithms on a larger PM dataset was not possible, as sharable databases of deceased faces with known identities and matching comparative data, similar to other living-face databases (see section III.6.2.7.2 Training Databases and Machine Bias), are not available for research at this time. Given the immense ethical and legal hurdles this would entail, it is not likely to change in the foreseeable future.

A considerable limitation of using Picasa is that the support for this application was discontinued by Google Inc. in 2015, and no software updates have been launched since. The company now encourages customers to use Google Photos instead, a cloud-based service, which was not an option for this study. The sensitive nature of the data, and fundamental identifiability of faces as a biometric, have considerably limited the choices for automated approaches in the current research. When using cloud-based services (e.g., Microsoft Azure, Google FaceNet, Facebook DeepFace) – although likely superior in performance to the ones used here (MATLAB®, Picasa) – involuntary data sharing with third parties would have to be expected but is not ethically nor legally acceptable in this context. Offline options are unfortunately extremely limited and generally less supported and advanced.

3 Overall Conclusion

This current research explored the application of manual face comparison and (semi-) automated facial recognition approaches to the identification of recently deceased individuals, utilising 3D vs. 2D and 2D vs. 2D PM and AM data. Two to three manual methods were used, depending on data format, in a combined, hierarchical approach, and although the methods themselves were not novel (feature-based analysis, facial superimposition, morphological comparison), their application in combination and to the identification of the deceased was. The two AFR algorithms MATLAB® and Google Picasa 3 were chosen on the basis of compliance with data security and ethical considerations, as well as relative operational ease due to the researcher's limited computer science background. Neither application had previously been applied to PM facial data.

The aims of the current study were to investigate the applicability of combined, manual face comparison and AFR algorithms – all of which developed for and generally applied to the identification of the living – to the deceased, to investigate their accuracy and reliability. Furthermore, this research set out to gain a better understanding of potential suitability and transferability of the aforementioned approaches to real-life identification scenarios, as well as evaluating different data formats (3D, 2D) and associated advantages and potential disadvantages thereof. The Artec Spider 3D handheld laser surface scanner was used for data collection, which brought its own challenges that ultimately resulted in a rather small PM subject cohort for 3D comparisons. AM facial photographs in casework scenarios are by default non-standardised, a factor that was considered and well represented in the current study because the AM data acquired was entirely random. The limited number of AM images (often one per person) posed a serious issue for automated methods resulting in poor recognition rates.

Research findings herein clearly indicate that human facial comparison ability in this context appears to be superior and more resilient to factors such as PM facial changes and limited data availability, compared to the two (semi-)automated options. Results from the manual comparison showed that all 2D PM vs. 2D AM comparisons were successful. Two out of three 3D PM subjects were also identified correctly in the 3D PM vs. 2D AM matching scenario. The third 3D PM subject was ultimately misidentified, however the correct match remained in the face pool until after the final identification decision, and data quality issues with the photogrammetry model should have excluded this subject from the study altogether. The first step in particular, using a preliminary, feature-based comparison, proved to be a highly effective and accurate tool to drastically narrow down the AM face pool without wrongfully excluding matching data.

MATLAB® and Picasa both preponderantly failed at recognising correct matches within the face pool. Although ultimately unsuccessful at the PM vs. AM face recognition task, the (semi-)automated methods highlighted some important issues. The element of mimicking 3D data inputs (with screenshots of the 3D model from multiple angles) is impaired by the applications' inability to recognise non-frontal images. Furthermore, Picasa at times failed at the face detection stage for PM images, which affected subsequent recognition functions. Pretend-dead faces (as explored in ancillary studies) posed no issue in this regard, suggesting a significant, yet undefined difference between pretend-dead and actual deceased faces.

The investigation of MATLAB® simple face recognition algorithm in the current context revealed that this application requires considerably more training data to function efficiently. Adapting the script/code to the current data was more challenging than anticipated, and preparations of the dataset (e.g., image augmentation, resizing) was time-consuming. The low accuracy rates likely result from the small number of images available to train the algorithm, as well as the non-standardised AM and additional PM element of the dataset used in this research. MATLAB® is also not suitable as a filtering tool for the face pool, as it became apparent that the correct AM matches did not reliably fall within a certain threshold (e.g., top 20%). At this stage, neither Picasa nor MATLAB® simple face recognition are deemed suitable to be applied on comparable datasets (e.g., PM, and non-standardised AM facial data), nor in casework scenarios. However, this does not equate to a blanket statement for all AFR systems, as better options may be available.

Obtaining ethical approval for this research and acquiring the necessary data proved to be extremely challenging and time-consuming. Several initially pursued research collaborations had to be abandoned, and data was ultimately acquired at the Forensic Anthropology Research Facility at Texas State University in San Marcos, USA, following a 2.5-year delay to the project timeline. Furthermore, the Artec Spider laser scanner was extremely sensitive to lighting conditions and reflective skin surfaces and presented some major calibration and warm-up time issues. It was therefore not deemed suitable for field conditions, and exploration of photogrammetry options instead are recommended. Additional limitations of the current study lie within the lack of quantifiability of manual methods, especially the chosen facial superimposition approach using Geomagic® Freeform® Modelling Plus software, the small PM subject cohort, as well as limited comparative AM data availability. To somewhat compensate for the small number of PM subjects, a higher number of comparisons was undertaken, but the evidentiary value of results remains limited.

The application of both manual and automated analyses in the context of facial identification is still an extremely neglected area of research, in part likely due to the difficulty in obtaining suitable data. The increasing prevalence of natural disasters and the ongoing Migrant crisis is giving rise to further scenarios were PM identification methods are required on a large scale. In many cases, the ideal AM comparative data for primary or secondary methods of identification are simply not available. Facial photographs are almost always attainable and should therefore be utilised as an important resource. However, facial identification methods are still predominantly developed for and applied in the context of the identification of the living, and their application to the deceased is only very scarcely evaluated and described to date. PM face data included in this current research focused on recently deceased individuals without facial trauma, with the exception of the post-autopsy condition of DM3. It is highly likely that facial trauma is encountered frequently in casework, which would cause additional challenges for feature comparison, and their impact on recognition rates and accuracy have yet to be explored.

This current study highlights important issues and challenges, whilst also contributing to validation efforts of existing methods. Nevertheless, future efforts should be conducted – whenever possible – on much larger datasets, as well as include inter-observer error studies for manual approaches. Automated methods still appear to struggle with non-standardised data, an issue that probably needs to be addressed prior to tackling the PM challenge. It must also be recognised and respected that ethically, morally, and legally questionable approaches to AFR, in which data security is not guaranteed, are not a viable option for the sensitive nature of PM data. It needs to be understood across a multitude of disciplines involved in human identification, that AM data tends to be highly circumstantial, in its general availability, quality, as well as quantity. PM data however can be much more controlled for, hence the importance of obtaining multiple high-quality images and/or 3D data at the scene/field or in the mortuary (with the best available equipment and conforming with standard PM data collection guidelines). This is a factor that is too often neglected or not understood, resulting in a missed opportunity which cannot be rectified retrospectively, but which has an enormous impact on data analyses.

The highly interdisciplinary nature of this field of research strongly suggests that collaborative efforts, as well as improved and continuous scientific exchange, are not only highly advantageous but rather indispensable in order to develop, validate, and advance both manual and automated facial imaging methods in the future.

VII. References

Abate, A., Nappi, M., Riccio, D. and Sabatino, G. (2007) ‘2D and 3D face recognition: A survey’, *Pattern Recognition Letters*, 28, pp. 1885–1906. Available at: https://doi.org/10.1016/j.patrec.2006.12.018.

Abaza, A., Ross, A., Hebert, C., Harrison, M.A.F. and Nixon, M. (2013) ‘A Survey on Ear Biometrics’, *ACM Computing Surveys*, 45(2), pp. 1-22:35. Available at: https://doi.org/10.1145/2431211.2431221.

Abdel-Mottaleb, M. and Zhou, J. (2006) ‘Human Ear Recognition from Face Profile Images’, in D. Zhang and A.K. Jain (eds) *ICB’06: Proceedings of the 2006 international conference on Advances in Biometrics*. Hong Kong, China: Springer Verlag, Berlin Heidelberg, pp. 786–792. Available at: https://doi.org/10.1007/11608288_105.

Abudarham, N., Shkiller, L. and Yovel, G. (2019) ‘Critical features for face recognition’, *Cognition*, 182, pp. 73–83. Available at: https://doi.org/10.1016/j.cognition.2018.09.002.

Abudarham, N. and Yovel, G. (2016) ‘Reverse engineering the face space: Discovering the critical features for face identification’, *Journal of Vision*, 16(3), pp. 1–18. Available at: https://doi.org/10.1167/16.3.40.

ACPO (2009) ‘Facial Identification Guidance 2009’, p. 45. Available at: http://library.college.police.uk/docs/acpo/facial-identification-guidance-2009.pdf.

Adams, G.W. (2016) *Utilizing Forensic Technologies for Unidentified Human Remains*. eBook. Taylor & Francis Group, LLC.

Adjabi, I., Ouahabi, A., Benzaoui, A. and Taleb-Ahmed, A. (2020) ‘Past, Present, and Future of Face Recognition: A Review’, *Electronics*, 9(8), pp. 1188-. Available at: https://doi.org/10.3390/electronics9081188.

Afzal, H.M.R., Luo, S., Afzal, M.K., Chaudhary, G., Khari, M. and Kumar, S.A.P. (2020) ‘3D Face Reconstruction From Single 2D Image Using Distinctive Features’, *IEEE Access*, 8, pp. 180681–180689. Available at: https://doi.org/10.1109/ACCESS.2020.3028106.

Aggrawal, A., Setia, P., Gupta, A. and Busuttil, A. (2010) ‘Age Evaluation after Growth Cessation’, in S. Black, A. Aggrawal, and J. Payne-James (eds) *Age Estimation in the Living*. Chichester, UK: John Wiley & Sons, Ltd., pp. 236–266. Available at: https://doi.org/10.1002/9780470669785.

AGIB (2011) *Standards für die Identifikation lebender Personen nach Bildern - Grundlagen, Kriterien und Verfahrensregeln für Gutachten*, *Standard*. Available at: http://bildidentifikation.de/agib-bildidentifikation-2/ (Accessed: 10 December 2021).

Ahmed, M.S., Ikram, S., Bibi, N. and Mir, A. (2017) ‘Hutchinson-Gilford Progeria Syndrome: A Premature Aging Disease’, *Molecular Neurobiology*, 55, pp. 4417–4427. Available at: https://doi.org/0.1007/s12035-017-0610-7.

AJL (2021) *Algorithmic Justice League - Unmasking AI harms and biases*. Available at: https://www.ajl.org/ (Accessed: 21 May 2021).

Al-Amad, S., McCullough, M., Graham, J., Clement, J. and Hill, A. (2006) ‘Craniofacial identification by computer-mediated superimposition’, *Journal of Forensic Odonto-*

Stomatology, 24(2), pp. 47–52.

Al-Meyah, K., Marshall, D. and Rosin, P.L. (2017) ‘4D Analysis of Facial Ageing Using Dynamic Features’, *Procedia Computer Science*, 112, pp. 790–799. Available at: https://doi.org/10.1016/j.procs.2017.08.037.

Alashkar, T., Amor, B. Ben, Daoudi, M. and Berretti, S. (2014) ‘A 3D Dynamic Database for Unconstrained Face Recognition’, *Proceedings of the 5th International Conference on 3D Body Scanning Technologies, Lugano, Switzerland, 21-22 October 2014*, (October), pp. 357–364. Available at: https://doi.org/10.15221/14.357.

Albert, A.M., Ricanek Jr., K. and Patterson, E. (2007) ‘A review of the literature on the aging adult skull and face: Implications for forensic science research and applications’, *Forensic Science International*, 172, pp. 1–9. Available at: https://doi.org/10.1016/j.forsciint.2007.03.015.

Albert, M., Sethuram, A. and Ricanek, K. (2011) ‘Implications of Adult Facial Aging on Biometrics’, in M. Albert (ed.) *Biometrics - Unique and Diverse Applications in Nature, Science, and Technology*. IntechOpen, pp. 89–106. Available at: https://doi.org/10.5772/645.

Alemán Aguilera, M.I. (2022) ‘European Academic Perspectives on MDVI’, in *Migrant Disaster Victim Identification Symposium*. Online.

Alenezi, H.M. and Bindemann, M. (2013) ‘The effect of feedback on face-matching accuracy’, *Applied Cognitive Psychology*, 27(6), pp. 735–753. Available at: https://doi.org/10.1002/acp.2968.

Ali, T., Spreeuwers, L. and Veldhuis, R. (2012) ‘Forensic Face Recognition: A Survey’, in A. Quaglia and C.M. Epifano (eds) *Face Recognition: Methods, Applications and Technology*. NOVA Science Publishers, pp. 187–238.

Allanson, J.E., Cunniff, C., Hoyme, H.E., McGaughran, J., Muenke, M. and Neri, G. (2009) ‘Elements of morphology: Standard terminology for the head and face’, *American Journal of Medical Genetics*, 149A(1), pp. 6–28. Available at: https://doi.org/10.1002/ajmg.a.32612.

Alonso, A., Martin, P., Albarrán, C., Garcia, P., Fernandez de Simon, L., Jesús Iturralde, M., Fernández-Rodriguez, A., Atienza, I., Capilla, J., García-Hirschfeld, J., Martinez, P., Vallejo, G., García, O., García, E., Real, P., Alvarez, D., León, A. and Sancho, M. (2005) ‘Challenges of DNA profiling in mass disaster investigations’, *Croatian medical journal*, 46(4), pp. 540–548. Available at: https://doi.org/10.1097/01.brs.0000224516.29747.52.

Alt, K.W. (2012) ‘Unverwechselbare Botschaften - Das Gesicht im Spiegel sozialer Wahrnehmung’, *MKG-Chirurg*, 5(1), pp. 7–15. Available at: https://doi.org/10.1007/s12285-011-0247-x.

Alzueta, E., Melcon, M., Poch, C. and Capilla, A. (2019) ‘Is your own face more than a highly familiar face?’, *Biological Psychology*, 142, pp. 100–107. Available at: https://doi.org/10.1016/j.biopsycho.2019.01.018.

Anderson, B.E. (2008) ‘Identifying the dead: methods utilized by the Pima County (Arizona) office of the medical examiner for undocumented border crossers: 2001-2006’, *Journal of Forensic Sciences*, 53(1), pp. 8–15. Available at: https://doi.org/10.1111/j.1556-4029.2007.00609.x.

Anderson, G.S. and Cervenka, V.J. (2002) 'Insects associated with the body: their use and analyses', in W.D. Haglund and M.H. Sorg (eds) *Advances in forensic taphonomy: method, theory and archaeological perspectives*. Boca Raton, Florida: CRC Press, pp. 193–200.

Andress, J. (2014) 'Identification and Authentication', in *The Basics of Information Security: Understanding the Fundamentals of Infosec in Theory and Practice*. 2nd edn. Oxford: Syngress, Elsevier Inc., pp. 23–38. Available at: https://doi.org/10.1016/B978-0-12-800744-0.00002-6.

De Angelis, D., Sala, R., Cantatore, A., Grandi, M. and Cattaneo, C. (2009) 'A new computer-assisted technique to aid personal identification', *International Journal of Legal Medicine*, 123(4), pp. 351–356. Available at: https://doi.org/10.1007/s00414-008-0311-x.

Anwarul, S. and Dahiya, S. (2019) 'A Comprehensive Review on Face Recognition Methods and Factors Affecting Facial Recognition Accuracy', in *Chamber Research and Information Center - ICRIC 2019*. Springer, Cham, pp. 495–514. Available at: https://doi.org/10.1007/978-3-030-29407-6_36.

Arnott, S. and Turner, B. (2008) 'Post-feeding larval behaviour in the blowfly, Calliphora vicina: effects on post-mortem interval estimates', *Forensic Science International*, 177(2–3), pp. 162–167. Available at: https://doi.org/10.1016/j.forsciint.2007.12.002.

Ashley, K.C. (2020) 'Data of the Dead: A Proposal for Protecting Posthumous Data Privacy', *William & Mary Law Review*, 62(2), pp. 649–682. Available at: https://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=3880&context=wmlr.

Asi, S.M., Ismail, N.H., Ahmad, R., Ramlan, E.I. and Rahman, Z.A.A. (2014) 'Automatic craniofacial anthropometry landmarks detection and measurements for the orbital region', *Procedia Computer Science*, 42, pp. 372–377. Available at: https://doi.org/10.1016/j.procs.2014.11.076.

ASTM (2018) *Standard Guide for Facial Image Comparison Feature List for Morphological Analysis E3149-18*. Available at: https://doi.org/10.1520/E3149-18.

Atlas, T. (1994) *DAYS OF THE JACKAL END IN KHARTOUM*, *Chicago Tribune newspaper article*. Available at: https://www.chicagotribune.com/news/ct-xpm-1994-08-16-9408160156-story.html (Accessed: 8 May 2020).

Atsuchi, M., Tsuji, A., Usumoto, Y., Yoshino, M. and Ikeda, N. (2013) 'Assessment of some problematic factors in facial image identification using a 2D/3D superimposition technique', *Legal Medicine*, 15(5), pp. 244–248. Available at: https://doi.org/10.1016/j.legalmed.2013.06.002.

Aulsebrook, W.A., Iscan, M.Y., Slabbert, J.H. and Becker, P. (1995) 'Superimposition and reconstruction in forensic facial identification: a survey', *Forensic Science International*, 75, pp. 101–120.

Austin-Smith, D. and Maples, W.R. (1994) 'The Reliability of Skull/Photograph Superimposition in individual Identification', *Journal of Forensic Sciences*, 39(2), pp. 446–455.

Avelar, L.E.T., Cardoso, M.A., Bordoni, L.S., Avelar, L.M. and Avelar, J.V.M. (2017) 'Aging and Sexual Differences of the Human Skull', *Plastic and Reconstructive Surgery Global Open*, 5(4), pp. e1297-1302. Available at:

https://doi.org/10.1097/GOX.0000000000001297.

Azouz, Z.B., Shu, C. and Mantel, A. (2006) 'Automatic Locating of Anthropometric Landmarks on 3D Human Models', in *IEEE Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 1–8.

Bacci, N., Briers, N. and Steyn, M. (2021) 'Assessing the effect of facial disguises on forensic facial comparison by morphological analysis', *Journal of Forensic Sciences*, 66(4), pp. 1120–1233. Available at: https://doi.org/10.1111/1556-4029.14722.

Bacci, N., Davimes, J.G., Steyn, M. and Briers, N. (2021) 'Forensic Facial Comparison: Current Status, Limitations, and Future Directions', *Biology*, 10(1269), pp. 1–26. Available at: https://doi.org/10.3390/biology10121269.

Bacci, N., Houlton, T.M., Briers, N. and Steyn, M. (2021) 'Validation of forensic facial comparison by morphological analysis in photographic and CCTV samples', *International Journal of Legal Medicine*, 135, pp. 1965–1981. Available at: https://doi.org/10.1007/s00414-021-02512-3.

Balazia, M., Happy, S.L., Bremond, F. and Dantcheva, A. (2021) 'How Unique Is a Face: An Investigative Study', in *2020 25th International Conference on Pattern Recognition (ICPR)*. Milan, Italy, pp. 7066–7071. Available at: https://doi.org/10.1109/ICPR48806.2021.9412446.

Balsdon, T., Summersby, S., Kemp, R.I. and White, D. (2018) 'Improving face identification with specialist teams', *Cognitive Research: Principles and Implications*, 3(25), pp. 1–13. Available at: https://doi.org/10.1186/s41235-018-0114-7.

Baraybar, J.P. (2022) 'Migrant Disaster Victim Identification (MDVI): Initiatives in Africa', in *Migrant Disaster Victim Identification Symposium*. Online.

Baraybar, J.P., Caridi, I. and Stockwell, J. (2020) 'A forensic perspective on the new disappeared: Migration revisited', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 101–115. Available at: https://doi.org/10.1002/9781119482062.ch7.

Barton, J.J.S., Davies-Thompson, J. and Corrow, S.L. (2021) 'Prosopagnosia and disorders of face processing', in L.J.S. Barton and A. Leff (eds) *Handook of Clinical Neurology*. Vol. 178. Elsevier B.V., pp. 175–193. Available at: https://www.sciencedirect.com/science/article/pii/B9780128213773000064.

Batres, C., Kramer, S.S., DeAngelis, C.G. and Russell, R. (2019) 'Examining the "cosmetics placebo effect"', *PLoS One*, 14(1), pp. 1–8. Available at: https://doi.org/10.1371/journal.pone.0210238.

Beaumont, J., Beveridge, J., Budde, W., Cross, M., Gutiérrez, A., Caffell, A., Hengelhaupt, I., Johnston, A., Koon, H., Mackenzie, L., Montgomery, J., Radini, A., Roberts, C., Roughley, M., Speller, C. and Wilkinson, C. (2018) *Lost Lives, New Voices: Unlocking the Stories of the Scottish Soldiers at the Battle of Dunbar 1650*. Edited by C. Gerrard, P. Graves, A. Millard, R. Annis, and A. Caffell. Oxford: Oxbow Books. Available at: https://doi.org/10.2307/j.ctvh1dk49.

Beauthier, J.-P., Lefèvre, P. and De Valck, E. (2011) 'Autopsy and Identification Techniques', in N.-A. Mörner (ed.) *The Tsunami Threat - Research and Technology*.

IntechOpen, pp. 691–714. Available at: https://doi.org/10.5772/13610.

Becker, B.C. and Ortiz, E.G. (2013) 'Evaluating Open-Universe Face Identification on the Web', in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 904–911. Available at: https://www.cv-foundation.org/openaccess/content_cvpr_workshops_2013/W16/html/Becker_Evaluating_Open-Universe_Face_2013_CVPR_paper.html.

Belhumeur, P.N., Hespanha, J.P. and Kriegman, D.J. (1997) 'Eigenfaces vs. Fisherfaces: recognition using class specific linear projection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), pp. 711–720. Available at: https://doi.org/10.1109/34.598228.

Belsey, S.L. and Flanagan, R.J. (2016) 'Postmortem biochemistry: Current applications', *Journal of Forensic and Legal Medicine*, 41, pp. 49–57. Available at: https://doi.org/10.1016/j.jflm.2016.04.011.

Bennett, C. (2020) 'Is DNA always the answer?', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 521–536. Available at: https://doi.org/10.1002/9781119482062.ch33.

Benta, K.-I. and Vaida, M.-F. (2015) 'Towards Real-Life Facial Expression Recognition Systems', *Advances in Electrical and Computer Engineering*, 15(2), pp. 93–102. Available at: https://doi.org/10.4316/AECE.2015.02012.

Berezowski, V., Mallett, X. and Moffat, I. (2020) 'Geomatic techniques in forensic science: A review', *Science & Justice*, 60(2), pp. 99–107. Available at: https://doi.org/10.1016/j.scijus.2019.10.006.

Berman, G.L. and Cutler, B.L. (1998) 'The influence of processing instructions at encoding and retrieval on face recognition accuracy', *Psychology, Crime & Law*, 4(2), pp. 89–106. Available at: https://doi.org/10.1080/10683169808401751.

Berretti, S., Ben Amor, B., Drira, H., Daoudi, M., Srivastava, A., del Bimbo, A. and Pala, P. (2013) 'Applications', in *3D Face Modeling, Analysis and Recognition*. 1st edn. Chicester, UK: Wiley & Sons Ltd, pp. 149–202.

Bertillon, A. (1890) *La Photographie Judiciaire*. Paris: Gauthier-Villars. Available at: https://wellcomelibrary.org/item/b21041787.

Besson, G., Barragan-Jason, G., Thorpe, S.J., Fabre-Thorpe, M., Puma, S., Ceccaldi, M. and Barbeau, E.J. (2017) 'From face processing to face recognition: Comparing three different processing levels', *Cognition*, 158, pp. 33–43. Available at: https://doi.org/10.1016/j.cognition.2016.10.004.

Betters, E. (2016) *Google is shutting down Picasa: What'll happen to your photos and videos?*, *Article*. Available at: https://www.pocket-lint.com/apps/news/google/136712-google-is-shutting-down-picasa-what-ll-happen-to-your-photos-and-videos (Accessed: 7 January 2021).

Beveridge, J.R., Givens, G.H., Phillips, P.J., Draper, B.A., Bolme, D.S. and Lui, Y.M. (2010) 'FRVT 2006: Quo Vadis face quality', *Image and Vision Computing*, 28, pp. 732–743. Available at: https://doi.org/10.1016/j.imavis.2009.09.005.

Bhanu, B. and Govindaraju, V. (2011) ‘Multibiometrics for human identification’, *Multibiometrics for Human Identification*, 9780521115, pp. 1–388. Available at: https://doi.org/10.1017/CBO9780511921056.

Bichsel, M. (1991) *Strategies of robust object recognition for the automatic identification of human faces*. ETH Zürich, Switzerland. Available at: https://doi.org/10.3929/ethz-a-000596296.

Biedermann, A., Bozza, S., Taroni, F. and Aitken, C. (2017) ‘The meaning of justified subjectivism and its role in the reconciliation of recent disagreements over forensic probabilism’, *Science & Justice*, 57(6), pp. 477–483. Available at: https://doi.org/10.1016/j.scijus.2017.08.005.

Bikker, J. (2014) ‘Identification of Missing Persons and Unidentified Remains in Disaster Victim Identification - Recommendations and Best Practice’, in X. Mallett, T. Blythe, and R. Berry (eds) *Advances in Forensic Human Identification*. 1st edn. Boca Raton, Florida: Taylor & Francis Group, LLC, pp. 37–58. Available at: https://doi.org/10.1201/b16509.

Bilge, Y., Kedici, P.S., Alakoc, Y.D., Ülküer, K.Ü. and Ilkyaz, Y.Y. (2003) ‘The identification of a dismembered human body: a multidisciplinary approach’, *Forensic Science International*, 137, pp. 141–146. Available at: https://doi.org/10.1016/S0379-0738(03)00334-7.

Bindemann, M., Attard, J. and Johnston, R.A. (2014) ‘Perceived ability and actual recognition accuracy for unfamiliar and famous faces’, *Cogent Psychology*, 1(1), pp. 1–15. Available at: https://doi.org/10.1080/23311908.2014.986903.

Bindemann, M., Attard, J., Leach, A. and Johnston, R.A. (2013) ‘The effect of image pixelation on unfamiliar-face matching’, *Applied Cognitive Psychology*, 27(6), pp. 707–717. Available at: https://doi.org/10.1002/acp.2970.

Bindemann, M., Avetisyan, M. and Rakow, T. (2012) ‘Who can recognize unfamiliar faces? Individual differences and observer consistency in person identification’, *Journal of Experimental Psychology: Applied*, 18(3), pp. 277–291. Available at: https://doi.org/10.1037/a0029635.

Bindemann, M., Burton, A.M. and Jenkins, R. (2005) ‘apacity limits for face processing’, *Cognition*, 98(2), pp. 177–197. Available at: https://doi.org/10.1016/j.cognition.2004.11.004.

Bindemann, M. and Burton, M. (2021) ‘Steps Towards a Cognitive Theory of Unfamiliar Face Matching’, in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 38–61.

Bindemann, M., Sandford, A., Gillatt, K., Avetisyan, M. and Megreya, A.M. (2012) ‘Recognising faces seen alone or with others: Why are two heads worse than one?’, *Perceptionption*, 41(4), pp. 415–435. Available at: https://doi.org/10.1068/p6922.

Birngruber, C.G., Kreutz, K., Ramsthaler, F., Krähahn, J. and Verhoff, M.A. (2010) ‘Superimposition technique for skull identification with Afloat® software’, *International Journal of Legal Medicine*, 124(5), pp. 471–475. Available at: https://doi.org/10.1007/s00414-010-0494-9.

Birngruber, C.G., Ramsthaler, F., Mattias, K. and Verhoff, M.A. (2011) ‘Superimposition of ante- and post-mortem photographs of tattoos as a means of identification - a case report', *Archiv für Kriminologie - Case Reports*, 227(1–2), pp. 48–54.

Bisgaard Munk, T. (2017) '100,000 false positives for every real terrorist: Why anti-terror algorithms don't work', *First Monday*, 22(9). Available at: https://doi.org/10.5210/fm.v22i9.7126.

Bisker, C. and Ralebitso-Senior, T.K. (2018) 'The Method Debate: A State-of-the-Art Analysis of PMI Estimation Techniques', in T.K. Ralebitso-Senior (ed.) *Forensic Ecogenomics - The Application of Microbial Ecology Analyses in Forensic Contexts*. Academic Press, Elsevier, pp. 61–86. Available at: https://doi.org/10.1016/B978-0-12-809360-3.00003-5.

Black, S. (2006) 'Introduction', in T. Thompson and S. Black (eds) *Forensic Human Identification: An Introduction*. 1st edn. Boca Raton, Florida: CRC Press, pp. xi–xvi. Available at: https://doi.org/10.1201/9781420005714.

Black, S. and Bikker, J. (2016) 'Forensic identification', in K.S. Greene and L. Alys (eds) *Missing Persons: A handbook of research*. Routledge, pp. 188–199.

Black, S., MacDonald-McMillan, B., Mallett, X., Rynn, C. and Jackson, G. (2014) 'The incidence and position of melanocytic nevi for the purposes of forensic image comparison', *International Journal of Legal Medicine*, 128(3), pp. 535–543. Available at: https://doi.org/10.1007/s00414-013-0821-z.

Black, S. and Thompson, T. (2007) 'Body Modification', in T. Thompson and S. Black (eds) *Forensic Human Identification: An Introduction*. Boca Raton, Florida: CRC Press, pp. 379–398.

Blanz, V. and Vetter, T. (2003) 'Face Recognition Based on Fitting a 3D Morphable Model', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), pp. 1063–1074.

Blau, S. (2016) 'More Than Just Bare Bones - Ethical Considerations for Forensic Anthropologists', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 593–606.

Blau, S. and Forbes, S. (2016) 'Anthropology: Taphonomy in the Forensic Context', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 227–235. Available at: https://doi.org/10.1016/B978-0-12-800034-2.0021-5.

Blau, S. and Hill, A. (2009) 'Disaster victim identification: a review', *Minerva Medicolegale*, 129(1), pp. 35–46.

Blau, S. and Ubelaker, D.H. (2016a) 'Forensic Anthropology and Archaeology - Moving Forward', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 1–9.

Blau, S. and Ubelaker, D.H. (2016b) 'International Perspectives on Issues in Forensic Anthropology', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 672–677.

Blauch, N.M., Behrmann, M. and Plaut, D.C. (2021) 'Computational insights into human perceptual expertise for familiar and unfamiliar face recognition', *Cognition*, 208, pp. 1–13. Available at: https://doi.org/10.1016/j.cognition.2020.104341.

Bledsoe, W.W. (1964) *The Model Method in Facial Recognition, Technical Report PRI 15*.

Palo Alto, California. Available at: https://www.historyofinformation.com/detail.php?entryid=2495.

Bledsoe, W.W. (1965) *A Man-Machine Facial Recognition System-Some Preliminary Results, Technical Report PRI 19A*. Palo Alto, California. Available at: https://www.historyofinformation.com/detail.php?entryid=2495.

Bledsoe, W.W. (1966) *Man-Machine Facial Recognition: Report on a Large-Scale Experiment, Technical Report PRI 22*. Palo Alto, California. Available at: https://www.historyofinformation.com/detail.php?entryid=2495.

Blythe, T. and Woodforde, S. (2007) 'Missing Persons in the United Kingdom', *Forensic Human Identification An Introduction*, pp. 425–443. Available at: https://doi.org/doi:10.1201/9781420005714.ch24.

de Boer, H.H., Roberts, J., Delabarde, T., Mundorff, A.Z. and Blau, S. (2020) 'Disaster victim identification operations with fragmented, burnt, or commingled remains: experience-based recommendations', *Forensic Sciences Research*, 5(3), pp. 191–201. Available at: https://doi.org/10.1080/20961790.2020.1751385.

Bolme, D.S., Srinivas, N., Brogan, J. and Cornett, D. (2020) 'Face Recognition Oak Ridge (FaRO): A Framework for Distributed and Scalable Biometrics Applications', in *IEEE International Joint Conference on Biometrics (IJCB)*. Houston TX ZSA: IEEE. Available at: https://doi.org/10.1109/IJCB48548.2020.9304933.

Bolme, D.S., Tokola, R.A., Boehnen, C.B., Saul, T.B., Sauerwein, K.A. and Steadman, D.W. (2016) 'Impact of environmental factors on biometric matching during human decomposition', in *IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. Niagara Falls, NY, USA: IEEE, pp. 1–7. Available at: https://doi.org/10.1109/BTAS.2016.7791177.

Bortolon, C. and Raffard, S. (2018) 'Self-face advantage over familiar and unfamiliar faces: A three-level meta-analytic approach', *Psychonomic Bulletin & Review*, 25, pp. 1287–1300. Available at: https://doi.org/10.3758/s13423-018-1487-9.

Boss, P. (2010) 'The trauma and complicated grief of ambiguous loss', *Pastoral Psychology*, 59(2), pp. 137–145. Available at: https://doi.org/10.1007/s11089-009-0264-0.

Bouguila, J. and Khochtali, H. (2020) 'Facial plastic surgery and face recognition algorithms: Interaction and challenges. A scoping reiew and future directions', *Journal of Stomatology, Oral and Maxillofacial Surgery*, 121, pp. 696–703. Available at: https://doi.org/10.1016/j.jormas.2020.06.007.

Bourlai, T., Ross, A. and Jain, A.K. (2011) 'Restoring degraded face images: A case study in matching faxed, printed, and scanned photos', *IEEE Transactions on Information Forensics and Security*, 6(2), pp. 371–384. Available at: https://doi.org/10.1109/TIFS.2011.2109951.

Bowyer, K.W., Chang, K. and Flynn, P. (2004) 'A Survey of 3D and Multi-Modal 3D+2D Face Recognition', in *International Conference on Pattern Recognition (ICPR)*.

Boyer, A.O. and Boyer, R.S. (1991) 'A Biographical Sketch of W. W. Bledsoe', in R.S. Boyer (ed.) *Automated Reasoning. Automated Reasoning Series, vol 1*. Dordrecht: Springer, pp. 1–19. Available at: https://doi.org/10.1007/978-94-011-3488-0_1.

Broach, J., Yong, R., Manuell, M.E. and Nichols, C. (2017) 'Use of Facial Recognition

Software to Identify Disaster Victims with Facial Injuries', *Disaster Medicine and Public Health Preparedness*, 11(5), pp. 568–572. Available at: https://doi.org/10.1017/dmp.2016.207.

Bromby, M.C. (2006) 'CCTV and Expert Evidence: Addressing the Reliability of New Sciences', *Archbold News*, 2 November, pp. 6–9.

Bronkhorst, D., Kenter, L. and Stratmann, H. (1998) *Excuses for the truth: Disappearances and their consequences*. 1st edn. Amsterdam, NL: Engels.

Brooks, D.N. (2017) *The Kardashian Effect: The Impact of Selfie Culture on Millenial Women*. Trinity Washington University.

Brough, A.L., Morgan, B. and Rutty, G.N. (2015) 'The basics of disaster victim identification', *Journal of Forensic Radiology and Imaging*, 3(1), pp. 29–37. Available at: https://doi.org/10.1016/j.jofri.2015.01.002.

Bruce, V. (2012) 'Familiar face recognition', *Craniofacial Identification*, pp. 1–10. Available at: https://doi.org/10.1017/CBO9781139049566.001.

Bruce, V. and Burton, M. (2002) 'Learning new faces', in M. Fahle and T. Poggio (eds) *Perceptual Learning*. MIT Press, pp. 317–334.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P.J.B., Burton, A.M. and Miller, P. (1999) 'Verification of face identities from images captured on video', *Journal of Experimental Psychology: Applied*, 5(4), pp. 339–360. Available at: https://doi.org/10.1037/1076-898X.5.4.339.

Bruce, V., Henderson, Z. and Newman, C. (2001) 'Matching Identities of Familiar and Unfamiliar Faces Caught on CCTV Images', *Journal of Experimental Psychology: Applied*, 7(3), pp. 207–218. Available at: https://doi.org/10.1037//1076-898X.7.3.207.

Bruce, V. and Young, A. (1986) 'Understanding face recognition', *British Journal of Psychology*, 77(3), pp. 305–327. Available at: https://doi.org/10.1111/j.2044-8295.1986.tb02199.x.

Brucker, M.J., Patel, J. and Sullivan, P.K. (2003) 'A morphometric study of the external ear: age- and sex-related differences', *Plastic and Reconstructive Surgery*, 112(2), pp. 647–652. Available at: https://doi.org/10.1097/01.PRS.0000070979.20679.1F.

Brumfield, B. (2013) *Selfie named word of the year for 2013*, *CNN News*. Available at: https://edition.cnn.com/2013/11/19/living/selfie-word-of-the-year/index.html (Accessed: 27 May 2020).

Brunelli, R. and Poggio, T. (1993) 'Face recognition: features versus templates', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10), pp. 1042–1052. Available at: https://doi.org/10.1109/34.254061.

Bruner, J.S. and Tagiuri, R. (1954) 'The Perception of People', in G. Lindzey (ed.) *The Handbook of Social Psychology*. 1st, Vol 2 edn. Reading, MA: Addison-Wesley, pp. 634–654. Available at: http://wjh-www.harvard.edu/~dtg/Ordinary Personology.pdf.

Bundesamt für Sicherheit in der Informationstechnik (2021) *Biometrische Verfahren*. Available at: https://www.bsi.bund.de/DE/Karriere/Einstieg/Studium-und-Ausbildung/Bachelor-und-Masterarbeiten/Arbeiten/Biometrische_Verfahren.html

(Accessed: 12 April 2021).

Bundesamt für Sicherheit in der Informationstechnik and Frauenhofer Institut (2006) 'Gesichtserkennung, dreidimensional - Über biometrische Verfahren', *Forschung & Lehre*, 8, pp. 444–445.

Bundeskriminalamt (2006) *Forschungsprojekt Foto-Fahndung - Biometrie Kompakt, Article*. Available at: https://www.bka.de/DE/UnsereAufgaben/Forschung/ForschungsprojekteUndErgebnisse/Foto-Fahndung/Biometrie/biometrieKompakt_node.html;jsessionid=56131F7FE66EE06579E3B241BFF9B550.live2302 (Accessed: 12 April 2021).

Bundeskriminalamt (2007) *Forschungsprojekt: Gesichtserkennung als Fahndungshilfsmittel Foto-Fahndung - Abschlussbericht*. Wiesbaden, Germany. Available at: file:///C:/Users/ISABEL~1/AppData/Local/Temp/fotofahndungAbschlussbericht.pdf%3Bjsessionid=56131F7FE66EE06579E3B241BFF9B550.pdf.

Buolamwini, J. and Gebru, T. (2018) 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification∗', in S.A. Friedler and C. Wilson (eds) *Proceedings of the 1st Conference on Fairness, Accountability and Transparency; 81*, pp. 77–91. Available at: http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

Burg, G. (2019) 'Changes in color of the skin and systemic disease', *Clinics in Dermatology*, 37, pp. 610–617. Available at: https://doi.org/10.1016/j.clindermatol.2019.07.033.

Burghardt, T. (2012) *A brief review of biometric identification*. Technical Report. Available at: https://research-information.bris.ac.uk/en/publications/a-brief-review-of-biometric-identification.

Burns, K.R. (2007) *Forensic Anthropology Training Manual*. 2nd edn. Pearson Education LTD.

Burton, A.M., Jenkins, R., Hancock, P.J.B. and White, D. (2005) 'Robust representations for face recognition: the power of averages', *Cognitive Psychology*, 51(3), pp. 256–284. Available at: https://doi.org/10.1016/j.cogpsych.2005.06.003.

Burton, A.M., Miller, P., Bruce, V., Hancock, P.J. and Henderson, Z. (2001) 'Human and automatic face recognition: a comparison across image formats', *Vision Research*, 41(24), pp. 3185–3195. Available at: https://doi.org/10.1016/s0042-6989(01)00186-9.

Burton, A.M., White, D. and McNeill, A. (2010) 'The Glasgow Face Matching Test', *Behavior Research Methods*, 42(1), pp. 286–291. Available at: https://doi.org/10.3758/BRM.42.1.286.

Busch, C., Brauckmann, M., Veldhuis, R., Deravi, F., Keevenaar, T., Nouak, A., Seibert, H., Weber, F. and Suchier, J.-M. (2012) 'Towards a more secure border control with 3D face recognition', in *5th Norsk Informasjons Sikkerhets Konferanse (NISK)*. Bodø, Norway, pp. 1–12. Available at: https://www.christoph-busch.de/files/Busch-3DFace-nisk-121012.pdf.

Busch, C., Daum, H. and Graf, F. (2003) 'Evaluierung von Gesichtserkennungssystemen - Projekt BioFace', in *IT-Sicherheit im verteilten Chaos. Tagungsband : 8. Deutscher IT-Sicherheitskongress des BSI*. Bonn, Germany: SecuMedia-Verlag, pp. 121–131. Available at: http://publica.fraunhofer.de/starweb/pub09/servlet.starweb.

Busch, C. and Nouak, A. (2007) ‘Das EU-Projekt »3D Face«. 3-D-Gesichtserkennung für die unbeaufsichtigte Grenzkontrolle. Beitrag zum 10. Deutschen IT-Sicherheitskongress des Bundesamtes für Sicherheit in der Informationstechnik im Themenbereich »2 Identitätsmanagement«’, in *10. Deutscher IT Sicherheitskongress des Bundesamtes für Sicherheit in der Informatik - Themenbereich 2 ‘Identitätsmanagement’*. Bad Godesberg, Germany: Bundesamtes für Sicherheit, pp. 1–14. Available at: https://atbc.de/pdf/BSI10Sicherheitskongress.pdf.

Buttle, H. and East, J. (2010) ‘Traditional facial tattoos disrupt face recognition processes’, *Perception*, 39(1), pp. 1672–1674. Available at: https://doi.org/10.1068/p6790.

Byard, R.W. (2016) ‘Tattoos: Medicolegal Significance – Forensic Issues’, in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 520–532. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00372-4.

Byard, R.W. (2020) ‘Estimation of the time since death in the early postmortem period (24–48 hours)’, in J. Hayman and M. Oxenham (eds) *Estimation of the Time since Death - Current Research and Future Trends*. Academic Press, Elsevier, pp. 11–27. Available at: https://doi.org/10.1016/B978-0-12-815731-2.00002-9.

Calleja-Agius, J., Muscat-Baron, Y. and Brincat, M.P. (2007) ‘Skin ageing’, *Menopause International*, 13, pp. 60–64.

Cambridge Dictionary (2021a) *IDENTITY | meaning in the Cambridge English Dictionary*. Available at: IDENTITY %7C meaning in the Cambridge English Dictionary (Accessed: 14 October 2021).

Cambridge Dictionary (2021b) *RECOGNITION | meaning in the Cambridge English Dictionary*. Available at: https://dictionary.cambridge.org/dictionary/english/recognition (Accessed: 14 October 2021).

Campobasso, C.P., Linville, J.G., Wells, J.D. and Introna, F. (2005) ‘Forensic Genetic Analysis of Insect Gut Contents’, *American Journal of Forensic Medicine & Pathology*, 26(2), pp. 161–165. Available at: https://doi.org/10.1097/01.paf.0000163832.05939.59.

Campomanes-Álvarez, B.R., Cordón, Ó., Damas, S. and Ibáñez, Ó. (2014) ‘Computer-based craniofacial superimposition in forensic identification using soft computing’, *Journal of Ambient Intelligence and Humanized Computing*, 5(5), pp. 683–697. Available at: https://doi.org/10.1007/s12652-012-0168-1.

Campomanes-Álvarez, B.R., Ibáñez, O., Navarro, F., Alemán, I., Botella, M., Damas, S. and Cordón, O. (2014) ‘Computer vision and soft computing for automatic skull-face overlay in craniofacial superimposition’, *Forensic Science International*, 245, pp. 77–86. Available at: https://doi.org/10.1016/j.forsciint.2014.10.009.

Campomanes-Álvarez, B.R., Ibáñez, O., Navarro, F., Alemán, I., Cordón, O. and Damas, S. (2015) ‘Dispersion assessment in the location of facial landmarks on photographs’, *International Journal of Legal Medicine*, 129(1), pp. 227–236. Available at: https://doi.org/10.1007/s00414-014-1002-4.

Campomanes-Alvarez, C., Ibanez, O., Cordon, O. and Wilkinson, C. (2018) ‘Hierarchical information fusion for decision making in craniofacial superimpostion’, *Information Fusion*, 39, pp. 25–40. Available at: http://researchonline.ljmu.ac.uk/id/eprint/6343.

Campomanes-Álvarez, C., Martos Fernández, R., Wilkinson, C., Ibáñez, O. and Cordón, O.

(2018) ‘Modeling Skull-Face Anatomical/Morphological Correspondence for Craniofacial Superimposition-Based Identification’, *IEEE Transactions on Information Forensics and Security*, 13(6), pp. 1481–1494. Available at: https://doi.org/10.1109/TIFS.2018.2791434.

Caple, J. and Stephan, C.N. (2016) ‘A standardized nomenclature for craniofacial and facial anthropometry’, *International Journal of Legal Medicine*, 103, pp. 863–879. Available at: https://doi.org/10.1007/s00414-015-1292-1.

Čaplová, Z. (2017) *Morphology of the Face as Post-mortem Personal Identifier*. Università degli Studi di Milano.

Caplova, Z., Compassi, V., Giancola, S., Gibelli, D.M., Obertová, Z., Poppa, P., Sala, R., Sforza, C. and Cattaneo, C. (2017) ‘Recognition of children on age-different images: Facial morphology and age-stable features’, *Science and Justice*, 57, pp. 250–256. Available at: https://doi.org/10.1016/j.scijus.2017.03.005.

Caplova, Z., Gibelli, D.M., Poppa, P., Cummaudo, M., Obertova, Z., Sforza. C. and Cattaneo, C. (2018) ‘3D quantitative analysis of early decomposition changes of the human face’, *International Journal of Legal Medicine*, 132, pp. 649–653. Available at: https://doi.org/10.1007/s00414-017-1647-x.

Caplova, Z., Obertova, Z., Gibelli, D.M., De Angelis, D., Mazzarelli, D., Sforza, C. and Cattaneo, C. (2018) ‘Personal Identification of Deceased Persons: An Overview of the Current Methods Based on Physical Appearance’, *Journal of Forensic Sciences*, 63(3), pp. 662–671. Available at: https://doi.org/10.1111/1556-4029.13643.

Caplova, Z., Obertova, Z., Gibelli, D.M., Mazzarelli, D., Fracasso, T., Vanezis, P., Sforza, C. and Cattaneo, C. (2017) ‘The Reliability of Facial Recognition of Deceased Persons on Photographs’, *Journal of Forensic Sciences*, 62(5), pp. 1286–1291. Available at: https://doi.org/10.1111/1556-4029.13396.

Cappella, A., De Angelis, D., Mazzarelli, D., Vitale, A., Caccia, G., Fracasso, T. and Cattaneo, C. (2022) ‘Rediscovering the value of images in supporting personal identification of missing migrants’, *Legal Medicine*, 54, pp. 1–6. Available at: https://doi.org/10.1016/j.legalmed.2021.101985.

Cappella, A., Gibelli, D., Cellina, M., Mazzarelli, D., Oliva, A.G., De Angelis, D., Sforza, C. and Cattaneo, C. (2019) ‘Three-dimensional analysis of sphenoid sinus uniqueness for assessing personal identification: a novel method based on 3D-3D superimposition’, *International Journal of Legal Medicine*, 133, pp. 1895–1901. Available at: https://doi.org/https://doi.org/10.1007/s00414-019-02139-5.

Cash, T.F., Dawson, K., Davis, P., Bowen, M. and Galumbeck, C. (1989) ‘Effects of Cosmetics Use on the Physical Attractiveness and Body Image of American College Women’, *The Journal of Social Psychology*, 129(3), pp. 349–355. Available at: https://www.tandfonline.com/doi/citedby/10.1080/00224545.1989.9712051?scroll=top&needAccess=true.

Castellini, R. (2012) *Importing photos into Picasa 3.5*, *Video*. Available at: https://www.youtube.com/watch?v=9reYYNiTLGM (Accessed: 26 January 2021).

Castelvecchi, D. (2020) *Is facial recognition too biased to be let loose?*, *Article*. Available at: https://doi.org/10.1038/d41586-020-03186-4.

Catanese, C. (2016) *Color atlas of forensic medicine and pathology*.

Catanese, C.A. and Labay, L.M. (2016) 'Substance Abuse and Poisoning', in C.A. Catanese (ed.) *Color Atlas of Forensic Medicine and Pathology*. 2nd edn. Boca Raton, Florida: CRC Press, pp. 105–143.

Catanese, C.A., Levy, B. and Catanese, G. (2021) 'Postmortem Change and Time of Death', in C. Catanese (ed.) *Color Atlas of Forensic Medicine and Pathology*. 2nd edn. Boca Raton, Florida: CRC Press, Taylor & Francis Group, pp. 145–191.

Catanese, C.A. and Rapkiewicz, A.V. (2016) 'Sudden Natural Death in a Forensic Setting', in C. Catanese (ed.) *Color Atlas of Forensic Medicine and Patholog*. 2nd edn. Boca Raton, Florida: CRC Press, pp. 1–71.

Cattaneo, C., De Angelis, D., Mazzarelli, D., Porta, D., Poppa, P., Caccia, G., D'Amico, M.E., Siccardi, C., Previderè, C., Bertoglio, B., Tidball-Binz, M., Ubelaker, D., Piscitelli, V. and Riccio, S. (2022) 'The rights of migrants to the identification of their dead: an attempt at an identification strategy from Italy', *International Journal of Legal Medicine*, online ahe, pp. 1–12. Available at: https://doi.org/10.1007/s00414-022-02778-1.

Cattaneo, C., Cantatore, A., Ciaffi, R., Gibelli, D., Cigada, A., De Angelis, D. and Sala, R. (2012) 'Personal Identification by the Comparison of Facial Profiles: Testing the Reliability of a High-Resolution 3D-2D Comparison Model', *Journal of Forensic Sciences*, 57(1), pp. 182–187. Available at: https://doi.org/10.111/j.1566-4029.2011.01944.x.

Cattaneo, C. and Gibelli, D. (2013) 'Identification', in J.A. Siegel, P.J. Saukko, and M.M. Houck (eds) *Encyclopedia of Forensic Sciences*. 2nd edn. Elsevier Ltd, pp. 158–165. Available at: https://doi.org/10.1016/B978-0-12-382165-2.00178-1.

Cattaneo, C. and Gibelli, D. (2014) 'Forensic Anthropology', in B. Madea (ed.) *Handbook of Forensic Medicine*. John Wiley & Sons, Ltd., pp. 1184–1192.

Cattaneo, C., Mazzarelli, D., Olivieri, L., De Angelis, D., Cappella, A., Vitale, A., Caccia, G., Piscitelli, V. and Iadicicco, A. (2020) 'The approach to unidentified dead migrants in Italy', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 559–570.

Chaikunrat, J., Pongpanitanon, P. and Petiju, M. (2011) 'Victim Identification in the Tsunami Disaster in Thailand', *Journal of Health Science*, 20, pp. 897–902.

Chellappa, R., Sinha, P. and Phillips, P.J. (2010) 'Face Recognition by Computers and Humans', *Computer*, 43(2), pp. 46–55. Available at: https://doi.org/10.1109/MC.2010.37.

Chellappa, R., Wilson, C.L. and Sirohey, S. (1995) 'Human and machine recognition of faces: a survey', *Proceedings of the IEEE*, 83(5), pp. 705–741. Available at: https://doi.org/10.1109/5.381842.

Chen, D., Xu, C., Yang, J., Qian, J., Zheng, Y. and Shen, L. (2018) 'Joint Bayesian guided metric learning for end-to-end face verification', *Neurocomputing*, 275, pp. 560–567. Available at: https://doi.org/10.1016/j.neucom.2017.09.009.

Chesson, L.A., Meier-Augenstein, W., Berg, G.E., Bataille, C.P., Bartelink, E.J. and Richards, M.P. (2020) 'Basic principles of stable isotope analysis in humanitarian forensic science', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 285–310. Available at: https://doi.org/10.1002/9781119482062.ch20.

Choi, S.H., Giu, J.H. and Kang, D.H. (2016) 'Analysis of Traffic Accident-Related Facial Trauma', *The Journal of Craniofacial Surgery*, 27(7), pp. 1682–1685. Available at: https://doi.org/10.1097/SCS.0000000000002916.

Chung, M.-S. and Thomson, D.M. (1995) 'Development of face recognition', *British Journal of Psychology*, 86(1), pp. 55–87. Available at: https://doi.org/10.1111/j.2044-8295.1995.tb02546.x.

Clark, M.A., Worrell, M.B. and Pless, J.E. (2006) 'Postmortem Changes in Soft Tissues', in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. 1st edn. Boca Raton, Florida: CRC Press LLC, pp. 151–164.

Clayton, J. (2022) *How facial recognition is identifying the dead in Ukraine - BBC News, Article*. Available at: https://www.bbc.co.uk/news/technology-61055319 (Accessed: 13 April 2022).

Clifford, C.W.G., Watson, T.L. and White, D. (2018) 'Two sources of bias explain errors in facial age estimation', *Royal Society Open Science*, 5(180841), pp. 1–10. Available at: https://doi.org/10.1098/rsos.180841.

Clutterbuck, R. and Johnston, R.A. (2002) 'Exploring Levels of Face Familiarity by Using an Indirect Face-Matching Measure', *Perception*, 31(8), pp. 985–994. Available at: https://doi.org/10.1068/p3335.

Coleman, S.R. and Grover, R. (2006) 'The anatomy of the aging face: volume loss and changes in 3-dimensional topography', *Aesthetic Surgery Journal*, 26(1), pp. S4–S9. Available at: https://doi.org/https://doi.org/10.1016/j.asj.2005.09.012.

Comfort, W.E., de Andrade, B.N., Wingenbach, T.S.H., Causeur, D. and Boggio, P.S. (2021) 'Implicit responses in the judgment of attractiveness in faces with differing levels of makeup', *Psychology of Aesthetics, Creativity, and the Arts* [Preprint]. Available at: https://doi.org/10.1037/aca0000408.

Connor, M., Baigent, C. and Hansen, E.S. (2017) 'Testing the Use of Pigs as Human Proxies in Decomposition Studies', *Journal of Forensic Sciences*, 63(5), pp. 1350–1355. Available at: https://doi.org/10.1111/1556-4029.13727.

Cook, S. (2020) 'Posthumous dignity and the importance in returning remains of the deceased', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 67–78. Available at: https://doi.org/10.1002/9781119482062.ch5.

Coppedè, F. (2018) 'Premature Aging Syndrome', in J. Ram and P.M. Conn (eds) *Conn's Handbook of Models for Human Aging*. 2nd edn. Elsevier Academic Press, pp. 21–34. Available at: https://doi.org/0.1016/B978-0-12-811353-0.00002-6.

Corcoran, P. and Costache, C. (2016) 'Smartphones, Biometrics, and a Brave New World', *IEEE Technology and Society Magazine*, pp. 59–66. Available at: https://doi.org/10.1109/MTS.2016.2593266.

Cornett, D.C., Bolme, D.S., Steadman, D.W., Sauerwein, K.A. and Saul, T.B. (2019) 'Effects of Postmortem Decomposition on Face Recognition', in *IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. Tampa, FL USA, pp. 1–8. Available at: https://doi.org/10.1109/BTAS46853.2019.9185971.

Cornwell, P. (2011) *The Body Farm*. Sphere.

Costa, M.C.F., Cavalcante, G.M.F., de Nóbrega, L.M., Oliveira, P.A.P. and d'Avila, S. (2014) 'Facial traumas among females through violent and non-violent mechanisms', *Brazilian Journal of Otorhinolaryngology*, 80(3), pp. 196–201. Available at: https://doi.org/10.1016/j.bjorl.2013.10.001.

Cotofana, S. (2018) 'Das alternde Gesicht - eine anatomische Zusammenfassung', *Ästhetische Dermatologie & Kosmetologie*, 03, pp. 20–22.

Cox, C.L. and Glick, W. (1986) 'Resume evaluations and cosmetics use: When more is not better', *Sex Roles*, 14, pp. 51–58. Available at: https://doi.org/10.1007/BF00287847.

Crabbe, S., Ambs, P., Black, S., Wilkinson, S., Bikker, J., Herz, N., Manger, D., Pape, R. and Seibert, H. (2013) *Results of the FASTID Project*. Available at: http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-2793589.pdf.

Crookes, K. and McKone, E. (2009) 'Early maturity of face recognition: no childhood development of holistic processing, novel face encoding, or face-space', *Cognition*, 111(2), pp. 219–247. Available at: https://doi.org/10.1016/j.cognition.2009.02.004.

Cross, P. and Williams, A. (2017) 'Taphonomy Facilities as Teaching Aids', in Williams A., J.P. Cassella, and P.D. Maskell (eds) *Forensic Science Education and Training: A Tool-kit for Lecturers and Oractitioner Trainers*. John Wiley & Sons Ltd, pp. 45–55. Available at: https://doi.org/10.1002/9781118689196.ch4.

Crosswhite, N., Byrne, J., Stauffer, C., Parkhi, O., Cao, Q. and Zisserman, A. (2018) 'Template adaptation for face verification and identification', *Image and Vision Computing*, 79, pp. 35–48. Available at: https://doi.org/10.1016/j.imavis.2018.09.002.

Crown Prosecution Service (2022) *DNA-17 Profiling | The Crown Prosecution Service*, *Legal Guidance*. Available at: https://www.cps.gov.uk/legal-guidance/dna-17-profiling (Accessed: 6 October 2022).

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B.C. and Pellicano, E. (2014) 'The Cambridge Face Memory Test for Children (CFMT-C): a new tool for measuring face recognition skills in childhood', *Neuropsychologia*, 62, pp. 60–67. Available at: https://doi.org/10.1016/j.neuropsychologia.2014.07.008.

Crumpler, W. (2020) *The Problem of Bias in Facial Recognition*, *CSIS - Center for Strategic & International Studies*. Available at: https://www.csis.org/blogs/technology-policy-blog/problem-bias-facial-recognition (Accessed: 15 June 2020).

Cui, J., Zhang, H., Han, H., Shan, S. and Chen, X. (2018) 'Improving 2D Face Recognition via Discriminative Face Depth Estimation', in *International Conference on Biometrics (ICB)*. Gold Coast, QLD Australia: IEEE, pp. 1–8. Available at: https://doi.org/10.1109/ICB2018.2018.00031.

Cummaudo, M., Guerzoni, M., Marasciuolo, L., Gibelli, D., Cigada, A., Obertovà, Z., Ratnayake, M., Poppa, P., Gabriel, P., Ritz-Timme, S. and Cattaneo, C. (2013) 'Pitfalls at the root of facial assessment on photographs: A quantitative study of accuracy in positioning facial landmarks', *International Journal of Legal Medicine*, 127(3), pp. 699–706. Available at: https://doi.org/10.1007/s00414-013-0850-7.

Damas, S., Cordón, O. and Ibáñez, O. (2020) *Handbook on Craniofacial Superimposition -*

The MEPROCS Project. 1st edn. Springer, Cham. Available at: https://doi.org/10.1007/978-3-319-11137-7.

Damas, S., Cordón, O., Ibáñez, O., Santamaría, J., Alemán, I., Botella, M. and Navarro, F. (2011) 'Forensic identification by computer-aided craniofacial superimposition', *ACM Computing Surveys*, 43(4), pp. 1–27. Available at: https://doi.org/10.1145/1978802.1978806.

Damas, S., Wilkinson, C., Kahana, T., Veselovskaya, E., Abramov, A., Jankauskas, R., Jayaprakash, P.T., Ruiz, E., Navarro, F., Huete, M.I., Cunha, E., Cavalli, F., Clement, J., Lestón, P., Molinero, F., Briers, T., Viegas, F., Imaizumi, K., Humpire, D. and Ibáñez, O. (2015) 'Study on the performance of different craniofacial superimposition approaches (II): Best practices proposal', *Forensic Science International*, 257, pp. 504–508. Available at: https://doi.org/10.1016/j.forsciint.2015.07.045.

Danziger, P.N. (2019) *6 Trends Shaping The Future Of The $532B Beauty Business*, *Article*. Available at: https://www.forbes.com/sites/pamdanziger/2019/09/01/6-trends-shaping-the-future-of-the-532b-beauty-business/?sh=1ea6145d588d (Accessed: 10 September 2020).

Davies, G.M. (2016) 'Identification of Familiar Faces after Long Intervals: The Tichborne Claimant Revisited', *Applied Cognitive Psychology*, 30, pp. 484–489. Available at: https://doi.org/10.1002/acp.3207.

Davis, G. and Thasen, S. (2000) 'Closed-circuit television: How effective an identification aid?', *British Journal of Psychology*, 91(3), pp. 411–426. Available at: https://doi.org/10.1348/000712600161907.

Davis, J.P., Maigut, A. and Forrest, C. (2019a) 'The wisdom of the crowd: A case of post-to ante-mortem face matching by police super-recognisers', *Forensic Science International*, 302, pp. 1–6. Available at: https://doi.org/10.1016/j.forsciint.2019.109910.

Davis, J.P., Maigut, A. and Forrest, C. (2019b) 'The wisdom of the crowd: A case of post-to ante-mortem face matching by police super-recognisers', *Forensic Science International*, 302, pp. 1–6. Available at: https://doi.org/10.1016/j.forsciint.2019.109910.

Davis, J.P. and Valentine, T. (2009) 'CCTV on trial: Matching video images with the defendant in the dock', *Applied Cognitive Psychology*, 23(4), pp. 482–505. Available at: https://doi.org/10.1002/acp.1490.

Davis, J.P., Valentine, T. and Wilkinson, C. (2012) 'Facial image comparison', in C. Wilkinson and C. Rynn (eds) *Craniofacial Identification*. Cambridge University Press, pp. 136–153. Available at: https://doi.org/10.1017/CBO9781139049566.012.

Decker, S. and Ford, J. (2017) *Management of 3D Image Data*, *Human Remains: Another Dimension The Application of Imaging to the Study of Human Remains*. Elsevier Inc. Available at: https://doi.org/10.1016/B978-0-12-804602-9.00014-X.

DelMonte, D.W. and Kim, T. (2011) 'Anatomy and physiology of the cornea', *Journal of Cataract & Refractive Surgery*, 37(3), pp. 588–598. Available at: https://doi.org/10.1016/j.jcrs.2010.12.037.

Department of Justice (2016) *John Dillinger*, *case story*. Available at: https://www.fbi.gov/history/famous-cases/john-dillinger (Accessed: 25 March 2019).

Devue, C., Wride, A. and Grimshaw, G.M. (2019) 'New insights on real-world human face recognition', *Journal of Experimental Psychology: General*, 148(6), pp. 994–1007. Available at: https://doi.org/10.1037/xge0000493.

Ding, C. and Tao, D. (2017) 'Pose-invariant face recognition with homography-based normalization', *Pattern Recognition*, 66, pp. 144–152. Available at: https://doi.org/10.1016/j.patcog.2016.11.024.

Dix, J. and Graham, M. (2000) *Time of Death, Decomposition and Identification*. 1st edn. Boca Raton, Florida: CRC Press. Available at: https://doi.org/10.1201/9781420048285.

Dixson, B.J. and Vasey, P.L. (2012) 'Beards augment perceptions of men's age, social status, and aggressiveness, but not attractiveness', *Behavioural Ecology*, 23(3), pp. 481–490. Available at: https://doi.org/10.1093/beheco/arr214.

Dobay, A., Ford, J., Decker, S., Ampanozi, G., Franckenberg, S., Affolter, R., Sieberth, T. and Ebert, L.C. (2020) 'Potential use of deep learning techniques for postmortem imaging', *Forensic Science, Medicine and Pathology*, 16, pp. 671–679. Available at: https://doi.org/10.1007/s12024-020-00307-3.

Doberentz, E. and Madea, B. (2010) 'Schätzung der Wasserliegezeit: Retrospektive Untersuchung zur Reliabilität', *Rechtsmedizin*, 20(5), pp. 393–399. Available at: https://doi.org/10.1007/s00194-010-0667-3.

Dodge, S. and Karam, L. (2016) 'Understanding how image quality affects deep neural networks', in *Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. Lisbon, Portugal: IEEE, pp. 1–6. Available at: https://doi.org/10.1109/QoMEX.2016.7498955.

Dorion, R.B. (1983) 'Photographic superimposition', *Journal of Forensic Sciences*, 28, pp. 724–734.

Downs, J.C.U. (2016) 'Carbon Monoxide Exposure: Autopsy Findings', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 444–460. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00058-6.

Dowsett, A.J. and Burton, A.M. (2015) 'Unfamiliar face matching : Pairs out-perform individuals and provide a route to training', *British Journal of Psychology*, 106(3), pp. 433–445. Available at: https://doi.org/10.1111/bjop.12103.

Drira, H., Amor, B.B., Srivastava, A., Daoudi, M. and Slama, R. (2013) '3D face recognition under expressions, occlusions, and pose variations', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(9), pp. 2270–2283. Available at: https://doi.org/10.1109/TPAMI.2013.48.

Dror, I.E., Charlton, D. and Péron, A.E. (2006) 'Contextual information renders experts vulnerable to making erroneous identifications', *Forensic Science International*, 156(1), pp. 74–78. Available at: https://doi.org/10.1016/j.forsciint.2005.10.017.

Duan, F., Yang, Y., Li, Y., Tian, Y., Lu, K., Wu, Z. and Zhou, M. (2014) 'Skull identification via correlation measure between skull and face shape', *IEEE Transactions on Information Forensics and Security*, 9(8), pp. 1322–1332. Available at: https://doi.org/10.1109/TIFS.2014.2332981.

Dunn, J.D., de Lima Varela, V.P., Nicholls, V.I., Papinutto, M., White, D. and Miellet, S.

(2021) ‘Visual information sampling in super-recognizers’, *PsyArXiv*, March 31, pp. 1–20. Available at: https://doi.org/10.31234/osf.io/z2k4a.

Dunn, J.D., Summersby, S., Towler, A., Davis, J.P. and White, D. (2020) ‘UNSW Face Test: A screening tool for super-recognizers’, *PLoS ONE*, 15(11), pp. 1–19. Available at: https://doi.org/10.1371/journal.pone.0241747.

Dunn, K.W. and Harrison, R.K. (1997) ‘Naming of parts: a presentation of facial surface anatomical terms’, *British Journal of Plastic Surgery*, 50, pp. 584–589.

Edelmann, H. (1938) ‘Die Profilanalyse. Eine Studie an photographischen und röntgenographischen Durchdringungsbildern’, *Zeitschrift für Morphologie und Anthropologie*, 37, pp. 166–188. Available at: https://www.jstor.org/stable/25749576.

Efremov, I.A. (1940) ‘Taphonomy: New branch of Paleontology’, *The Pan-American Geologist*, 74(2), pp. 81–93. Available at: http://iae.makorzh.ru/science/taph.htm.

Eggerstedt, M., Rhee, J., Urban, M.J., Mangahas, A., Smith, R.M. and Revenaugh, P.C. (2020) ‘Beauty is in the eye of the follower: Facial aesthetics in the age of social media’, *American Journal of Otolaryngology*, 41(6), pp. 1–5. Available at: https://doi.org/10.1016/j.amjoto.2020.102643.

Elbashir, R.M. and Yap M.H. (2020) ‘Evaluation of Automatic Facial Wrinkle Detection Algorithms’, *Journal of Imaging*, 6(4), pp. 1–10. Available at: https://doi.org/10.3390/jimaging6040017.

Ellingham, S.T.D., Perich, P. and Tidball-Binz, M. (2017) ‘The fate of human remains in a maritime context and feasibility for forensic humanitarian action to assist in their recovery and identification’, *Forensic Science International*, 279, pp. 229–234. Available at: https://doi.org/10.1016/j.forsciint.2017.07.039.

Ellis, H.D., Davies, G.M. and Shepherd, J.W. (1978) ‘Remembering pictures of real and “unreal” faces: Some practical and theoretical considerations’, *British Journal of Psychology*, 69, pp. 467–474.

Ellis, H.D., Shepherd, J.W. and Davies, G.M. (1979) ‘Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition’, *Perception*, 8(4), pp. 431–439. Available at: https://doi.org/10.1068/p080431.

Ellson, A. (2020) *Nose job ads? Not in front of the children*, *The Times*. Available at: https://www.thetimes.co.uk/article/nose-job-ads-not-in-front-of-the-children-gdml2f5wp?fbclid=IwAR2eMewk1jeZ0xoyf-hcp-7QSIhaWZ9BE784b-4NHgH30fatHFBGF93K6lI (Accessed: 11 September 2020).

Elmahmudi, A. and Ugail, H. (2019) ‘Deep face recognition using imperfect facial data’, *Future Generation Computer Systems*, 99, pp. 213–225. Available at: https://doi.org/10.1016/j.future.2019.04.025.

Encyclopedia Britannica (2020) *Abū Niḍāl*, *Article*. Available at: https://www.britannica.com/biography/Abu-Nidal (Accessed: 7 June 2020).

ENFSI (2018) *Best Practice Manual for Facial Image Comparison*. Available at: http://enfsi.eu/wp-content/uploads/2017/06/ENFSI-BPM-DI-01.pdf (Accessed: 12 April 2018).

Erickson, W.B. (2016) *Recognition Training for Faces Across Age Gaps*. University of Arkansas, Fayetteville.

Escoffier, C., de Rigal, J., Rochefort, A., Vassalet, R., Léve, J.-L. and Agache, P. (1989) 'Age-related mechanical propoerties of human skin: an in vivo study', *Journal of Investigative Dermatology*, 93(3), pp. 353–357. Available at: https://doi.org/10.1016/0022-202X(89)90058-4.

Estudillo, A.J. and Bindemann, M. (2014) 'Generalization across view in face memory and face matching', *i-Perception*, 5(7), pp. 589–601. Available at: https://doi.org/10.1068/i0669.

Etcoff, N.L., Stock, S., Haley, L.E., Vickery, S.A. and House, D.M. (2011) 'Cosmetics as a feature of the extended human phenotype: modulation of the perception of biologically important facial signals', *PLoS ONE*, 6(10), p. e25656. Available at: https://doi.org/10.1371/journal.pone.0025656.

European Commission (2021) *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*. Brussels. Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF.

Evans, R. (2014) 'Image Analysis Forensic Facial Comparison: Issues and Misconceptions', in X. Mallett, T. Blythe, and R. Berry (eds) *Advances in Forensic Human Identification*. CRC Press, Taylor & Francis Group, pp. 213–234.

Evison, M. and Vorder Bruegge, R.W. (2010) 'Problems and Prospects', in P. Evison and R.W. Vorder Bruegge (eds) *Computer-aided Forensic Facial Comparison*. Boca Raton, Florida: CRC Press, pp. 157–168.

Fancher, J.P., Aitkenhead-Peterson, J.A., Farris, T., Mix, K., Schwab, A.P., Wescott, D.J. and Hamilton, M.D. (2017) 'An evaluation of soil chemistry in human cadaver decomposition islands: Potential for estimating postmortem interval (PMI)', *Forensic Science International*, 279, pp. 130–139. Available at: https://doi.org/10.1016/j.forsciint.2017.08.002.

Farage, M.A., Miller, K.W., Elsner, P. and Maibach, H.I. (2008) 'Intrinsic and extrinsic factors in skin ageing: a review', *International Journal of Cosmetic Science*, 30, pp. 87–95.

Farazdaghi, E. and Nait-Ali, A. (2017) 'Backward face ageing model (B-FAM) for digital face image rejuvenation', *IET Biometrics*, 6(6), pp. 478–486. Available at: https://doi.org/10.1049/iet-bmt.2016.0079.

Fardouly, J. and Rapee, R.M. (2019) 'The impact of no-makeup selfies on young women's body image', *Body Image*, 28, pp. 128–134. Available at: https://doi.org/10.1016/j.bodyim.2019.01.006.

Farkas, L.G. (1994) 'Examination', in L.G. Farkas (ed.) *Anthropometry of the Head and Face*. 2nd edn. New York: Raven Press, pp. 3–56.

Farkas, L.G. (1996) 'Accuracy of anthropometric measurements: Past, present, and future', *Cleft Palate-Craniofacial Journal*, pp. 10–18. Available at: https://doi.org/10.1597/1545-1569(1996)033<0010:AOAMPP>2.3.CO;2.

Farrell, A.-M. (2018) ‘Managing the dead in disaster response: a matter for health security in the Asia-Pacific region?’, *Australian Journal of International Affairs*, 72(6), pp. 551–566. Available at: https://doi.org/10.1080/10357718.2018.1534940.

Favelle, S., Hill, H. and Claes, P. (2017) ‘About Face: Matching Unfamiliar Faces Across Rotations of View and Lighting’, *i-Perception*, 8(6), pp. 1–22. Available at: https://doi.org/10.1177/2041669517744221.

Featherstone, M. (1999) ‘Body Modification’, *Body & Society*, 5(2–3), pp. 1–13.

Fenton, T.W., Heard, A.N. and Sauer, N.J. (2008) ‘Skull-Photo Superimposition and Border Deaths: Identification Through Exclusion and the Failure to Exclude’, *Journal of Forensic Sciences*, 53(1), pp. 34–40. Available at: https://doi.org/10.111/j.1556-4029.2007.00624.x.

Fiedler, S. and Graw, M. (2003) ‘Decomposition of buried corpses, with special reference to the formation of adipocere’, *Naturwissenschaften*, 90(7), pp. 291–300. Available at: https://doi.org/10.1007/s00114-003-0437-0.

FISWG (2012) *Guidelines for Facial Comparison Methods*. Available at: https://www.fiswg.org/FISWG_GuidelinesforFacialComparisonMethods_v1.0_2012_02_02.pdf.

FISWG (2018a) *Facial Image Comparison Feature List for Morphological Analysis*. Available at: https://fiswg.org/FISWG_Morph_Analysis_Feature_List_v2.0_20180911.pdf.

FISWG (2018b) *Standard Guide for Postmortem Facial Image Capture*, *Guidelines*.

FISWG (2019a) *Facial Comparison Overview and Methodlogy Guidelines - version 1.0*. Available at: https://www.fiswg.org/fiswg_facial_comparison_overview_and_methodology_guidelines_V1.0_20191025.pdf.

FISWG (2019b) *Guide for Mentorship of Facial Comparison Trainees in Role Based Facial Comparison*.

FISWG (2020) *Guide for Role-Based Training in Facial Comparison*.

FISWG (2021) *Physical Stability of Facial Features of Adults*. Available at: https://www.fiswg.org/fiswg_physical_stability_of_facial_features_of_adults_v2.0_2021.05.28.pdf.

Flament, F., Amar, D. and Bazin, R. (2018) ‘Evaluating age-related changes of some facial signs aming men of four different ethnic groups’, *International Journal of Cosmetic Science*, 40, pp. 502–515. Available at: https://doi.org/10.1111/ics.12492.

Fletcher, K.I., Butavicius, M.A. and Lee, M.D. (2008) ‘Attention to internal face features in unfamiliar face matching’, *British Journal of Psychology*, 33(pt3), pp. 379–394. Available at: https://doi.org/10.1348/000712607X235872.

Foltyn, J.L. (2008) ‘Dead famous and dead sexy: Popular culture, forensics, and the rise of the corpse’, *Mortality*, 13(2), pp. 153–173. Available at: https://doi.org/10.1080/13576270801954468.

Forbes, S. (2008) ‘Forensic Chemistry: Appliations to Decomposition and Preservation’, in

M. Oxenham (ed.) *Forensic Approaches to Death, Disaster and Abuse*. Australian Academic Press, Australia, pp. 233–262.

Forbes, S. and Nugent, K. (2016) 'Dating of Anthropological Skeletal Remains of Forensic Interest', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 213–225.

Forbes, S., Wilson, M.E.A. and Stuart, B.H. (2011) 'Examination of adipocere formation in a cold water environment', *International Journal of Legal Medicine*, 125(5), pp. 643–650. Available at: https://doi.org/10.1007/s00414-010-0460-6.

Forensic Science Regulator (2016a) *Codes of Practice and Coduct for Forensic Science Providers and Practitioners FSR-C-100 Ossie 7*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/968638/100_Codes_of_Practice_and_Conduct_-_Issue_7.pdf.

Forensic Science Regulator (2016b) *Forensic Image Comparison and Interpretation Evidence: Guidance for Prosecutors and Investigators - Issue 2*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/912880/Image_Comparison_and_Interpretation_Guidance_Issue_2.pdf.

Forensic Science Regulator (2020) *Guidance - Cognitive Bias Effects Relevant to Forensic Science Examinations- Issue 2*.

Forsberg, A., Kullberg, J., Agartz, I., Ahlström, H., Johansson, L. and Hendriksson, K.M. (2009) 'Landmark-based software for anatomical measurements: a precision study', *Clinical Anatomy*, 22(4), pp. 456–462. Available at: https://doi.org/Landmark-based software for anatomical measurements: a precision study.

Fosdick, R.B. (1915) 'The Passing of the Bertillon System of Identification', *Journal of the American Institute of Criminal Law and Criminology*, 6(3). Available at: https://www.jstor.org/stable/1132744.

Freeberg, M.A.T., Kallenbach, J.G. and Awad, H.A. (2019) 'Assessment of Cellular Responses of Tissue Constructs in vitro in Regenerative Engineering', in R. Narayan (ed.) *Encyclopedia of Biomedical Engineering, Vol .1*. Elsevier, pp. 414–426. Available at: https://doi.org/10.1016/B978-0-12-801238-3.99898-2.

FRONTEX (2016) *Best Practice Operational Guidelines for Automated Border Control (ABC) Systems*. Available at: https://frontex.europa.eu/assets/Publications/Research/Best_Practice_Operational_Guidelines_ABC.pdf.

Frowd, C. (2012) 'Facial recall and computer composites', *Craniofacial Identification*, pp. 42–56. Available at: https://doi.org/10.1017/CBO9781139049566.004.

Frowd, C., Bruce, V., McIntyre, A. and Hancock, P. (2007) 'The relative importance of external and internal features of facial composites', *British Journal of Psychology*, 98(1), pp. 61–77. Available at: https://doi.org/10.1348/000712606X104481.

Furlong, W.B. (1984) *Thirty Years Ago—; The summer of 1934 marked the high point in the improbable career of one John Dillinger. It also marked its end.*, *New York Times newspaper article*. Available at: https://www.nytimes.com/1964/08/09/archives/thirty-years-ago-the-summer-of-1934-marked-the-high-point-in-the.html (Accessed: 28 May 2020).

Fysh, M.C. (2021) ‘Factors Limiting Face Matching at Passport Control and in Police Investigations’, in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 15–37.

Fysh, M.C. and Bindemann, M. (2017a) ‘Effects of time pressure and time passage on face-matching accuracy’, *Royal Society Open Science*, 4(6), pp. 1–13. Available at: https://doi.org/10.1098/rsos.170249.

Fysh, M.C. and Bindemann, M. (2017b) ‘Forensic face matching : A review’, in M. Bindemann and A.M. Megreya (eds) *Face processing: Systems, Disorders and Cultural Differences*. New York: Nova Science Publishing, Inc, pp. 1–20.

Fysh, M.C. and Bindemann, M. (2018) ‘Human–Computer Interaction in Face Matching’, *Cognitive Science*, 42, pp. 1714–1732. Available at: https://doi.org/10.1111/cogs.12633.

Fysh, M.C., Stacchi, L. and Ramon, M. (2020) ‘Differences between and within individuals, and subprocesses of face cognition: implications for theory, research and personnel selection’, *Royal Society Open Science*, 7(200233), pp. 1–17. Available at: https://doi.org/10.1098/rsos.200233.

Galton, F. (1896) ‘The Bertillon System of Identification (Book Review)’, *Nature*, 54(1407), pp. 569–570.

Galton, F. (1907) ‘Vox Populi - The wisdom of crowds’, *Nature*, 75, pp. 450–451. Available at: https://doi.org/10.1038/075450a0.

Ganor, B. (2019) ‘Artificial or Human: A New Era of Counterterrorism Intelligence?’, *Studies in Conflict & Terrorism*, 44(7), pp. 605–624. Available at: https://doi.org/10.1080/1057610X.2019.1568815.

Gao, Y. and Leung, M.K.H. (2002) ‘Face recognition using line edge map’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6), pp. 764–779. Available at: https://doi.org/10.1109/TPAMI.2002.1008383.

Gapert, R. and Tsokos, M. (2013) ‘Anthropological analysis of extensive rodent gnaw marks on a human skull using post-mortem multislice computed tomography (pmMSCT)’, *Forensic Science, Medicine, and Pathology*, 9(3), pp. 441–445. Available at: https://doi.org/10.1007/s12024-012-9363-9.

García-Zurdo, R., Frowd, C.D. and Manzenero, A.L. (2018) ‘Effects of facial periphery on unfamiliar face recognition’, *Current Psychology*, 39, pp. 1767–1773. Available at: https://doi.org/10.1007/s12144-018-9863-1.

Gaudio, D., Olivieri, L., De Angelis, D., Poppa, P., Galassi, A. and Cattaneo, C. (2016) ‘Reliability of Craniofacial Superimposition Using Three-Dimension Skull Model’, *Journal of Forensic Sciences*, 61(1), pp. 5–11. Available at: https://doi.org/10.1111/1556-4029.12856.

Germine, L.T., Duchaine, B. and Nakayama, K. (2011) ‘Where cognitive development and aging meet: face learning ability peaks after age 30’, *Cognition*, 118(2), pp. 201–210. Available at: https://doi.org/10.1016/j.cognition.2010.11.002.

Ghiass, R.S., Arandjelovic, O., Bendada, A. and Maldague, X. (2014) ‘Infrared face recognition: A comprehensive review of methodologies and databases’, *Pattern Recognition*, 47(9), pp. 2807–2824. Available at:

https://doi.org/10.1016/j.patcog.2014.03.015.

Gibelli, D., DeAngelis, D., Poppa, P., Sforza, C. and Cattaneo, C. (2017) 'A View to the Future: A Novel Approach for 3D-3D Superimposition and Quantification of Differences for Identification from Next-Generation Video Surveillance Systems', *Journal of Forensic Science*, 62(2), pp. 457–461. Available at: https://doi.org/10.1111/1556-4029.13290.

Gibelli, D., Obertová, Z., Ritz-Timme, S., Gabriel, P., Arent, T., Ratnayake, M., De Angelis, D. and Cattaneo, C. (2016) 'The identification of living persons on images: A literature review', *Legal Medicine*, 19, pp. 52–60. Available at: https://doi.org/10.1016/j.legalmed.2016.02.001.

Gilbert, D.T. (1998) 'Ordinary Personology', in D.T. Gilbert, S.T. Fiske, and G. Lindzey (eds) *The Handbook of Social Psychology*. 4th, Vol. edn. New York: McGraw-Hill Education, pp. 89–150.

Gill-King, H. (2006) 'Chemical and ultrastructural aspects of decomposition', in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. Boca Raton, Florida: CRC Press, pp. 93–108.

Girardin, G.R. and Helmer, W.J. (1994) *Dillinger: The Untold Story*. Indiana University Press.

Giroux, H.A. (2015) 'Selfie Culture in the Age of Corporate and State Surveillance', *Third Text*, 29(3), pp. 155–164. Available at: https://doi.org/10.1080/09528822.2015.1082339.

Glaister, J. and Brash, J. (1937) *Medico-legal aspects of the Ruxton case*. Edinburgh: The Darien Press.

Glassman, D.M. (2001) 'Methods of Superimposition', in K.T. Taylor (ed.) *Forensic Art and Illustration*. Boca Raton, Florida: CRC Press LLC, pp. 477–498.

Goff, M.L. (2010) 'Early Postmortem Changes and Stages of Decomposition', in J. Amendt, M.L. Goff, C.P. Campobasso, and M. Grassberger (eds) *Current Concepts in Forensic Entomology*. 1st edn. Springer Netherlands, pp. 1–24. Available at: https://doi.org/10.1007/978-1-4020-9684-6.

Goff, M.L. (2016) 'Entomology', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 428–434. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00167-1.

GoMeasure3D (2017) *Artec 3D Scanners*, *Article*. Available at: https://gomeasure3d.com/artec/ (Accessed: 8 May 2017).

Gómez, O., Ibañez, O., Vascecchi, A., Cordón, O. and Kahana, T. (2018) '3D-2D silhouette-based image registration for comparative radiography-based forensic identification', *Pattern Recognition*, 83, pp. 469–480. Available at: https://doi.org/10.1016/j.patcog.2018.06.011.

Goos, M.I.M., Alberink, I.B. and Ruifrok, A.C.C. (2006) '2D/3D image (facial) comparison using camera matching', *Forensic Science International*, 163(1–2), pp. 10–17. Available at: https://doi.org/10.1016/j.forsciint.2005.11.004.

Gordon, G.M. and Steyn, M. (2012) 'An investigation into the accuracy and reliability of skull-photo superimposition in a South African sample', *Forensic Science International*, 216, pp. 198.e1-198.e6. Available at: https://doi.org/10.1016/j.forsciint.2011.09.008.

Gordon, G.M. and Steyn, M. (2016) 'A discussion of current issues and concepts in the practice of skull-photo/craniofacial superimposition', *Forensic Science International*, 262, pp. 287.e1-287.e4. Available at: https://doi.org/http://dx.doi.org/10.1016/j.forsciint.2016.01.030.

Gowland, R. and Thompson, T. (2013a) 'The skin', in *Human Identity and Identification*. Cambridge University Press, pp. 37–70. Available at: https://doi.org/10.1017/CBO9781139029988.003.

Gowland, R. and Thompson, T. (2013b) 'The skin', in *Human Identity and Identification*. 1st edn. Cambridge University Press.

Grm, K., Štruc, V., Artiges, A., Caron, M. and Ekenel, H.K. (2018) 'Strengths and weaknesses of deep learning models for face recognition against image degradations', *IET Biometrics*, 7(1), pp. 81–89. Available at: https://doi.org/10.1049/iet-bmt.2017.0083.

Grother, P. and Ngan, M. (2015) *Child exploitation image analystics: Face recognition evaluation*. Maryland, USA.

Grother, P., Ngan, M. and Hanaoka, K. (2021) *Face Recognition Vendor Test (FRVT) Part 1: Verification (Draft)*. Available at: https://pages.nist.gov/frvt/reports/11/frvt_11_report.pdf.

Grother, P., Ngan, M. and Hanaoka, K. (2022) *NISTIR 8271 Draft Supplement - Face Recognition Vendor Test (FRVT) - Part 2: Identification*. Available at: https://doi.org/10.6028/NIST.IR.8271.

Grother, P., Ngan, M., Hanaoka, K., Boehnen, C. and Ericson, L. (2017) *The 2017 IARPA Face Recognition Prize Challenge(FRPC)*. Available at: https://doi.org/10.6028/NIST.IR.8197.

Grüner, O. (1993) 'Identification of Skulls: A Historical Review and Practical Applications', in M.Y. İşcan and R.P. Helmer (eds) *Forensic Analysis of the Skull*. New York: Wiley-Liss, pp. 29–45.

Guisantes, E. (2019) 'Beauty and Aging', in H. Pinto and J. Fontdevila (eds) *Regenerative Medicine Procedures for Aesthetic Physicians*. Springer, Cham, pp. 33–43. Available at: https://doi.org/10.1007/978-3-030-15458-5_4.

Gunn, A. (2011) *Essential Forensic Biology*. 2nd edn. Oxford: Wiley & Sons Ltd.

Guo, B., Lam, K.-M., Lin, K.-H. and Siu, W.-C. (2003) 'Human face recognition based on spatially weighted Hausdorff distance', *Pattern Recognition Letters*, 24(1–3), pp. 499–507. Available at: https://doi.org/10.1016/S0167-8655(02)00272-6.

Guo, Y., Lei, Y., Liu, L., Wang, Y., Bennamoun, M. and Sohel, F. (2016) 'EI3D: Expression-invariant 3D face recognition based on feature and shape matching', *Pattern Recognition Letters*, 83(3), pp. 403–412. Available at: https://doi.org/10.1016/j.patrec.2016.04.003.

Gupta, M. and Mysore, V. (2016) 'Classifications of Patterned Hair Loss: A Review', *Journal of Cutaneous and Aesthetic Surgery*, 9(1), pp. 3–12. Available at: https://doi.org/10.4103/0974-2077.178536.

Gupta, R. (2019) *Breaking Down Facial Recognition: The Viola-Jones Algorithm*, *Article*. Available at: https://towardsdatascience.com/the-intuition-behind-facial-detection-the-viola-jones-algorithm-29d9106b6999 (Accessed: 13 January 2021).

Gupta, S., Markey, M.K. and Bovik, A.C. (2007) ‘Advances and Challenges in 3D and 2D+3D Human Face Recognition’, in E.A. Zoeller (ed.) *Pattern Recognition: Theory and Application*. Nova Science Publishing, Inc, pp. 143–182.

de Haas, B. (2021) ‘What’s a super-recogniser?’, *Neuropsychologia* [Preprint]. Available at: https://doi.org/10.1016/j.neuropsychologia.2021.107805.

Hach, W. and Hach-Wunderle, V. (2016) ‘Symbolik und Physiognomie, Semiotik und Pathognomie’, *Geschichte der Gefäßchirurgie*, 21, pp. 43–50. Available at: https://doi.org/10.1007/s00772-015-0103-9.

Hackman, L., Delabarde, T., Roberts, J. and de Boer, H. (2022) *Forensic Anthropology: A primer for courts*. 1st edn. Edited by Judical Primer. London: The Royal Society. Available at: https://discovery.dundee.ac.uk/ws/portalfiles/portal/71696309/DES7700_Science_and_the_Law_Forensic_Anthropology_primer_WEB_ToPublishOnline_3_.pdf.

Hadi, H. (2014) ‘Using Facial Creases for Identification’, *Malaysian Journal of Forensic Sciences*, 5(1), pp. 53–68.

Hadi, H. and Wilkinson, C. (2014) ‘The post-mortem resilience of facial creases and the possibility for use in identification of the dead’, *Forensic Science International*, 237, pp. 149.e1-149.e7. Available at: https://doi.org/https://doi.org/10.1016/j.forsciint.2013.12.014.

Hadi, H. and Wilkinson, C. (2017) ‘Categorizing facial creases: A review’, *Journal of Cosmetic Dermatology*, 16, pp. 180–185. Available at: https://doi.org/10.1111/jocd.12305.

Haglund, W.D. and Sorg, M.H. (2006a) ‘Introduction to Forensic Taphonomy’, in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. 2nd edn. Boca Raton, Florida: CRC Press, pp. 1–9.

Haglund, W.D. and Sorg, M.H. (2006b) ‘Method and Theory of Forensic Taphonomy Research’, in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. 2nd edn. Boca Raton, Florida: CRC Press, pp. 13–26.

Hamra, S.T. (1995) ‘Arcus marginalis release and orbital fat preservation in midface rejuvenation’, *Plastic and Reconstructive Surgery*, 96(2), pp. 354–362. Available at: https://doi.org/10.1097/00006534-199508000-00014.

Hancock, P.J., Bruce, V. and Burton, A.M. (2000) ‘Recognition of unfamiliar faces’, *Trends in Cognitive Sciences*, 4(9), pp. 330–337. Available at: https://doi.org/10.1016/s1364-6613(00)01519-9.

Hans, M.G. and Enlow, D.H. (2008) *Essentials of Facial Growth*. end. Ann Arbor, MI: Needham Press.

Harrison, M.T. and Strother, L. (2020) ‘Does right hemisphere superiority sufficiently explain the left visual field advantage in face recognition?’, *Attention, Perception & Psychophysics*, 82, pp. 1205–1220. Available at: https://doi.org/10.3758/s13414-019-01896-0.

Hartman, D., Drummer, O., Eckhoff, C., Scheffer, J.W. and Stringer, P. (2011) 'The contribution of DNA to the disaster victim identification (DVI) effort', *Forensic Science International*, 205(1–3), pp. 52–58. Available at: https://doi.org/10.1016/j.forsciint.2010.09.024.

Hassan, M.Y., Khalifa, O.O., Talib, A.A. and Abdulla, A.H. (2015) 'Unconstrained Facial Recognition Systems: A Review', *Asian Journal of Applied Sciences*, 3(2), pp. 346–354.

Havard, C. and Memon, A. (2014) 'Facial recognition from identificatin parades', in C. Wilkinson and C. Rynn (eds) *Craniofacial Identification*. Cambridge University Press, pp. 86–100.

Hayflick, L. and Moorhead, P.S. (1961) 'The serial cultivation of human diploid cell strains', *Experimental Cell Research*, 25, pp. 585–621. Available at: https://doi.org/10.1016/0014-4827(61)90192-6.

Hayman, J. and Oxenham, M. (2016) *Human Body Decomposition*. Academic Press, Elsevier.

Hefenbrock, D., Oberg, J., Thanh, N.T.N., Kastner, R. and Baden, S.B. (2010) 'Accelerating Viola-Jones Face Detection to FPGA-Level Using GPUs', *18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines, Charlotte*, pp. 11–18. Available at: https://doi.org/10.1109/FCCM.2010.12.

Hehman, E., Sutherland, C.A.M., Flake, J.K. and Slepian, M.L. (2017) 'The unique contributions of perceiver and target characteristics in person perception', *Journal of Personality and Social Psychology*, 113(4), pp. 513–529. Available at: https://doi.org/10.1037/pspa0000090.

Helmer, R. (1984) *Schädelidentifizierung durch elektronische Bildmischung: Zugleich ein Beitrag zur Konstitutionsbiometrie und Dickenmessung der Gesichtsweichteile*. Heidelberg, Germany: Kriminalistik Verlag.

Helmer, R. (1987) 'Identification of the cadaver remains of Josef Mengele', *Journal of Forensic Sciences*, 32(6), pp. 1622–1644.

Henderson, Z., Bruce, C. and Burton, A.M. (2001) 'Matching the faces of robbers captured on video', *Applied Cognitive Psychology*, 15(4), pp. 445–464. Available at: https://doi.org/10.1002/acp.718.

Heyer, R. and Semmler, C. (2013) 'Forensic confirmation bias: The case of facial image comparison', *Journal of Applied Research in Memory and Cognition*, 2(1), pp. 68–70. Available at: https://doi.org/10.1016/j.jarmac.2013.01.008.

Heyer, R., Semmler, C. and Hendrickson, A.T. (2018) 'Humans and Algorithms for Facial Recognition: The Effects of Candidate List Length and Experience on Performance', *Journal of Applied Research in Memory and Cognition*, 7(4), pp. 597–609. Available at: https://doi.org/10.1016/j.jarmac.2018.06.002.

Hills, P.J., Cooper, R.E. and Pake, J.M. (2013) 'Removing the own-race bias in face recognition by attentional shift using fixation crosses to diagnostic features: An eye-tracking study', *Visual Cognition*, 21(7), pp. 876–898. Available at: https://doi.org/10.1080/13506285.2013.834016.

Hills, P.J. and Pake, J.M. (2013) 'Eye-tracking the own-race bias in face recognition:

revealing the perceptual and socio-cognitive mechanisms', *Cognition*, 129(3), pp. 586–597. Available at: https://doi.org/10.1016/j.cognition.2013.08.012.

Hinchliffe, J. (2011) 'Forensic odontology, part 1. Dental identification', *British Dental Journal*, 210(5), pp. 219–224. Available at: https://doi.org/10.1038/sj.bdj.2011.199.

Hinchliffe, J.A. (2007) 'Disaster dentistry', *British Dental Journal*, 202(8), pp. 493–494. Available at: https://doi.org/10.1038/bdj.2007.339.

Hochmeister, M., Grassberger, M. and Stimpfl, T. (2007) *Forensische Medizin*. 2nd edn. Vienna, Austria: Wilhelm Maudrich Nfg. GmbH & Co KG, Verlag für medizinische Wissenschaften.

Hogue, J.V. and Mills, J.S. (2019) 'The effects of active social media engagement with peers on body image in young women', *Body Image*, 28, pp. 1–5. Available at: https://doi.org/10.1016/j.bodyim.2018.11.002.

Home Office (2013) *Guidance on dealing with fatalities in emergencies*. London, UK. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/61191/fatalities.pdf.

Huang, G., Yang, M.-H., Learned-Miller, E., Ma, Y., Turk, M., Kriegman, D.J. and Huang, T.S. (2011) 'Introduction to the special section on real-world face recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10), pp. 1921–1924. Available at: https://doi.org/10.1109/TPAMI.2011.182.

Huang, G.B., Ramesh, M., Berg, T. and Learned-Miller, E. (2007) *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*. Amherst.

Huete, M.I., Ibáñez, O., Wilkinson, C. and Kahana, T. (2015) 'Past, present, and future of craniofacial superimposition: Literature and international surveys', *Legal Medicine*, 17(4), pp. 267–278. Available at: https://doi.org/10.1016/j.legalmed.2015.02.001.

Hurley, D.J., Arbab-Zavar, B. and Nixon, M.S. (2008) 'The Ear as a Biometric', in A.K. Jain, P. Flynn, and A.A. Ross (eds) *Handbook of Biometrics*. Boston: Springer, pp. 1–20. Available at: https://doi.org/10.1007/978-0-387-71041-9_7.

Hurst, D. (2017) 'Chinese women fail to get past South Korean immigration officials after plastic surgery', *The Times*, p. 1. Available at: https://www.thetimes.co.uk/article/chinese-women-fail-to-get-past-immigration-officials-after-plastic-surgery-jgcr7g836.

Huttenlocher, D.P., Klanderman, G.A. and Rucklidge, W.J. (1993) 'Comparing images using the Hausdorff distance', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9), pp. 850–863. Available at: https://doi.org/10.1109/34.232073.

Iannarelli, A.V. (1964) *The Iannarelli System of Ear Identification*. Brooklyn: Foundation Press.

Iannarelli, A.V. (1989) *Ear Identification*. California: Paramont Publishing Company.

Ibáñez, O., Ballerini, L., Cordón, O., Damas, S. and Santamaría, J. (2009) 'An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification', *Information Sciences*, 179(23), pp. 3998–4028. Available at:

https://doi.org/10.1016/j.ins.2008.12.029.

Ibáñez, O., Cavalli, F., Campomanes-Álvarez, B.R., Campomanes-Álvarez, C., Valsecchi, A. and Huete, M.I. (2015) 'Ground truth data generation for skull–face overlay', *International Journal of Legal Medicine*, 129(3), pp. 569–581. Available at: https://doi.org/10.1007/s00414-014-1074-1.

Ibáñez, O., Cordón, O. and Damas, S. (2012) 'A cooperative coevolutionary approach dealing with the skull-face overlay uncertainty in forensic identification by craniofacial superimposition', *Soft Computing*, 16(5), pp. 797–808. Available at: https://doi.org/10.1007/s00500-011-0770-8.

Ibáñez, Oscar, Cordón, O., Damas, S. and Santamaría, J. (2012) 'An advanced scatter search design for skull-face overlay in craniofacial superimposition', *Expert Systems with Applications*, 39(1), pp. 1459–1473. Available at: https://doi.org/10.1016/j.eswa.2011.08.034.

Ibáñez, O., Valsecchi, A., Cavalli, F., Huete, M.I., Campomanes-Alvarez, B.R., Campomanes-Alvarez, C., Vicente, R., Navega, D., Ross, A., Wilkinson, C., Jankauskas, R., Imaizumi, K., Hardiman, R., Jayaprakash, P.T., Ruiz, E., Molinero, F., Lestón, P., Veselovskaya, E., Abramov, A., Steyn, M., Cardoso, J., Humpire, D., Lusnig, L., Gibelli, D., Mazzarelli, D., Gaudio, D., Collini, F. and Damas, S. (2016) 'Study on the criteria for assessing skull-face correspondence in craniofacial superimposition', *Legal Medicine*, 23, pp. 59–70. Available at: https://doi.org/10.1016/j.legalmed.2016.09.009.

Ibáñez, O., Vicente, R., Navega, D., Campomanes-Álvarez, C., Cattaneo, C., Jankauskas, R., Huete, M.I., Navarro, F., Hardiman, R., Ruiz, E., Imaizumi, K., Cavalli, F., Veselovskaya, E., Humpire, D., Cardoso, J., Collini, F., Mazzarelli, D., Gibelli, D. and Damas, S. (2016) 'MEPROCS framework for Craniofacial Superimposition: Validation study', *Legal Medicine*, 23, pp. 99–108. Available at: https://doi.org/10.1016/j.legalmed.2016.10.007.

Ibáñez, O., Vicente, R., Navega, D.S., Wilkinson, C., Jayaprakash, P.T., Huete, M.I., Briers, T., Hardiman, R., Navarro, F., Ruiz, E., Cavalli, F., Imaizumi, K., Jankauskas, R., Veselovskaya, E., Abramov, A., Lestón, P., Molinero, F., Cardoso, J., Cagdir, A.S., Humpire, D., Nakanishi, Y., Zeuner, A., Ross, A.H., Gaudio, D. and Damas, S. (2015) 'Study on the performance of different craniofacial superimposition approaches (I)', *Forensic Science International*, 257, pp. 496–503. Available at: https://doi.org/10.1016/j.forsciint.2015.05.030.

Ichibori, R., Fujiwara, T., Tanigawa, T., Kanazawa, S., Shingaki, K., Torii, K., Tomita, T., Yano, K., Osaka Twin Research Group, Sakai, Y. and Hosokawa. K. (2014) 'Objective assessment of facial skin aging and the associated environmental factors in Japanese monozygotic twins', *Journal of Cosmetic Dermatology*, 13(2), pp. 158–163. Available at: https://doi.org/10.1111/jocd.12081.

ICO (2021) *What is personal data? - ICO*, *Website*. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/what-is-personal-data/what-is-personal-data/ (Accessed: 2 February 2021).

ICO (2022) *ICO fines facial recognition database company Clearview AI Inc more than £7.5m and orders UK data to be deleted*, *Article*. Available at: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2022/05/ico-fines-facial-recognition-database-company-clearview-ai-inc/.

ICRC (1949) *Geneva Convention Relative to the Protection of Civilian Persons in Time of War (Fourth Geneva Convention)*, *12 August 1949, 75 UNTS 287*. International Committee of the Red Cross (ICRC). Available at: https://www.refworld.org/docid/3ae6b36d2.html (Accessed: 28 January 2022).

ICRC (2008) *Missing People, DNA Analysis and Identification of Human Remains: A Guide to Best Practice in Armed Conflicts and Other Situations of Armed Violence*. Geneva, Switzerland. Available at: https://www.icrc.org/en/publication/4010-missing-people-dna-analysis-and-identification-human-remains-guide-best-practice.

ICRC (2016) *Management of Dead Bodies after Disasters: A Field Manual for First Responders*. Second (re. Edited by S. Cordner, R. Coninx, H.-J. Kim, D. van Alphen, and M. Tidball-Binz. Washington, D.C.: PAHO. Available at: https://shop.icrc.org/management-of-dead-bodies-after-disasters-a-field-manual-for-first-responders-pdf-en.html.

Iino, M., Fujimoto, H., Yoshida, M., Matsumoto, H. and Fujita, M.Q. (2016) 'Identification of a jawless skull by superimposing post-mortem and ante-mortem CT', *Journal of Forensic Radiology and Imaging*, 6, pp. 31–37. Available at: https://doi.org/10.1016/j.jofri.2015.08.005.

IJIS Institute and IACP (2019) *Law Enforcement Facial Recognition Use Case Catalog*.

Ikeda, T., Tani, N., Aoki, Y., Shida, A., Morioka, F., Oritani, S. and Ishikawa, T. (2019) 'Effects of postmortem positional changes on conjunctival petechiae', *Forensic Science, Medicine and Pathology*, 15, pp. 13–22. Available at: https://doi.org/10.1007/s12024-018-0032-5.

Ilkankovan, V. (2014) 'Anatomy of ageing face', *British Journal of Oral and Maxillofacial Surgery*, 52, pp. 195–202. Available at: https://doi.org/10.1016/j.bjoms.2013.11.013.

Imwinkelried, E.J. (2016) 'Expert Witness: Daubert and Beyond', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 456–462. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00173-7.

Indriati, E. (2009) 'Historical perspectives on forensic anthropology in Indonesia', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archeology*. Left Coast Press, pp. 115–125.

Ings, S. (2019) 'Call AI by a better name', *New Scientist*, 241(3212), pp. 44–45. Available at: https://doi.org/10.1016/S0262-4079(19)30078-8.

Innocence Project (2022) *Eyewitness Identification Reform: Mistaken Identifications are the Leading Factor In Wrongful Convictions*, *Article*. Available at: https://innocenceproject.org/eyewitness-identification-reform/ (Accessed: 3 March 2022).

Interpol (1996) 'General Assembly on Disaster Victim Identification', in *Interpol GA-1996-65-RES-13 DVI*. Lyon, France, pp. 23-29 October.

Interpol (2018) *INTERPOL Disaster Victim Identification Guide*. Available at: https://www.interpol.int/en/How-we-work/Forensics/Disaster-Victim-Identification-DVI.

Interpol (2021) *Facial Recognition*, *Notice*. Available at: https://www.interpol.int/How-we-work/Forensics/Facial-Recognition (Accessed: 10 April 2021).

Ireland, J. and Beaumont, J. (2015) ‘Admitting scientific expert evidence in the UK: reliability challenges and the need for revised criteria – proposing an Abridged Daubert’, *The Journal of Forensic Practice*, 17(1), pp. 3–12. Available at: https://doi.org/10.1108/JFP-03-2014-0008.

İşcan, M.Y. (1993) ‘Introduction of Techniques for Photographic Comparison: Potential and Problems’, in M.Y. İşcan and R.P. Helmer (eds) *Forensic Analysis of the Skull*. New York: Wiley-Liss, pp. 57–70.

Ishii, M., Yayama, K., Motani, H., Sakuma, A., Yasijma, D., Hayakawa, M., Yamamoto, S. and Iwase, H. (2011) ‘Application of Superimposition-Based Identification Using Skull Computed Tomography Images’, *Journal of Forensic Sciences*, 56(4), pp. 960–966. Available at: https://doi.org/10.111/j.1556-4029.2011.01797.x.

ISO - International Organization for Standardization (2017) *ISO/IEC 2382-37:2017 Information technology — Vocabulary — Part 37: Biometrics*, *Standard*. Available at: https://www.iso.org/standard/66693.html (Accessed: 27 May 2021).

Ito, I., Imada, M., Ikeda, M., Sueno, K., Arikuni, T. and Kida, A. (2001) ‘A Morphological Study of Age Changes in Adult Human Auricular Cartilage With Special Emphasis on Elastic Fibers’, *The Laryngoscope*, 111(5), pp. 881–886. Available at: https://doi.org/10.1097/00005537-200105000-00023.

Jain, A.K. (2004) ‘Biometric recognition: how do I know who you are?’, in *IEEE 12th Signal Processing and Communications Applications Conference*. Kusadasi, Turkey: IEEE, pp. 1–3. Available at: https://doi.org/10.1109/SIU.2004.1338241.

Jain, A.K., Klare, B. and Park, U. (2012) ‘Face Matching and Retrieval in Forensics Applications’, *IEEE MultiMedia*, 19, pp. 20–28. Available at: https://doi.org/10.1109/MMUL.2012.4.

Jain, A.K., Nadnakumar, K. and Ross, A. (2016) ‘50 years of biometric research: Accomplishments, challenges, and opportunities’, *Pattern Recognition Letters*, 79, pp. 80–105. Available at: https://doi.org/10.1016/j.patrec.2015.12.013.

Jain, A.K., Ross, A. and Pankanti, S. (2006) ‘Biometrics: A Tool for Information Security’, *IEEE Transactions on Information Forensics and Security*, 1(2), pp. 125–143. Available at: https://doi.org/10.1109/IC3I.2014.7019771.

Jain, A.K., Ross, A. and Prabhakar, S. (2004) ‘An Introduction to Biometric Recognition’, *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image. and Video-Based Biometrics*, 14(1), pp. 1–29. Available at: https://doi.org/10.1109/TCSVT.2003.818349.

Janaway, R.C., Percival, S.L. and Wilson, A.S. (2009) ‘Decomposition of Human Remains’, in S.L. Percival (ed.) *Microbiology and Aging*. Humana Press, pp. 313–334. Available at: https://doi.org/10.1007/978-1-59745-327-1_14.

Jayaprakash, P.T., Singh, B., Yusop, R.A.A.M. and Asmundi, H.S. (2010) ‘Skull-Photo Superimposition : A Remedy to the Problem of Unidentified Dead in Malaysia’, *Malaysian Journal of Forensic Sciences*, 1(1), pp. 34–41.

Jayaprakash, P.T., Srinivasan, G.J. and Amravaneswaran, M.G. (2001) ‘Cranio-facial morphanalysis: A new method for enhancing reliability while identifying skulls by photo superimposition', *Forensic Science International*, 117(1–2), pp. 121–143. Available at: https://doi.org/10.1016/S0379-0738(00)00455-2.

Jdid, R., Latreille, J., Soppelsa, F., Tschachler, E. and Morizot, F. (2018) 'Validation of digital photographic reference scales for evaluating facial aging signs', *Skin Research and Technology*, 24, pp. 196–202. Available at: https://doi.org/10.1111/srt.12413.

Jeckeln, G., Hahn, C.A., Noyes, E., Cavaos, J.G. and O'Toole, A. (2018) 'Wisdom of the social versus non-social crowd in face identification', *British Journal of Psychology*, 109(4), pp. 724–735. Available at: https://doi.org/10.1111/bjop.12291.

Jenkins, R. and Burton, A.M. (2008) 'Limitations in Facial Identification: The Evidence', *Justice of the Peace*, 172, pp. 4–6. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.9728&rep=rep1&type=pdf.

Jenkins, R., White, D., Van Montfort, X. and Burton, A.M. (2011) 'Variability in photos of the same face', *Cognition*, 121, pp. 313–323. Available at: https://doi.org/10.1016/j.cognition.2011.08.001.

Jensen, R.A. (1999) *Mass Fatality and Casualty Incidents: A Field Guide*. 1st edn. CRC Press.

Jin, X. and Tan, X. (2016) 'Face alignment in-the-wild: A survey', *Computer Vision and Image Understanding*, 162(9), pp. 1–38. Available at: https://doi.org/10.1016/j.cviu.2017.08.008.

Johnson, B.T. and Riemen, J.A.J.M. (2019) 'Digital capture of fingerprints in a disaster victim identification setting: a review and case study', *Forensic Science Research*, 4(4), pp. 293–302. Available at: https://doi.org/10.1080/20961790.2018.1521327.

Johnson, M.H. and Morton, J. (1991) *Biology and cognitive development: The case of face recognition*. Oxford: John Wiley and Sons Ltd.

Johnston, R.A. and Edmonds, A.J. (2009) 'Familiar and unfamiliar face recognition: A review', *Memory*, 17(5), pp. 577–596. Available at: https://doi.org/10.1080/09658210902976969.

Jones, M. and Viola, P. (2003) 'Fast Multi-view Face Detection', in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.145.7598.

Joukal, M. and Frišhons, J. (2015) 'A facial reconstruction and identification technique for seriously devastating head wounds', *Forensic Science International*, 252, pp. 82–86. Available at: https://doi.org/10.1016/j.forsciint.2015.04.027.

Jourabloo, A. and Liu, X. (2015) 'Pose-Invariant 3D Face Alignment', in *IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: IEEE, pp. 3694–3702. Available at: https://doi.org/10.1109/ICCV.2015.421.

Kadlec, T.D. (2021) *Why do Daubert and Frye standards matter to expert witnesses?* Available at: envista-whitepaper-daubert-and-frye-standards_final.pdf.

Kahana, T., Almog, J., Levy, J., Shmeltzer, E., Spier, Y. and Hiss, J. (1999) 'Marine taphonomy: Adipocere formation in a series of bodies recovered from a single shipwreck', *Journal of Forensic Sciences*, 44(5), pp. 897–901. Available at: https://doi.org/10.1520/JFS12012J.

Kahn, D.M. and Shaw Jr, R.B. (2008) 'Aging of the boy orbit: a three-dimensional computed tomographic study', *Aesthetic Surgery Journal*, 28(3), pp. 258–264. Available at: https://doi.org/10.1016/j.asj.2008.02.007.

Kalina, N. (2020) *Noah takes a photo of humself every day for 20 years*, *YouTube*. Available at: https://www.youtube.com/watch?v=wAIZ36GI4p8 (Accessed: 13 October 2020).

Kamar, N., Berg, A.C., Belhumeur, P.N. and Nayar, S.K. (2009) 'Attribute and Simile Classifiers for Face Verification', in *12th International Conference on Computer Vision - ICCV*. Kyoto, Japan: IEEE. Available at: https://ieeexplore.ieee.org/document/5459250.

Kanade, T. (1973) *Picture Processing System by Computer Complex and Recognition of Human Faces*. KYoto University. Available at: https://doi.org/10.14989/doctor.k1486.

Kanathur, S., Sarvajnyamurthy, S. and Somaiah, S.A. (2013) 'Characteristic facies: An index of the disease', *Indian Journal of Dermatology, Venereology, and Leprology*, 79(3), pp. 439–443. Available at: https://doi.org/10.4103/0378-6323.110801.

Kantayya, S. (2020) *Coded Bias©*. United States: 7th Empire Media. Available at: https://www.codedbias.com/about (Accessed: 6 May 2021).

Karayanidis, F., Kelly, M., Chapman, P., Mayes, A. and Johnston, P. (2009) 'Facial identity and facial expression matching in 5–12-year-old children and adults', *Infant and Child Development*, 18(5), pp. 404–421. Available at: https://doi.org/10.1002/icd.615.

Kaur, M., Garg, R.K. and Singla, S. (2015) 'Analysis of facial soft tissue changes with aging and their effects on facial morphology: A forensic perspective', *Egyptian Journal of Forensic Sciences*, 5, pp. 46–56. Available at: https://doi.org/10.1016/j.ejfs.2014.07.006.

Kaur, P., KRishan, K., Sharma, S.K. and Kanchan, T. (2020) 'Facial-recognition algorithms: A literature review', *Medicine, Science and the Law*, 60(2), pp. 131–139. Available at: https://doi.org/10.1177/0025802419893168.

Kealy, G., Gapert, R., Buckley, L., Cassidy, M., Mcnulty, J., Foyle, R., Meier-, W., Kemp, H. and Wilkinson, C. (2014) 'Multi-Disciplinary Approach toward the Identification of a Human Skull Found 55 km off the Southeast Coast of Ireland', in X. Mallett, T. Blythe, and R. Berry (eds) *Advances in Forensic Human Identification*. 1st edn. Ta, pp. 193–210.

Keil, W. (2021) *Rechtsmedizin Basics*. Urban & Fischer Verlag/Elsevier GmbH.

Kemp, R., Towell, N. and Pike, G. (1997) 'When seeing should not be believing: Photographs, credit cards and fraud', *Applied Cognitive Psychology*, 11(3), pp. 211–222. Available at: https://doi.org/10.1002/(SICI)1099-0720(199706)11:3<211::AID-ACP430>3.0.CO;2-O.

Kemp, R.I., Caon, A. l., Howard, M. and Brooks, K.R. (2016) 'Improving unfamiliar face matching by masking the external facial features', *Applied Cognitive Psychology*, 30(4), pp. 622–627. Available at: https://doi.org/10.1002/acp.3239.

Kenny, C.T. and Fletcher, D. (1973) 'Effects of Beardedness on Person Perception', *Perceptual and Motor Skills*, 37(2), pp. 413–414. Available at: https://doi.org/10.2466/pms.1973.37.2.413.

Keough, N., Myburgh, J. and Steyn, M. (2017) ‘Scoring of Decomposition: A ProposedAmendment to the Method When Using a PigModel for Human Studies*’, *Journal of Forensic Sciences*, 62(4), pp. 986–993. Available at: https://doi.org/10.1111/1556-4029.13390.

Kersten, T.P., Przybilla, H.-J. and Lindstaedt, M. (2016) ‘Investigations of the Geometrical Accuracy of Handheld 3D Scanning Systems’, *Photogrammetrie - Fernerkundung - Geoinformation*, 5–6, pp. 271–283. Available at: https://doi.org/10.1127/pfg/2016/0305.

Keshtgar, S., Keshtgar, A., Mistry, P. and Shakib, K. (2019) ‘Assessing facial recognition after orthognathic surgery at automated border controls in airports’, *British Journal of Oral and Maxillofacial Surgery*, 57(6), pp. 536–538. Available at: https://doi.org/10.1016/j.bjoms.2018.12.021.

Khoo, L.S., Lai, P.S., Hasmi, A.H. and Mahmood, M.S. (2016) ‘Secondary identifier for positive identification in DVI’, *Forensic Science and Criminology*, 1(1), pp. 1–3. Available at: https://doi.org/10.15761/FSC.1000102.

Khoo, L.S. and Mahmood, M.S. (2020) ‘Application of facial recognition technology on identification of the dead during large scale disasters’, *Forensic Science International: Synergy*, 2, pp. 238–239. Available at: https://doi.org/10.1016/j.fsisyn.2020.07.001.

Khudomoma, S.A.N. (2017) ‘The first case of mummified face and skull-ohoto superimposition in the United Arab Emirates’, in *Proceedings - American Academy of Forensic Sciences 69th annual scientific meeting*, p. 1381. Available at: https://csidds.files.wordpress.com/2017/02/2017aafs-proceedings.pdf.

Kisyova, R., Rahman, S.M., Dheansa, B. and Karkhi, A. (2019) ‘Ethics in aesthetics: social media’, *Journal of Aesthetic Nursing*, 8(7), pp. 310–312. Available at: https://doi.org/10.12968/joan.2019.8.7.310.

Kleinberg, K.F., Vanezis, P. and Burton, A.M. (2007) ‘Failure of anthropometry as a facial identification technique using high-quality photographs’, *Journal of Forensic Sciences*, 52(4), pp. 779–783. Available at: https://doi.org/10.1111/j.1556-4029.2007.00458.x.

Knobel, Z., Ueland, M., Nizio. K.D., Patel, D. and Forbes, S.L. (2019) ‘A comparison of human and pig decompositionrates and odour profiles in an Australianenvironment’, *Australian Journal of Forensic Sciences*, 51(5), pp. 557–572. Available at: https://doi.org/10.1080/00450618.2018.1439100.

Knussmann, R. (1988) ‘Methoden des morphologischen Vergleichs in der forensischen Anthropologie’, in R. Knussmann (ed.) *Anthropologie - Handbuch der vergleichenden Biologie des Menschen - Band 1*. 4th edn. Stuttgart, Germany: Gustav Fischer Verlag, pp. 368–407.

Knußmann, R. (ed.) (1992) *Anthropologie - Handbuch der vergleichenden Biologie des Menschen. Band I/2 Physiologische, psychologische, genetische und mathematische Methoden*. 4th edn. Stuttgart, Germany: Gustav Fischer Verlag.

Kobelinsky, C. and Furri, F. (2020) *Hosting the dead by migration. The treatment of lifelessbodies in Catania (Sicily). COMPAS Special Working Paper series on Migrations in Latin America andthe Mediterranean compared: Violence, state cruelty and (un-)institutional resistance*. MFO LAC 1.9. Oxford. Available at: https://halshs.archives-ouvertes.fr/halshs-03038850/document.

Kobus, H.J., Kirkbride, K.P. and Raymond, M.A. (2016) 'Identification: Fingerprints a Key Identification Parameter –Detection, Identification, and Interpretation', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 65–73. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00208-1.

Kortli, Y., Jridi, M., Al Falou, A. and Atri, M. (2020) 'Face Recognition Systems: A Survey', *Sensors*, 20(2), pp. 342-. Available at: https://doi.org/10.3390/s20020342.

Kosut, M. (2015) 'Tattoos and Body Modification', *International Encyclopedia of the Social & Behavioral Sciences*, 24(2), pp. 32–38. Available at: https://doi.org/10.1016/B978-0-08-097086-8.64027-8.

Kramer, R.S.S., Mohamed, S. and Hardy, S.C. (2019) 'Unfamiliar Face Matching With Driving Licence and Passport Photographs', *Perception*, 48(2), pp. 175–184. Available at: https://doi.org/10.1177/0301006619826495.

Kramer, R.S.S. and Reynolds, M.G. (2018) 'Unfamiliar Face Matching With Frontal and Profile Views', *Perception*, 47(4), pp. 414–431. Available at: https://doi.org/10.1177/0301006618756809.

Kramer, R.S.S., Young, A.W. and Burton, A.M. (2018) 'Understanding face familiarity', *Cognition*, 172, pp. 46–58. Available at: https://doi.org/10.1016/j.cognition.2017.12.005.

Kravitz, N.D. and Bowman, S.J. (2016) 'A Paradigm Shift in Orthodontic Marketing', *Seminars in Orthodontics*, 22(4), pp. 297–300. Available at: https://doi.org/10.1053/j.sodo.2016.08.010.

Kremer, D. (2017) *What is The Kardashian Effect?*, *blog*. Available at: https://www.harleystreetaesthetics.com/blog/dr-kremers-blog/2017/01/27/what-is-the-kardashian-effect (Accessed: 22 July 2020).

Kreutz, K. and Verhoff, M.A. (2004) 'Gesichtserkennung oder der etwas andere Sinn des Menschen', *Gießener Universitätsblätter*, 37, pp. 23–29.

Krishan, K., Kanchan, T. and Thakur, S. (2019) 'A study of morphological variations of the human ear for its applications in personal identification', *Egyptian Journal of Forensic Sciences*, 9(6), pp. 1–11. Available at: https://doi.org/10.1186/s41935-019-0111-0.

Kyllonen, K.M. and Monson, K.L. (2020) 'Depiction of ethnic aging by forensic artists and preliminary assessment of the applicability of facial averages', *Forensic Science International*, 313, pp. 1–13. Available at: https://doi.org/10.1016/j.forsciint.2020.110353.

Labati, R.D., De Angelis, D., Bertoglio, B., Cattaneo, C., Scotti, F. and Piuri, V. (2021) 'Automatic Face Recognition for Forensic Identification of Persons Deceased in Humanitarian Emergencies', in *2021 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. Hong Kong, China: IEEE, pp. 1–6. Available at: https://doi.org/10.1109/CIVEMSA52099.2021.9493678.

Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R.P. and Konen, W. (1993) 'Distortion invariant object recognition in the dynamic link architecture', *IEEE Transactions on Computers*, 42(3), pp. 300–311. Available at: https://doi.org/10.1109/12.210173.

Lain, R., Griffiths, C. and Hilton, J.M.N. (2003) ‘Forensic dental and medical response to the Bali bombing – a personal perspective’, *Medical Journal of Australia*, 179(7), pp. 362–365. Available at: https://doi.org/10.5694/j.1326-5377.2003.tb05594.x.

Lakotta, B. (2003) ‘Noch mal leben vor dem Tod’, *Der Spiegel - 26*, pp. 126–132. Available at: https://www.spiegel.de/spiegel/a-254960.html.

Law Commission (2009) *The Admissability of Expert Evidence in Criminal Proceedings in England and Wales: A New Approach to the Determination of Evidentiary Reliability - A Consultation Paper*. London.

Lawler, W. (1992) ‘Bodies recovered from water: a personal approach and consideration of difficulties’, *Journal of Clinical Pathology*, 45(8), pp. 654–659. Available at: https://doi.org/10.1136/jcp.45.8.654.

Lawrence, C. (2020) *5 Women on The Clever Ways They’re Future-Proofing Their Skin*, *Article*. Available at: https://www.womenshealthmag.com/uk/beauty/a33336933/loreal-future-proof-skin-tips/ (Accessed: 3 August 2020).

Lee, J.-H. and Hong, G. (2018) ‘Definitions of groove and hollowness of the infraorbital region and clinical treatment using soft-tissue filler’, *Archives of Plastic Surgery*, 45, pp. 214–221. Available at: https://doi.org/10.5999/asp.2017.01193.

Lee, P.B., Miano, D.I., Sesselmann, M., Johnson, J., Chung, M.T., Abboud, M., Johnson, A.P. and Zuliani, G.F. (2022) ‘RealSelf Social Media Analysis of Rhinoplasty Patient Reviews’, *Journal of Plastic, Reconstructive & Aesthetic Surgery*, (Article in Press), pp. 1–7. Available at: https://doi.org/10.1016/j.bjps.2022.02.060.

Lee, W.J., Kim, D.M., Lee, U.Y., Cho, J.H., Kim, M.S., Hong, J.H. and Hwang, Y.I. (2019) ‘A Preliminary Study of the Reliability of Anatomical Facial Landmarks Used in Facial Comparison’, *Journal of Forensic Sciences*, 64(2), pp. 519–527. Available at: https://doi.org/https://doi.org/10.1111/1556-4029.13873.

Lee, W.J., Mackenzie, S. and Wilkinson, C. (2011) ‘Facial Identification of the Dead’, in S. Black and E. Ferguson (eds) *Forensic Anthropology 2000 to 2010*. 1st edn. Boca Raton, Florida: CRC Press, pp. 363–394.

Lee, W.J., Wilkinson, C., Memon, A. and Houston, K. (2009) ‘Matching unfamiliar faces from poor quality closed-circuit television (CCTV) footage: An evaluation of the effect of training on facial identification ability’, *AXIS Online J CAHid*, 1(1), pp. 19–28.

Lei, Y., Bennamoun, M., Hayat, M. and Guo, Y. (2014) ‘An efficient 3D face recognition approach using local geometrical signatures’, *Pattern Recognition*, 47(2), pp. 509–524. Available at: https://doi.org/10.1016/j.patcog.2013.07.018.

Leipner, A., Obertová, Z., Wermuth, M., Thali, M., Ottiker, T. and Sieberth, T. (2019) ‘3D mug shot—3D head models from photogrammetry for forensic identification’, *Forensic Science International*, 300, pp. 6–12. Available at: https://doi.org/10.1016/j.forsciint.2019.04.015.

Lenferink, L.I.M., de Keijser, J., Piersma, E. and Boelen, P.A. (2018) ‘I’ve changed but I’m not less happy: Interview study among nonclinical relatives of long-term missing persons’, *Death Studies*, 42(6), pp. 346–355. Available at: https://doi.org/10.1080/07481187.2017.1347213.

Lesciotto, K.M. (2015) 'The Impact of Daubert on the Admissibility of Forensic Anthropology Expert Testimony', *Journal of Forensic Sciences*, 60(3), pp. 549–555. Available at: https://doi.org/10.1111/1556-4029.12740.

Levine, S.C., Banich, M.T. and Koch-Weser, M.P. (1988) 'Face recognition: a general or specific right hemisphere capacity?', *Brain and Congnition*, 8(3), pp. 303–325. Available at: https://doi.org/10.1016/0278-2626(88)90057-7.

Lin, L., Wang, K., Meng, D., Zuo, W. and Zhang, L. (2017) 'Active Self-Paced Learning for Cost-Effective and Progressive Face Identification', *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 40(1), pp. 7–19. Available at: https://doi.org/10.1109/TPAMI.2017.2652459.

Liu, C.Y.J. (2018) *Facial identification from online images for use in the prevention of child trafficking and exploitation*. Liverpool John Moores University. Available at: https://doi.org/10.24377/LJMU.t.00009901.

Lobmaier, J.S. and Mast, F.W. (2007) 'Perception of Novel Faces: The Parts Have it!', *Perception*, 36(11), pp. 1660–1673. Available at: https://doi.org/10.1068/p5642.

Londoño Romanowsky, X. and Silva Chau, M. (2020) 'The protection of the missing and the dead under international law', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 25–35. Available at: https://doi.org/10.1002/9781119482062.ch2.

Longmore, C.A., Liu, C.H. and Young, A.W. (2015) 'The importance of internal facial features in learning new faces', *The Quarterly Journal of Experimental Psychology*, 68(2), pp. 249–260. Available at: https://doi.org/10.1080/17470218.2014.939666.

Lopatina, O.L., Komleva, Y.K., Gorina, Y.V., Higashida, H. and Salmina, A.B. (2018) 'Neurobiological Aspects of Face Recognition: The Role of Oxytocin', *Frontiers in Behavioral Neuroscience*, 12(195), pp. 1–11. Available at: https://doi.org/10.3389/fnbeh.2018.00195.

Le Louarn, C. (2009) 'Muscular aging and its involvement in facial aging: the Face Recurve concept', *Annales de dermatologie et de venereologie*, 136(4), pp. 67–72. Available at: https://doi.org/10.1016/S0151-9638(09)74530-2.

Lu, X., Jain, A.K. and Colbry, D. (2006) 'Matching {2.5D} Face Scans to {3D} Models', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1), pp. 31–43.

Lucas, T. and Henneberg, M. (2015) 'Are human faces unique? A metric approach to finding single individuals without duplicates in large samples', *Forensic Science International*, 257, pp. 514.e1-514.e6. Available at: https://doi.org/10.1016/j.forsciint.2015.09.003.

Lynnerup, N. (2001) 'Cranial thickness in relation to age, sex and general body build in a Danish forensic sample', *Forensic Science International*, 117(1–2), pp. 45–51. Available at: https://doi.org/10.1016/s0379-0738(00)00447-3.

Lynnerup, N., Clausen, M.L., Kristoffersen, A.M. and Steglich-Arnholm, H. (2009) 'Facial recognition and laser surface scan: A pilot study', *Forensic Science, Medicine, and Pathology*, 5(3), pp. 167–173. Available at: https://doi.org/10.1007/s12024-009-9094-8.

MacKinnon, G. and Harrison, K. (2016) 'Forensic Anthropology and Archaeology in the United Kingdom - Are We Nearly There Yet?', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 13–26.

Madea, B. (2016) 'Methods for determining time of death', *Forensic Science, Medicine, and Pathology*, 12(4), pp. 451–485. Available at: https://doi.org/10.1007/s12024-016-9776-y.

Madea, B., Schmidt, P., Stenzinger, A. and Dietel, M. (2014) 'Praktische Durchführung der ärztlichen Leichenschau - Aufgabenkomplexe', in B. Madea (ed.) *Die ärztliche Leichenschau*. 3rd edn. Berlin Heidelberg, Germany: Springer, pp. 67–142.

Maguire, M. (2009) 'The birth of biometric security', *Anthropology Today*, 25(2), pp. 9–15. Available at: http://mural.maynoothuniversity.ie/3014/1/The_Birth_of_Biometric_Security.pdf.

MailOnline (2020) *"My skin is so different, it's unbelievable!": How an anti-ageing cream is making REAL women postpone scary cosmetic procedures!*, *Article*. Available at: https://www.dailymail.co.uk/femail/article-8498365/How-anti-ageing-cream-making-REAL-women-postpone-scary-cosmetic-procedures.html (Accessed: 3 August 2020).

Makrantonaki, E., Bekou, V. and Zouboulis, C.C. (2012) 'Genetics and skin aging', *Dermato-Endocrinology*, 4(3), pp. 280–284. Available at: https://doi.org/10.4161/derm.22372.

Makrantonaki, E., Pfeifer, G.P. and Zouboulis, C.C. (2016) 'Endogene Faktoren, Gene und Hautalterung', *Hautarzt*, 67, pp. 103–106.

Maliniak, J.W. (1935) 'The Plastic Surgeon and Crime', *American Institute of Criminal LAw and Criminology*, 26(4), pp. 594–600.

Mallett, X. and Evison, M.P. (2013) 'Forensic facial comparison: issues of admissibility in the development of novel analytical technique', *Journal of Forensic Sciences*, 58(4), pp. 859–865. Available at: https://doi.org/10.1111/1556-4029.12127.

Maltais Lapointe, G., Lynnerup, N. and Hoppa, R.D. (2016) 'Validation of the New Interpretation of Gerasimov's Nasal Projection Method for Forensic Facial Approximation Using CT Data', *Journal of Forensic Sciences*, 61(January), pp. 193–200. Available at: https://doi.org/10.1111/1556-4029.12920.

Marionnet, C., Tricaud, C. and Bernerd, F. (2015) 'Exposure to Non-Extreme Solar UV Daylight: Spectral Characterization, Effects on Skin and Photoprotection', *International Journal of Molecular Sciences*, 16, pp. 68–90. Available at: https://doi.org/10.3390/ijms16010068.

Marks, M.K., Love, J.C. and Dadour, I.R. (2009) 'Taphonomy and Time: Estimating the Postmortem Interval', in D.W. Steadman (ed.) *Hard Evidence: Case Studies in Forensic Anthropology*. Upper Saddle River, JR: Pearson Education LTD, pp. 167–180.

Marks, M.K. and Tersigni-Tarrant, M.A. (2016) 'Death Investigation Systems: Decomposition, Patterns, and Rates', in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 8–12. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00135-X.

Márquez-Grant, N. and Roberts, J. (2021) 'Redefining forensic anthropology in the 21st century and its role in mass fatality investigations', *European Journal of Anatomy*, 25(Suppl. 2), pp. 19–34. Available at: https://dspace.lib.cranfield.ac.uk/handle/1826/16093.

Marshall, J. (2014) *Who do we think we are?*, *Article*. Available at: http://www.eric-group.co.uk/philosophy/who-do-we-think-we-are.

De Marsico, M., Nappi, M., Riccio, D. and Wechsler, H. (2013) 'Robust Face Recognition for Uncontrolled Pose and Illumination Changes', *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(1), pp. 149–163. Available at: https://doi.org/10.1109/TSMCA.2012.2192427.

De Marsico, M., Nappi, M., Riccio, D. and Wechsler, H. (2015) 'Robust face recognition after plastic surgery using region-based approaches', *Pattern Recognition*, 48(4), pp. 1261–1276. Available at: https://doi.org/10.1016/j.patcog.2014.10.004.

Martires, K.J., Fu, P., Polster, A.M., Cooper, K.D. and Baron, E.D. (2009) 'Factors That Affect Skin Aging: A Cohort-Based Survey on Twins', *Archives of Dermatology*, 145(12), pp. 1375–1379. Available at: https://doi.org/https://doi.org/10.1001/archdermatol.2009.303.

Martos Fernández, R. (2021) *Identificación Craneofacial: Cuantificación y Comparación Morfológica*. Universidad de Granada.

Martos, R., Guerra, R., Navarro, F., Peruch, M., Neuwirth, K., Valsecchi, A., Jankauskas, R. and Ibáñez, O. (2022) 'Computer-aided craniofacial superimposition validation study: the identification of the leaders and participants of the Polish-Lithuanian January Uprising (1863–1864)', *International Journal of Legal Medicine*, pp. 1–15. Available at: https://doi.org/10.1007/s00414-022-02929-4.

Masi, I., Wu, Y., Hassner, T. and Natarajan, P. (2019) 'Deep Face Recognition: a Survey', in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. Parana, Brazil: IEEE. Available at: https://doi.org/10.1109/SIBGRAPI.2018.00067.

Mathworks UK (2014) *Face Recognition with MATLAB*, *Website*. Available at: https://uk.mathworks.com/videos/face-recognition-with-matlab-98076.html (Accessed: 9 November 2020).

Mathworks UK (2021) *What Is Image Recognition? - MATLAB & Simulink*, *Website*. Available at: https://uk.mathworks.com/discovery/image-recognition-matlab.html (Accessed: 8 September 2021).

Maurer, D., Le Grand, R. and Mondloch, C.J. (2002) 'The many faces of configural processing', *Trends in Cognitive Sciences*, 6(6), pp. 255–260. Available at: https://doi.org/10.1016/S1364-6613(02)01903-4.

Mazzarelli, D., Bertoglio, B., Boscacci, M., Caccia, G., Ruffetta, C., De Angelis, D., Fracasso, T., Baraybar, J.P., Riccio, S., Marzagalia, M.M. and Cattaneo, C. (2021) 'Ambiguous loss in the current migration crisis: A medico-legal, psychological, and psychiatric perspective', *Forensic Science International: Mind and Law*, 2, pp. 1–11. Available at: https://doi.org/10.1016/j.fsiml.2021.100064.

McCaffery, J.M. and Burton, A.M. (2016) 'Passport checks: Interactions between matching faces and biographical details', *Applied Cognitive Psychology*, 30(6), pp. 925–933. Available at: https://doi.org/10.1002/acp.3281.

Mckenna, J.J.L., Jablonski, N.G. and Fearnhead, R.W. (1984) 'A Method of Matching Skulls with Photographic Portraits Using Landmarks and Measurements of the Dentition', *Journal of Forensic Sciences*, 29(3), pp. 787–797.

McKone, E., Crookes, K. and Kanwisher, N. (2009) 'The cognitive and neural development of face recognition in humans', in M.S. Gazzaniga, E. Bizzi, L.M. Chalupa, S.T. Grafton, T.F. Heatherton, C. Koch, J.E. LeDoux, S.J. Luck, G.R. Mangan, J.A. Movshon, H. Neville, E.A. Phelbs, P. Rakic, D.L. Schacter, M. Sur, and B.A. Wandell (eds) *The Cognitive Neurosciences*. Massachusetts Institute of Technology, pp. 467–482.

McNeill, A., McNeill M. and Strathie, A. (2015) 'Expert facial comparison evidence: Science versus pseudo science', *Psychology & Law*, 5(4), pp. 127–140. Available at: https://doi.org/10.17759/psylaw.2015050411.

Mealey, L. (2000a) 'Female Strategies and Tactics', in *Sex Differences: Development and Evolutionary Strategies*. Elsevier Inc., pp. 117–142.

Mealey, L. (2000b) 'Sexual Differentiation', in *Sex Differences: Development and Evolutionary Strategies*. Elsevier Inc., pp. 11–38. Available at: https://doi.org/10.1016/B978-0-12-487460-2.X5000-4.

Mediati, N. (2016) *PC World: Google kills off Picasa to focus its efforts on Google Photos, Article*. Available at: https://www.pcworld.com/article/3033156/google-kills-off-picasa-to-focus-its-efforts-on-google-photos.html (Accessed: 7 September 2021).

Megreya, A.M. (2018) 'Feature-by-feature comparison and holistic processing in unfamiliar face matching', *PeerJ*, 6(e4437), pp. 1–10. Available at: https://doi.org/10.7717/peerj.4437.

Megreya, A.M. and Bindemann, M. (2015) 'Developmental improvement and age-related decline in unfamiliar face matching', *Perception*, 44(1), pp. 5–22. Available at: https://doi.org/10.1068/p7825.

Megreya, A.M. and Bindemann, M. (2018) 'Feature instructions improve face-matching accuracy', *PLos ONE*, 13(3), pp. 1–16. Available at: https://doi.org/10.1371/journal.pone.0193455.

Megreya, A.M. and Burton, A.M. (2006) 'Unfamiliar faces are not faces: evidence from a matching task', *Memory & Cognition*, 34(4), pp. 865–876. Available at: https://doi.org/10.3758/bf03193433.

Megreya, A.M. and Burton, A.M. (2008) 'Matching faces to photographs: poor performance in eyewitness memory (without the memory)', *Journal of Experimental Psychology: Applied*, 14(4), pp. 364–372. Available at: https://doi.org/10.1037/a0013464.

Megreya, A.M., Sandford, A. and Burton, A.M. (2013) 'Matching Face Images Taken on the Same Day or Months Apart: the Limitations of Photo ID', *Applied Cognitive Psychology*, 27, pp. 700–706. Available at: https://doi.org/10.1002/acp.2965.

Meijerman, L., van der Lugt, C. and Maat, G.J.R. (2007) 'Cross-sectional anthropometric study of the external ear', *Journal of Forensic Sciences*, 52(2), pp. 286–293. Available at: https://doi.org/10.1111/j.1556-4029.2006.00376.x.

Meinhardt-Injac, B., Boutet, I., Persike, M., Meinhardt, G. and Imhof, M. (2017) 'From development to aging: Holistic face perception in children, younger and older adults', *Cognition*, 158, pp. 134–146. Available at: https://doi.org/10.1016/j.cognition.2016.10.020.

Mendelson, B. and Wong, C.-H. (2012) 'Changes in the Facial Skeleton With Aging: Implications and Clinical Applications in Facial Rejuvenation', *Aesthetic Plastic Surgery*, 36, pp. 753–760. Available at: https://doi.org/10.1007/s00266-012-9904-3.

Michel, C., Caldara, R. and Rossion, B. (2006) 'Same-race faces are perceived more holistically than other-race faces', *Visual Cognition*, 14(1), pp. 55–73. Available at: https://doi.org/10.1080/13506280500158761.

Microsoft (2015) *Windows 10 / Picasa problems*, *Customer Support Forum*. Available at: https://answers.microsoft.com/en-us/windows/forum/all/windows-10-picasa-problems/aafe741d-e188-4f33-90aa-80747dc8761b (Accessed: 23 January 2021).

Miller, D., Brossard, E., Seitz, S. and Kemelmacher-Shlizerman, I. (2015) 'MegaFace: A Million Faces for Recognition at Scale', *arXiv: Computer Vision and Pattern Recognition*, v2, pp. 1–10. Available at: https://doi.org/10.48550/arXiv.1505.02108.

Milroy, C.M. (2017) 'A Brief History of the Expert Witness', *Academic Forensic Pathology*, 7(4), pp. 516–526. Available at: https://doi.org/10.23907/2017.044.

Miranda, G.E., de Freitas, S.G., de Abreu Maia, L.V. and Haltenhoff Melani, R.F. (2016) 'An unusual method of forensic human identification: use of selfie photographs', *Forensic Science International*, 263, pp. e14–e17. Available at: https://doi.org/10.1016/j.forsciint.2016.04.028.

Moreton, R. (2021a) *Expertise in applied face matching: training, forensic examiners, super matchers and algorithms*. The Open University. Available at: https://doi.org/10.21954/ou.ro.00013240.

Moreton, R. (2021b) 'Forensic Face Matching - Procedures and Application', in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 144–173.

Moreton, R. and Morley, J. (2011) 'Investigation into the use of photoanthropometry in facial image comparison', *Forensic Science International*, 212(1–3), pp. 231–237. Available at: https://doi.org/10.1016/j.forsciint.2011.06.023.

Morgan, O.W., Sribanditmongkol, P., Perera, C., Sulasmi, Y., Van Alphen, D. and Sondorp, E. (2006) 'Mass fatality management following the south Asian tsunami disaster: Case studies in Thailand, Indonesia, and Sri Lanka', *PLoS Medicine*, 3(6), pp. 0809–0815. Available at: https://doi.org/10.1371/journal.pmed.0030195.

Mukherjee, S., Singh, P., Tuccia, F., Pradelli, J., Giordani, G. and Vanin, S. (2019) 'DNA characterization from gut content of larvae of Megaselia scalaris (Diptera, Phoridae)', *Science & Justice*, 59, pp. 654–659. Available at: https://doi.org/10.1016/j.scijus.2019.06.006.

Mulawka, M. (2014) *Postmortem Fingerprinting and Unidentified Human Remains*. 2nd edn. Elsevier Inc. Available at: https://www.sciencedirect.com/book/9780323266178/postmortem-fingerprinting-and-unidentified-human-remains#book-description.

Mullins, J. (2012) 'Age progression and regression', in C. Wilkinson and C. Rynn (eds)

Craniofacial Identification. Cambridge: Cambridge University Press, pp. 68–75.

Munson, R. (2020a) *The L'Oréal products you need in your anti-ageing skincare routine*, *Article*. Available at: https://www.goodhousekeeping.com/uk/fashion-beauty/g33378074/loreal-anti-ageing-products/ (Accessed: 3 August 2020).

Munson, R. (2020b) *Want Fresh, Glossy Skin of Dreams? Take Note Of These 3 Youth-Boosting Ingredients*, *Article*. Available at: https://www.elle.com/uk/beauty/a33323481/loreal-youth-boosting-ingredients/ (Accessed: 3 August 2020).

Nagi, J., Ahmed, S.K. and Nagi, F. (2008) 'A MATLAB based face recognition system using image processing and neural networks', in *4th IEEE International Colloquium on Signal Processing and its Applications*. Kuala Lumpur, Malaysia: IEEE, pp. 83–88.

Nappi, M., Ricciardi, S. and Tistarelli, M. (2016) 'Deceiving faces: When plastic surgery challenges face recognition', *Image and Vision Computing*, 54, pp. 71–82. Available at: https://doi.org/10.1016/j.imavis.2016.08.012.

NCA (2021) *UK Missing Persons Unit - Missing Persons Data Report 2019/20*.

Neave, N. and Shields, K. (2008) 'The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces', *Personality and Individual Differences*, 45(5), pp. 373–377. Available at: https://doi.org/10.1016/j.paid.2008.05.007.

Neave, R. (1998) 'Age changes to the face in adulthood', in J.G. Clement and D.L. Ranson (eds) *Craniofacial Identification in Forensic Medicine*. Arnold, pp. 225–231.

Nelson, C.A. (2001) 'The development and neural bases of face recognition', *Infant and Child Development1*, 10(1–2), pp. 3–18. Available at: https://doi.org/10.1002/icd.239.

Nelson, C.A. and Ludemann, P.M. (1989) 'Past, current, and future trends in infant face perception research', *Canadian Journal of Psychology*, 43(2), pp. 183–193. Available at: https://doi.org/10.1037/h0084221.

Next Generation Media (2018) *Snoop Dogg's Hollywood Walk Of Fame Speech! 'I Wanna Thank Me For Never Quitting'*, *Video*. Available at: https://www.youtube.com/watch?v=ozwntTM509c (Accessed: 8 July 2019).

Nickerson, B., Fitzhorn, P., Koch, S. and Charney, M. (1991) 'A methodology for near-optimal computational superimposition of two-dimensional digital facial photographs and three-dimensional cranioal surface meshes', *Journal of Forensic Sciences*, 36, pp. 480–500.

Nikolakis, G., Zoschke, C., Makrantonaki, E., Hausmann, C., Schäfer-Korting, M. and Zouboulis, C.C. (2016) 'Experimentelle Modelle humaner Hautalterung', *Hautarzt*, 67(2), pp. 93–98. Available at: https://doi.org/10.1007/s00105-015-3747-1.

Nixon, M.S., Bouchrika, I., Arbab-Zavar, B. and Carter, J.N. (2010) 'On use of biometrics in forensics: Gait and ear', *European Signal Processing Conference*, pp. 1655–1659. Available at: https://doi.org/10.1016/j.neuron.2015.10.013.

Van Noorden, R. (2020) 'The ethical questions that haunt facial-recognition research', *Nature*, 587(7834), pp. 354–358. Available at: https://doi.org/10.1038/d41586-020-

03187-3.

Norwood, O.T. (1975) ‘Male pattern baldness: classification and incidence’, *Southern Medical Journal*, 68(11), pp. 1359–1365. Available at: https://doi.org/10.1097/00007611-197511000-00009.

Nuzzolese, E., Lupariello, F. and Di Vella, G. (2018) ‘Selfie identification app as a forensic tool for missing and unidentified persons’, *Journal of Forensic Dental Sciences*, 10(2), pp. 75–78. Available at: https://doi.org/10.4103/jfo.jfds_80_17.

Nysten, P.-H. (1811) *Recherches de physiologie et de chimie pathologiques, suite à celles de Bichat sur la vie et la mort*. Paris: Brosson.

O’Brien, T.G. (1994) *Human Soft-Tissue Decomposition in an Aquatic Environment and its Transformation into Adipocere*. University of Tennessee, Knoxville.

O’Brien, T.G. (2006) ‘Movement of Bodies in Lake Ontario’, in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. Boca Raton, Florida: CRC Press LLC, pp. 559–565.

O’Brien, T.G. and Kuehner, A.C. (2007) ‘Waxing Grave About Adipocere: Soft Tissue Change in an Aquatic Context’, *Journal of Forensic Science*, 52(2), pp. 294–301. Available at: https://doi.org/10.1111/j.1556-4029.2006.00362.x.

Obertová, Z., Adalian, P., Baccino, E., Cunha, E., De Boer, H.H., Fracasso, T., Kranioti, E., Lefévre, P., Lynnerup, N., Petaros, A., Ross, A., Steyn, M. and Cattaneo, C. (2019) ‘The Status of Forensic Anthropology in Europe and South Africa: Results of the 2016 FASE Questionnaire on Forensic Anthropology’, *Journal of Forensic Sciences*, 64(4), pp. 1017–1025. Available at: https://doi.org/10.1111/1556-4029.14016.

Ohlrogge, S., Arent, T., Huckenbeck, W., Gabriel, P. and Ritz-Timme, S. (2009) *Anthropological Atlas of Female Facial Features*. 1st edn. Frankfurt, Germany: Verlag für Polizeiwissenschaft.

Ohlrogge, S., Nohrden, D., Schmitt, R., Drabik, A., Gabriel, P. and Ritz-Timme, S. (2008) *Anthropological Atlas of Male Facial Features*. 2nd edn. Frankfurt, Germany: Verlag für Polizeiwissenschaft.

Okuda, I., Yoshioka, N. and Akita, K. (2019) ‘Basic analysis of facial ageing: The relationship between the superficial musculoaponeurotic system and age’, *Experimental Dermatology*, 28(51), pp. 38–42. Available at: https://doi.org/10.1111/exd.13827.

Olivieri, L., Mazzarelli, D., Bertoglio, B., De Angelis, D., Previderè, C., Grignani, P., Cappella, A., Presciuttini, S., Bertuglia, C., Di Simone, P., Polizzi, N., Iadicicco, A., Piscitelli, V. and Cattaneo, C. (2018) ‘Challenges in the identification of dead migrants in the Mediterranean: The case study of the Lampedusa shipwreck of October 3rd 2013’, *Forensic Science International*, 285, pp. 121–128. Available at: https://doi.org/10.1016/j.forsciint.2018.01.029.

Oloyede, M.O., Hancke, G.P. and Myburgh, H.C. (2020) ‘A review on face recognition systems: recent approaches and challenges’, *Multimedia Tools and Applications*, 79, pp. 27891–27922. Available at: https://doi.org/10.1007/s11042-020-09261-2.

Oostra, R.-J., Gelderman, T., Groen, W.J.M., Uiterdijk, G., Cammeraat, E.L.H., Krap, T., Wilk, L.S., Lüschen, M., Morrien, E.E., Wobben, F., Duijst, W.L.J.M. and Aalders, M.C.G.

(2020) 'Amsterdam Research Initiative for Sub-surface Taphonomy and Anthropology (ARISTA) - A taphonomic research facility in the Netherlands for the study of human remains', *Forensic Science International*, 317, pp. 1–6. Available at: https://doi.org/10.1016/j.forsciint.2020.110483.

Ord, L. (2018) 'Facial depictions of WW1 German soldier excavated from a battlefield near the French Village of Bullecourt - featured on Digging for Britain: "East", Series 7, Episode 3'. United Kingdom: BBC Two. Available at: https://www.bbc.co.uk/programmes/m0001jg7.

OSAC (2021) *OSAC 2022-S-0007 Standard Guide for Facial Comparison Overview and Methodology Guidelines (DRAFT version 1.0)*. Available at: https://www.nist.gov/system/files/documents/2021/12/06/OSAC_2022-S-0007_Standard_Guide_for_Facial_Comparison_Overview_and_Methodology_Guidelines_DRAFT_OSAC_PROPOSED.pdf.

Oscar, G.R. (2020) *'Black box'. There's no way to determine how the algorithm came to your decision*, *Article*. Available at: https://towardsdatascience.com/black-box-theres-no-way-to-determine-how-the-algorithm-came-to-your-decision-19c9ee185a8 (Accessed: 3 January 2021).

Osman, A.M. and Viriri, S. (2018) 'Face verification across age progression: A survey of the state-of-the-art', in *Conference on Information Communications Technology and Society (ICTAS)*. Durban, South Africa: IEEE, pp. 1–6. Available at: https://doi.org/10.1109/ICTAS.2018.8368755.

Oxlee, G. (2007) 'Facial Recognition and Imagery Analysis', in T. Thompson and S. Black (eds) *Forensic Human Identification*. Boca Raton, London, New York: Taylor & Francis Group, LLC, pp. 257–270. Available at: https://doi.org/10.1016/B978-0-7506-8419-4.50019-X.

La Padula, S., Hersant, B., Bompy, L. and Meningaud, J.P. (2019) 'In search of a universal and objective method to assess facial aging: The new face objective photo-numerical assessment scale', *Journal of Cranio-Maxillo-Facial Surgery*, 47, pp. 1209–1215. Available at: https://doi.org/10.1016/j.jcms.2019.03.014.

Pala, P., Seidenari, L., Berretti, S. and Del Bimbo, A. (2019) 'Enhanced skeleton and face 3D data for person re-identification from depth cameras', *Computers & Graphics*, 79, pp. 69–80. Available at: https://doi.org/10.1016/j.cag.2019.01.003.

Palermo, R., Rossion, B., Rhodes, G., Laguesse, R., Tez, T., Hall, B., Albonico, A., Malaspina, M., Daini, R., Irons, J., Al-Janabi, S., Taylor, L.C., Rivolta, D. and McKone, E. (2017) 'Do People Have Insight into their Face Recognition Abilities?', *Quarterly Journal of Experimental Psychology*, 70(2), pp. 218–233. Available at: https://doi.org/10.1080/17470218.2016.1161058.

Panacea Cooperative Research (2021) *Skeleton-ID - Forensic Human Identification with AI*, *Website*. Available at: https://skeleton-id.com/ (Accessed: 8 September 2021).

Panis, G., Lanitis, A., Tsapatsoulis, N. and Cootes, T.F. (2016) 'Overview of research on facial ageing using the FG-NET ageing database', *IET Biometrics*, 5(2), pp. 37–46. Available at: https://doi.org/10.1049/iet-bmt.2014.0053.

Papesh, M.H. (2018) 'Photo ID verification remains challenging despite years of practice', *Cognitive Research: Principles and Implications*, 3(19), pp. 1–9. Available at: https://doi.org/10.1186/s41235-018-0110-y.

Parra, R.C., Anstett, E., Perich, P. and Buikstra, J.E. (2020) ‘Unidentified deceased persons: Social life, social death and humanitarian action’, in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons, Inc., pp. 79–99. Available at: https://doi.org/10.1002/9781119482062.ch6.

Pascalis, O. and de Schonen, S. (1994) ‘Recognition memory in 3- to 4-day-old human neonates’, *NeuroReport*, 5(14), pp. 1721–1724. Available at: https://doi.org/10.1097/00001756-199409080-00008.

Passalacqua, N.V., Pilloud, M.A. and Congram, D. (2021) ‘Forensic Anthropology as a Discipline’, *Biology*, 10(691), pp. 1–17. Available at: https://doi.org/10.3390/biology10080691.

Pecora, N.G., Baccetti, T. and McNamara Jr, A. (2008) ‘The aging craniofacial complex: a longitudinal cephalometric study from late adolescence to late adulthood’, *American Journal of Orthodontics and Dentofacial Orthopedics*, 134(4), pp. 496–505. Available at: https://doi.org/10.1016/j.ajodo.2006.11.022.

Pendolino, A.L. and Ottaviano, G. (2019) ‘Social Networks and the Rhinoplasty Patient’, *JAMA Facial Plastic Surgery*, 21(3), pp. 264–265. Available at: https://doi.org/10.1001/jamafacial.2018.1391.

Perper, J.A. (1980) ‘Time of death and changes after death: Chapter II Part 1 - Anatomical Considerations’, in W.U. Spitz and R.S. Fisher (eds) *Medicolegal Investigation of Death: Guidelines for the Application of Pathology to Crime Investigation*. 2nd edn. Springfield, IL: Charles C.Thomas, pp. 14–49.

Pessa, J.E. (2000) ‘An agorithm of facial aging: verification of Lambro’s theory by three-dimensional stereolithography, with reference to the pathogenesis of midfacial aging, scleral show, and the lateral suborbital trough deformity’, *Plastic and Reconstructive Surgery*, 106(2), pp. 179–488; discussion 489-490. Available at: https://doi.org/10.1097/00006534-200008000-00040.

Phillip, P.J., Grother, P., Michaelis, R.J., Blackburn, D.M., Tabassi, E. and Bone, M. (2003) *FRVT 2002: Overview and Summary*. Available at: https://www.nist.gov/itl/iad/image-group/face-recognition-vendor-test-frvt-documents.

Phillips, M. and Knoppers, B.M. (2019) ‘Whose Commons? Data Protection as a Legal Limit of Open Science’, *The Journal of Law, Medicine & Ethics*, 47(1), pp. 106–111. Available at: https://doi.org/10.1177/1073110519840489.

Phillips, P., Flynn, P., Scruggs, W., Bowyer, K., Chang, J., Hoffman, K., Marques, J., Min, J. and Worek, W. (2005) ‘Overview of the Face Recognition Grand Challenge’, in *IEEE Computer Society International Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE. Available at: https://doi.org/10.1109/CVPR.2005.268.

Phillips, P.J. (2001) ‘Support vector machines applied to face recognition’, *Advances in Neural Information Processing Systems*, 11(3).

Phillips, P.J., Yates, A.N., Hu, Y., Hahn, C.A., Noyes, E., Jackson, K., Cavazos, J.G., Jeckeln, G., Ranjan, R., Sankaranarayanan, S., Chen, J.C., Castillo, C.D., Chellappa, R., White, D. and

O'Toole, A.J. (2018) 'Face Recognition Accuracy of Forensic Examiners, Superrecognizers, and Face Recognition Algorithms', in T.D. Albright (ed.) *Proceedings of the National Academy of Sciences; June 2018, 115 (24)*, pp. 6171–6176. Available at: https://doi.org/10.1073/pnas.1721355115.

Pohran, A. (2009) *Just Plain Obvious - Picasa 3.5 face recognition tips*, *blog*. Available at: http://justplainobvious.blogspot.com/2009/10/picasa-35-face-recognition-tips.html (Accessed: 25 February 2021).

Pongakkasira, K. and Bindemann, M. (2015) 'The shape of the face template: Geometric distortions of faces and their detection in natural scenes', *Vision Research*, 109(Part A), pp. 99–106. Available at: https://doi.org/10.1016/j.visres.2015.02.008.

Potente, S., Ramsthaler, F., Kettner, M., Ikeda, T. and Schmidt, P. (2021) 'Application of the "bubbling" procedure to dead body portraits in forensic identification', *International Journal of Legal Medicine*, pp. 1–5. Available at: https://doi.org/10.1007/s00414-021-02515-0.

Powers, J.D., Malingen, S.A., Regnier, M. and Daniel, T.L. (2021) 'The Sliding Filament Theory Since Andrew Huxley: Multiscale and Multidisciplinary Muscle Research', *Annual Review of Biophysics*, 50, pp. 373–400. Available at: https://doi.org/10.1146/annurev-biophys-110320-062613.

Pressler, M.P., Kislevitz, M.L., Davis, J.J. and Amirlak, B. (2022) 'Size and Perception of Facial Features with Selfie Photographs, and Their Implication in Rhinoplasty and Facial Plastic Surgery', *Plastic and Reconstructive Surgery*, 149(4), pp. 859–867. Available at: https://doi.org/10.1097/PRS.0000000000008961.

Preuss, R., Renz-Polster, H., Reuter, H. and Wellhoehner, P. (2001) 'Leber, Galle, Pankreas', in *Basislehrbuch Innere Medizin*. 2nd edn. Munich: Urban & Fischer Verlag, p. 1220.

Prokop, O. and Göhler, W. (1975) *Forensische Medizin*. 3. überarb. Berlin: Volk und Gesundheit.

Rahman, M., Islam, R., Bhuiyan, N.I., Ahmed, B. and Islam, A. (2007) 'Person Identification Using Ear Biometrics', *International Journal of The Computer, the Internet and Management*, 15(2), pp. 1–8.

Rajanala, S., Maymone, M.B.C. and Vashi, N.A. (2018) 'Selfies—Living in the Era of Filtered Photographs', *JAMA Facial Plastic Surgery*, 20(6), pp. 443–444. Available at: https://doi.org/10.1001/jamafacial.2018.0486.

Ramage, J. (2022) *Facial recognition company Clearview AI permanently banned from selling data to private companies - EuroNews*, *Article*. Available at: https://www.euronews.com/next/2022/05/10/facial-recognition-company-clearview-ai-permanently-banned-from-selling-data-to-private-co (Accessed: 12 May 2022).

Ramanathan, N., Chellappa, R. and Roy Chowdhury, A.K. (2004) 'Facial similarity across age, disguise, illumination and pose', in *International Conference on Image Processing, ICIP '04*. Singapore: IEEE, pp. 1999–2002. Available at: https://doi.org/10.1109/ICIP.2004.1421474.

Ramírez Páez, D.E. (2020) 'Integration of information on missing persons and unidentified human remains: Best practices', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 159–169. Available at: https://doi.org/10.1002/9781119482062.ch11.

Ramon, M. (2021) 'Super-Recognizers – a novel diagnostic framework, 70 cases, and guidelines for future work', *Neuropsychologia*, 158, pp. 1–11. Available at: https://doi.org/10.1016/j.neuropsychologia.2021.107809.

Randolph-Quinney, P., Haines, S.D. and Kruger, A. (2018) 'The Use of Three-Dimensional Scanning and Surface Capture Methods in Recording Forensic Taphonomic Traces: Issues of Technology, Visualisation, and Validation', in P.M. Barone and M. Groen (eds) *Multidisciplinary Approaches to Forensic Archaeology. Soil Forensics*. Springer International Publishing, pp. 115–130. Available at: https://doi.org/doi.org/10.1007/978-3-319-94397-8_8.

Ranson, D. (2016) 'Legal Aspects of Identification', in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. New York: Routledge, pp. 642–659. Available at: https://doi.org/10.4324/9781315528939.

Rathgeb, C., Dogan, D., Stockhardt, F., De Marsico, M. and Busch, C. (2020) 'Plastic Surgery: An Obstacle for Deep Face Recognition?', in *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle, WA USA: IEEE, pp. 1–8. Available at: https://doi.org/10.1109/CVPRW50498.2020.00411.

Rattenbury, A.E. (2018) 'Forensic Taphonomy', in T.K. Ralebitso-Senior (ed.) *Forensic Ecogenomics - The Application of Microbial Ecology Analyses in Forensic Contexts*. Academic Press, Elsevier, pp. 37–59. Available at: https://doi.org/0.1016/B978-0-12-809360-3.00002-3.

Raviv, S. (2020) *The Secret History of Facial Recognition - Wired*, *Article*. Available at: https://www.wired.com/story/secret-history-facial-recognition/?fbclid=IwAR3ZBo0GncRn08eG5bR1JXY7KP-FSkgufsZV46ncB50VvqjEjPbqZH4iY9w (Accessed: 26 June 2020).

Reh, H., Haarhoff, K. and Vogt, C.D. (1977) 'Die Schätzung der Todeszeit bei Wasserleichen', *Zeitschrift für Rechtsmedizin*, 79(4), pp. 261–266. Available at: https://doi.org/10.1007/BF00201166.

Remnick, N. (2016) *Comombian Cocaine Kingpin Sentenced to 35 Years in Prison*, *The New York Times newspaper article*. Available at: https://www.nytimes.com/2016/07/26/nyregion/colombian-cocaine-kingpin-sentenced-to-35-years-in-prison.html (Accessed: 7 May 2020).

Rhine, S. (1990) 'Non-metric skull racing', in G.W. Gill and S. Rhine (eds) *Skeletal Attribution of Race: Methods for Forensic Anthropology*. Albuquerque, NM: Maxwell Museum of Anthropology, Anthropological Papers Number 4, pp. 9–20.

Ricanek Jr., K. and Tesafaye, T. (2006) 'MORPH: A Longitudinal Image Database of Normal Adult Age-Prgression', in *7th International Conference on Automatic Face and Gesture Recognition (FGR'06)*, pp. 1–5.

Richler, J.J., Cheung, O.S. and Gauthier, I. (2011) 'Holistic Processing Predicts Face Recognition', *Psychological Science*, 22(4), pp. 464–471. Available at: https://doi.org/10.1177/0956797611401753.

Richler, J.J. and Gauthier, I. (2014) 'A Meta-Analysis and Review of Holistic Face

Processing', *Psychological Bulletin*, 140(5), pp. 1281–1302. Available at: https://doi.org/10.1037/a0037004.

De Rigal, J., Des Mazis, I., Diridollou, S., Querleux, B., Yang, G., Leroy, F. and Holloway Barbosa, V. (2010) 'The effect of age on skin color and color heterogeneity in four ethnic groups', *Skin Research and Technology*, 16(2), pp. 168–178. Available at: https://doi.org/10.1111/j.1600-0846.2009.00416.x.

Ritchie, K.L. and Burton, A.M. (2017) 'Learning faces from variability', *Quarterly Journal of Experimental Psychology*, 70(5), pp. 897–905. Available at: https://doi.org/10.1080/17470218.2015.1136656.

Ritchie, K.L., Smith, F.G., Jenkins, R., Bindemann, M., White, D. and Burton, A.M. (2015) 'Viewers base estimates of face matching accuracy on their own familiarity: Explaining the photo-ID paradox', *Cognition*, 141, pp. 161–169. Available at: https://doi.org/10.1016/j.cognition.2015.05.002.

Ritchie, K.L., White, D., Kramer, R.S.S., Noyes, E., Jenkins, R. and Burton, A.M. (2018) 'Enhancing CCTV: Averages improve face identification from poor-quality images', *Applied Cognitive Psychology*, 32, pp. 671–680. Available at: https://doi.org/10.1002/acp.3449.

Ritz-Timme, S., Gabriel, P., Obertová, Z., Boguslawski, M., Mayer, F., Drabik, A., Poppa, P., De Angelis, D., Ciaffi, R., Zanotti, B., Gibelli, D. and Cattaneo, C. (2011) 'A new atlas for the evaluation of facial features: advantages, limits, and applicability', *International Journal of Legal Medicine*, 125, pp. 301–306.

Roberts, D. (2015) 'Modified People: Indicators of a Body Modification Subculture in a Post-Subculture World', *Sociology*, 49(6), pp. 1096–1112. Available at: https://doi.org/10.1177/0038038514554672.

Robertson, D., Middleton, R. and Burton, A.M. (2015) 'From policing to passport control: the limitations of photo ID', *Keesing: The Journal of Documents and Identity*, 46, pp. 3–8.

Robertson, J. (2008) 'The Role of an International Law Enforcement Agency in the Identification of Deceased Persons and Remains', in M. Oxenham (ed.) *Forensic Approaches to Death, Disaster and Abuse*. Australian Academic Press, Australia, pp. 263–280.

Rodriguez, W.C. (2006) 'Decomposition of Buried and Submerged Bodies', in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. Boca Raton, Florida: CRC Press LLC, pp. 459–467.

Roll, S. and Verinis, J.S. (1971) 'Stereotypes of Scalp and Facial Hair as Measured by the Semantic Differential', *Psychological Reports*, 28(3), pp. 975–980. Available at: https://doi.org/10.2466/pr0.1971.28.3.975.

Romero Moreno, F. (2022) 'The Conversation UK - Facial recognition technology: how it's being used in Ukraine and why it's still so controversial', *Article*. Available at: https://theconversation.com/facial-recognition-technology-how-its-being-used-in-ukraine-and-why-its-still-so-controversial-183171 (Accessed: 18 August 2022).

Rossion, B. (2018) 'Humans Are Visual Experts at Unfamiliar Face Recognition', *Trends in Cognitive Sciences*, 22(6), pp. 471–472. Available at: https://doi.org/10.1016/j.tics.2018.03.002.

Rossion, B. (2021) ‘What are superior face identity recognizers (SFIR) made of?’, *Neuropsychologia*, 158, pp. 1–2. Available at: https://doi.org/10.1016/j.neuropsychologia.2021.107807.

Rowe, D.J. and Guyuron, B. (2010) ‘Environmental and Genetic Factors in Facial Aging in Twins’, in M.A. Farage, K.W. Miller, and H.I. Maibach (eds) *Textbook of Aging Skin*. Berlin Heidelberg, Germany: Springer, pp. 441–446. Available at: https://doi.org/10.1007/978-3-540-89656-2_45.

Rule, N.O., Krendl, A.C., Ivcevic, Z. and Ambady, N. (2013) ‘Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates’, *Journal of Personality and Social Psychology*, 104(3), pp. 409–426. Available at: https://doi.org/10.1037/a0031050.

Russell, R., Batres, C., Courrèges, S., Kaminski, G., Soppelsa, F., Morizot, F. and Porcheron, A. (2019) ‘Differential effects of makeup on perceived age’, *British Journal of Psychology*, 110, pp. 87–100.

Rynn, C., Balueva, T. and Veselovskaya, E. (2012) ‘Relationships between the skull and face’, in C. Wilkinson and C. Rynn (eds) *Craniofacial Identification*. 1st edn. Cambridge, UK: Cambridge University Press, pp. 193–202. Available at: https://doi.org/10.1017/CBO9781139049566.016.

Sandford, A. and Ritchie, K.L. (2021) ‘Unfamiliar face matching, within-person variability, and multiple-image arrays’, *Visual Cognition*, 29(3), pp. 143–157. Available at: https://doi.org/10.1080/13506285.2021.1883170.

Sandulescu, T., Franzmann, M., Jast, J., Blaurock-Sandulescu, T., Spilker, L., Klein, C., Naumova, E.A. and Arnold, W.H. (2019) ‘Facial Fold and Crease Development: a New Morphological Approach and Classification’, *Clinical Anatomy*, 32, pp. 573–584.

Sandulescu, T., Stoltenberg, F., Buechner, H., Schmidt-Park, H., Linnerz, F., Jast, J., Franzmann, M., Blaurock-Sandulescu, T., Naumova, E.A. and Arnold, W.H. (2020) ‘Platysma and the cervical superficial musculoaponeurotic system - Comparative analysis of facial crease and platysmal band development’, *Annals of Anatomy - Anatomischer Anzeiger*, 227, pp. 1–10. Available at: https://doi.org/10.1016/j.aanat.2019.151414.

Santamaría, J., Cordon, O., Damas, S. and Ibáñez, O. (2009) ‘3D-2D image registration for craniofacial superimposition in forensic medicine using covariance matrix adaptation evolution strategy’, in *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine, ITAB 2009*. Larnaca, Cyprus, pp. 5–7. Available at: https://doi.org/10.1109/ITAB.2009.5394387.

Sauer, N.J., Michael, A.R. and Fenton, T.W. (2012) ‘Human Identification Using Skull-Photo Superimposition and Forensic Image Comparison’, in D. Dirkmaat (ed.) *A Companion to Forensic Anthropology*. 1st edn. John Wiley & Sons, Inc., pp. 432–446.

Sauer, N.J., Wankmiller, J.C. and Hefner, J.T. (2016) ‘Assessment of Ancestry and the Concept of Race’, in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. Routledge, pp. 243–260. Available at: https://www.routledge.com/Handbook-of-Forensic-Anthropology-and-Archaeology/Blau-Ubelaker/p/book/9781629583853.

Sauerwein, K., Saul, T.B., Steadman, D.W. and Boehnen, C.B. (2017) ‘The Effect of

Decomposition on the Efficacy of Biometrics for Positive Identification', *Journal of Forensic Sciences*, 62(6), pp. 1599–1602. Available at: https://doi.org/10.1111/1556-4029.13484.

Schmidlin, E.J., Steyn, M., Houlton, T.M.R. and Briers, N. (2018) 'Facial ageing in South African adult males', *Forensic Science International*, 289, pp. 277–286. Available at: https://doi.org/10.1016/j.forsciint.2018.05.006.

Schotsmans, E.M.J., Márquez-Grant, N. and Forbes, S.L. (2017) 'Introduction', in E.M.J. Schotsmans, N. Márqzez-Grant, and S.L. Forbes (eds) *Taphonomy of Human Remains: Forensic Analysis of the Dead and the Depositional Environment: Forensic Analysis of the Dead and the Depositional Environment*. 1st edn. John Wiley & Sons Ltd, pp. 1–8. Available at: https://doi.org/10.1002/9781118953358.

Schrödinger, E. (1947) 'The Foundation of the Theory of Probability: I', *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, 51(1945–1948), pp. 51–66. Available at: https://www.jstor.org/stable/20488471.

Schroff, F., Kalenichenko, D. and Philbin, J. (2015) 'FaceNet: A Unified Embedding for Face Recognition and Clustering', in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 815–823. Available at: https://doi.org/10.1109/CVPR.2015.7298682.

Schüler, G. and Obertová, Z. (2020) 'Visual identification of persons: Facial image comparison and morphological comparative analysis', in Z. Obertová, A. Stewart, and C. Cattaneo (eds) *Statistics and Probability in Forensic Anthropology*. Elsevier Inc., pp. 313–330. Available at: https://doi.org/10.1016/b978-0-12-815764-0.00025-3.

Schulz, I., Schneider, P.M., Olek, K., Rothschild, M.A. and Tsokos, M. (2006) 'Examination of postmortem animal interference to human remains using cross-species multiplex PCR', *Forensic Science, Medicine, and Pathology*, 2(2), pp. 95–101. Available at: https://doi.org/10.1385/FSMP:2:2:95.

Schweinberger, S.R., Baird, L.M., Blümler, M., Kaufmann, J.M. and Mohr, B. (2003) 'Interhemispheric cooperation for face recognition but not for affective facial expressions', *Neuropsychologia*, 41(4), pp. 407–414. Available at: https://doi.org/10.1016/s0028-3932(02)00173-2.

Schwidetzky, I. and Knussmann, R. (1988) 'Morphognose und Typognose', in R. Knussmann (ed.) *Anthropologie - Handbuch der vergleichenden Biologie des Menschen - Band 1*. 4th edn. Stuttgart, Germany: Gustav Fischer Verlag, pp. 359–368.

Schwikart, G. (2010) *Tod und Trauer in den Weltreligionen*. 2nd edn. Kevelaer, Germany: Verlagsgemeinschaft topos plus.

Scully, B. and Nambiar, P. (2002) 'Determining the validity of Furue's method of craniofacial superimposition for identification', *Annals of Dentistry University of Malaya*, 9(1), pp. 17–22. Available at: https://doi.org/10.22452/adum.vol9no1.4.

Sekharan, P.C. (1971) 'A Revised Superimposition Technique for Identification of the Individual from the Skull and Photograph Author (s): P . Chandra Sekharan Source : The Journal of Criminal Law , Criminology , and Police Science , Vol . 62 , N', *The Journal of Criminal Law, Criminology and Police Science*, 62(1), pp. 107–113. Available at: http://www.jstor.org/stable/1142133.

Seneviratne, S.I., Nicholls, N., Easterling, C.M., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C. and Zhang, X. (2012) 'Changes in climate extremes and their impacts on the natural physical environment', in C.B. Field, C. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds) *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation - A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge, UK: Cambridge University Press, pp. 109–230.

Senthilkumar, R. and Gnanamurthy, R.K. (2013) 'Face Databases for 2D and 3D Facial Recognition: A Survey', *IOSR Journal of Engineering*, 3(4), pp. 43–48. Available at: https://d1wqtxts1xzle7.cloudfront.net/31410295/F03414348.pdf?1371516107=&response-content-disposition=inline%3B+filename%3DIOSR_Journal_of_Engineering_IOSRJEN.pdf&Expires=1621988372&Signature=TitrqZQELxuMPnX4lJ5yXOxNlA7Cvcu7yK0cycY39CtqsXntbyQVCjzsDmYRDxz.

Sforza, C., Grandi, G., Binelli, M., Dolci, C., De Menezes, M. and Rerrario, V. (2010) 'Age- and sex-related changes in three-dimensional lip morphology', *Forensic Science International*, 200, pp. 182.e1-182.e7. Available at: https://doi.org/10.1016/j.forsciint.2010.04.050.

Sforza, C., Grandi, G., Binelli, M., Tommasi, D., Rosati, R. and Ferrario, V. (2009) 'Age- and sex-related changes in the normal human ear', *Forensic Science International*, 187, pp. 110.e1-110.e7. Available at: https://doi.org/:10.1016/j.forsciint.2009.02.019.

Sforza, C., Grandi, G., Catti, F., Tommasi, D., Ugolini, A. and Ferrario, V. (2009) 'Age- and sex-related changes in the soft tissues of the orbital region', *Forensic Science International*, 185, pp. 155.e1-155.e8. Available at: https://doi.org/10.1016/j.forsciint.2008.12.010.

Sforza, C., Grandi, G., De Menezes, M., Tartaglia, G. and Ferrario, V. (2010) 'Age- and sex-related changes in the normal human external nose', *Forensic Science International*, 204, pp. 205.e1-205.e.9. Available at: https://doi.org/10.1016/j.forsciint.2010.07.027.

Sforza, C., Ulaj, E., Gibelli, D.M., Allevi, F., Pucciarelli, V., Tarabbia, F., Ciprandi, D., Dell' Aversana Orabona, G., Dolci, C. and Biglioli, F. (2018) 'Three-dimensional superimposition for patients with facial palsy: an innovative method for assessing the success of facial reanimation procedures', *British Journal of Oral and Maxillofacial Surgery*, 56, pp. 3–7. Available at: https://doi.org/https://doi.org/10.1016/j.bjoms.2017.11.015.

Shapiro, P.N. and Penrod, S. (1986) 'Meta-analysis of facial identification studies', *Psychological Bulletin*, 100(2), pp. 139–156. Available at: https://doi.org/10.1037/0033-2909.100.2.139.

Sharma, P., Arora, A. and Valiathan, A. (2014) 'Age Changes of Jaws and Soft Tissue Profile', *The Scientific World Journal*, pp. 1–7. Available at: https://doi.org/10.1155/2014/301501.

Shaw, R.B. and Kahn, D.M. (2007) 'Aging of the midface bony elements: a three-dimensional somputer tomographic study', *Plastic and Reconstructive Surgery*, 119(2), pp. 675–681. Available at: https://doi.org/10.1097/01.prs.0000246596.79795.a8.

Sheehan, M.J. and Nachman, M.W. (2014) 'Morphological and population genomic evidence that human faces have evolved to signal individual identity', *Nature communications*, 5(4800), pp. 1–10. Available at: https://doi.org/10.1038/ncomms5800.

Simmons, T. and Cross, P.A. (2013) ‘Forensic Taphonomy’, *Encycolpedia of Forensic Sciences*, pp. 12–17. Available at: https://doi.org/10.1016/B978-0-12-382165-2.00004-0.

Simmons, T. and Heaton, V. (2013) ‘Postmortem Interval: Submerged Bodies’, in A. Jamieson (ed.) *Wiley Encyclopedia of Forensic Science*. John Wiley & Sons, Ltd., pp. 1–10. Available at: https://doi.org/10.1002/9780470061589.fsa1080.

Simpson, E.A., Maylott, S.E., Leonard, K., Lazo, R.J. and Jakobsen, K.V. (2019) ‘Face detection in infants and adults: Effects of orientation and color’, *Journal of Experimental Child Psychology*, 186, pp. 17–32. Available at: https://doi.org/10.1016/j.jecp.2019.05.001.

Sinclair, R., Jolley, D., Mallari, R. and Magee, J. (2004) ‘The reliability of horizontally sectioned scalp biopsies in the diagnosis of chronic diffuse telogen hair loss in women’, *Journal of the American Academy of Dermatology*, 51(2), pp. 189–199. Available at: https://doi.org/10.1016/S0190-9622(03)00045-8.

Singh, R., Vatsa, M., Bhatt, H.S., Bharadwaj, S., Noore, A. and Nooreyezdan, S.S. (2010) ‘Plastic Surgery: A New Dimension to Face Recognition’, *IEEE Transactions on Information Forensics and Security*, 5(3), pp. 441–448. Available at: https://doi.org/10.1109/TIFS.2010.2054083.

Sinha, P., Balas, B., Pstrovsky, Y. and Russell, R. (2006) ‘Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About’, *Proceedings of the IEEE*, 94(11), pp. 1948–1962. Available at: https://doi.org/10.1109/JPROC.2006.884093.

Sirovich, L. and Kirby, M. (1987) ‘Low-dimensional procedure for the characterization of human faces’, *The Journal of the Optical Society of America*, 4(3), pp. 519–524. Available at: https://doi.org/10.1364/josaa.4.000519.

Sledzik, P. and Micozzi, M. (1997) ‘Autopsied, Embalmed, and Preserved Human Remains: Distinguishing Features in Forensic and Historic Contexts’, *Geography* [Preprint].

Sledzik, P. and Mundorff, A.Z. (2016) ‘Forensic Anthropology in Disaster Response’, in S. Blau and D.H. Ubelaker (eds) *Handbook of Forensic Anthropology and Archaeology*. 2nd edn. New York: Routledge, pp. 477–495. Available at: https://doi.org/10.4324/9781315528939.

Sledzik, P.S. and Rodriguez, W.C. (2002) ‘Damnum Fatale: The Taphonomic Fate of Human Remains in Mass Disasters’, in W.D. Haglund and M.H. Sorg, (eds) *Advances in Forensic Taphonomy - Method, Theory, and Archaeological Perspectives*. Boca Raton, Florida: CRC Press LLC, pp. 321–330.

Smart, J.L. and Kaliszan, M. (2012) ‘The post mortem temperature plateau and its role in the estimation of time of death. A review’, *Legal Medicine*, 14(2), pp. 55–62. Available at: https://doi.org/10.1016/j.legalmed.2011.11.002.

Solomon, C. and Gibson, S. (2014) ‘Developments in Forensic Facial Composites’, *Advances in Forensic Human Identification*, pp. 235–270. Available at: https://doi.org/doi:10.1201/b16509-15.

Sood, A., Quintal, V. and Phau, I. (2017) ‘Keeping Up with the Kardashians: Consumers’ Intention to Engage in Cosmetic Surgery’, *Journal of Promotion Management*, 23(2), pp. 185–206. Available at: https://doi.org/10.1080/10496491.2016.1267677.

Soomer, H., Ranta, H. and Penttilä, A. (2001) 'Identification of victims from the M/S Estonia', *International Journal of Legal Medicine*, 11(4–5), pp. 259–262. Available at: https://doi.org/10.1007/s004140000180.

Sorg, M.H., Dearborn, J.H., Monahan, E.I., Ryan, H.E., Sweeney, K.G. and David, E. (2006) 'Forensic Taphomony in Marine Contexts', in W.D. Haglund and M.H. Sorg (eds) *Forensic Taphonomy - The Postmortem Fate of Human Remains*. Boca Raton, Florida: CRC Press, pp. 567–604.

de Souza, M.A., de Oliveira Urtiaga, G., Grangeiro Ferreira, R.C., da Silva, L.M., Umbelino, J.K.G., de Melo, F.R. and de Jesus, S. (2021) 'Friction ridge analysis in disaster victim identification (DVI): Brazilian case studies', *Forensic Science Research*, pp. 1–8. Available at: https://doi.org/10.1080/20961790.2021.1882745.

Spradley, K. and Gocha, T.P. (2020) 'Migrant deaths along the Texas/Mexico border: A collaborative approach to forensic identification of human remains', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 537–548.

Stacchi, L., Huguenin-Elie, E., Caldara, R. and Ramon, M. (2020) 'Normative data for two challenging tests of face matching under ecological conditions', *Cognitiv Research: Principles and Implications*, 5(8), pp. 1–17. Available at: https://doi.org/10.1186/s41235-019-0205-0.

Stark, L. and Hoffman, A.L. (2019) 'Data is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture', *Journal of Cultural Analytics*, May 2, pp. 1–22. Available at: https://doi.org/10.22148/16.036.

Starkie, A., Birch, W., Ferllini, R. and Thompson, T.J.U. (2011) 'Investigation into the merits of infrared imaging in the investigation of tattoos postmortem', *Journal of Forensic Sciences*, 56(6), pp. 1569–1573. Available at: https://doi.org/10.1111/j.1556-4029.2011.01869.x.

Stavrianos, C., Zouloumis, L., Papadopoulos, C., Emmanouil, J., Petalotis, N. and Tsakmalis, P. (2012) 'Facial Mapping: Review of Current Methods', *Research Journal of Medical Sciences*, 6(2), pp. 77–82. Available at: https://doi.org/10.3923/rjmsci.2012.77.82.

Stephan, C.N. (2003) 'Facial approximation: an evaluation of mouth-width determination', *American Journal of Physical Anthropology*, 121(1), pp. 48–57. Available at: https://doi.org/10.1002/ajpa.10166.

Stephan, C.N. (2015) 'Perspective distortion in craniofacial superimposition: Logarithmic decay curves mapped mathematcally and by practical experiment', *Forensic Science International*, 257, pp. 520.e1-520.e8. Available at: http://dx.doi.org/10.1016/j.forsciint.2015.09.009.

Stephan, C.N. (2017) 'Estimating the Skull-to-Camera Distance from Facial Photographs for Craniofacial Superimposition', *Journal of Forensic Sciences*, 62(4), pp. 850–860. Available at: https://doi.org/10.1111/1556-4029.13353.

Stephan, C.N., Caple, J.M., Guyomarc'h, P. and Claes, P. (2019) 'An overview of the latest developments in facial imaging', *Forensic Science Research*, 4(1), pp. 10–28. Available at: https://doi.org/10.1080/20961790.2018.1519892.

Stephan, C.N., Taylor, R.G.. and Taylor, J.A. (2008) 'Methods of Facial Approximation and

Skull-Face Superimposition, With Special Consideration of Method Development in Australia', in M. Oxenham (ed.) *Forensic Approaches to Death, Disaster and Abuse*. Queensland: Australian Academic Press, Australia, pp. 133–154.

Stevens, C. (2021) 'Person Identification at Airports During Passport Controls', in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 1–14.

Steyn, M., Pretorius, M., Briers, N., Bacci, N., Johnson, A. and Houlton, T.M.R. (2018) 'Forensic facial comparison in South Africa: State of the science', *Forensic Science International*, 287, pp. 190–194. Available at: https://doi.org/10.1016/j.forsciint.2018.04.006.

Stonham, T.J. (1986) 'Practical Face Recognition and Verification with Wisard', in H.D. Ellis, M.A. Jeeves, F. Newcombe, and A. Young (eds) *Aspects of Face Processing. NATO ASI Series (Series D: Behavioural and Social Sciences)*. vol. 28. Dordrecht: Springer, pp. 426–441. Available at: https://doi.org/10.1007/978-94-009-4420-6_44.

Strathie, A. and McNeill, A. (2016) 'Facial Wipes don't Wash: Facial Image Comparison by Video Superimposition Reduces the Accuracy of Face Matching Decisions', *Applied Cognitive Psychology*, 30, pp. 504–513. Available at: https://doi.org/10.1002/acp.3218.

Strathie, A., McNeill, A. and White, D. (2011) 'In the dock: chimeric image composites reduce identification accuracy', *Applied Cognitive Psychology*, 26(1), pp. 140–148. Available at: https://doi.org/10.1002/acp.1806.

Stuart, B. (2013) 'Decomposition Chemistry: Overview, Analysis, and Interpretation', in J.A. Siegel, P.J. Saukko, and M.M. Houck (eds) *Encycolpedia of Forensic Sciences*. 2nd edn. Elsevier Ltd, pp. 11–15. Available at: https://doi.org/10.1016/B978-0-12-382165--2.00120-3.

Sundal, J.A., Mæhle, B.O., Eide, G.E. and Morild, I. (2020) 'Petechial Hemorrhages in Suicide by Hanging: Possible Contributing Variables in Petechial Development', *The American Journal of Forensic Medicine and Pathology*, 41(2), pp. 90–96. Available at: https://doi.org/10.1097/PAF.0000000000000541.

Suri, S., Sankaran, A., Vasta, M. and Singh, R. (2019) 'On Matching Faces with Alterations due to Plastic Surgery and Disguise', in *IEEE 9th Internatinal Conference on Biometrics Theory, Applications and Systems*. Redondo Beach, CA, USA, pp. 1–7. Available at: https://doi.org/10.1109/BTAS.2018.8698571.

Surowiecki, J. (2004) *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations.* New York: Doubleday & Co.

Survey (2007) 'NIST test results unveiled', *Biometric Technology Today*, 15(4), pp. 10–11. Available at: https://doi.org/10.1016/S0969-4765(07)70102-9.

Tag That Photo (2021) *Tag That Photo - FAQs*, *Website*. Available at: https://tagthatphoto.com/faqs/ (Accessed: 17 February 2021).

Tagai, K., Ohtaka, H. and Nittono, H. (2016) 'Faces with Light Makeup Are Better Recognized than Faces with Heavy Makeup', *frontiers in Psychology*, 7(226), pp. 1–8. Available at: https://doi.org/10.3389/fpsyg.2016.00226.

Taigman, Y., Yang, M., Ranzato, M.A. and Wolf, L. (2014) ‘DeepFace: Closing the Gap to Human-Level Performance in Face Verification’, in *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH USA: IEEE, pp. 1701–1708. Available at: https://doi.org/10.1109/CVPR.2014.220.

Taister, M.A. and Holliday, S.D. (2000) ‘Comments on Facial Aging in Law Enforcement Investigations’, *Forensic Science Communications*, 2(2). Available at: https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/april2000/taister.htm#introduction.

Takács, B. (1998) ‘Comparing face images using the modified Hausdorff distance’, *Pattern Recognition*, 31(12), pp. 1873–1880. Available at: https://doi.org/10.1016/S0031-3203(98)00076-4.

Tamura, S., Kawai, H. and Mitsumoto, H. (1996) ‘Male/female identification from 8 × 6 very low resolution face images by neural network’, *Pattern Recognition*, 29(2), pp. 331–335. Available at: https://doi.org/10.1016/0031-3203(95)00073-9.

Tanaka, J.W. and Farah, M.J. (1993) ‘Parts and wholes in face recognition’, *The Quarterly Journal of Experimental Psychology*, 46(2), pp. 225–245. Available at: https://doi.org/10.1080/14640749308401045.

Tanaka, J.W. and Simonyi, D. (2016) ‘The “parts and wholes” of face recognition: a review of the literature’, *The Quarterly Journal of Experimental Psychology*, 69(10), pp. 1876–1889. Available at: https://doi.org/10.1080/17470218.2016.1146780.

Tarasova, E.Y. and Mintii, I.S. (2021) ‘Web application for facial wrinkle recognition’, in *CS&SE@SW 2021: 4th Workshop for Young Scientists in Computer Science & Software Engineering, December 18th 2021*. Kryvyi Rih, Ukraine: CEUR Workshop Proceedings, pp. 198–210. Available at: https://ceur-ws.org/Vol-3077/paper19.pdf.

Taroni, F., Garbolino, P., Biedermann, A., Aitken, C. and Bozza, S. (2018) ‘Reconciliation of subjective probabilities and frequencies in forensic science’, *Law, Probability & Risk*, 17(3), pp. 243–262. Available at: https://doi.org/10.1093/lpr/mgy014.

Taylor, J.A. and Brown, K.A. (1998) ‘Superimposition Techniques’, in J.G. Clement and D.L. Ranson (eds) *Craniofacial Identification in Forensic Medicine*. London, UK: Arnold, pp. 151–164.

Taylor, K.T. (2001) *Forensic Art and Illustration*. Boca Raton, Florida: CRC Press LLC.

Temple-Raston, D. and Powers, S. (2022) *At war with facial recognition: Clearview AI in Ukraine - The Record by Recorded Future*, *Article*. Available at: https://therecord.media/at-war-with-facial-recognition-clearview-ai-in-ukraine/ (Accessed: 16 June 2022).

Texas State University (2018) *Forensic Anthropology Research Facility : Forensic Anthropology Center : Texas State University*, *University Website*. Available at: https://www.txstate.edu/anthropology/facts/labs/farf.html (Accessed: 25 January 2019).

Thakur, S., Sing, J.K., Basu, D.K. and Nasipuri, M. (2009) ‘Face Recognition Using Fisher Linear Discriminant Analysis and Support Vector Machine’, in S. Ranka and et al. (eds) *Contemporary Computing. IC3 2009. Communications in Computer and Information Science; vol 40*. Berlin,Heidelberg: Springer, pp. 318–326.

Thomas, C.D.L. (1998) 'Quantification of facial shape and form', in J.G. Clement and D.L. Ranson (eds) *Craniofacial Identification in Forensic Medicine*. London: Arnold, pp. 165–176.

Tibbett, M. (2008) 'The Basics of Forensic Taphonomy: Understanding Cadaver Decomposition in Terrestrial Gravesites', in M. Oxenham (ed.) *Forensic Approaches to Death, Disaster and Abuse*. Australian Academic Press, Australia, pp. 29–36.

Tidball-Binz, M. (2020) 'Using forensic science to care for the dead and search for the missing: In conversation with Morris Tidball-Binz', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 3–23. Available at: https://doi.org/10.1002/9781119482062.ch1.

Tillotson, A.C. (2011) *Could craniofacial analysis have a role to play in disaster victim identification?* University of Dundee.

Tinsley, J.A. (1918) 'Physiognomy in Diagnosis', *Journal of the National Medical Association*, 10(2), pp. 74–75.

Tobin, D.J. (2017) 'Introduction to skin aging', *Juornal of Tissue Viability*, 26(1), pp. 37–46. Available at: https://doi.org/10.1016/j.jtv.2016.03.002.

Tokkari, N., Verdaasdonk, R.M., Liberton, N., Wolff, J., den Heijer, M., van der Veen, A. and Klaessens, J.H. (2017) 'Comparison and use of 3D scanners to improve the quantification of medical images (surface structures and volumes) during follow up of clinical (surgical) procedures', in *SPIE 10054, Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XV, 100540Z*. Available at: https://doi.org/10.1117/12.2253241.

Tolba, A.S., El-Baz, A.H. and El-Harby, A.A. (2006) 'Face Recognition: A Literature Review', *International Journal of Signal Processing*, 2(2), pp. 88–103.

Towler, A., Kemp, R.I., Burton, A.M., Dunn, J.D., Wayne, T., Moreton, R. and White, D. (2019) 'Do professional facial image comparison training courses work?', *PLos ONE*, 14(2), pp. 1–17. Available at: https://doi.org/10.1371/journal.pone.0211037.

Towler, A., Kemp, R.I. and White, D. (2021) 'Can Face Identification Ability Be Trained? - Evidence for Two Routes to Expertise', in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 89–114.

Towler, A., White, D. and Kemp, R.I. (2014) 'Evaluating training methods for facial image comparison: the face shape strategy does not work', *Perception*, 43(2–3), pp. 214–218. Available at: https://doi.org/10.1068/p7676.

Towler, A., White, D. and Kemp, R.I. (2017) 'Evaluating the feature comparison strategy for forensic face identification', *Journal of Experimental Psychology: Applied*, 23(1), pp. 47–58. Available at: https://doi.org/10.1037/xap0000108.

Trauring, M. (1963) 'Automatic comparison on finger ridge patterns', *Nature*, 197, pp. 938–940. Available at: https://doi.org/10.1038/197938a0.

Trojahn, C., Dobos, G., Lichterfeld, A., Blume-Peytavi, U. and Kottner, J. (2015) 'Characterizing facial skin ageing in humans: Disentangling extrinsic from intrinsic phenomena', *Biomedical Research International*, 2015, pp. 1–10. Available at:

https://doi.org/10.1155/2015/318586.

Trokielewicz, M., Czajka, A. and Maciejewicz, P. (2020) ‘Post-mortem Iris Decomposition and its Dynamics in Morgue Conditions’, *Journal of Forensic Sciences*, 65(5), pp. 1530–1538. Available at: https://doi.org/10.1111/1556-4029.14488.

Trüeb, R.M., Rezende, H.D. and Gavazzoni Dias, M.F.R. (2018) ‘A Comment on the Science of Hair Aging’, *International Journal of Trichology*, 10(6), pp. 245–254. Available at: https://doi.org/10.4103/ijt.ijt_56_18: 10.4103/ijt.ijt_56_18.

Tsokos, M. and Byard, R.W. (2016) ‘Postmortem Changes: Overview’, in J. Payne-James and R.W. Byard (eds) *Encyclopedia of Forensic and Legal Medicine*. 2nd edn. Elsevier, pp. 10–31. Available at: https://doi.org/10.1016/B978-0-12-800034-2.00312-8.

Tsokos, M., Lessig, R., Grundmann, C. and Peschel, O. (2006) ‘Experiences in tsunami victim identification’, *International Journal of Legal Medicine*, 120(3), pp. 185–187. Available at: https://doi.org/10.1007/s00414-005-0031-4.

Tsokos, M., Matschke, J., Gehl, A., Koops, E. and Püschel, K. (1999) ‘Skin and soft tissue artifacts due to postmortem damage caused by rodents’, *Forensic Science International*, 104(1), pp. 47–57. Available at: https://doi.org/10.1016/s0379-0738(99)00098-5.

Tsukahara, K., Moriwaki, S., Hotta, M., Fujimara, T. and Kitahara, T. (2004) ‘A study of diurnal variation in wrinkles on the human face’, *Archives of Dermatological Research*, 296(4), pp. 169–174. Available at: https://doi.org/10.1007/s00403-004-0500-5.

Turk, M.A. and Pentland, A.P. (1991) ‘Face recognition using eigenfaces’, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Maui, HI USA: IEEE, pp. 586–591. Available at: https://doi.org/10.1109/CVPR.1991.139758.

Ubelaker, D.H. (2018) ‘A history of forensic anthropology’, *American Journal of Physical Anthropology*, 165(4), pp. 915–923. Available at: https://doi.org/10.1002/ajpa.23306.

Ubelaker, D.H., Wu, Y. and Cordero, Q.R. (2019) ‘Craniofacial photographic superimposition: New developments’, *Forensic Science International: Synergy*, 1, pp. 271–274. Available at: https://doi.org/https://doi.org/10.1016/j.fsisyn.2019.10.002.

Ueda, S. and Koyama, T. (2010) ‘Influence of make-up on facial recognition’, *Perception*, 39(2), pp. 260–264. Available at: https://doi.org/10.1068/p6634.

Ulguim, P.F. (2017) *Recording In Situ Human Remains in Three Dimensions: Applying Digital Image-Based Modeling*, *Human Remains: Another Dimension The Application of Imaging to the Study of Human Remains*. Elsevier Inc. Available at: https://doi.org/10.1016/B978-0-12-804602-9.00007-2.

Ulrich, V.S. (2017) ‘Sicherheit biometrischer Systeme -Gesichtserkennung’, *Seminararbeit an der Hochschule Bonn-Rhein-Sieg*, pp. 1–10. Available at: https://www.h-brs.de/files/20171215_fbinf_mclab_17_01_08_ulrich_mk.pdf.

UN (2017) *Unlawful death of refugees and migrants - A/72/335*. Available at: https://drive.google.com/file/d/0ByFv9rzlqJaBdklJTEY3ZjZ0cEU/view?resourcekey=0-Qsgq8QxzTde5pqvZvWJQcw.

UN (2021) *Climate and weather related disasters surge five-fold over 50 years, but early warnings save lives - WMO report*, *Article*. Available at:

https://news.un.org/en/story/2021/09/1098662 (Accessed: 19 December 2021).

University of Dundee (2019) *Selfies could help spot missing persons a smile away*, *Press Release*. Available at: https://www.dundee.ac.uk/stories/selfies-could-help-spot-missing-persons-smile-away (Accessed: 2 March 2020).

University of Dundee Press Office (2010) *Dundee forensic expert helps produce 'true likeness' of Shakespeare*, *Article*. Available at: http://app.dundee.ac.uk/pressreleases/2010/prsept10/shakespearemask.htm (Accessed: 27 October 2017).

Upadhyay, R.B., Upadhyay, J., Adrawal, P. and Rao, N.O. (2012) 'Analysis of gonial angle in relation to age, gender, and dentition status by radiological and anthropometric methods', *Journal of Forensic Dental Sciences*, 4(1), pp. 29–33.

Urbanová, P., Eliášová, H. and Dostálová, T. (2022) 'How accurate is forensic facial identification of surgically altered faces?', *Australian Journal of Forensic Sciences*, pp. 1–20. Available at: https://doi.org/10.1080/00450618.2022.2032341.

Urbanová, P., Hejna, P. and Jurda, M. (2015) 'Testing photogrammetry-based techniques for three-dimensional surface documentation in forensic pathology', *Forensic Science International*, 250, pp. 77–86. Available at: https://doi.org/10.1016/j.forsciint.2015.03.005.

Utsuno, H., Kanoh, T., Tadokoro, O. and Inoue, K. (2005) 'Preliminary study of post mortem identification using lip prints', *Forensic Science International*, 149(2–3), pp. 129–132. Available at: https://doi.org/10.1016/j.forsciint.2004.05.013.

Valsecchi, A., Damas, S. and Cordon, O. (2018) 'A Robust and Efficient Method for Skull-Face Overlay in Computerized Craniofacial Superimposition', *IEEE Transactions on Information Forensics and Security*, 13(8), pp. 1960–1974.

Vanezis, P. and Brierley, C. (1996) 'Facial image comparison of crime suspects using video superimposition', *Science & Justice*, 36(1), pp. 27–33. Available at: https://doi.org/10.1016/S1355-0306(96)72551-0.

Vanezis, P., Lu, D., Cockburn, J., Gonzalez, A., McCombe, G., Trujillo, O. and Vanezis, M. (1996) 'Morphological Classification of Facial Features in Adult Caucasian Males Based On an Assessment of Photographs of 50 Subjects', *Journal of Forensic Sciences*, 41(5), pp. 786–791. Available at: https://doi.org/10.1520/JFS13998J.

Vary Jr., J.V. (2015) 'Selected Disorders of Skin Appendages - Acne, Alopecia, Hyperhidrosis', *Medical Clinics of North America*, 99(6), pp. 1195–1211. Available at: https://doi.org/10.1016/j.mcna.2015.07.003.

Vashi, N.A., Maymone, M.B.C. and Kundu, R.V. (2016) 'Aging Differences in Ethnic Skin', *The Journal of Clinical and Aesthetic Dermatology*, 9(11), pp. 31–38.

Vega Dulanto, M.C. (2020) 'Between darts and bullets: A bioarcheological view on the study of human rights and IHL violations', in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 49–65. Available at: https://doi.org/10.1002/9781119482062.ch4.

Venkatesh, S., Maymone, M.B.C. and Vashi, N.A. (2019) 'Aging in skin of color', *Clinics in*

Dermatology, 37, pp. 351–357. Available at: https://doi.org/10.1016/j.clindermatol.2019.04.010.

Verhelst, H.M., Stannat, A.W. and Mecacci, G. (2020) ‘Machine Learning Against Terrorism: How Big Data Collection and Analysis Influences the Privacy-Security Dilemma’, *Science and Engineering Ethids*, 26, pp. 2975–2984. Available at: https://doi.org/10.1007/s11948-020-00254-w.

Verosky, S.C., Porter, J., Martinez, J.E. and Todorov, A. (2018) ‘Robust Effects of Affective Person Learning on Evaluation of Faces’, *Journal of Personality and Social Psychology*, 114(4), pp. 516–528. Available at: https://doi.org/10.1037/pspa0000109.

Vetter, T. and Poggio, T. (1997) ‘Linear object classes and image synthesis from a single example image’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), pp. 733–742. Available at: https://doi.org/10.1109/34.598230.

Vidak, S. and Foisner, R. (2016) ‘Molecular insights into the premature aging disease progeria’, *Histochemistry and Cell Biology*, 145, pp. 401–417. Available at: https://doi.org/10.1007/s00418-016-1411-1.

Vince, J. (2020) *6 effective skincare alternatives to booking a treatment*, *Article*. Available at: https://www.harpersbazaar.com/uk/beauty/skincare/a33356891/skincare-revitalift-laser-renew/ (Accessed: 3 August 2020).

Viola, P., Jones, M. and Snow, D. (2003) ‘Detecting pedestrians using patterns of motion and appearance’, in *The International Conference on Computer Vision ICCV*. Nice, France, pp. 1–10. Available at: https://doi.org/10.1109/ICCV.2003.1238422.

W.R., T. (1985) ‘The Mengele report’, *The American Journal of Forensic Medicine and Pathology*, 6(4), pp. 279–283.

Wang, M. and Deng, W. (2021) ‘Deep face recognition: A survey’, *Neurocomputing*, 429, pp. 215–244. Available at: https://doi.org/10.1016/j.neucom.2020.10.081.

Wang, Y. and Olson, I.R. (2018) ‘The Original Social Network: White Matter and Social Cognition’, *Trends in Cognitive Sciences*, 22(6), pp. 504–516. Available at: https://doi.org/10.1016/j.tics.2018.03.005.

Ward, B., Ward, M., Fried, O. and Paskhover, B. (2018) ‘Nasal Distortion in Short-Distance Photographs: The Selfie Effect’, *JAMA Facial Plastic Surgery*, 20(4), pp. 333–335.

Webster, G. (1955) ‘Photography as an Aid in Identification’, *Police Journal*, 185, pp. 185–192.

Webster, J.W. and Bemis, G. (1850) *Report of the Case of John W. Webster, indicted for the murder of George Parkman before the Supreme Judicial Court of Massachusetts*. Boston: Charles C. Little and James Brown. Available at: https://archive.org/details/reportcasejohnw00courgoog/page/n11/mode/2up?q=bemis+webster.

Weiler, S.M., Tetzlaff, B.-O., Herzberg, P.Y. and Jacobsen, T. (2021) ‘When personality gets under the skin: Need for uniqueness and body modifications’, *PLoS ONE*, 16(3), pp. 1–13. Available at: https://doi.org/10.1371/journal.pone.0245158.

Welcker, H. (1883) *Schiller’s Schädel und Todtenmaske, nebst Mittheilungen über Schädel* *und Todtenmaske Kant's*. Braunschweig: Friedrich Vieweg und Sohn.

Werner, G. (2018) 'The Kardashian Effect - Radio 1 Stories'. United Kingdom: BBC. Available at: https://www.bbc.co.uk/programmes/p05y7zsr.

White, D., Burton, A.M., Jenkins, R. and Kemp, R.I. (2014) 'Redesigning photo-ID to improve unfamiliar face matching performance', *Journal of Experimental Psychology: Applied*, 20(2), pp. 166–173. Available at: https://doi.org/10.1037/xap0000009.

White, D., Burton, A.M., Kemp, R. and Jenkins, R. (2013) 'Crowd effects in unfamiliar face matching', *Applied Cognitive Psychology*, 27, pp. 769–777. Available at: https://doi.org/10.1002/acp.2971.

White, D., Burton, A.M. and Kemp, R.I. (2014) 'Redesigning photo-ID to improve unfamiliar face matching performance', *Journal of Experimental Psychology: Applied*, 20(2), pp. 166–173. Available at: https://doi.org/10.1037/xap0000009.

White, D., Kemp, R.I., Jenkins, R. and Burton, A.M. (2014) 'Feedback training for facial image comparison', *Psychonomic Bulletin & Review*, 21(1), pp. 100–106. Available at: https://doi.org/10.3758/s13423-013-0475-3.

White, D., Kemp, R.I., Jenkins, R., Matheson, M. and Burton, A.M. (2014) 'Passport Officers' Errors in Face Matching', *PLoS ONE*, 9(8), pp. 1–6. Available at: https://doi.org/10.1371/journal.pone.0103510.

White, D., Norell, K., Philips, P.J. and O'Toole, A.J. (2017) 'Human Factors in Forensic Face Identification', in M. Tistarelli and C. Champod (eds) *Handbook of Biometrics for Forensic Science. Advances in Computer Vision and Pattern Recognition*. Springer, Cham, pp. 195–219. Available at: https://doi.org/10.1007/978-3-319-50673-9_9.

White, D., Phillips, J., Hahn, C.A., Hill, M. and O'Toole, A.J. (2015) 'Perceptual expertise in forensic facial image comparison', *Royal Society - Proceedings Biological Sciences*, 282(1814), pp. 1–6. Available at: https://doi.org/10.1098/rspb.2015.1292.

White, D., Towler, A. and Kemp, R.I. (2021) 'Understanding Professional Expertise in Unfamiliar Face Matching', in M. Bindemann (ed.) *Forensic Face Matching - Research and Practice*. 1st edn. New York: Oxford University Press, pp. 62–88.

Whiting, D.A. (1998) 'Male pattern hair loss: current understanding', *International Journal of Dermatology*, 37, pp. 561–566.

Whittaker, D.K., Richards, B.H. and Jones, M.L. (1998) 'Orthodontic Reconstruction in a Victim of Murder', *British Journal of Orthodontics*, 25, pp. 11–14.

Wiese, H., Komes, J. and Schweinberger, S.R. (2013) 'Ageing faces in ageing minds: A review on the own-age bias in face recognition', *Visual Cognition*, 21(9–10), pp. 1337–1363. Available at: https://doi.org/10.1080/13506285.2013.823139.

Wilber, M.J., Shmatikov, V. and Belongie, S. (2016) 'Can we still avoid automatic face detection?', in *IEEE Winter Conference on Applications of Computer Vision (WACV)*. Lake Placid, NY, USA: IEEE, pp. 1–9. Available at: https://doi.org/10.1109/WACV.2016.7477452.

Wilkinson, C. (2008) *Forensic Facial Reconstruction*. Cambridge, UK: Cambridge University Press.

Wilkinson, C. (2014) 'A review of the changing culture and social context relating to forensic facial depiction of the dead', *Forensic Science International*, 245, pp. 95–100. Available at: https://doi.org/10.1016/j.forsciint.2014.10.007.

Wilkinson, C. (2015) 'Craniofacial Analysis and Identification', in T. Valentine and J.P. Davis (eds) *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV*. Wiley Blackwell, pp. 93–126.

Wilkinson, C. and Castaneyra-Ruiz, M. (2021) 'The current status of Migrant Disaster Victim Identification in the Canary Islands', *Journal of the British Academy*, 9(s8), pp. 115–135. Available at: https://doi.org/10.5871/jba/009s8.115.

Wilkinson, C. and Evans, R. (2009) 'Are facial image analysis experts any better than the general public at identifying individuals from CCTV images?', *Science & Justice*, 49(3), pp. 191–196. Available at: https://doi.org/10.1016/j.scijus.2008.10.011.

Wilkinson, C. and Lofthouse, A. (2015) 'The use of craniofacial superimposition for disaster victim identification', *Forensic Science International*, 252, pp. 187.e1-187.e6. Available at: https://doi.org/10.1016/j.forsciint.2015.03.023.

Wilkinson, C. and Tillotson, A. (2012) 'Post-mortem prediction of facial appearance', in C. Wilkinson and C. Rynn (eds) *Craniofacial Identification*. 1st edn. Cambridge: Cambridge University Press, pp. 166–183. Available at: https://doi.org/10.1016/B978-0-7506-8419-4.50019-X.

Wilkinson, C.M. (2004) *Forensic Facial Reconstruction*. Cambridge, UK: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781107340961.

Williams, A., Rogers, C.J. and Cassella, J.P. (2019) 'Why does the UK need a Human Taphonomy Facility?', *Forensic Science International*, 296, pp. 74–79. Available at: https://doi.org/10.1016/j.forsciint.2019.01.010.

Willsher, K. (2011) *'Carlos the Jackal' goes on trial in France*, *Los Angeles Times newspaper article*. Available at: https://www.latimes.com/world/la-xpm-2011-nov-07-la-fg-france-jackal-trial-20111108-story.html (Accessed: 9 May 2020).

Wilson, C. (2010) *Vein pattern recognition: A privacy-enhancing biometric*. 1st edn. Boca Raton, Florida: CRC Press. Available at: https://doi.org/10.1201/9781439821381.

Wirth, B. and Carbon, C. (2017) 'An Easy Game for Frauds? Effects of Professional Experience and Time Pressure on Passport-Matching Performance', *Journal of Experimental Psychology: Applied*, 23(2), pp. 138–157. Available at: https://doi.org/10.1037/xap0000114.

Wiskott, L., Krüger, N., Kuiger, N. and von der Malsburg, C. (1997) 'Face recognition by elastic bunch graph matching', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), pp. 775–779. Available at: https://doi.org/10.1109/34.598235.

Wogalter, M.S. and Hosie, J.A. (1991) 'Effects of cranial and facial hair on perceptions of age and person', *The Journal of Social Psychology*, 131(4), pp. 589–591. Available at: https://doi.org/10.1080/00224545.1991.9713892.

Wong, R., Geyer, S., Weninger, W., Guimberteau, J.-C. and Wong, J.K. (2016) 'The dynamic anatomy and pattering of skin', *Experimental Dermatology*, 25, pp. 92–98. Available at: https://doi.org/10.1111/exd.12832.

Woodhead, M.M., Baddeley, A.D. and Simmonds, D.C. (1979) 'On training people to recognize faces', *Ergonomics*, 22(3), pp. 333–343. Available at: https://doi.org/10.1080/00140137908924617.

World Population Review (2022) *Texas Population 2022*, *Website*. Available at: https://worldpopulationreview.com/states/texas-population (Accessed: 12 August 2022).

Wright, D.B. and Sladden, B. (2003) 'An own gender bias and the importance of hair in face recognition', *Acta Psychologica*, 114(1), pp. 101–114. Available at: https://doi.org/10.1016/s0001-6918(03)00052-0.

Wright, J. and Ma, Y. (2010) 'Dense Error Correction Via ℓ1-Minimization', *IEEE Transactions on Information Theory*, 56(7), pp. 3540–3560. Available at: https://doi.org/10.1109/TIT.2010.2048473.

Wright, J., Yang, A.Y., Ganesh, A., Shankar Sastry, S. and Ma, Y. (2008) 'Robust Face Recognition via Sparse Representation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2), pp. 210–227. Available at: https://doi.org/10.1109/TPAMI.2008.79.

Wright, K., Mundorff, A., Chaseling, J., Forrest, A., Maguire, C. and Crane, D.I. (2015) 'A new disaster victim identification management strategy targeting "near identification-threshold" cases: Experiences from the Boxing Day tsunami', *Forensic Science International*, 250(January 2005), pp. 91–97. Available at: https://doi.org/10.1016/j.forsciint.2015.03.007.

Yang, M., Zhang, L., Xang, J. and Zhang, D. (2011) 'Robust sparse coding for face recognition', in *CVPR*. Colorado Springs, CO USA: IEEE, pp. 625–632. Available at: https://doi.org/10.1109/CVPR.2011.5995393.

Yoshino, M. (2012) 'Craniofacial superimposition', in C. Wilkinson and C. Rynn (eds) *Craniofacial Identification*. Cambridge University Press, pp. 238–253.

Yoshino, M., Imaizumi, K., Miyasaka, S. and Seta, S. (1995) 'Evaluation of anatomical consistency in craniofacial superimposition images', *Forensic Science International*, 74(1–2), pp. 125–134. Available at: https://doi.org/10.1016/0379-0738(95)01742-2.

Yoshino, M., Matsuda, H., Kubota, S., Imaizumi, K. and Miyasaka, S. (2000) 'Computer-assisted facial image identification system using a 3-D physiognomic range finder', *Forensic Science International*, 109(3), pp. 225–237. Available at: https://doi.org/10.1016/S0379-0738(00)00149-3.

Young, A.W. and Burton, A.M. (2017) 'Recognizing Faces', *Current Directions in Psychological Science*, 26(3), pp. 212–217. Available at: https://doi.org/10.1177/0963721416688114.

Young, A.W. and Burton, A.M. (2018) 'Are We Face Experts?', *Trends in Cognitive Sciences*, 22(2), pp. 100–110. Available at: https://doi.org/10.1016/j.tics.2017.11.007.

Young, A.W., Hay, D.C., McWeeny, K.H., Flude, B.M. and Ellis, A.W. (1985) 'Matching familiar and unfamiliar faces on internal and external features', *Perception*, 14(6), pp. 737–746. Available at: https://doi.org/10.1068/p140737.

Young, A.W., Hellawell, D. and Hay, D.C. (1987) 'Configurational information in face perception', *Perception*, 16(6), pp. 747–759. Available at:

https://doi.org/10.1068/p160747.

Zankl, A., Eberle, L., Molinari, L. and Schinzel, A. (2002) ‘Growth Charts for Nose Length, Nasal Protrusion, and Philtrum Length From Birth to 97 Years’, *American Journal of Medical Genetics*, 111, pp. 388–391. Available at: https://doi.org/10.1002/ajmg.10472.

Zejdlik, K., Passalacqua, N.V. and Williams, J.A. (2018) *A Force for Good: The what, where, and why of human decomposition facilities*, *Article*. Available at: https://thepathologist.com/outside-the-lab/a-force-for-good (Accessed: 20 February 2019).

Zewail, R., Elsafi, A., Saeb, M. and Hamdy, N. (2004) ‘Soft and hard biometrics fusion for improved identity verification’, in *47th IEEE International Midwest Symposium on Circuits and Systems*. Hiroshima, Japan: IEEE, pp. 225–228. Available at: https://doi.org/10.1109/MWSCAS.2004.1353967.

Zhang, X. and Gao, Y. (2009) ‘Face recognition across pose: A review’, *Pattern Recognition*, 42, pp. 2876–2896. Available at: https://doi.org/10.1016/j.patcog.2009.04.017.

Zhao, J., Han, J. and Shao, L. (2018) ‘Unconstrained Face Recognition Using a Set-to-Set Distance Measure on Deep Learned Features’, *IEEE Transactions on Circuits and Systems for Video Technology*, 28(10), pp. 2679–2689. Available at: https://doi.org/10.1109/TCSVT.2017.2710120.

Zhao, W., Yang, C., Ye, J., Yan, Y., Yang, X. and Huang, K. (2022) ‘From 2D images to 3D model: Weakly supervised multi-view face reconstruction with deep fusion’, *Computer Vision and Pattern Recognition*, pp. 1–11. Available at: https://doi.org/10.48550/arXiv.2204.03842.

Zheng, X., Mondloch, C.J. and Segalowitz, S.J. (2012) ‘The timing of individual face recognition in the brain’, *Neuropsychologia*, 50(7), pp. 1451–1461. Available at: https://doi.org/10.1016/j.neuropsychologia.2012.02.030.

Zorba, G.K., Eleftheriou, T., Engin, I., Hartsioti, S. and Zenonos, C. (2020) ‘Forensic identification of human remains in Cyprus: The humanitarian work of the Committee on Missing Persons in Cyprus (CMP)’, in R.C. Parra, S.C. Zapico, and D.H. Ubelaker (eds) *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. 1st edn. John Wiley & Sons Ltd, pp. 609–623.

Zuo, K.J., Saun, T.J. and Forrest, C.R. (2019) ‘Facial Recognition Technology: A Primer for Plastic Surgeons’, *Plastic and Reconstructive Surgery*, 143(6), pp. 1298e-1306e. Available at: https://doi.org/10.1097/PRS.0000000000005673.

VIII. Appendices

Appendix A – *Pilot Study*

Dataset images, full method, and results of pilot study.

Appendix B – *Automated Comparison – Ancillary Studies*

Ancillary studies for automated and semi- automated comparison using MATLAB® and Google Picasa 3, with data from the pilot study and two additional living individuals.

Appendix C –*Manual Comparison – Notes*

Detailed notes for one-to-one comparisons from the hierarchical, combined-methods approach of the manual comparison sections.

Appendix D – *Manual Comparison – Superimposition Workflow and Morphological Feature List*

Step-by-step workflow and explanations for data analysis of main study in Geomagic® Freeform® Modelling Plus and full feature list template for morphological comparison.

Appendix E – *Automated Comparison – MATLAB® Scripts*

MATLAB® scripts for re-sizing and augmenting images, as well as an adapted code for the automated simple face recognition task as used in chapter IV, Part 2, section 2A.

Appendix F – *Ethics*

Ethics application to LJMU for pilot and main studies, participant information sheets and consent forms, living donor and next of kin donation forms for TXST.

Appendix G – *Dataset images*

Dataset images for pilot and main studies; 2D photographs for AM face pools and 2D PM subjects, and 2D screenshots of pPM and PM 3D face models.

Appendix H – *Supplemental Material – Guide to Digital files*

A brief guide explaining the order and context of all digital files submitted in conjunction with this thesis.

Appendix I – *Automated Comparison - Picasa Workflow*

Overview of the Google Picasa 3 workflow as used for the semi-automated comparison, with additional notes further explaining the approach.

Appendix J – *Image Copyrights and Permissions*

Disclaimers and correspondence related to copyright and use of images, figures, and tables in this thesis.