Intelligent strategies for sheep monitoring and management

Natasa Kleanthous

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the Degree of Doctor of Philosophy

July 2021

Dedication

I would like to dedicate this thesis to the memory of my loving parents …

Acknowledgements

I am profoundly grateful to my director of studies, Prof. Abir Hussain, for granting me the opportunity to start my research career at Liverpool John Moores University by offering me a PhD scholarship. I appreciate her support, encouragement, and assistance through the PhD. All of her support and advice helped me to face the challenges during the research with success. I consider myself very fortunate to be one of her students. Having many unfortunate events in my life during my PhD, I wouldn't be able to continue without her support and motivation. *Prof. Abir - Thank you not only for the academic guidance but, more importantly, for your moral support in every single meltdown I experienced over these years!*

I am also very grateful, and I would like to express my gratitude to my second supervisor, Dr. Jennifer Sneddon, for her help and support. During the PhD, a huge challenge was collecting gait signals from sheep, and Dr. Jenny Sneddon offered me the opportunity to visit her farm and manage to collect all the data needed. *Dr. Jenny - Thank you for all your help during the experiments and your precious support.*

Also, my special thank goes to Prof. Panos Liatsis for his precious time reviewing my research work that ended up in journal papers waiting for acceptance. Prof. Panos Liatsis encouraged me to continue and motivate me when I lost track. All his help and support were more than enough to encourage me to complete this work. *Prof. Pano - Thank you for all your motivation and precious feedback during the whole time. Your humour helped a lot as well!*

My appreciation also goes to Tricia Waterson for all the help during some personal issues I faced through the PhD. *Tricia – I really can't help you enough for your assistance!*

I am also deeply obligated to Dr. Wasiq Khan for his time reviewing and giving feedback on the conference papers and journal preparations extracted from this thesis.

Additionally, I would like to acknowledge and thank the staff and technicians of the Faculty of Engineering and Technology, Liverpool John Mores University, for assisting when needed and excellent computing facilities.

Abstract

With the growth in world population, there is an increasing demand for food resources and better land utilisation, e.g., domesticated animals and land management, which in turn brought about developments in intelligent farming. Modern farms rely upon intelligent sensors and advanced software solutions, to optimally manage pasture and support animal welfare. A very significant aspect in domesticated animal farms is monitoring and understanding of animal activity, which provides vital insight into animal well-being and the environment they live in. Moreover, "virtual" fencing systems provide an alternative to managing farmland by replacing traditional boundaries.

This thesis proposes novel solutions to animal activity recognition based on accelerometer data using machine learning strategies, and supports the development of virtual fencing systems via animal behaviour management using audio stimuli. The first contribution of this work is four datasets comprising accelerometer gait signals. The first dataset consisted of accelerometer and gyroscope measurements, which were obtained using a Samsung smartphone on seven animals. Next, a dataset of accelerometer measurements was collected using the MetamotionR device on 8 Hebridean ewes. Finally, two datasets of nine Hebridean ewes were collected from two sensors (MetamotionR and Raspberry Pi) comprising of accelerometer signals describing active, inactive and grazing activity of the animal. These datasets will be made publicly available as there is limited availability of such datasets. In respect to activity recognition, a systematic study of the experimental setup, associated signal features and machine learning methods was performed. It was found that Random Forest using accelerometer measurements and a sample rate of 12.5Hz with a sliding window of 5 seconds provides an accuracy of above 96% when discriminating animal activity. The problem of sensor heterogeneity was addressed with transfer learning of Convolutional Neural Networks, which has been used for the first time in this problem, and resulted to an accuracy of 98.55%, and 96.59%, respectively, in the two experimental datasets. Next, the feasibility of using only audio stimuli in the context of a virtual fencing system was explored.

Specifically, a systematic evaluation of the parameters of audio stimuli, e.g., frequency and duration, was performed on two sheep breeds, Hebridean and Greyface Dartmoor ewes, in the context of controlling animal position and keeping them away from a designated area. It worth noting that the use of sounds is different to existing approaches, which utilize electric shocks to train animals to adhere within the boundaries of a virtual fence. It was found that audio signals in the frequencies of 125Hz-440Hz, 10kHz-17kHz and white noise are able to control animal activity with accuracies of 89.88%, and 95.93%, for Hebridean and Greyface Dartmoor ewes, respectively. Last but not least, the thesis proposes a multifunctional system that identifies whether the animal is active or inactive, using transfer learning, and manipulates its position using the optimized sound settings achieving a classification accuracy of over 99.95%.

Publications, presentations, and abstracts

Journal Papers

Natasa Kleanthous, Abir Hussain, Wasiq Khan, Jennifer Sneddon, Ahmed Al-Shamma'a, Panos Liatsis, "Deep Transfer Learning in Sheep Activity Recognition using Accelerometer Data" Under review in the Elsevier Journal Expert Systems With Applications 2020

Natasa Kleanthous, Abir Hussain, Wasiq Khan, Jennifer Sneddon, Panos Liatsis, "A survey on machine learning approaches in agriculture" Under review in the Elsevier Journal Neurocomputing 2020

Natasa Kleanthous, Abir Hussain, Wasiq Khan, Jennifer Sneddon, Panos Liatsis, "The Sound of the Lambs: Moving Sheep Using Audio" Under review in Plos One journal 2021

Conference Papers

Kleanthous N., Hussain A., Mason A., Sneddon J., Shaw A., Fergus P., Chalmers C., Al-Jumeily D. (2018) "Machine Learning Techniques for Classification of Livestock Behavior" In: Cheng L., Leung A., Ozawa S. (eds) Neural Information Processing. ICONIP 2018. Lecture Notes in Computer Science, vol 11304. Springer, Cham.

Kleanthous N., Hussain A., Mason A., Sneddon J. (2019) "Data Science Approaches for the Analysis of Animal Behaviours" In: Huang DS., Huang ZK., Hussain A. (eds) Intelligent Computing Methodologies. ICIC 2019. Lecture Notes in Computer Science, vol 11645. Springer, Cham.

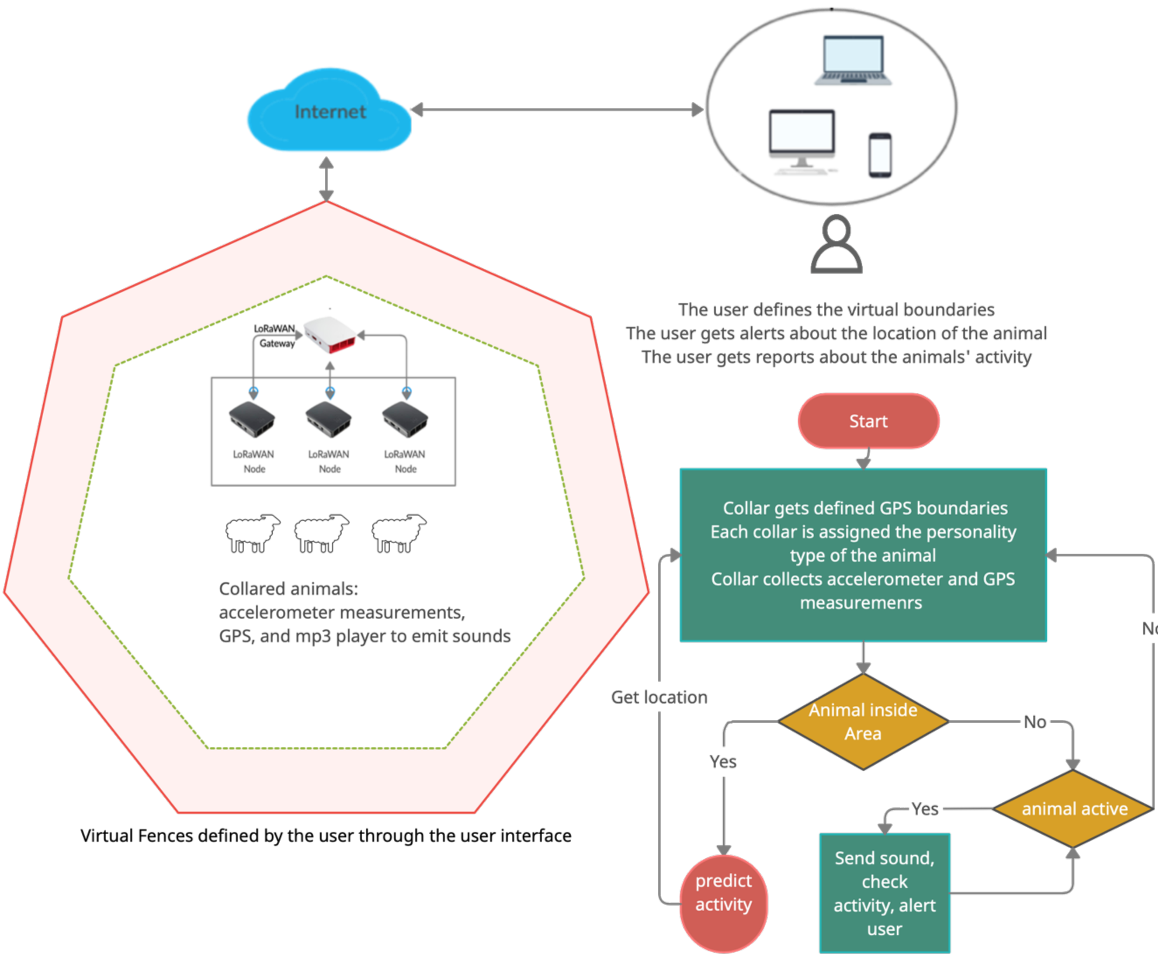
Kleanthous N., Hussain A., Khan W., Sneddon J., Mason A. (2020) "Feature Extraction and Random Forest to Identify Sheep Behavior from Accelerometer Data" In: Huang DS., Premaratne P. (eds) Intelligent Computing Methodologies. ICIC 2020. Lecture Notes in Computer Science, vol 12465. Springer, Cham.

Presentations and Abstracts

N Kleanthous, A Hussain, J Sneddon, A Shaw, P Fergus and C Chalmers "A virtual fencing system for real-time monitoring and controlling animal position and behaviour": Call for Abstracts - 2017-18 Faculty Research Week, publication and oral presentation at Liverpool John Moores University.

Natasa Kleanthous "Virtual Fencing System experiment: Testing the response of sheep on various audio signals", Call for Abstracts 2019 - Faculty Research Week: presented and won 3rd place award for best presentation pitch.

Contents


Internet
The user defines the virtual boundaries
The user gets alerts about the location of the animal
The user gets reports about the animals' activity
LoRaWAN Gateway
LoRaWAN Node
LoRaWAN Node
LoRaWAN Node
Collared animals:
accelerometer measurements,
GPS, and mp3 player to emit sounds
Virtual Fences defined by the user through the user interface
Start
Collar gets defined GPS boundaries
Each collar is assigned the personality type of the animal
Collar collects accelerometer and GPS measuremenrs
Get location
Animal inside Area
Yes
No
No
animal active
Yes
predict activity
Send sound, check activity, alert user

List of Figures

List of Tables

Abbreviations

AAR	Animal Activity Recognition
ACC	Accelerometer
AdaBoost	Adaptive Boosting
ANN	Artificial Neural Networks
CART	Classification and Regression Trees
CNN	Convolutional Neural Network
CNN	Convolutional Neural Networks
DNN	Deep Neural Networks
DT	Decision Trees
FCM	Fecal Cortisol Metabolite
GFS	Greedy Feature Selection
GPS	Global Positioning System
GYR	Gyroscope
KNN	k-Nearest Neighbours
LDA	Linear Discriminant Analysis
LR	Linear Regression
MAGN	Magnetometer
ML	Machine Learning
MLP	Multilayer Perceptron
NB	Naïve Bayes
OOB	Out-of-bag
PC	Principal Components
PCA	Principal Component Analysis
QDA	Quadradic Discriminant Analysis
RF	Random Forest
RFs	Random Forest Feature selection
RMS	Root Mean Square
SAR	Sheep Activity Recognition
SD	Standard Deviation

SFS	Sequential Forward Selection
SMA	Signal Magnitude Area
SVM	Support Vector Machines
TL	Transfer Learning
VF	Virtual Fence
XGB	Extreme Gradient Boosting

Chapter 1 Introduction

This chapter introduces the importance of Sheep Activity Recognition (SAR) as this information provides a valuable source of knowledge regarding the health and the environment the sheep live in. Additionally, the Virtual Fence (VF) system will be introduced, and how it can benefit the Agricultural community and the environment. Information on the motivation of the thesis will be presented as well as the challenges, aims and objectives. Contributions of the thesis and an overview of the thesis structure will be provided.

1.1 Introduction

Animal activity recognition (AAR) is a vital agricultural subject, which can help understand animal behaviour, where the wellbeing of animals can be estimated and classified [1]. Various studies suggested AAR could be used to indicate animal health [2]–[4]. It is noted that the daily monitoring of animal activities and locomotion utilising sensor technology can provide information regarding stress and diseases, such as lameness [5]–[10], in addition to animal daily nutritional consumption [11]. Furthermore, a decrease in animal activities or hyper activities can provide evidence of animal disease and distress [7]. Information gathering through human observation is time-consuming and labour-intensive, and costly. Therefore, devices to measure daily animal behaviour have been proposed and used over the past two decades [12], [13]. It was also noted that the monitoring of animals with a human observer could influence the biological activity of animals on the pasture [14], [15], which may not be the case when using technological devices. Therefore the use of computerised methods for animal monitoring and controlling can be used as a reporting tool for the farm managers to help them improve animal welfare and also avoid economic losses due to the animals' health. Labour costs can also reduced, as animal management requires considerable economic and labour efforts.

Animal food consumption, the presence of diseases and general level of activity or inactivity can be estimated and identified using devices with embedded machine learning (ML) algorithms [16]. These provide the ability to monitor and diagnose animal welfare [17]. Additionally, the position of animals and activity information can be used to nominate pasture utilisation patterns and animal distribution for pursued animal behaviour [18]. Thus, intelligent technologies can play a valuable part in animal health management [19] and provide vital insights to individuals and concerned bodies (e.g., farm managers).

For activity recognition problem, accelerometers are the most commonly used sensors due to their ability to provide information related to animal gait patterns. Due to their small size, lightweight, and low power requirements, these sensors are widely used in various applications for animal behaviour monitoring [20]–[22]. On the other hand, several studies illustrated the use of computer vision [23], [24] and image analysis [24]–[26] for monitoring such behaviours. Moreover, several works demonstrated the collection of tracking data from livestock using Global Positioning Systems (GPS) collars [27]–[30]. Additionally, recent research efforts focused on measuring food intake through sound analysis [31], [32]. These studies illustrate the importance of computer science in intelligent monitoring of animal behaviours and providing support for the research community and their industrial counterparts.

A VF system is a computerised method for creating spatial boundaries of any geometric size and shape without any physical fences or barriers [33]. Virtual Fence (or Fenceless) systems have been discussed over the last 30-40 years due to the inflexibility, cost and heavy maintenance demands required upon traditional fences. Physical fences are 100% stock-proof; however, agricultural fencing was one of the most significant expenses of the 19th century [34][35], and therefore, new means of livestock containment is required. Consequently, VF is considered by the Agricultural community as the next step to replace physical fences. Such a solution will require the physical barriers to be removed or reduced, and visual, auditory, and possibly olfactory cues could replace them as means to control animal's movement and position on the field [36].

The use of VF will offer immense potential for a more flexible and cost-effective way to manage livestock and manage the land they graze. In addition, VF systems will advance the decision-making approach of the farmers and stakeholders by replacing intense manual labour with cognitive labour [37]. For instance, real-time data that reports animal distribution and pasture usage will be available to land managers, providing them with opportunities to efficiently decide how to manipulate the shape and size of the boundaries based on sward characteristics using smart devices. Furthermore, an intelligent system embedded in the animal's collar might automatically make the decisions without the need for human interaction. Decision making by these means is not possible with physical fences, which depend upon manual labour and human scrutiny alone.

An intelligent system is needed to provide information on where the animals are and what they are doing, where they mostly graze, and their nutritional habits during the day [37]. Having this vast amount of data, decisions about animal health, animal position monitoring, distribution control, and efficient land utilisation can help prevent soil erosion, soil contamination, water pollution and the spread of animal diseases [38] [39]. VF system can be used to move the animals among the land they graze using intermittent acoustic cues to manipulate their location as it is noted in the literature that there is potential to train animals to respond to such stimulus [36], [40]. The most commonly used method of preparing animals to learn the virtual fence system is by the use of audio warning signals, followed by an electric shock. This method of training is considered stressful for the animal, and therefore, the interest of this thesis is to test whether it can train animals using audio cues only and act as an alternative to electric shocks.

In this thesis, firstly, the application of ML using accelerometer measurements collected from sheep will be provided. Several ML algorithms are tested and used to recognize patterns from various sheep activities using multiclass, and binary classification. Specifically, the whole process of data mining using data collected from the collar of the animal is investigated. The goal is to propose solutions to SAR problems and suggest solutions for robust performance. Additionally, a method to address the heterogeneity of accelerometer orientation and sensor devices is presented. Regarding the VF system, the main goal is to examine if the sheep's position can be controlled and manipulated using only audio stimuli emitted from the animal's collar. Thus, the purpose of this thesis is to investigate opportunities and suggest possible solutions on how a multifunctional fencing system based on machine learning and audio stimulus for automatic monitoring and control of animals can be utilised.

1.2 Motivation

In this section, a list of the challenges of SAR will be provided, which are the focus of this thesis. Additionally, challenges regarding the implementation of virtual fencing systems based on acoustic cues will be presented.

1.2.1 Challenges

Energy efficiency and memory usage of an embedded system are critical challenges. It should be noted that energy efficiency and computation in the data mining process is considered. Additionally, in the thesis, sensor position and orientation, heterogeneity, feature extraction, ML selection, data annotation and data labelling challenges are also discussed.

For data collection, the animals are fitted with sensor devices. Those devices might shift and rotate due to the animals' movement. For example, the animal might rub on a bush or tree, and the sensor might change position and orientation. This might result in errors if the extracted variables in the data mining process are sensitive to the orientation and position of the sensor. Hence, they can affect the classification performance [41], [42]. Consequently, classification solutions that are independent of the orientation of the accelerometer should be considered.

Another challenge is the heterogeneity of the sensor devices. For example, some accelerometer sensors might exhibit more noise than other accelerometers. Therefore, once a sensor device must be replaced with a new one, this might mean that further data collection and model training must be conducted, which will result in a time-consuming process. Thus, solutions considering the heterogeneity of the device must be investigated. This will result in a time-saving technique of data processing and classification.

Regarding the sheep, there is a problem of animal heterogeneity as well. For example, collecting data from sheep that comprise behaviours from older animals may be different from the data collected from younger animals, as older sheep might be less active or exhibit different gait patterns. Therefore, training models on such data might result in poor performance if the new unseen data has more activities involved and might cause misclassifications.

The extraction of features in the data mining process is a challenging task as the performance of ML heavily depends on the data representation [43]. The selection of an ML algorithm is a challenging task as each method has different parameters and properties that affect the model's overall performance. In the SAR, it is difficult to identify a one-fits-all solution since each application of interest might focus on diverse activities of the animals.

Lastly, data collection and labelling is one of the most time-consuming and laborious stages of the SAR. At the same time, the animals need to be video-recorded, and an observer needs to be present. The videos must be time-synchronised to serve as ground truth for data labelling.

Additionally, The training of animals based on sounds is challenging. Previous research used audio cues as a warning, and electric shocks as punishment to train the animals to associate the sounds to the electric shocks and learn the VF [44], [45]. In the scenario of discarding completely electric shocks, a thorough investigation of the auditory awareness of the sheep and the audio to be used must be conducted to train the animals to respond to the audio cue alone.

1.3 Research Aims and objectives

This thesis aims to propose a "smart" virtual fencing algorithm able to monitor the behaviour of the animals using ML models while considering orientation-independence of the sensors, and heterogeneity of the animals and the sensor device, so it can be used across a variety of accelerometer sensors and reused when new animals and new devices are introduced. Additionally, a key component is to investigate the response of the animals on acoustic cues to evaluate the feasibility of a virtual fence system for sheep without the need for electric shocks.

1.3.1 Objectives

1. Thorough understanding of current technologies with respect to virtual fencing systems and animal activity recognition using ML models
2. Application of machine learning techniques and evaluation of the performance of various models on a dataset containing accelerometer, gyroscope, and magnetometer measurements from sheep and goats to identify the most promising features and ML algorithm to classify animal behaviours.
3. Collection of primary datasets to evaluate the use of ML for animal activity recognition.
4. Proposing a Convolutional Neural Network (CNN) Transfer Learning (TL) technique for monitoring sheep activity with robustness to sensor position and orientation. Also, the primary datasets will be used to address the challenge of heterogeneity of accelerometer sensors.
5. Design and development of a sound system to test whether sheep position can be manipulated based solely on acoustic cues
6. Identification of the most effecting sounds based on the animals' response
7. Statistical analysis on the time the animals respond to the sounds based on their personality type, their motivation, the frequency used, and breed
8. Intelligent system proposal, able to detect animal behaviour and send sounds to manipulate the animals' position on the pasture

1.4 Contributions

This PhD research project provided the following novel aspects:

- An approach that aims at finding the optimal feature set and ML model to identify sheep behaviour based on various window sizes and sample rate was proposed. This contribution provides guidance to the research community on how to approach SAR problems based on the researchers' requirements, as the solution is not a one-fits-all.
- A real-time data-driven approach for animal activity recognition is proposed comprising a combination of CNN and hand-crafted features, which significantly improves classification performance.
- Four primary datasets are acquired through different sensors comprising accelerometer gait signals from sheep, and will be made publicly available to the research community as there is limited availability for such datasets.
- A new method which uses deep Transfer learning to evaluate the generalisation properties of the system is proposed and is used for the first time in SAR
- Designing a system based on acoustic stimulus to manipulate the animals' position. VF systems are mostly tested on cattle, however, limited studies focused on sheep, which is the interest of this thesis. Additionally, to the best of the author knowledge, this is the first study that used only audio sounds to train sheep.
- A custom sound system is designed and tested sounds on sheep to identify if audio can replace electrick shocks for a VF system. This is the first study which focused on sheep to the best of the authors' knowledge.
- An intelligent system is proposed which will be able to identify whether the sheep is active or inactive and send audio signals to the animal through its collar to control and manipulate their position. To the best of the authors knowledge, this system is a novel contribution to the research community as limited studies exists in relation to sheep and solely acoustic stimuli for a VF.

1.5 Organisation of the thesis

The remainder of this thesis is organised as follows:

Chapter 2 : this chapter provides background information on SAR. It mainly specifies an overview of data mining methods used over the past ten years to classify sheep behaviours. Additionally, background information regarding the development and requirements of a VF system is provided.

Chapter 3 : this chapter provides information regarding three experiments conducted concerned with SAR using ML methods. The purpose is to provide suggestions for implementing intelligent devices to monitor the activities of the animals. More specifically, this chapter investigates ML methods' use with various combinations of feature sets to identify the optimal solution for specific SAR problems.

Chapter 4 : this chapter proposes the use of Deep Transfer Learning (DTL) to monitor sheep's activities using accelerometer measurements. This chapter aims to test the potential of a SAR method independent of the type and orientation of the accelerometer device to indicate the importance of DTL in terms of generalisation.

Chapter 5 : this chapter presents experiments conducted for testing sounds on sheep and whether their position could be manipulated and controlled using acoustic cues. This chapter demonstrates the potential of replacing electric shocks with audio in a virtual fence scenario.

Chapter 6 : this chapter evaluates the performance of the proposed CNN TL model from Chapter 4. Furthermore, it presents the algorithm for the implementation of a multifunctional VF system.

Chapter 7 : this chapter provides the summary and conclusions obtained from the thesis. Additionally, it highlights the thesis contributions and limitations. Also, it provides suggestions for future work.

Chapter 2 Related work: Sheep activity recognition and virtual fencing systems

In this Chapter, background information on sheep activity recognition will be provided, as well as virtual fencing system techniques. This chapter is divided into two subsections. The first section (Section 2.1) focuses on recent advancements in sheep activity recognition based on ML and DL using accelerometer signals. More specifically, it provides a thorough overview of various window segmentation, feature extraction, feature selection, and classification algorithms, which have been used over the past ten years to identify sheep activities. In Section 2.2, information regarding animal responses to external stimuli such as sounds and visual cues for VF systems will be provided.

2.1 Sheep Activity Recognition (SAR)

This section surveys the research studies addressing sheep activity recognition based on (ML) using accelerometer measurements. While several existing studies propose solutions to this problem, over the last decade, there is a significant research gap between understanding the relationships between the focus of a research and the specific solution parameters, i.e., window size, feature set and significance level, and choice of ML techniques [21]. Furthermore, current research on sheep activity recognition (SAR) using ML is limited and provides an opportunity to systematically analyse data processing and analysis protocols [46], as addressed in this contribution.

The remainder of this section is organized as follows. Firstly, an overview of a typical sheep activity recognition problem will be illustrated. Furthermore, existing works and challenges in each of the technical aspects of sheep activity recognition will be summarized and discussed.

2.1.1 Sheep Activity Recognition using Motion Sensors

This section describes the requirements and typical characteristics of a sheep activity recognition problem, as illustrated in Figure 2-1. Existing works along with the techniques proposed in these studies are presented. A typical activity recognition problem includes: (1) Sensor mounting on the animals' body to capture the activities: Sensors are placed on the collar, ear, leg, or under the jaw. (2) Data labelling: Once the raw data is collected, data labelling is manually performed, and video recordings are time-synchronised to serve as ground truth; (3) Pre-processing of the acquired data based on the specified problem; (4) Selection of window size and feature extraction from the processed data; (5) Feature selection to identify the most information-rich feature set; (6) Model development and training using ML algorithms; (7) Performance evaluation of the models and selection of the most appropriate method.

Figure 2-1 presents a typical diagrammatic overview of the core stages followed in related research studies to tackle questions about animal behaviour based on the application requirements, e.g., identifying when an animal is active or inactive, classifying whether an animal suffers from lameness. Based on the nature of the problem and type of data, different options for window size, feature extraction, and ML models can be used.

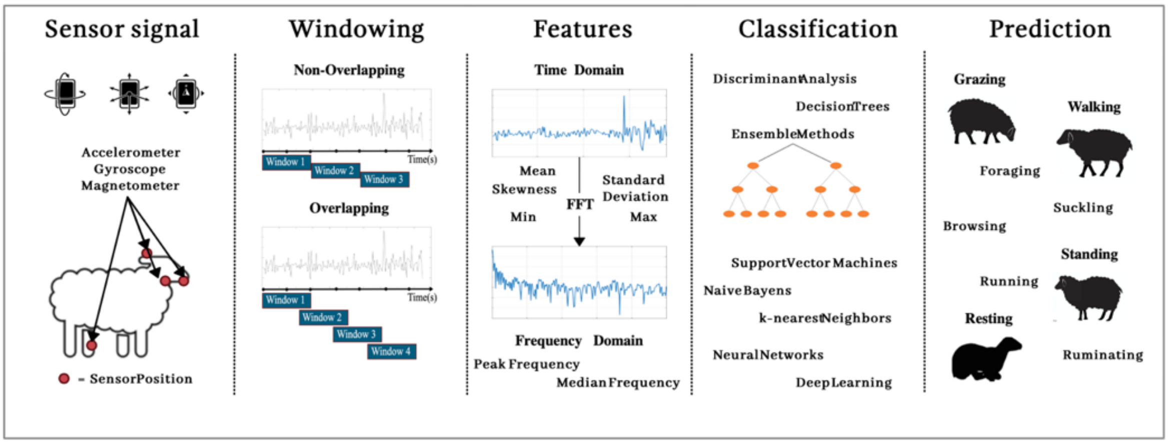


Figure 2-1 Illustration of the animal activity recognition problem

2.1.2 Accelerometers

In this research, we focus on studies related to sheep activity recognition using accelerometer signals. Accelerometers measure the acceleration of motion and are a ubiquitous type of sensor in activity recognition problems since they are light, small, inexpensive, and have reduced power consumption requirements [18], [21], [47]–[51]. Activities such as walking, grazing, scratching, lying, and standing can be easily recognised, yielding overall accuracies over 98% using only accelerometers [52], [53]. Additionally, running was detected with an accuracy of 96.62% using only one acceleration axis [54]. Several research studies also used gyroscopes and magnetometers combined with accelerometers to understand better domestic and wild animal behaviour [1]–[4]. Combining accelerometer and gyroscope features yielded an accuracy of 98% in identifying lying behaviour [55]. Moreover, walking behaviour was correctly predicted with an accuracy of 99% using only three features extracted from accelerometer data [56]. Other works also use accelerometer features to discriminate between active and inactive states with 98.10% accuracy [21].

2.1.3 Sensor placement

To collect gait patterns from the animals, accelerometer sensors are attached to the animal body, usually in the collar and ear [5], [6], [21], [46], [55]–[58]. On the other hand, there are some studies, where accelerometers are mounted on the leg and under the jaw [59]–[63].

Table 2-1 presents the four most common sensor placements reported in research studies over the last decade. The position is determined based on the activity problem under investigation. For example, a recent study showed that an accelerometer attached to the animal's ear could discriminate lame walking from grazing, standing and walking [5]. However, the same technique used on data collected from the animal's collar and leg failed to detect normal walking and lame walking. Placing the sensor under the jaw was used when the research focused on feeding behaviour, e.g., chewing and biting, yielding sensitivity and specificity of 97.4% and 97.7%, respectively [59]. Additionally, other types of behaviour were discriminated from biting and chewing with 100% sensitivity, when the sensors were attached under the jaw [59].

Table 2-1 Sensor placement vs animal activity

Sensor Placement	**References**	**Animal activity domain**
Collar	[2], [5], [46], [52], [53], [55]–[57], [64]–[71]	Collar-borne devices were used to classify “active” vs “inactive” behaviour, or “grazing” vs “non-grazing”. Additionally, collar-borne devices were used in multiclass classification to discriminate behaviours such as “grazing”, “browsing”, “foraging”, “standing”, “walking”, “running”, “resting”, “lame walking”.
Leg	[5], [46], [56], [63]	Leg-borne devices were used to identify behaviours such as “walking”, “lame walking”, “trotting”, galloping”, “running”, “resting”, “grazing”.
Ear	[5], [6], [21], [46], [55]–[58]	Ear-borne sensors were used to identify behaviours such as “lame” vs “not-lame”, posture (upright vs prostrate), “grazing”, “lying”, “standing”, “walking”.

Under the jaw	[59]–[62]	Jaw-based sensors were used to identify “biting”, “chewing”, “grazing”, “lying”, and “standing” behaviours.

2.1.4 Sheep Activities

Various studies related to sheep activity recognition problems using ML attempted to identify different types of activities. In Table 2-2, we define and present the activities studied in the associated studies. Those activities differ in their complexity. The most common behaviours found in research studies are grazing, walking, standing, resting, and lying. However, other types of behaviour, such as biting, chewing, ruminating, and foraging, are also studied and referenced in Table 2-2.

Classification of animal activity to active or inactive states has a low degree of complexity and can be easily performed by utilizing conventional ML methods with high accuracy. It is noted in the literature that decreased animal activity or hyperactivity could be an indicator of disease and distress [7]. This kind of information is valuable for farm managers and related individuals. On the other hand, detection of speed [65] and running direction is necessary when the aim is to identify when a thief or a predator is pursuing an animal, especially in remote locations; therefore, classifying trotting/running is essential. A variety of studies indicated that trotting is generally the most challenging gait to determine [63], [72], [73].

Identifying real-time foraging activity is essential for sheep farmers working in extensive agricultural hill systems [74]. These grazing systems characterize the bulk of the sheep farming industry in the UK and other parts of the world. Changes in the eating behaviour of sheep could indicate health or management problems, e.g., quality of pasture [75]. Continuous monitoring of food intake in real-time could provide better estimations of carcass value at market and grazing impact on the sward. It could also be a valuable land management tool, preventing the occurrence of dangerous ecological tipping points, leading to overgrazing, soil erosion and water contamination, particularly in sensitive upland ecosystems [76], [77].

Additionally, studies are conducted to identify lameness, one of the most common and persistent health problems in sheep flocks around the world [5]–[10]. Unusual amounts of lying time were shown to indicate lameness in cattle and sheep [78]–[80], and therefore lying is a critical activity to be detected in monitoring systems.

Table 2-2 Description of sheep activity and behaviour

Behaviour	Description	Reference
Grazing	Eating sward at ground level with the head down.	[21] [46] [70][71] [64] [5] [65] [60] [2] [57] [52] [66] [67] [61] [62] [53] [56] [69]
Infracting	Eating from branches above a certain height.	[70]
Browsing	Eating the leaves of shrubs or trees with the head up off the ground.	[71]
Chewing	Rotation of the lower jaw after a bite activity in any head position (up or down).	[59]
Biting	Gathering forage (browse or grass) with incisor teeth.	[59]
Ruminating	Usually performed with the body lying in the sternal position. Fermentation of digesta in the reticulo-rumen complex frequently accompanied by cud-chewing.	[2] [57] [61] [62]
Foraging	A general term for the acquisition of nutrients with the ingestive apparatus: teeth, lips and tongue.	[68]
Walking	Four-time slow quadrupedal locomotion in sheep; speed 1.1-1.3 $m.s^{-1}$.	[21] [46] [71] [64] [6] [5] [65] [60] [2] [52] [66] [67] [68] [55] [56] [58] [69] [63]
Moving	This is an intended movement from one place to another. Naturally, the sheep is not looking for nutrition.	[70]
Running/ Trotting	Two-time quadrupedal locomotion in sheep; speed 1.41-2.41 $m.s^{-1}$ [81].	[70] [65] [60] [66] [68] [69] [63]
Galloping	Four-time rapid quadrupedal locomotion in sheep; speed 2.28-3.56 $m.s^{-1}$ [81].	[70] [68] [63]
Scratching	Rubbing body surface against a solid object.	[71] [64] [52]
Standing	Standing with all four feet on the ground.	[21] [46] [70] [71] [64] [6] [5] [65] [60] [2] [66] [67] [68] [55] [56] [58] [69]

Resting	Lying in the absence of rumination. Usually performed with the body lying in sternal position or infrequently with the body lying horizontally in a lateral position.	[21] [46] [71] [64] [6] [5] [65] [60] [2] [66] [67] [68] [61] [55] [56] [58] [69]
Active	There is body movement, e.g., locomotion, foraging, scratching.	[21] [67] [2]
Inactive	There is no movement; the sheep are lying down to ruminate or are asleep in sternal or lateral recumbence.	[52] [67][2]
Upright	The body is standing in vertical position.	[21]
Prostrate	The body is lying in horizontal position.	[21]
Lame	Asynchronous gait commonly due to lameness in one or more limbs; usually, the hoof.	[6][5]
Not Lame	Normal gait, related to normal walking, trotting and galloping.	[6]

2.1.5 Data collection and labelling

Data collection and labelling of the raw dataset are essential for the methods related to the identification of sheep activities and behaviour. Obtaining the data and labelling is an extremely time consuming and labour-intensive task. Usually, the process involves the animals in their natural environment, having the sensors logging motion signals from the collar, leg, ear, or mounted under the jaw. Video recordings of animals are being observed to label behaviour during data collection. There are tools for labelling the data, such as ELAN_5.7_AVFX Freeware tool [82], however most authors perform the labelling manually having the data measurements and video recordings timestamped to serve as a ground truth during the labelling step. The camera is set to record also the time in HH(hour):MM(minute):SS(second) so it can be easily synchronised with the timestamped data measurements. The camera is usually placed in the pasture having a clear view of the selected animal or all the animals. During the video recordings, an observer is present and is responsible to move the camera if the animals are out of view. The animal is either recognised by the colour of the device or is numbered by a spray on its body, so the observer can recognise the corresponding animal. The various behaviours are labelled based on an expert's knowledge (refer to Table 2-2 in the previous section).

The number of animals involved in data collection is crucial since animals exhibit different characteristics such as age, height, and health status. However, metadata such as age, height, and health status is not integrated in the datasets, but it can be considered during the selection of the animals for the data collection. For example, an animal suffering from lameness will differ in gait patterns from a healthy animal [7]. Additionally, younger sheep might be more active than older ones and behaviours such as walking and running may vary. Therefore, the animal selection plays a crucial role in data collection since having multiple behaviours from a variety of animals can ensure more representative training data, resulting in improved predictive model characteristics, e.g., adapting better when new animals are added to the flock.

2.1.6 Windowing and sample rate

Accelerometer measurements are collected in time intervals (milliseconds, seconds, minutes etc.), forming a time series dataset that needs to be analysed in overlapping intervals. Therefore, a technique to slice the signals is commonly used, known as 'windowing' [83]. This step is critical since accelerometer signals provide valuable information about motion patterns in blocks of data and not as a single variable measure. The windows are of the same size and are either disjoint [2], [5], [21], [46], [52], [59], [62], [64], [71] or overlapping, typically at 50% [83]. Overlapping windows are suggested because of their ability to capture the transitions of activities more precisely. On the other hand, very small disjoint windows can avoid transitions, but may lose important information from the signal [83]. Therefore, research studies analysed the effects of varying window size to identify the appropriate size, containing sufficient features to discriminate the gait patterns, while simultaneously avoiding false classifications. This can result from comparatively longer windows due to activity transitions, i.e. walking to grazing, standing to lying. Additionally, windowing affects the computational complexity of the feature extraction process, which must also be taken into account [84].

In relation to SAR, various window sizes were tested and evaluated in the literature, as shown in Table 2-3. It can be observed that 5s and 10s windows have commonly been used. Similar to window size, the sampling rate is also important since the quantity/amount of samples in each window also depends on this. The choice of sampling rate influences accelerator signal information, and subsequently, feature extraction, while playing an important role in terms of the time complexity and power consumption of the device. Various choices of sampling rate were reported in the literature, e.g., 1Hz [67], 4Hz [54], 8Hz [55], 10Hz [60], [71], 12Hz [5], [46], [56], 12.5Hz [52], [21], 16Hz [6], [55], [57], [58], 20Hz [53], 25Hz [59], 32Hz [2][55], 33Hz [63], 50Hz [70], 100Hz [65], [66], [69], 62.5Hz [61], [62] and 200Hz [64], [68]. An illustration of the choice of sampling rate vs window size selection in SAR is shown in Figure 2-2. The plot shows that the window size range, which has been mostly used, is between 3s to 10s, with sampling rates between 10Hz to 20Hz.

Table 2-3 Variation of window sizes and sample rate used in the literature

Window size (seconds)	**Sample Rate and Reference**
1	25Hz [59]
3	12Hz [46], 8Hz, 16Hz, 32Hz [55], 25Hz [59], 10Hz [60]
5	12.5Hz [21], 12Hz [46], 20Hz [53], 8Hz, 16Hz, 32Hz [55], 25Hz [59], 10Hz [60], 62.5Hz [62], 33Hz [63], 100Hz [65], 100Hz [66], 100Hz [71]
6.4	100Hz [69]
7	16Hz [6], 8Hz, 16Hz, 32Hz [55], 16Hz [57], 10Hz [60]
10	12Hz [5], 12.5Hz [21], 12Hz [46], 20Hz [53], 12Hz [56], 16Hz [58], 62.5Hz [62], 10Hz [65]
15	20Hz [53]
25	32Hz [2]
30	12.5Hz [21], 62.5Hz [62], 100Hz [66]
60	62.5Hz [61], 62.5Hz [62]
120	62.5Hz [62]
180	62.5Hz [62]
300	62.5Hz [62]

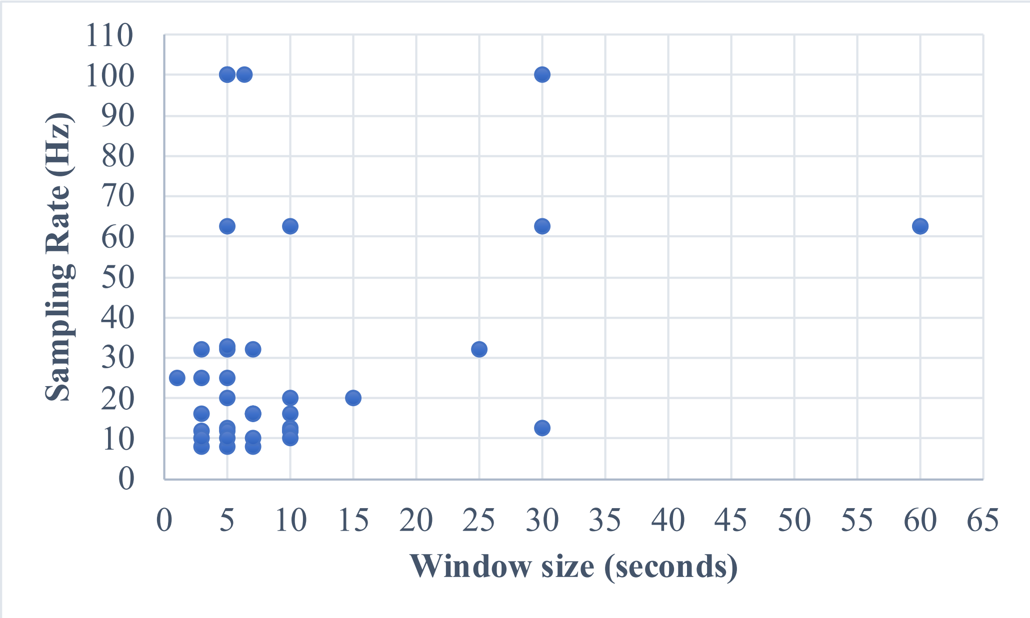


Figure 2-2 Window size and sample rate used in sheep activity recognition studies

2.1.7 Time and frequency domain features

Over the last years, DL outperforms the classic ML methods and it does not require the feature extraction process in the pre-processing step, as the features are learned during the training phase of the raw data. However, for the DL to outperform the ML methods and get optimal results, it requires large datasets which are not always available in SAR field and also the collection of such data is very time-consuming. Therefore, the ML methods are still an option, and thus, the feature extraction process is required and is considered a vital step in classification problems [85], [86].

Several continuous accelerometer measurements (i.e., time windows) are required in order to be able to characterize activity patterns, in contrast with other measurement modalities, which can provide information from a single value. Therefore, there is a need for feature extraction. A variety of techniques have been suggested to represent the information in raw accelerometer signals to be used with ML algorithms in the classification of gait activities [87]–[89]. Previous research in the field considered the use of an extensive number of time- and frequency-domain features (refer to Table 2-4 and Table 2-5). Examples of time-domain features include statistical parameters such as the mean, variance, correlation [65], [69], higher-order moments, and sensor-based measurements such as pitch, yaw, roll and inclination angles [59], [70]. The advantage of time-domain features is that they are straightforward to calculate, and thus, in most cases, they are computationally efficient [50], [90]. However, they are affected by measurement and calibration errors [91]. Frequency-domain features, on the other hand, e.g., signal area, spectral entropy, peak frequency, etc., often require additional processing, i.e., windowing, filtering, and the application of the Fourier transform, however, they are able to robustly represent the information in the signal. Thus, they are more computationally expensive than time-domain features [92]–[94]. In Table 2-4 and Table 2-5, we present an extensive list of time- and frequency-domain features, respectively, used in SAR.

Table 2-4 Time-domain features in SAR

Time-Domain Features		
Name	**Description and Formula**	**Reference**
Mean/Average	$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{j,i}\,, j = 1,2,3$	[5], [6], [21], [46], [52], [53], [55]–[59], [61], [62], [64]–[66], [68]–[71]
Average all-axis	$\bar{x} = \frac{1}{m}\sum_{j=1}^{3} x_j$	[21]
Standard deviation (sd)	$s_j = \sqrt{\sum_{i=1}^{n}\frac{(a_i - \bar{a})^2}{n}}$	[6], [21], [52], [53], [55], [57]–[59], [64]–[66], [68], [69], [71]
Variance	s_j^2	[61], [62], [65], [66], [69], [70]
Inverse coefficient of variation	$\frac{\bar{x}_j}{s_j} \times 100\%$	[61], [62]

Average Standard deviation	$\bar{s} = \frac{1}{3}\sum_{j=1}^{3} s_j$	[21]
Median[1]	$med(x_j) = \begin{cases} x_{j,\frac{n}{2}}, & if\ n\ is\ even \\ \frac{x_{j,\frac{n-1}{2}} + x_{j,\frac{n+1}{2}}}{2}, & if\ n\ is\ odd \end{cases}$	[68]
Minimum[1]	$\min(x_j) = x_{j,1}$	[5], [6], [21], [55]–[59], [65], [66], [68]–[70]
Maximum[1]	$\max(x_j) = x_{j,n}$	[5], [6], [21], [55]–[57], [59], [65], [66], [68]–[70]
Pairwise correlation between axes j and k	$corr(x_j, x_k) = \frac{cov(x_j, x_k)}{s_j s_k}, where\ cov(j,k)\ is\ the\ covariance$	[65], [66], [69]
Mean distance between axes j and k	$dist(x_j, x_k) = \frac{1}{n}\sum_{i=1}^{n}(x_{j,i} - x_{k,i})$	[69]
25th/75% percentile[1]	$m_P = \left[\frac{P}{100} \times n\right], whereP = 25{,}75$	[68]
		[68]
Movement Intensity / Average Signal Magnitude Vector	$\frac{1}{n}\sum_{i=1}^{n}\sqrt{{x_{1,i}}^2 + {x_{2,i}}^2 + {x_{3,i}}^2}$	[5], [21], [46], [56], [58], [65], [66], [69]
Movement Variation	$\frac{1}{n}\left(\sum_{i=1}^{n-1}\lvert x_{1,i+1} - x_{1,i}\rvert + \sum_{i=1}^{n-1}\lvert x_{2,i+1} - x_{2,i}\rvert + \sum_{i=1}^{n-1}\lvert x_{3,i+1} - x_{3,i}\rvert\right)$	[5], [21], [46], [56], [59], [60], [70]
Signal Magnitude Area (SMA)	$\frac{1}{n}\left(\sum_{i=1}^{n}\lvert x_{1,i}\rvert + \sum_{i=1}^{n}\lvert x_{2,i}\rvert + \sum_{i=1}^{n}\lvert x_{3,i}\rvert\right)$	[5], [21], [46], [53], [56], [59], [60]

[1] Observations have been ranked in ascending order.

Signal Vector Magnitude	$\sqrt{x_{1,i}{}^2 + x_{2,i}{}^2 + x_{3,i}{}^2}$	[6], [53], [59], [60], [70]
skewness	$\frac{1}{n}\sum_{i=1}^{n}\frac{(x_{j,i} - \bar{x}_j)^3}{s_j^3}$	[6], [52], [53], [65], [66], [68], [69]
kurtosis	$\frac{1}{n}\sum_{i=1}^{n}\frac{(x_{j,i} - \bar{x}_j)^4}{s_j^4}$	[6], [52], [53], [55], [57], [58], [65], [66], [68], [69]
Interquartile Range[1]	$m_{75} - m_{25}$	[6], [55], [57], [58]
Root mean square (RMS) signal value	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_{j,i}{}^2}$	[52], [64], [71]
RMS velocity value	$\sqrt{\frac{1}{n}\sum_{i=1}^{n-1}\frac{1}{(x_{j,i+1} - x)^2}}$	[52], [64], [71]
Zero crossing rate (per window)	$count\left((x_{j,i} - \bar{x}_j) == 0\right)$	[6], [52], [55], [57], [68], [70]
Crest factor	$\frac{max\left(\sqrt{x_{1,i}{}^2 + x_{2,i}{}^2 + x_{3,i}{}^2}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{1,i}{}^2 + x_{2,i}{}^2 + x_{3,i}{}^2\right)}}$	[52]
Peak to Peak	$max(x_{j,i}) - min(x_{j,i})$	[52]
Pitch (degrees)	$tan^{-1}\left(\frac{-x_{1,i}}{\sqrt{x_{2,i}^2 + x_{3,i}^2}}\right)$	[59], [60], [70]
Roll (degrees)	$atan2(x_{2,i}, x_{3,i}) \times \frac{180}{\pi}$	[59], [60], [70]
		[70]

Yaw (degrees)	$atan2(x_{1,i}, x_{2,i}) \times \frac{180}{\pi}$	[59], [60]
Inclination	$tan^{-1}\left(\frac{\sqrt{x_{1,i}^2 + x_{2,i}^2}}{x_{3,i}}\right)$	
Sum of changes	$\sum_{i=1}^{n-1} x_{j,i+1} - x_{j,i}$	[52], [64], [71]
Mean of absolute changes	$\frac{1}{n}\sum_{i=1}^{n-1}\lvert x_{j,i+1} - x_{j,i}\rvert$	[52], [64], [71]
Integrals [95]	$\int_{t=0}^{T}\lvert x_1(t)\rvert\, dt + \int_{t=0}^{T}\lvert x_2(t)\rvert\, dt + \int_{t=0}^{T}\lvert x_3(t)\rvert\, dt$	[52], [64], [71]
Squared integrals [95]	$\left(\int_{t=0}^{T}\lvert x_1(t)\rvert\, dt\right)^2 + \left(\int_{t=0}^{T}\lvert x_2(t)\rvert\, dt\right)^2 + \left(\int_{t=0}^{T}\lvert x_3(t)\rvert\, dt\right)^2$	[52], [64], [71]
Madogram [96]	$\frac{1}{2}\mathrm{E}[x_{j,i} - x_{j,t+u}]$ where *t=lag, E[.]=expectation*	[52], [64], [71]
Energy	$(x_{1,i}^2 + x_{2,i}^2 + x_{3,i}^2)$	[5], [21], [53], [56], [59], [60], [64]–[66], [68], [69], [71]
Entropy	$(1 + (X_{1,i} + X_{2,i} + X_{3,i}))^2 \times log(1 + (X_{1,i} + X_{2,i} + X_{3,i}))^2$	[21][46][59] [5] [60] [52] [53] [56]

Table 2-5 Frequency-domain features in SAR

Name	Description and Formula	Reference
Energy in 1Hz bins	$\frac{1}{N_b}\sum_{n=BN_b+1}^{BN_b+N_b}\lvert X_{j.n}\rvert^2$ where X is the Fourier transform of x, Nb is the number of	[69]

	samples in each bin, and B={0,…,9}.	
Spectral entropy [97]	$$\sum_{n=1}^{N} P(X_{j,n}) \times log \frac{1}{P(X_{j.n})}$$ where P(X) is the normalized power spectrum of X.	[6] [69][57][65][68][55][66]
Signal Area	$$SA = \sum mag \,. \frac{1}{f_s}$$ where mag is the magnitude and f_s is the sampling frequency.	[55]
Absolute signal area	$$ASA = \sum \lvert mag \rvert \,. \frac{1}{f_s}$$	[55]
Peak frequency	$$\mathrm{ar\,g}\left(\frac{f_s}{n} max\left(P(X_{j,n})\right)\right)$$	[71][64][52][68][6][57] [55]
Frequency Magnitudes	Magnitudes of the first six components of the Fourier-transformed signal	[68]
Spectral Area	$$2\sum_{n=1}^{N} S(f_n) \times \Delta f$$ where $S(f_n)$ is the power spectral density at frequency n.	[6]
Harmonic frequency (2^{nd} and 3^{rd})	Frequencies were the Fourier-transformed signal has its second and third highest power values	[6]
Harmonic ratio	$$\frac{\sum_{n=1}^{n/2} f_{2n}}{\sum_{n=1}^{\frac{n}{2}-1} f_{2n+1}}$$	[6]

2.1.8 Feature selection and dimensionality reduction

Feature selection and dimensionality reduction are of major importance in activity recognition. Extracted features may contain irrelevant, duplicate, or misleading information, which could affect the predictive or classification tasks [98], [99]. While exhaustive search algorithms may be useful in identifying distinctive features, the deployment of exhaustive search is impractical in most of the cases, specifically, when in big, high dimensionality datasets. To handle this, a variety of feature selection and dimensionality reduction algorithms have been used to identify the optimal set of features, which would be used in the classification/ predictive model in various fields [100]. The most commonly used algorithms can be classified into filters, wrappers, and hybrid approaches, as described in the following subsections.

2.1.8.1 Filter Methods

Filter methods use proxy measures for dimensionality reduction in high dimensional spaces, which mostly include the amount of information, statistical attributes such as variance, similarity score, consistency etc., [98], [101], [102]. Likewise, there exist variety of filter methods to be used depending on the nature of data and hence the task in hand, such as prediction or classification. Various studies have used different information-based filters, e.g., Information Gain [103][104], Gain ratio [105], fast correlation-based filter [106] and Symmetrical uncertainty [106]. On the other hand, the Chi-square test [104], Fisher score [107] and feature weighting k-Means [108] are examples of works that use statistical filters. Similarly, Relief and ReliefF [106], [109]–[111] are examples of similarity-based filter methods used in classification and regression problems. A recent work uses filter methods to identify the candidate feature set for human activity recognition with an inertial sensing unit [112]. The algorithm identifies 48 candidate features out of a set of 585 temporal and spectral features, concluding the effectiveness of the selected features and classification algorithm. Additionally, several works related to SAR used the Relief method to reduce the number of features for the specific problem [6], [53], [57], [68]. Overall, filter-based methods are comparatively better than wrapper and hybrid methods [113], specifically, in high dimensional feature spaces, due to lower execution times and generalization abilities, as they are independent of the employed supervised algorithm. On the other hand, filter methods are unable to eliminate inter-related features due to univariate analysis [114], [115].

2.1.8.2 Wrapper Methods

Wrapper methods use a subset of feature space recursively to train a predictive or classification model and evaluate the performance over unseen data for the candidate feature set. Selection of subset in each round can be performed through various algorithms such as hold out (forward, backward), selection [116], and heuristic search methods [117]. Finally, the best performing feature set is identified using the test data for the corresponding trained model. One of the major issues with wrapper methods is time complexity specifically, in high dimensional feature spaces [118]. Recursive training and performance validation for large size datasets and high dimensional feature spaces is computationally expensive, and thus deemed impractical. Therefore, these methods are useable only for small size feature sets, while at the same time using simple ML models such as Naïve Bayes, SVM and Random Forest for faster training and validation over a recursive subset of candidate features.

An example of wrapper methods is the Boruta [119], [120], which deploys the Random Forest algorithm for recursive selection of candidate features. Work related to SAR applied the Boruta for feature selection, prior to fitting the data to the predictive model [64]. Likewise, [121] presented a detailed comparison of RF-based feature selection and standard chemometric methods, when classifying spectral data. Suto et al., [122] presented an interesting study involving wrapper and filter methods in human activity recognition. Specifically, they presented a naïve Bayesian wrapper method, which outperformed filter methods, including Chi-Square, Fisher score and T-test for the task in hand.

Several techniques have been introduced to overcome time complexity issues in wrapper methods, specifically, for the subset selection task. For instance, Bayesian network [123], sequential search using aggregation [124], expectation maximization [125], and beam search [126] are some of the example works towards the optimization of feature search in wrapper methods. While these methods yield high classification accuracy, they work better only for the specific models adopted for feature selection. In other words, these methods are computationally expensive as well as lacking in terms of generalization [127].

Some works related to SAR applied wrapper methods in feature selection, i.e., Boruta [64], Sequential Forward Selection (SFS) [69][128], and Recursive Feature Elimination [71].

2.1.8.3 Embedded and Hybrid Methods

Embedded feature selection methods partly use supervised learning and hence, they are relatively faster compared to wrapper methods. These techniques utilize automated pruning, regularization or a built-in strategy to select the candidate feature set. For instance, the SVM model can be used to recursively prune a feature with associated variance less than a set threshold. Likewise, decision tree-based approaches such as CART [129], C4.5 [130] and XGBoost [131] are other commonly used embedded methods for feature selection. A study presented in [132] introduced an embedded method for recursive feature elimination using SVM. Feature significance is measured through the associated weights in the trained model, and then used to iteratively eliminate the least important features [133].

Alternatively, a variety of hybrid methods were introduced by integrating the properties of wrappers and filter methods [102]. For instance, a filter method (e.g., based on variance in PCA or alternative statistical metrics) is applied over the entire feature space to identify significant features, which are then forwarded to wrapper methods with the reduced feature set. In this way, the overall complexity and execution times can be reduced to support the use of recursive procedures within wrapper methods. Thus, a hybrid approach tends to be faster and more general than wrapper methods, but slower and less general than filter methods.

A widely used technique for SAR in the literature is the Random Forest feature selection [46], [5], [60], [56].

2.1.9 Machine Learning Algorithms

Various ML techniques have been used to classify the activities of animals [134]. In this subsection, some popular classification algorithms used in SAR will be described.

2.1.9.1 Instance-based Algorithms

The k-nearest neighbours (KNN) classifier is one of the most popular classifier algorithms in machine learning [135], [136]. KNN is a nonparametric model, where the classification process is based on the similarity between the training and testing samples. Because of the effectiveness and simplicity of the KNN algorithm, it is widely popular in various disciplines, e.g., data science [137]–[140]. The algorithm determines the nearest k neighbours for an unseen sample, and then provides its category based on the maximum frequency label in the k nearest neighbours [141]. Consider a set X of n labelled samples, KNN performs the classification task, as shown in Algorithm 1.

Algorithm 1: KNN Algorithm

Let y represent the unknown sample
  Let $k \in [1, n]$
  Repeat
    Calculate similarly between y and x_i
    If ($i <= k$)
      $x_i \in$ into k nearest neighbour of y
    Else if (x_i is close to y than one of the nearest neighbours)
      Eliminate the farthest neighbour in the k nearest neighbour set
      $x_i \in$ into k nearest neighbour of y
    End
    i++
Until (i > No of Training data)

Support Vector Machine (SVM) is a robust classifier, which utilizes the kernel trick in conjunction with supervised learning [142][143]. The algorithm was first introduced by Boser et al., [144] and further detailed by Cortes [145]. The decision hyperplane generated through SVM depends on the so-called support-vectors. A description of the SVM classifier is provided in Algorithm 2.

Algorithm 2. Support Vector Machine Algorithm [146]

Let S represent a set of m data points where $s = \{\{(x_i, y_i)|i = 1..m\}\}$

Where $x \in R^n$ and $y \in \{1, -1\}$ for binary classification

Let φ be a map function, where $Z = \varphi(x)$ and φ maps the input space to a high-dimensional dot-product feature space. Determine the hyperplane define by $w.z + b = 0$ where $w \in R^m$ and $b \in R$.

$\exists\ (w, b)$

If (S is linearly separable)

$w.z_i + b \geq 1, y_i = 1$

$w.z_i + b < 1, y_i = -1$

Else

$w.z_i + b \geq 1 - \varepsilon_i$

$$\min_{w,b,\varepsilon}\left(\frac{1}{2}w^T w + c\sum_i^m \varepsilon_i\right)$$

$y_i\left(w^T\varphi(x_i) + b\right) \geq 1 + \varepsilon_i, \forall i = 1, \dots, m$

$\varepsilon_i \geq 0, \forall i = 1, \dots, m$

c is a constant

$K(u, v) = \varphi(u)\varphi(v)$

K is the kernel

2.1.9.2 Logistic Regression

Logistic regression (LR) is a supervised learning algorithm that is a generalization of linear regression. LR is used for prediction as well as binary and multi-class classification. Logistic regression involves the calculation of the prediction function, building the loss function, and determining the regression parameters that are capable of minimizing the loss function. Optimal parameters are determined using iterative optimization techniques [147]. Algorithm 3 illustrates the LR steps.

Algorithm 3: LR algorithm [148]

Let S be the prediction function (a sigmoid function)

Let x be the variable

$$S = \frac{1}{1 + e^{-x}}$$

Let g to be the prediction function and L the Loss function

$$g_w(x) = S(w^T x) = \frac{1}{1 + e^{-w^T x}}$$

$$L(w) = \frac{1}{m}\left[\sum_{i=1}^{m} (y_i \log g_w(x_i) + (1 - y_i) \log(1 - g_w(x_i)))\right]$$

Determine w using gradient descent to minimize the loss function

2.1.9.3 Decision Tree and ensemble learning Algorithms

Decision trees are well-known machine learning algorithms used in classification and regression [149][150]. There are various algorithms utilising decision trees (DT) for the classification of data, including ID3 by Quinlan et al., [150], C4.5 by Quinlan [130] and the classification and regression tree (CART) by Breiman et al., [129]. DT algorithms recursively partition the data into subsets, then assigning decision rules to their nodes, as presented in Algorithm 4, which shows an overview of the DT learning process.

Algorithm 4. A general method for learning the DT [129]

Let (X, Y), $\{x_1, x_2 \ldots\ldots, x_d\}$, K_{nots} depth, $depth_{max}$ and $Gain_{split}^{min}$ as the inputs

Let N_0 the initial node

Calculate $Gain_{split}$

For $\alpha \in \{1, \ldots, d\}$

For $\beta \in k_{nots}$

Find $Gain_{split}$

Find $Gain_{split}^{*}$=max($Gain_{split}$, $Gain_{split}^{*}$)

Record optimal split variable x_{v^*} and split the knot β^*

End

End

Update the nodes based on x_{v^*} and β^*

depth ++

if $Gain_{split}^{*} > Gain_{split}^{min} Or\ depth > depth_{max}$

exit

end

Algorithm 4 applied to $(X,Y)_{left}$ and $(X, Y)_{right}$

Ensemble models further extend the functionality and hence the ability of conventional DT, while utilizing the bagging concept. Random Forest is a commonly used ensemble ML model that was initially proposed by Breiman [151]. The algorithm uses the Random Subspaces method [152] and bagging [153] to combine several weak classifiers leading to a robust classification. The algorithm is successfully applied to both prediction and classification tasks. Using the RF, training data are randomly received using subsets to form trees based on a random algorithm [154]. Using the RF algorithm, bootstrap samples (new training set) are selected by substituting the original data set for the tree, allowing a number of training data to be excluded, which can then be reused i.e., out of bag samples. Figure 2-3 demonstrates the steps used in the RF algorithm utilizing the bagging approach for decision making.

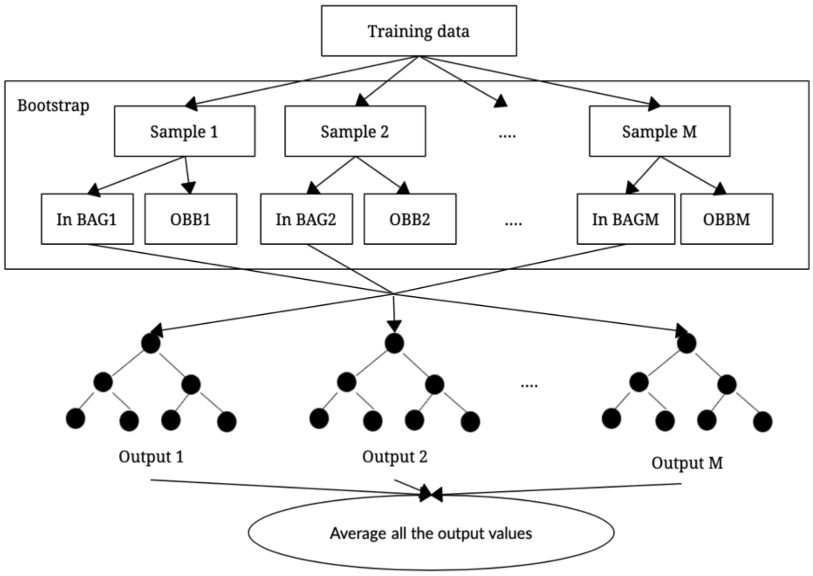


Figure 2-3 The RF Algorithm [306]

2.1.9.4 Naïve Bayes

The Naïve Bayes (NB) classifier has been widely employed in a variety of data mining and classification tasks [155]. The algorithm assumes that the probability of a new data sample can belong to a particular class, when the attributes are independent of each other [156]. The algorithm works as follows. Let D_{train} represent a set of training samples of t classification objects. In this case, let the probability P(y|x) for a new data sample X = <x_1, x_2….x_a> to belong to the class $y \in \{1, \dots, c\}$, where x_i represents the value of the attribute. Algorithm 5 shows the basic steps in NB classification.

Algorithm 5: NB algorithm

Let P(y|x) be the posterior probability of the target class

Let P(y) be the probability of the class

Let P(x) be the probability of the predictor

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

$$P(y|X) = P(x_1|y) \times P(x_2|y) \times \ldots\ldots.. \times P(x_a|y) \times P(y)$$

2.1.9.5 Multilayer Perceptron

Multilayer perceptron (MLP) or artificial neural networks (ANNs) have been commonly used in a variety of domains, including classification and regression [157]–[164] and successfully solved many computation problems of signal processing [165]. Additionally ANNs have been applied to applications such as face recognition [166], texture classification [167], shape recognition [168],and image segmentation [169]. ANNs have also been used to improve generalization [170] The external inputs are presented to the network through the input neurons, while the outputs are shown in the output layer. All other layers are called hidden layers. Each layer has its own weights, biases and transfer functions. The use of more than one layers of nonlinear units makes the network more powerful than a single layer network [171]. As an example, multilayer networks can predict many functions using two layers with sigmoid and linear functions in the first and the second layers, respectively. Multilayered neural networks can be used for pattern classification and function approximation, as well as modelling and prediction [172]. For instance, (Nadimi et al., 2012) used MLP with sheep accelerometer measurements to predict grazing, lying down, walking, standing, and other activities, obtaining an accuracy of 76.2%. Figure 2-4 shows the structure of a MLP network, where f denotes the transfer function. Algorithm 6 summarizes the forward propagation process in a MLP.

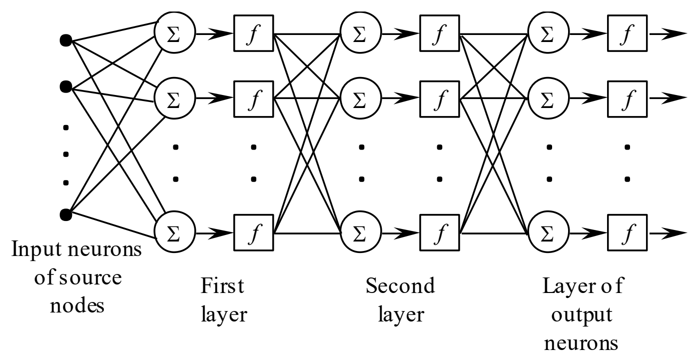


Figure 2-4 MLP network structure

Algorithm 6: The input-output equations for MLP network

Let M to be the number of inputs

Let N is the number of outputs

S is the number of hidden units

Let y represents the N-tuple outputs of the output layer,

Let x represents the M-tuple inputs to the network hidden layer, $n_j = \sum_{k=1}^{M} W^1 x_k$

n_j represents the net sum at the hidden neuron j

The output of this unit is:

$$v_j(n_{j)} = f\left(\sum_{k=1}^{M} W^1 x_k\right)$$

where f is a nonlinear transfer function.

The output of the hidden layer is the input to the next layer and the net input to the output unit i is:

$$n_i = \sum_{j=1}^{S} W^2 V_j$$

and the output unit i produces the following output value:

$$y_i = f(n_i) = f\left(\sum_{j=1}^{S} W^2 V_j\right)$$

2.1.9.6 Deep Learning

Deep neural networks (DNN) is considered as a standard neural network with more hidden layers (i.e., a deeper structure). In this case, the depth of the network is defined by the number of hidden layers. It should be noted that there is no present number of hidden layers for a neural network to be defined as deep, however, Schmidhuber [173] considered that if their credit assignment paths exceed 10, then the corresponding ANN is very deep. The aim of the DNN is to be trained to model complex nonlinearities in the input data by mining unique features. Each of the layers of the DNN aims to extract certain features. Examples of DNN include CNN that are widely used in image processing [174] and deep belief networks. In the literature for SAR, deep neural networks were applied to predict activities such as stationary, foraging, walking, trotting, and running resulting in an overall accuracy of 94% [68].

2.1.9.7 Discriminant Analysis

The sparsity and the high dimensional properties of the data will result in increased complexity. Linear discriminant analysis (Fisher, 1936) has been used for dimensionality and sparsity reduction and represents one of the most favourable tools for the projection into a low dimensional space. There are various applications utilizing linear discriminant analysis (LDA), including image retrieval, speech recognition, and microarray data analysis [176]. Traditional LDA assumes that the dispersion matrices are identical for all classes. When this is not possible, quadratic discriminant analysis (QDA) is used. Algorithm 7 presents the basic steps of the LDA algorithm.

Algorithm 7: Linear discriminant analysis (LDA) [177]

Let $x_{ij} \in R^N$ to be the training sample

Let c to be the number of unknown classes

i, j are the class numbers

Let m_i to be the mean vector of class c_i

Let S_b and S_w to be the between and within class scatter matrices, respectively, where

$$S_b = \frac{1}{n}\sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^T$$

$$S_w = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(x_{i:j} - m\right)\left(x_{i:j} - m\right)^T$$

LDA aims to find a projection that is optimal in separating data classes in a low-dimensional space If U is a set of projection vector, then U is selected to maximize the ratio between S_b and S_w

$$U^* = \underset{U}{\arg max} \; \frac{tr(U^T S_b U)}{tr(U^T S_w U)}$$

Discriminant analysis was used in a study with accelerometer measurements collected from sheep to discriminate between grazing, ruminating, and resting [61]. Furthermore, Discriminant analysis and embedded statistical classifiers were used to study measurements from accelerometer data collected from [63], [178]. Various studies used LDA [21][65] [66] [68] [53] [69] [62] [61] [63] and QDA [21] [46] [5] [65] [56] [69] for classification of animal behaviour.

2.1.10 Discussion

Tabular summaries of the research works surveyed in this Chapter with respect to sensor placement are provided in Tables 2-6 to 2-9 in descending order by Accuracy. These provide information about the type of activities, sample rate, window size, feature selection method, learning model, and accuracy.

Table 2-6 Collar-borne sensors in sheep activity recognition

ACC=accelerometer, GYR=gyroscope, MAGN=magnetometer, RFs = random forest feature selection, SFS= Sequential forward selection, GFS=greedy feature selection

Model	Sensor	Sensor Placement	Activities	Sample Rate	Window (s)	Feature Selection	Accuracy	Ref
RF	ACC	collar	grazing, walking, scratching,	12.5Hz	5	-	99.43%	[52]

			inactive					
LDA	ACC	collar	grazing, not-grazing (tall pasture)	20Hz	10	Relief	98.20%	[53]
LDA	ACC	collar	grazing, not-grazing (short pasture)	20Hz	10	Relief	97.80%	[53]
LDA	ACC	collar	grazing, not-grazing (medium pasture)	20Hz	10	Relief	97.40%	[53]
CART	ACC, ultrasound module	collar	Running, not-running	4Hz	-	-	96.62%	[54]
RF	ACC, GYR, MAGN	collar	grazing, lying, standing, walking, scratching	200Hz	30	Boruta	96.47%	[64]
RF	ACC, GYR	collar	grazing, lying, standing, walking, browsing, scratching	10Hz	10	-	96.43%	[71]
RF	ACC	collar	grazing, lying, standing, walking, browsing, scratching	10Hz	10	-	96.03%	[71]
CART	ACC, ultrasound module	collar	Posture (Infracting and Not Infracting)	4Hz	-	-	95.95%	[54]

XGB	ACC, GYR, MAGN	collar	grazing, lying, standing, walking, scratching	200Hz	30	Boruta	95.85%	[64]
RF	ACC, GYR	collar	Walking, standing, lying	32Hz	5	-	95.00%	[55]
RF	ACC, GYR	collar	Walking, standing, lying	32Hz	7	-	95.00%	[55]
MLP	ACC, GYR, MAGN	collar	grazing, lying, standing, walking, scratching	200Hz	30	Boruta	94.40%	[64]
DNN	ACC, GYR	collar	stationary, foraging, walking, trotting, running	200Hz	1	-	94.00%	[68]
RF	ACC, GYR	collar	Walking, standing, lying	32Hz	3	-	94.00%	[55]
KNN	ACC, GYR, MAGN	collar	grazing, lying, standing, walking, scratching	200Hz	30	Boruta	93.57%	[64]
RF	ACC, GYR	collar	Walking, standing, lying	16Hz	7	-	93.00%	[55]
MLP	ACC	collar	active, inactive	1Hz	-	-	92.30%	[67]
RF	ACC,	collar	grazing,	16Hz	7	Relief	92.00%	[57]

	GYR		ruminating, non-eating (walking, standing, lying)					
CART	ACC	collar	infracting, grazing, standing, moving, running,	50Hz	0.5	oneR	91.78%	[70]
RF	ACC, GYR	collar	Walking, standing, lying	16Hz	5	-	91.00%	[55]
RF	ACC, GYR	collar	Walking, standing, lying	8Hz	5	-	91.00%	[55]
RF	ACC, GYR	collar	Walking, standing, lying	16Hz	3	-	90.00%	[55]
RF	ACC, GYR	collar	Walking, standing, lying	8Hz	7	-	90.00%	[55]
QDA	ACC	collar	grazing, lying, standing, walking, running	100Hz	5.12	GFS	89.70%	[65]
RF	ACC, GYR	collar	Walking, standing, lying	8Hz	3	-	89.00%	[55]
Linear SVM	ACC	collar	grazing, lying, standing, walking, running	100Hz	6.4	SFS	88.40%	[69]
LDA	ACC	collar	grazing,	100Hz	5.3	SFS	82.40%	[66]

			lying, standing, walking, running					
CART	ACC, ultrasound module	collar	Resting vs Not resting	4Hz	-	-	81.31%	[54]
MLP	ACC	collar	grazing, lying, standing, walking and others	1Hz	-	-	76.20%	[67]
QDA	ACC	collar	grazing, standing, walking, resting	12Hz	10	RFs	54%-96%	[56]
QDA	ACC	collar	grazing, standing, walking, resting, lame walking	12Hz	10	RFs	35%-95%	[5]
QDA	ACC	collar	grazing, lying, standing, walking	12Hz	3	RFs	6%-88%	[46]
QDA	ACC	collar	grazing, lying, standing, walking	12Hz	5	RFs	6%-88%	[46]
QDA	ACC	collar	grazing, lying, standing, walking	12Hz	10	RFs	6%-90%	[46]

Table 2-7 Ear-borne sensors in sheep activity recognition

ACC=accelerometer, GYR=gyroscope, RFs = random forest feature selection

Model	Sensor	Sensor Placement	Activities	Sample Rate	Window (s)	Feature Selection	Accuracy	Ref
CART	ACC	ear	active, inactive	12.5Hz	30	-	98.10%	[21]
RF	ACC, GYR	ear	standing, walking, lying	32Hz	5	-	95.00%	[55]
RF	ACC, GYR	ear	standing, walking, lying	32Hz	7	-	95.00%	[55]
RF	ACC, GYR	ear	standing, walking, lying	32Hz	3	-	94.00%	[55]
QDA	ACC	ear	grazing, standing, walking	12Hz	10	RFs	94%-99%	[56]
RF	ACC, GYR	ear	standing, walking, lying	8Hz	5	-	91.00%	[55]
RF	ACC, GYR	ear	grazing, ruminating, non-eating	16Hz	7	Relief	91.00%	[57]
RF	ACC, GYR	ear	standing, walking, lying	16Hz	7	-	91.00%	[55]
RF	ACC, GYR	ear	standing, walking, lying	8Hz	7	-	91.00%	[55]
LDA	ACC	ear	posture (upright and prostrate)	12.5Hz	30	-	90.60%	[21]
RF	ACC,	ear	standing,	16Hz	5	-	90.00%	[55]

	GYR		walking, lying					
RF	ACC, GYR	ear	standing, walking, lying	8Hz	3	-	89.00%	[55]
QDA	ACC	ear	grazing, standing, walking	12Hz	10	RFs	89%-93%	[46]
QDA	ACC	ear	grazing, standing, walking	12Hz	5	RFs	86%-93%	[46]
QDA	ACC	ear	grazing, standing, walking	12Hz	3	RFs	83%-92%	[46]
QDA	ACC	ear	grazing, standing, walking, lame walking	12Hz	10	RFs	82%-96%	[5]
Linear SVM	ACC	ear	grazing, lying, standing, walking	12.5Hz	10	-	76.90%	[21]
RF	ACC, GYR	ear	lame, not lame (within Walking)	16Hz	7	Relief	76.83%	[6]

Table 2-8 Jaw-based sensors in sheep activity recognition

ACC=accelerometer, RFs = random forest feature selection, SDA=stepwise discriminant analysis

Model	Sensor	Sensor Placement	Activities	Sample Rate	Window (s)	Feature Selection	Accuracy	Ref
CART	ACC	Under the jaw	bite, chewing, other	25Hz	5	RFs	96.70%	[59]

			(grazing pasture plots, different sward height treatments combined)					
CART	ACC	Under the jaw	bite, chewing, other (while grazing micro-sward boxes)	25Hz	5	RFs	96.6%	[59]
CART	ACC	Under the jaw	bite, chewing, other (grazing pasture plots, different sward height treatments combined)	25Hz	3	RFs	93.30%	[59]
LDA	ACC	Under the jaw	grazing, ruminating and resting	62.5Hz	60	SDA	93.00%	[61]
CART	ACC	Under the jaw	bite, chewing, other (while grazing micro-	25Hz	3	RFs	90.80%	[59]

			sward boxes)					
LDA	ACC, FORCE	Under the jaw	grazing, ruminating , other	62.5Hz	30	S DA	8 9.70%	[62]
CART	ACC	Under the jaw	bite, chewing, other (grazing pasture plots, different sward height treatments combined)	25Hz	1	RFs	86.10%	[59]
CART	ACC	Under the jaw	grazing, lying, standing, walking, running	10Hz	5	RFs	85.50%	[60]
CART	ACC	Under the jaw	grazing, lying, standing, walking, running	10Hz	10	RFs	83.40%	[60]
CART	ACC	Under the jaw	grazing, lying, standing, walking, running	10Hz	3	RFs	82.90%	[60]
CART	ACC	Under the jaw	bite, chewing, other (while grazing	25Hz	1	RFs	80.40%	[59]

			micro-sward boxes)					

Table 2-9 Leg-based sensors in sheep activity recognition

ACC=accelerometer, RFs = random forest feature selection, SDA=stepwise discriminant analysis

Model	**Sensor**	**Sensor Placement**	**Activities**	**Sample Rate**	**Window (s)**	**Feature Selection**	**Accuracy**	**Ref**
LDA	ACC	leg	walking, trotting, galloping (only horizontal axis was used)	33Hz	3	SDA	90.91%	[63]
QDA	ACC	leg	grazing, standing, walking, resting, lame walking	12Hz	10	RFs	58%-100%	[5]
QDA	ACC	leg	grazing, standing, walking, resting	12Hz	10	RFs	56%-100%	[56]
QDA	ACC	leg	grazing, lying, standing, walking	12Hz	3	RFs	38%-93%	[46]
QDA	ACC	leg	grazing, lying, standing, walking	12Hz	10	RFs	35%-94%	[46]
QDA	ACC	leg	grazing,	12Hz	5	RFs	29%-94%	[46]

			lying, standing, walking					

As mentioned above, accelerometers have been widely used in animal activity recognition due to their ability to distinguish between various behavioural patterns with high accuracy. It is important to note that when using accelerometers, the absolute acceleration feature must be considered as it diminishes the effect of sensor orientation, which can adversely affect the performance of predictive models [46]. Indeed, it has been previously reported that even a small change in sensor positioning could alter the results [179]. Some studies combined accelerometers with gyroscopes and magnetometers, however, from the above tables, e.g., [55], [64], [71], it is observe that accuracy was not significantly increased and therefore, it is suggested that an accelerometer sensor suffices in accurately identifying animal behavior.

There is a common trend in attaching the sensor on the collar and ear. A sensor attached on the collar successfully classified grazing, walking, scratching, and inactivity with accuracies above 99.13% for all activities [52]. On the other hand, a collar sensor is not recommended when the purpose of the study is concerned with lameness [5]. Lame walking was classified with an accuracy of 87% when the sensor was attached to the leg, and 82% when the sensor was attached to the ear [5]. On the other hand, grazing behaviour can be identified with an accuracy in excess of 97%, when the sensor is attached to the collar [52], [53]. It should be noted that high performance does not solely depend on sensor placement, but rather it is achieved through a combination of factors, including the employed features, ML techniques, and window size. When reviewing the locations of sensor placement, the advantage of ear-based sensors is that they can be integrated into the already existing ear-tags on animals. Based on the reported results of the reviewed studies in Tables 6-9 [21], [52], [53], [59], we can conclude that one sensor per animal suffices in producing high predictive results in identifying behavioural patterns.

The selection of the window size and sample rate is vital in achieving high classification rates. It should be noted that the choice of window size always depends on the activity we wish to detect. Smaller windows can distinguish better between some behaviours. For example, a study reported that the classification of biting, chewing, and other feeding behaviours, through signals collected from under the jaw of the animal, increased accuracy when the window size was increased [59]. On the other hand, another study evaluated 3, 5, and 10-second windows and noted that there was no significant difference in the classification accuracy of grazing behaviour; however, the highest accuracy for running was predicted using the 10s window [60]. Having a larger window size in real-time animal activity classification may lead to mislabelling, as the animal may exhibit more than one behaviour in a short time interval. Therefore, a 5-second window is considered an appropriate selection [52].

Additionally, it was noted that a lower sampling frequency improves memory usage and is less power demanding [180]. A study evaluated the effect of sampling frequency and window size on power consumption to identify sheep behaviour and suggested that a sampling frequency of 16Hz using 7s windows offers benefits due to a decreased demand in power [55]. The study yielded higher accuracy using a 32Hz sampling rate, however, the results were close to those reported for 16Hz [55]. In a previous study, spectral analysis was performed on a sheep dataset collected with a sample rate of 100Hz, and it was observed that there is limited spectral information above 10Hz [66]. To summarize, the sampling rate and window size need to be chosen based on the specific animal activity recognition problem, in the context of memory and power consumption application constraints.

Additionally, during the course of this research study, it became evident that there is no universal method for feature selection in sheep activity recognition using accelerometer signals. From the reviewed studies (Section 2.1.7: Table 2-4 and Table 2-5), it is observed that the majority of works use time-domain features as they are computationally efficient. Frequency-domain features are robust to noise, however, they are more computationally expensive, thus requiring more power [92]. A variety of time and frequency domain features have been considered in the literature, including their importance in the classification task, using feature selection methods. The most commonly used feature selection approach is Random Forest. Regarding model selection for classification, it was observed that the most commonly used methods are LDA, QDA, RF, CART, and KNN (refer to Table 2-6, Section

2.8). However, from Tables 6-9, it can be observed that the highest accuracies are achieved using RF, LDA, and CART. Given the different settings of each study, i.e., type of animal activities, and sensor position, a direct comparison in regards to performance is not straightforward. Tables 2-10 to 2-13 provide an overview of the applied feature selection methods and final set of features, depending on the location of the sensors. These are sorted in terms of descending accuracy. It is envisaged that the information presented in these tables could be used to provide guidance in future research studies subject to the sensor position, activity type and overall system requirements.

Table 2-10 Top performances in collar-borne sensors: Feature selection and feature set

ACC=accelerometer, GYR=gyroscope, MAGN=magnetometer, FS = feature selection, acc=accuracy

Model	**Sensor**	**FS**	**Final Features**	**Activities**	**Results per activity**	**Accuracy**	**Ref**
RF	ACC	correlation >80	mean, crest factor, root mean square velocity, skewness, kurtosis, madogram, zero crossing rate, squared integrals, and signal entropy	grazing, walking, scratching, inactive	acc 99.08% acc 99.13% acc 99.90% acc 99.85%	99.43%	[52]
LDA	ACC	Relief	entropy az, mean az, mean gy, mean gx, entropy ay	grazing, not-grazing (tall pasture)	acc 98.2%	98.20%	[53]
LDA	ACC	Relief	mean az, entropy az, mean gy, entropy ay, mean gx	grazing, not-grazing, (short pasture)	97.80%	97.80%	[53]

LDA	ACC	Relief	entropy az, mean gy, mean az, entropy ay, mean gx	grazing, not-grazing (medium pasture)	97.40%	97.40%	[53]
CART	ACC	No	the 3-axis dynamic acceleration results only considered one axis to differentiate activity	Running/not-running	96.62%	96.62%	[54]
RF	ACC, GYR, MAGN	Boruta	mean, sd,RMS,RMS velocity, energy, sum of changes, mean absolute change, integrals , squared integrals, madogram, peak frequency	grazing, lying, standing, walking, scratching	sens 97.66% spec 97.74% sens 93.22% spec 99.76% sens 97.32% spec 98.50% sens 96.23% spec 99.53% sens 95.70% spec 99.74%	96.47%	[64]
RF	ACC, GYR	NA	mean, sd, RMS, RMS velocity, energy, sum of changes, mean absolute change, integrals , squared	grazing, lying, standing, walking, browsing, scratching	sens 94.90% spec 98.21% sens 97.29% spec 99.34% sens 95.48% spec 97.46% sens and spec 100% sens 78.91%	96.43%	[71]

			integrals, madogram, peak frequency		spec 99.99% sens 90.91% spec 100%		

Table 2-11 Top performances in ear-borne sensors: Feature selection and feature set

ACC=accelerometer, FS = feature selection

Model	Sensor	FS	Final Features	Activities	Results per activity	Accuracy	Ref
CART	ACC	No	average, average all axis, minimum, maximum, sd, average sd, movement intensity, SMA, energy, entropy, movement variation	Active, inactive	sens 97.4% spec 98.5% sens 98.5% spec 97.4%	98.10%	[21]
QDA	ACC	RF	Movement Variation, Average Intensity, Average y-axis of accelerometer	grazing, standing, walking	acc 94% acc 96% acc 99%	94-99%	[56]

Table 2-12 Top performances in leg-mounted sensors : Feature selection and feature set

ACC=accelerometer, FS = feature selection

Model	Sensor	FS	Final Features	Activities	Results per activity	Accuracy	Ref
QDA	ACC	RF	Average x, SMA, Average Intensity	grazing, lying, standing, walking, lame walking	acc 89% acc 100% acc 58% acc 64% acc 87%	58%-100%	[5]
QDA	ACC	RF	Average x, SMA, Movement Variation	grazing, lying, standing, walking	acc 91% acc 100% acc 56% acc 100%	56%-100%	[56]
QDA	ACC	RF	Average x , SMA and Movement Variation	grazing, lying, standing, walking	acc 76% acc 38% acc 48% acc 93%	38%-93%	[46]
QDA	ACC	RF	Average x, SMA and Movement Variation	grazing, lying, standing, walking	acc 81% acc 35% acc 47% acc 94%	35%-94%	[46]
QDA	ACC	RF	Average x, SMA and Movement Variation	grazing, lying, standing, walking	acc 80% acc 29% acc 48% acc 94%	29%-94%	[46]

Table 2-13 Top performances in jaw-mounted sensors: Feature selection and feature set

ACC=accelerometer, FS = feature selection

Model	Sensor	FS	Final Features	Activities	Results per activity	Accuracy	Ref

CART	ACC	RF	mean of Movement Variation and Energy	Bite, chewing, other (grazing pasture plots, different sward height treatments combined)	sens 97.4% spec 97.7% sens 96.3% spec 96.8% sens 95.4% spec 100%	96.70%	[59]
CART	ACC	RF	mean of Movement Variation and Energy	Bite, chewing, other (while grazing micro-sward boxes)	sens 96.3% spec 98.4% sens 95.1% spec 97.4% sens 100.0% spec 99.0%	96.60%	[59]

Studies with the sensors attached to the animal collar reported high accuracy in identifying various types of behaviour. The best results, i.e., in the range of 97.40%-99.43%, were obtained from RF and LDA. RF yielded an overall accuracy of 99.43% for grazing, walking, scratching, and inactive [52]. On the other hand, LDA yielded an accuracy of 98.20% in binary classification, i.e., grazing, and non-grazing [53]. These two studies used different combinations of features. For example, the RF approach used features such as the mean, crest factor, root mean square velocity, skewness, kurtosis, Madogram, zero crossing rate, squared integrals, and signal entropy to obtain an accuracy of 99.08% for grazing [52]. Guo et al., [53], used entropy and mean with LDA and achieved an overall performance of 98.20% for grazing and non-grazing. Additionally, Cardoso et al., [54] using only information from one accelerometer axis, correctly classified running with an accuracy of 96.62%.

When the sensor was attached to the ear and CART was used for classification with a feature set including mean of each axis, mean of all axis, minimum, maximum, sd, average sd, movement intensity, SMA, energy, entropy, the system resulted to an accuracy of 98.10% in distinguishing between active and inactive behaviours [21]. On the other hand, a study gathered measurements from a leg-mounted sensor to discriminate between grazing, standing and walking, using QDA and using only three features, namely, movement variation, average intensity, and average of y-axis, achieved accuracies for the activities in the range of 94%-99% [56]. This indicates that there are complex relationships between the location of the sensor, the activities that we seek to distinguish and the selected features, which impact on the performance of the ML system.

Only a handful of research works have been reported for sensors mounted on the leg and under the jaw. The classification of activities using measurements obtained from the leg yielded unbalanced accuracies (refer to Table 2-12). For example, accuracies of 89%, 100%, 58%, 64%, and 87%, respectively, were reported for grazing, lying, standing, walking, and lame walking, using QDA and only three features, i.e., average of x-axis, SMA, and average intensity. Thus, it can be inferred that these three features play an important role in discriminating lying (100%), however, they demonstrate poor performance in discriminating standing and walking [5]. Studies with the sensor placed under the jaw show great potential in analysing and understanding feeding behaviour. For instance, Alvarenga et al., achieved accuracies of 96.60%, and 96.70% using CART, and identified that the most important features were the mean of Movement Variation and Energy to discriminate between biting, chewing, and other activities. Energy is claimed to successfully discriminate between sedentary activities [181]–[183].

However, as previously stated, there is no universal approach for animal activity recognition. Instead, each activity recognition system needs to be designed according to the specific aim, objectives, environment, available datasets, etc. Depending on the problem settings, each set of sensor configuration, feature extraction and ML techniques have associated advantages and disadvantages, which primarily relate to the context of the investigated activities.

Automatic monitoring and detection of sheep behaviour using accelerometers and ML is an important research topic, since the provided information offers potential for more efficient decision-making in terms of animal welfare as well as land utilization. In this Chapter, the problem of sheep activity recognition was considered in terms of its essential building blocks. Several system aspects were the focus of this contribution, including sensor type and positioning, window size and sampling rate, feature extraction, feature selection, and classification methods in relation to sheep activity recognition. Additionally, an overview of the foundations of the utilized techniques and the opportunities and challenges was presented.

Based on the surveyed works, this extensive literature review indicated that each of the challenges can be solved differently, depending on the problem at hand and the availability of data. For example, an essential aspect in system design is the choice of activities to be identified, as this will inform the solution methodology, i.e., sensor placement, window size, feature selection, and choice of the classification algorithm. From the review of the state-of-the-art, it was identified that lameness recognition in sheep using accelerometer signals is a problem that has not been sufficiently studied, thus offering opportunities for novel contributions in this field.

Most research works focus on using collar-borne and ear-borne accelerometers. Animal activity prediction results indicated that these two sensor positions result in higher accuracy, when compared to sensors mounted on the leg and under the jaw. Hence, more research should be conducted to further explore the advantages of using leg and under the jaw based sensors. Indeed, a leg-borne sensor can provide valuable information regarding movement activities, while the signals collected from the jaw of the animal could provide more information regarding feeding activities, such as grazing, biting, chewing, and ruminating, which are critical behaviours for the sheep industry as well as for conservation purposes.

Section 2.2 will provide details regarding virtual fencing systems and animal auditory awareness.

2.2 Virtual fencing systems

In the following sections, a brief background on VF vs traditional fences will be provided. Additionally, the key aspects for the development of a VF will be introduced as well as discussion about opportunities and challenges. The background research was conducted to gain an understanding on how sheep could be treated and trained for the use of the VF. The following information help build an understanding on the ability of the animals to learn, and how they react as a group or individually. The information formed the basis of how the sound experiments on the animals were conducted (the experiments are described in Chapter 5).

2.2.1 Virtual vs Traditional fencing

A VF system can improve the way the lands are utilised by controlling the position of the animals while they graze. Additionally, some areas are difficult to be fenced using traditional fencing methods, due to terrain or size, as it could result in excessive costs. Thus, considering VF solution, might result in the promotion of conservation grazing which can create habitats needed for rare species and improve biodiversity [184]–[186].

One more advantage is that a VF could replace traditional fences and improve animal welfare and environment. For example, traditional fences can impact negatively the environment as maintenance of the fences and construction can interrupt the natural habitat or cause injuries to wildlife [187]. Moreover, barred wires may also injure the animal and can even cause death.

Fencing large areas or moving the fences is costly, therefore, a VF system can reduce the costs needed and a rough estimate of the costs could be comparable favourably with traditional fencing tools [33]–[35].

Another advantage of a VF system is the flexibility they can offer in comparison with traditional fences. A VF can be managed and adapted fast in new seasonal requirements and can easily be moved. Also, animals can be gathered remotely, as a VF can also offer the opportunity for animals to be virtually herded, resulting in replacing labour intensive tasks into cognitive labour [33], [188]. Remote monitoring and controlling can reduce overgrazing, preventing soil erosion and contamination [39].

On the other hand, VF are not 100% stock-proof and cannot be considered when absolute control of the animals is required. For example, if the animals are in an area where humans and central roads are nearby, the use of traditional fence is necessary since it could be dangerous for both animals and humans. Additionally, VF cannot keep predators away as is the case with a traditional fence.

A summary of advantages and disadvantages of VF vs TF is presented in Table 2-14. In the next section, the development approaches for virtual fencing systems are presented.

Table 2-14 Virtual vs Traditional fencing

	Advantages	Disadvantages
Traditional fence (conventional and electric fence)	Have control over animals and is 100% stock-proof Electric fence, more flexible than conventional fence	Expensive Conventional fence is not flexible Can cause injuries to animals (wildlife and livestock) Can be damaged if animals go against it Regular maintenance and monitor Can negatively impact the environment and the welfare of the animals
Virtual fence	Cost effective Flexible Reduce overgrazing and prevent soil erosion if used as virtual herding means Can fence areas that are difficult to be fenced	Not 100% stock-proof Cannot keep predators away Are not visible and there are risks of accidents if humans are near an animal while a stimulus is given

2.2.2 Virtual fencing: past and present development and approaches

Lot of research is conducted for the investigation of the viability of virtual fences. In 1973, Peck [189] patented a system to keep dogs within a specific area. This was the first invention intended to keep animals in a specific area without visible fences. The system used a wire placed on the ground, and a collar or blanket that was worn by the dog. When the animal approached the wire, the collar provided electric shock. However, this patent raises issues of inconvenience when the animal is outside the wire and receive the same response once it tries to move back to the allowed area.

In 1989, Fay *et al*. [190] evaluated the feasibility of electric shock collars to contain goats within a designated test area, noting that the electric-collars had the potential to contain animals without a fence. The research indicated that collared goats stayed within the test area and the remaining goats stayed closed to the collared goats due to herd instinct. The goats learned to avoid the shock, after 30 minutes of training.

In 1990, Quigley *et al*. [191] used dog training collars on 4 cattle. The trial lasted for 4 days and the correct responses of the animals to stay within the defined area under the stimulation of the collar was 83%, 93%, 97% and 100%, for day 1, 2, 3 and 4, respectively. The authors noted that the animals learned to associate the audio with the electric stimulator and suggest that it might be an option to use only audio to train cattle in the future.

A year later, Rose [192] invented the “Animal Control Device”. The device consisted of signal transmitter/receiver (low power radio frequency (RF)) and an animal implant Electrical Simulator. The simulator was implanted into the animal’s nose or upper lip through surgery. When the animal approach the boundary, electric shock was transited through the implant. This patent was abandoned and continued in 1994 and 1996 where the researcher worked on a different patent in which nose clip and solar cell were invented to improve the battery life of the previous system[193].

Anderson *et al*. [194] tested the abovementioned VF and concluded that the system was 100% effective in controlling the animal’s distribution without any detectable stress which was determined from heart rate measurements. Since the animal’s behaviour cannot be 100% predictable and the VF depends also on the behaviour of the animal to be effective, the authors indicated that the system should not be considered for controlling animals if health or safety issues of the animals or humans would be compromised. However, it could be considered for managing animal distribution. Anderson *et al*. [195] tested again the Directional Virtual Fencing system (DVF) in two cows. Both studies indicated that the system could monitor the direction of the animal. The authors suggested that for more animals, further investigation of the system and animal training is necessary.

Associative learning by cattle to enable the use of VF was also examined by Lee et al. [196]. The authors tested the hypothesis that the animals would exhibit associative learning by responding to a sound alone to stay in the virtual boundary.

Umstatter et al. [187] conducted a study to investigate whether an electric shock could be replaced with aversive sounds. Initially they investigated 20 different sounds and for the final testing they used 2 sounds which generated the highest alerting response. For the final testing, 8 cows worn collars with attached loudspeakers which generated sounds when the animals approached the boundary. The authors found good responses, but these were not consistent enough since only 53% of the cows responded successfully. However, this was only a preliminary study and more research is needed in order to develop a system which alerts the animals without electric shocks.

Umstatter *et al.* [197] considered the control of cattle location using broadcast audio instead electric stimuli. The experiment used 38 animals and placed loudspeakers around a small paddock. The loudspeakers played irritating sounds in the range of 8kHz, and 8-10kHz, and 1 acute sound. Movement sensors were located close and linked to the loudspeakers and triggered the sounds when the animals approached the restricted area. According to the authors, irritating and acute sounds have a potential to control cattle location, however it is not sufficiently effective to replace conventional fences and more research is needed.

Umstatter et al. [198] tested a new commercially available virtual fencing system, namely Boviguard, to investigate its efficacy and to measure its impact on animal behaviour. "The Boviguard (Agrifence, Henderson Products Ltd., Gloucester, UK) consists of cow collars, a battery-based transformer, and an induction cable laid or buried on the ground. As the Boviguard collar comes close to the induction cable, a warning sound is triggered and if the animal continues to move closer, an electrical stimulus is triggered. The system was tested on 10 cows wearing the collars and the authors concluded that this system can be used as an animal management solution. Though, issues exist when the animal leaves the wired area because when it tries to move back, the warning sound and electric shock are stimulated again.

On the other hand, Campbell et al. [199] noted that the animals can learn the VF system and associate their learning to the auditory cues instead of the spatially restricted area. For their study, 6 angus heifers carried a virtual fencing collar device on their necks, which was based on GPS monitoring (eShepherd™, Agersens, Melbourne, VIC, Australia). The boundary was transmitted to the collar through radio frequency link and reacted with auditory stimuli when necessary and was followed by electric stimuli if the animal did not respond successfully to the sound. The authors experimented with several exclusion boundaries and noted that the VF is possible to be moved and that the animals can associate their learning based on the stimulus.

Limited studies were concerned with the implementation of a VF for sheep. Brunberg et al. [200] investigated a virtual fencing system consisting of collars which stimulate sound and a weak electric shock when the sheep crosses the border. The “Nofence” collar [201] is especially developed for sheep and is based on GPS technology. The study conducted 2 experiments in which the first experiment included 24 ewes while the second included 3 groups of 3 ewes in each group. From the 1st experiment, out of the 24 ewes, only nine associated the sound with the electric shock. Some animals run uncontrollably, and others didn’t react at all. When the authors conducted the 2nd experiment, the 9 ewes from the first experiment were used. The first group (3 ewes) run away from the experimental arena. The authors concluded that it is challenging for the animals to generalise their learning and suggest that new training might be needed each time the boundaries are moved.

In a similar study, Brunberg *et al.* [202] investigated the ability of ewes with lambs to learn a virtual fencing system. The technology used, namely Nofence, is a GPS virtual fencing system designed to keep sheep within a predefined area. The collars worn by the sheep provide a sound signal when the animal crosses the boundary and a weak electric shock if they continue to walk towards the restricted area. During the experiments, there were some technical problems with the collars that may have affected the results. Therefore, the authors concluded that the Nofence prototype was unable to keep the sheep within the predefined area and thus cannot replace the conventional fences.

Campbell et al. [45] fitted cattle with automated collars that emit audio (785 Hz ± 15Hz) for 2 seconds followed by an electric shock, to restrict animal’s access from hay. The results showed that the cattle successfully stayed away from the hay attractant, however there was high variation in the animals’ learning rate. The authors suggested that the learning may be influenced by group dynamics and the animals’ temperament, therefore they suggested further investigation. Campbell et al. also examined the effects of a VF in comparison with an electric tape fence on cattle by Fecal Cortisol Metabolite (FCM[2]) concentrations [203]. The authors used pre-commercial prototype collars (eShepherd™, Agersens, Melbourne, VIC) that emit audio followed by an electric shock in the animal training. Results showed that there were no differences in FCM levels between the two types of fences. Additionally, the results indicated that the VF could contain the animals for four weeks in the defined area without any welfare or behavioural impact on the animals.

The prototype VF system collar (eShepherd™, Agersens, Melbourne, VIC), was also used by Lomax et al. [44]. The study evaluated how individual cattle learn the VF. The results showed that the system could contain cattle in the defined areas for 99% of the time; however, there were significant variations between individuals of the time of interactions. The same prototype collar was also tested again by Campbell et al. in order to restrict cattle from entering a defined boundary [204]. Results showed that the animals were almost excluded from the restricted areas. All animals received audio and electric stimuli; however, the approach to the fence varied within animals. The authors noted that the group's behaviour might have facilitated associative learning for each animal. Campbell in another study, used the same collar (eShepherd™, Agersens, Melbourne, VIC) to examine if the system could exclude cattle from an environmentally sensitive area [205]. Results indicated that the cattle learned fast to respond to the VF, being able to respond to the audio cue alone (74.5% of 4378 audio signals), and the animals were excluded from the restricted area at the rate of 99.8%.

Muminov et al. [206], fitted goats with collars able to emit sounds, followed by an electric shock to assess whether the system can keep the goats away from a restricted area. The system combines a VF system and animal behaviour monitoring. The VF used various sounds to avoid giving electric shocks to the animals. The dog and emergency sounds were the most successful, resulting in reducing the number of electric shocks during each day of training.

2

Many implementations of virtual fencing systems relied on a perimeter cable to establish the boundary line. Therefore, McSweeney et al. [188] implemented a wearable GPS system that does not need such cabling, and attempted to train cows to associate an audio warning with boundary encroachment. In their study they tested 1) audio + electric shock, 2) vibration + electric shock, 3) audio only, and 4) vibration only. The authors observed a reduction in grazing and ruminating activity during training. The animals developed a learning association between the warning and the consequence (electric shock). The results suggest that the cows can learn the VF boundary, however, further experimental work is needed to assess best implementation protocols.

A GPS-based dog training equipment was used that administers audio (70-80 dB, 2.7 kHz), followed by an electric shock in order to exclude sheep from the exclusion zone was used by Marini et al. [207]. The study investigated the effects of a VF to control a small flock of sheep, including leaders, middle, or followers, to move them through a laneway. The VF was successful at keeping animals away from the exclusion zone when 100%, and 66% of the animals that had the VF implemented. The results indicated that the VF could control two-thirds of the flock and is equally effective as virtually fencing the whole flock. On the other hand, controlling one-third of the flock was not effective.

The temperament of sheep in the learning of the VF and the importance of an audio warning was examined by Marini et al. [208]. The sheep were fitted with manually controlled collars that emit an audio warning sound for 2 seconds, followed by an electric shock, when the animal approaches the VF boundary. The results showed that the animals were able to learn the audio warning to avoid the electric shock, however the authors were not able to determine whether the animal's temperament could be associated with their learning ability and they suggested testing larger groups in the future.

Kearton et al. [209] aimed to assess the stress responses of sheep based on various stimuli. The authors applied several treatments involving; 1) collars emitting a beep tone (45-55 dB, 2.7 kHz) for two seconds with two-second intervals, 2) Speakers placed in the area played dog barking (58–68 dB, 6.1 kHz), 3) manual restraint of the sheep, and 4) an electric shock triggered from collars. The results suggest that dog bark, the beep tone, and the electric shock, are aversive than the manual restraint of the animals. Additionally, the electric shock was extremely aversive in comparison with the beep tone and the dog bark.

Colusso et al. [210] used the pre-commercial prototype collar (eShepherd™, Agersens, Melbourne, VIC) to assess whether the learning and response to the VF differs when the cattle are trained as individuals or in groups. The results indicated an influence of group members on the response of individuals, as cows trained in a group were more likely to interact with the boundary, compared with the cattle who were trained as individuals.

In the past couple of years, there is also an interest in the use of drones that emit sounds to alert animals move away from a virtual boundary, as drones can be perceived as a thread to the animal [211], [212]. Aversive stimuli such as olfactory cues have been proposed to make animals stay in the desired area [211], however the research in herding animals using drones is very recent and limited [212].

According to the abovementioned studies, training, and teaching animals to respond and learn VFs based on several means of stimulation is a feasible but a challenging task. Additionally, studies investigated temperament and group dynamics in the associative learning of the system. Results revealed that the VF system can be successful, however more research is needed. Additionally, it is noted that this area is extensively studied with cattle, but limited studies exist with goats and sheep. Table 2-15 provides brief descriptions of VF experiments from the literature in chronological order. Two examples of already commercialised VF systems are illustrated in Figure 2-5 and Figure 2-6.

Table 2-15 Brief description of VF experiments

Ref	Year	Animal	Stimuli	Description
[189]	1973 (patent)	Cats and dogs	Audio + electric shock	The animal wears a collar or a blanket. A wire surrounds an area, and the device emits the stimulus once the animal is approaching the wire
[190]	1989	Goats	Beep tone + electric shock	Collars emitting a beep tone followed by an electric shock when the animals entered the transmission zone
[213]	1990 (patent)	Dogs	Warning sounds + electric pulse	Transmitter/ receiver borne on the collar or the harness of the animal. There is no wire necessary. The sounds are emitted when the animal is entering

				a warning zone. There are different entrance zones
[191]	1990	Cattle	Audio + electric shock	Collar containing a radio receiver and an electrical stimulator. The electric shock is applied when the animal approaches the aversion area.
[192]	1991 (patent)	Cattle and sheep	shock	An animal implant electrical stimulator implanted in the animal's nose or upper lip. Once the animal approaches the boundary, a radio signal causes a minor shock
[193]	1996 (patent)	Cattle	shock	A nose clip that pulses shock when the animal approaches a restricted area
[214]	1995 (patent)	cows, sheep, pigs, goats and horses	Audio + electric shock	Ear-tags emitting audio (preferably at 850Hz) and electric shock once the animal enters the exclusion zone
[215]	1997 (patent)	Livestock	Audio + electric shock	A portable unit with a GPS receiver sends an audio signal and electric shock when the animal approaches the defined boundary. The GPS designated location is compared with the device's location.
[216]	1999	Cattle	Audio + electric shock	Ear-tag wore by cattle that emits audio and electric stimulus when the animal approaches the exclusion area
[217]	1999 (patent)	Livestock	Release of a substance	The device would release a substance as a corrective measure
[218]	2001 (patent)	Livestock	Audio + electric shock	A device is worn externally or inserted in the ear canal of the animal. The stimulus is applied to the left or right side of the animal's ear to provide a directional movement when the animal goes through a defined boundary.
[219]	2003	Cow	Audio + electric shock	The Directional Virtual Fencing system (DVF) used was the one referred at [218]
[195]	2004	Cows	Audio + electric shock	Refer to [218]
[220]	2004	Cows	Audio	A collar comprising of GPS, wi-fi, and sound amplifier. The collar emitted sound when the animal was approaching the restricted area
[221]	2006	Cows	Audio	The collar randomly emits sounds that are scary to the animals (a roaring tiger, a barking dog, a hissing snake) to move the animals to the desired location

[222]	2007	Cattle	Electric shock	Electronic shock collars used to deliver electric cues to the animal
[223]	2007	Cattle	1) electric shock 2) audio + shock 3) vibration+ shock 4) light + shock 5) electric fence + shock	The collar administered cues for 3 seconds followed by an electric shock to prevent the animals from cross the fence line
[224]	2008	Cows	Audio (human voice and sounds from gas-powered all-terrain vehicle)	The Directional virtual fencing system [218] was used to gather two groups of cows into a corner corral using only audio
[225]	2008	Bulls	Electric shock	Collars containing GPS and providing an electric shock to control bulls during mating. The shock was emitted when the bull approached the nonallowed animal.
[196]	2009	Cattle	1) Audio + shock 2) only shock	GPS collars administering audio warning and shock to the animals when approaching a restricted zone
[187]	2009	Cattle	Audio (various irritating and acute sounds)	Randomly selected sounds were administered from the collar when the animal was approaching the VF.
[226]	2012	Sheep	Audio + electric shock	A commercially available dog training collar was used, which is connected with a wire lying on the ground. It emits a warning sound for 2 seconds, followed by a shock if the animal approaches the wire.
[197]	2013	Cattle	Audio (8 kHz, mix of 8-10 kHz, and acute alarming sounds)	Loudspeakers were placed on the ground, and sounds were triggered when the animals approached the restricted area
[198]	2015	Cows	Audio + electric shock	The authors used the commercially available VF system, "The Boviguard (Agrifence, Henderson Products Ltd., Gloucester, UK). An induction cable is placed on the ground and is connected to the animal's collar. Once the animal approaches the line, a warning sound followed by an electric stimulus is triggered.
[200]	2015	Sheep	Audio + electric	A GPS-based collar emitting sounds followed by

			shock	electric shock once the animals approach the restricted area.
[199]	2017	Cattle	Audio + electric shock	GPS-based collar device worn by cattle triggering audio warning followed by an electric shock if the animals approached the restricted boundary
[202]	2017	Ewes and lambs	Audio + electric shock	GPS-based collar (Nofence collar developed for sheep) emitting sounds followed by an electric shock to restrict access to the animal
[227]	2018	Sheep	Audio + electric shock	Commercial dog training collars were mounted on the sheep and emitted audio for 2 seconds, followed by an electric shock to keep the animals in the desired area
[228]	2018	Sheep	Audio + electric shock	Manually controlled collars deliver audio and an electric shock to the animals to keep the animals away from an attractant
[45]	2018	Cattle	Audio + electric shock	Automated collars provided an audio and electric shock to animals to restrict them from hay attractant
[208]	2019	Sheep	Audio + electric shock	Manually controlled collars emitted audio for 2 seconds, followed by an electric shock when the animal approached the defined VF
[209]	2019	Sheep	1. Audio 2. electric shock	Several treatments were conducted to assess stress responses: 1. Remotely controlled collars emitting a beep tome (45-55 dB, 2.7 kHz) for two seconds with two-second intervals 2. Speakers placed in the area played dog barking (58–68 dB, 6.1 kHz) 3. manual restraint 4. electric shock was triggered from dog control collars
[203]	2019	Cattle	Audio + electric shock	The pre-commercial prototype collar (eShepherd™, Agersens, Melbourne, VIC) system was used that emits audio followed by an electric shock when the animal approaches the restricted area
[44]	2019	Cattle	Audio + electric shock	Cattle were fitted with a prototype collar (eShepherd™, Agersens, Melbourne, VIC) and emitted audio followed by electric shock in the animal approached the restricted area
[206]	2019	Goats	Audio + electric shock	The goats were fitted with collars able to emit sounds, followed by an electric shock to keep the animals in the predefined area. Sounds included

				dog bark, ultrasound, emergency sound, lion, and tiger sound.
[204]	2019	Cattle	Audio + electric shock	The pre-commercial prototype collar (eShepherd™, Agersens, Melbourne, VIC) system emits audio followed by an electric shock when the animal approaches the restricted area.
[188]	2020	Cows	1) Audio + electric shock 2) vibration + electric shock 3) audio 4) vibration	A wearable GPS-based collar was used. The cows received audio warnings for 2 seconds, followed by an electric shock when the animals approached the defined boundary
[207]	2020	Sheep	Audio + electric shock	A GPS-based dog training equipment was used that administers audio (70-80 dB, 2.7 kHz), followed by an electric shock in order to exclude sheep from the exclusion zone.
[205]	2020	Cattle	Audio + electric shock	The pre-commercial prototype collar (eShepherd™, Agersens, Melbourne, VIC) system emits audio followed by an electric shock. The trial conducted to assess whether the VF could exclude the cattle from an area of regenerating saplings
[210]	2020	Cattle	Audio + electric shock	The pre-commercial prototype collar (eShepherd™, Agersens, Melbourne, VIC) system emits audio followed by an electric shock. The trial conducted to assess whether the learning and response to the VF differs when the animals are trained as individuals or in groups

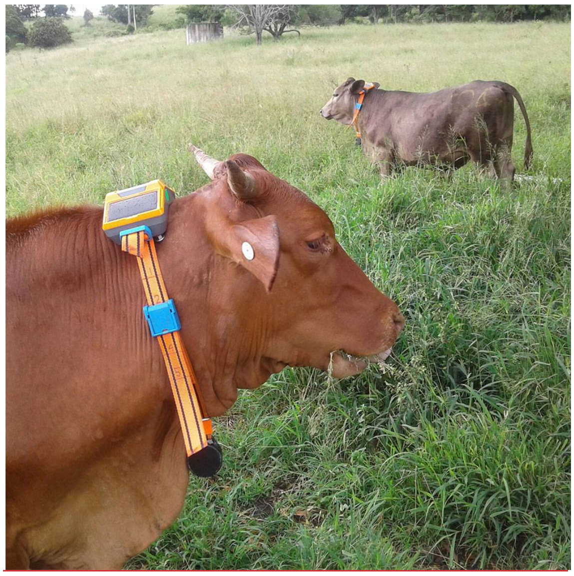


Figure 2-5 The eShepherd® virtual fencing system for cattle The system is commercialised by Agersens (Melbourne, VIC, Australia) and uses licensed IP (Intellectual Property) developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO). The system uses GPS technology, a base station, and collars worn by cattle. The fences are created using an online interface. The device sends an audio tone when the animal approaches the boundary, followed by an electric pulse if the animal continues moving forward. More information can be found in [205]	Figure 2-6 NoFence® virtual fencing for cattle, and sheep/goats, respectively [201] The Nofence is GPS-based virtual fencing system. The animals wearing a Nofence collar receive a sound signal when crossing the virtual boundary followed by an electric shock if they continue to walk from the virtual enclosure [202].

2.2.2.1 Main means of aversive cues and alternatives

As mentioned in the previous section, the main way that a VF works is by sending an audio signal that acts as a warning and then followed by an electric shock that acts as the punishment. However, one problem is that electric shocks can have a negative impact on the animals' welfare and is not ethically accepted worldwide [33] and are banned in countries such as Wales, Germany, Switzerland, and are likely to be banned in the whole UK [33], [229]. For example, a collar that uses electric shock as a punishment might be faulty and might shock the animal repeatedly, resulting in a lot of stress. Additionally, a faulty collar might interfere with the learning of the animal as it might not give the chance to correctly learn since an audio warning is missing. Therefore, new means of aversive stimulus are needed.

As a need for replacement of electric stimulus due to the legal situation of the European union and based on animal welfare, different methods are introduced[33]. In 1999, Boyd patented an alternative concept that describes the release of a substance as a stimulus to control the animals [217]. Additionally, Butler et al. [220] developed a VF algorithm for controlling cows with the use of sound stimulus.

2.2.2.2 Emotional and Auditory Awareness in sheep

Sheep behaviour in its natural state is based on group foraging and group anti-predation drives and is passed on through direct learning from other animals [230]. These behaviours are most clearly evidenced in UK hill breeds. Hill sheep make use of resource bases in heterogeneous landscapes when offered the opportunity. Vegetation mosaics, trees and other shelter resources are a strong influence on the position in a grazing area [231]–[233].

Hunter & Woodgush made some long term studies of hefting behaviour of Scottish Hill flocks [231], [232]. Hill sheep are highly gregarious, living in female-dominated family subgroups. This has also reported in other breeds if grazing in heterogeneous landscapes [231], [232], [234]. There were hefts or preferred areas of hillside based on detailed knowledge of foraging patches passed on through generations ewes to lambs [231], [232]. Grazing and camping (lying down to ruminate) are highly social activities. Grazing hours vary seasonally with day length and are based on the patterns of rumen fill followed by rumination. Summer ranges were larger than winter ranges [231], [232]. They generally break into two main periods: early morning and late afternoon. The period from mid-morning to mid-afternoon is the least active [231], [234].

Sheep form strong social groups that are stable; thus, the social organisation of the flock influences grazing patterns. Dominance-based aggression is more common in single age group flocks commonly seen in commercial settings, than in mixed-aged group flocks [235], [236]. Family or groups with long term stability are less likely to display fighting and attention-seeking behaviours, indicating that a socially stable flock may be less affected by environmental heterogeneity than groups that are not as well integrated socially [235], [237]. Where sub-groupings occur, they are dynamic [232], [238], however, this has no overall effect on overall flock cohesion. It is very important to understand flock structure in detail when managing sheep ,as animals from different sources do not readily integrate into a socially stable group [238]. Position in a moving flock is highly correlated with social dominance, but there is no definite study to show consistent voluntary leadership by an individual sheep [236]. Animal response to external stimulus

Drinking, foraging and habitat selection of grazing livestock together with predator avoidance behaviours can be influenced using visual and olfactory cues, in combination with positive or negative associations [36]. Thus, opportunities arise for managing and controlling animal spatial distribution using external stimuli [36], [40], [239]. Grazing herbivores including sheep are often found in sizeable groups for the purpose of predator avoidance and can therefore influence each other's behavioural responses through social learning [240], [241]. Consequently, it might be possible for farmers to train only a few animals instead of the whole flock for a VF system. Sheep have a concept of other individuals that is invariant with context and even with visual perspective. They are also sensitive to emotional status in another sheep using finely discriminated facial cues associated with facial musculature, the position of the ears [242] and altered appearance of the eyes [243]. Advanced capability for discriminating among faces of conspecifics is an important component of social cognition and empathy, as it is the basis for the formation of relationships and social hierarchies of increasing social complexity [241]. These social and cognitive attributes of sheep lend themselves to rapid flock-level learning to respond to cues that direct spatial position within a VF system.

Previous studies (Sebe et al. 2008; Shillito et al. 1981) noted that sheep could distinguish between their lamb and the lambs of other ewes from the pitch of and frequency of auditory signals. Auditory cues and discriminatory power is also of importance especially in the ewe lamb relationship [230], [243]. There is a lack of published research on the ability of sheep to perform a specific task in response to an auditory stimulus only. The ability of sheep to complete a task based on auditory and visual cues was studied by Morris et al. [40], [246], [247]. The study performed by Morris et al. used 20 merino ewes, where half were tested with visual cues and the other half with auditory cues. A continuous audio cue of 392 Hz originated from speakers positioned above the feed buckets of the sheep. They reported that sheep could not learn to respond to the audio cues since the probability of learning did not increase over the testing period. The authors suggested that there was potential for sheep learning from audio cues and the need for further research. [40].

Heffner [248] generated data on frequencies that several animals and humans can hear. They observed that sheep could hear in the range between 125 Hz and 42,000 Hz, with the most sensitive hearing frequency at 10,000 Hz. Moreover, they noted that sheep display the best sensitivity at 6 dB. In the study mentioned above of Morris *et al.,* [40], the sound frequency chosen for the experiments was 392 Hz. There is a need to investigate further the intensity and frequency of sounds needed to train animals. Moreover, using intermittent sound pulses instead of a continuous stimulus may prove more appropriate since it has been shown that discontinuous sound produces better response [249], [250].

For the current research, the aim is to test audio signals and analyse how the animals respond to each of them. The goal is to identify sounds which can discourage the animals from going or staying in the restricted areas. The desired response is for the animal to either stop or turn back from the boundary.

The existence of individual personality in animals, stable across different contexts, is now well established, including in sheep [241] Whether they are bold or shy, sheep and other grazing animals mostly associate unknown auditory signals with predators and, consequently, have an aversion to them [251]. Animal position could be choreographed by sound to stay within the boundaries of a restricted area of any size and shape. Careful consideration of how auditory cues might impact an animal or flock behaviour is highly relevant, and such knowledge can be applied when training animals or flocks to learn to respond to any VF system [36]. Position in a moving flock is highly correlated with social dominance in sheep, but there is no definite study to show consistent voluntary leadership by an individual sheep [235], [236]. Sheep form strong social groups are stable thus the social organization of the flock influences grazing patterns [235]. Sheep like humans exposed to a standard system of appraisal evaluate situations according their suddenness, familiarity, expectations, predictability and the consistency of these events with their own expectations and, in particular, the control they have over the events [242], [252], [253]. A VF solution cannot be used, in ethical and therefore welfare terms, if animals cannot be in control of their ability and capacity to avoid an acoustic cue [34]. Thus, VF can be applicable only when "leaky boundaries" are acceptable, [34].

2.2.3 Discussion

In the literature, advantages of the use of VF can be found in many areas, especially on improving the management and utilisation of areas that are difficult to be fenced. From the literature, it is clear that there is no universal agreement upon the technology to be used for the creation of VF systems. Additionally, there is no general method on the training of animals to learn the VF. Yet, most studies that deal with the VF, assumes that the animals are commonly domesticated without any physical barriers on the land they graze. A VF can reduce costs concerned with building or erecting fences, it is flexible, and can be applied to areas that otherwise could not be fenced. Additionally, a VF can reduce overgrazing and therefore prevent soil erosion and contamination. In the literature, many reports focused on using VF systems as herding tools and thus reducing the need of labour-intensive tasks. Moreover, as noted by Umstatter et al. [33], with climate change, there is a possibility of flooding in the grazing areas which can be life threatening to the fenced animals. With the use of a virtual fence, this could be avoided by switching off the system and let the animals make their own decision. Then, the animals could be tracked back with the GPS module incorporated on the system devices, situation which could not be avoided in a fenced area.

On the other hand, many disadvantages are also present in the literature regarding VF. The most important is that a VF cannot be 100% stock-proof and cannot be applied to areas where animals are in close proximity of busy roads, or in areas where humans are present once the stimulus is applied to the animal. Additionally, a VF cannot keep predators away, and this is an area that must be investigated further.

Another challenging and important problem is the training of animals using electric shocks. This have a negative impact on the animal welfare and many countries banned the electric shocks, therefore, alternative ways to train the animals must be considered. In the literature, most researches used audio followed by electric shocks to train the animals and proved successful at containing the animals away from the restricted areas [45], [189], [208], [210], [214]–[216], [226], [228], [190], [191], [198]–[200], [202], [205], [207]. On the other hand, alternative solutions were suggested such as the use of an odour substance [217], or the use of audio only solution [187], [197], [220], [221]. Such alterative solutions are very limited as it is evidenced in the literature and there is a call for more research regarding the use of audio only training methods. Additionally, the literature has shown that limited experiments exist for sheep, in contrast with cattle. Therefore, experiments using sheep is an area which needs further investigation.

Furthermore, the literature has shown that there is still lot of areas to focus on, such as, position of the device, training method, robustness of the hardware, and energy supply due to the high power consumption required from the GPS module incorporated on the device. However, there are potential benefits of VF systems and could help improve farm management, reduce farm costs, and improve biodiversity. Therefore, there is a need for further research on the abovementioned areas to be investigated to develop and create a robust VF system to benefit the Agricultural community, animal welfare, and environment.

2.3 Summary

Section 2.1 – Sheep activity recognition (SAR): Animal activity recognition (AAR) is an important topic that facilitates understanding of animal behaviour, where the animals' wellbeing can be analysed and classified. Extensive research showed that animal activity could be utilised as a useful indicator of health state. In this Chapter, the focus is on recent advancements in machine intelligence utilizing wearable devices for sheep activity recognition. Existing works were summarised with focus on the various types of sensors used in agricultural sheep activity recognition. Data segmentation methods used in each study were addressed, followed by potential recommendations on window size and sample rate selection. Finally, the features identified as significant were presented, followed by an overview of machine and deep learning algorithms in the domain of sheep activity recognition using accelerometer signals.

Section 2.2 – Virtual fencing systems: VF system is a computerised method for creating spatial boundaries of any geometric size and shape without the use of any physical fences or barriers. VF (or Fenceless) systems have been discussed over the last 30-40 years due to the inflexibility, cost and heavy maintenance demands required upon traditional fences. In this Chapter, a brief background on VF systems was provided. Advantages and disadvantages are also introduced and highlighted the potential of advancements in the decision-making approach regarding land utilisation, overgrazing, costs and labor-intensive tasks. Furthermore, this chapter presented previous and current VF development and approaches, followed by the main means and alternative ways of training animals to learn a VF system. Information regarding emotional and auditory awareness of sheep was introduced. This information will help build an understanding on how to treat sheep and train them on a VF system, as it is one of the main focus of this thesis.

Chapter 3 will demonstrate the application of ML in SAR. Specifically, the chapter is concerned with the feature extraction, selection, window size and ML methods in order to investigate and identify optimal setups for the classification of various sheep activities.

Chapter 3 Sheep Activity Recognition using Machine Learning

The purpose of Chapter 3 is to propose new methods of SAR using ML. Three experiments are performed to identify an optimal feature set and machine learning algorithm to classify behaviours of sheep to minimise the window size and sample rate to minimise the device's energy expenditure. The first experiment is concerned with identifying an optimal feature set and machine learning algorithm to classify behaviours of sheep and goats and select the algorithm with the best performance using a publicly available dataset. The second experiment aimed to evaluate the algorithm's performance from the previous experiment and apply it to a new dataset collected from Hebridean ewes wearing collars with mounted accelerometer and gyroscope sensors using a smaller window size. Lastly, the third Section provides the methods used in our third experiment to improve the previously suggested technique on a newly collected dataset containing only accelerometer measurements from Hebridean ewes, minimising the window size again while keeping the computational cost to a minimum. The experiments are presented in the following three sections, which follow a similar structure (refer to Figure 3-1).

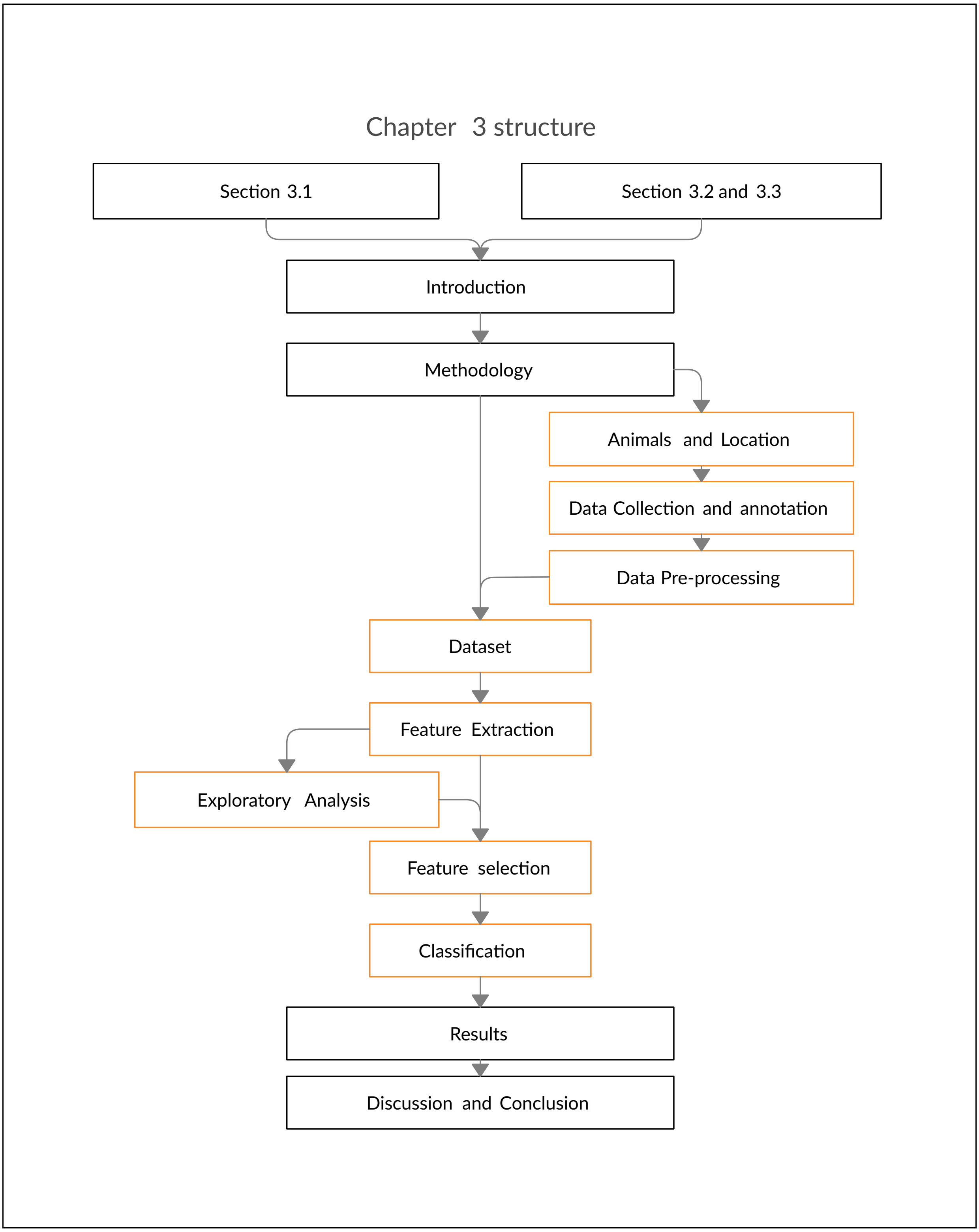


Figure 3-1 Chapter 3 structure

3.1 Experiment A: Machine Learning Techniques for the Classification of Livestock Behavior

This Section reports the findings from using different features and ML methods for the animal activity recognition problem. An online open-source dataset [254] was used as a testbed for an extensive evaluation and in-depth comparison of the most promising features and ML algorithms for SAR problems, to implement the findings in a larger body of work concerned with monitoring and controlling sheep without the need for human observation. This work's main novelty is the different combination of feature extraction, feature selection, and ML techniques, which allows identifying an optimal method to improve sensitivity, specificity, kappa value, and accuracy of the classifier. The main contributions of this experiment are:

- the suggestion of features such as RMS velocity, Integrals, Squared Integrals, and Madogram in SAR problems as they were not used before
- the Boruta feature selection algorithm is used for the first time in SAR problems to the best of the authors knowledge
- High performance is achieved using this setup which shows improvement over previous research

3.1.1 Methodology

This Section comprises the methods and techniques used for the detection of animal behaviour using intelligent systems. It also contains a description of the dataset, data pre-processing steps, feature extraction and feature selection techniques. Additionally, it defines the machine learning algorithms and the evaluation metrics used to assess the classifiers' performance. Figure 3-2 shows the stages used during the analysis process.

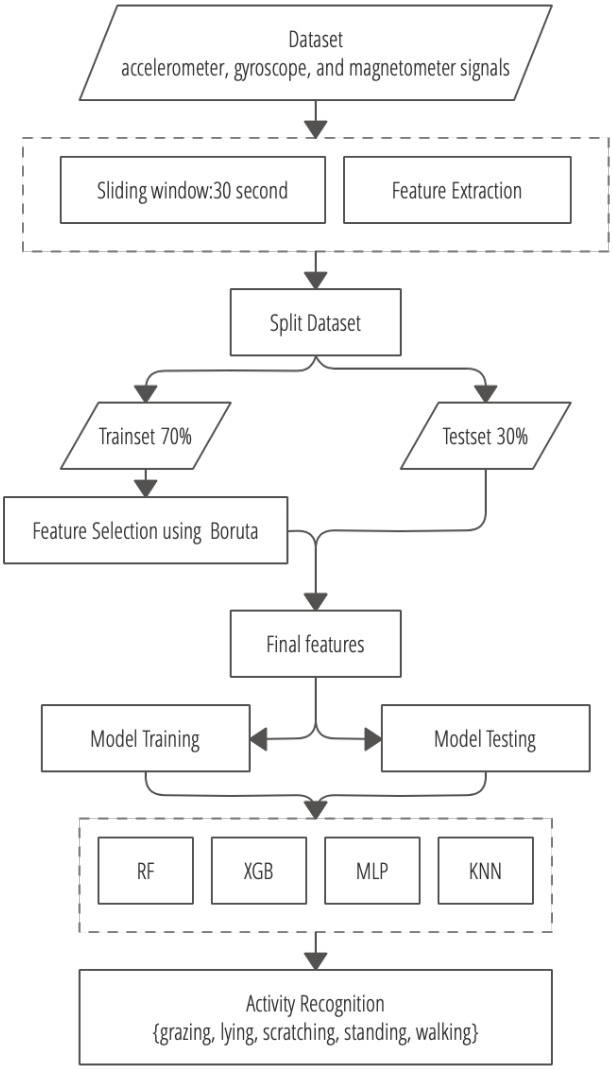


Figure 3-2 A breakdown of the stages used during the analysis process

3.1.1.1 Dataset

The dataset used for this study is available online from Data Archiving and Networked Services (DANS) website [14, 23] and contains labelled behavioural data from four goats and two sheep. Each data file includes a record of daily activity for each animal. The file holds nine different behaviours: grazing, scratching or biting, standing, walking, fighting, running, trotting, shaking, and lying.

The measurements were obtained from sensors placed on the collars of the animals. The parameters used in this study are described in Table 3-1, however, the original dataset included other measurements such as temperature and pressure, but are not included in the study as there were of no use in this setting. Additionally, four other activities are omitted: fighting, running, trotting, and shaking because these behaviours occurred too infrequently to draw statistically significant conclusions. Consequently, five activities are included; grazing, lying, scratching or biting, standing, and walking. The activities in the datasets are labelled per sample (refer to Figure 3-3).

Table 3-1. Annotation and description of the parameters used from each dataset

Parameters	**Description of raw data**
ax	accelerometer x-axis. Sampling Rate:200 Hz
ay	accelerometer y-axis. Sampling Rate: 200 Hz
az	accelerometer z-axis. Sampling Rate: 200 Hz
cx	magnetometer x-axis. Sampling Rate: 100 Hz
cy	magnetometer y-axis. Sampling Rate: 100 Hz
cz	magnetometer z-axis. Sampling Rate: 100 Hz
gx	gyroscope x-axis. Sampling Rate: 200 Hz
gy	gyroscope y-axis. Sampling Rate: 200 Hz
gz	gyroscope z-axis. Sampling Rate: 200 Hz

The final dataset used for this research study consists of time-series measurements from an accelerometer (acc), a magnetometer (magn), and a gyroscope (gyr). Additionally, all sheep and goat datasets were used to examine the performance of different algorithms to learn their activities and identify whether the algorithms could generalise well between other animals and activities. The number of samples per activity is presented in Table 3-2. An example of the dataset is presented in Figure 3-3, and the distribution of each activity is presented in Figure 3-4.

Table 3-2 Number of samples per activity

Grazing	Lying	Scratching or biting	Standing	Walking
2623146.32	2389977.76	2502675.9	3202181.585	2237122.81

label	timestamp_ms	ax	ay	az	cx	cy	cz	gx	gy	gz
walking	1	-0.0694318	-4.70939	-6.90248	1.995	0.075	0.315	65.3049	4.45122	23.4756
walking	6	-0.296881	-4.43885	-6.60321	NaN	NaN	NaN	63.9024	10.1829	27.5
walking	11	-0.562637	-4.2521	-6.38055	2.0085	0.0825	0.3075	61.4634	14.8171	30.9756
walking	16	-0.830788	-4.13718	-6.13634	NaN	NaN	NaN	57.6829	18.3537	33.5976
walking	21	-1.0726	-4.10845	-5.96156	2.01	0.087	0.294	52.8659	20.5488	35.5488
walking	26	-1.29766	-4.07254	-5.85861	NaN	NaN	NaN	47.2561	21.2195	36.7683
walking	31	-1.51074	-4.05338	-5.9472	2.01	0.087	0.294	40.4268	20.6707	37.5
walking	36	-1.74537	-4.04141	-6.27281	NaN	NaN	NaN	32.7439	19.5122	37.9268
walking	41	-1.96085	-4.02944	-6.87136	2.0145	0.0825	0.3015	24.0244	17.9878	37.8049
walking	46	-2.13084	-4.12282	-7.61356	NaN	NaN	NaN	14.2073	17.0122	37.378
walking	51	-2.20027	-4.34308	-8.37731	2.013	0.078	0.297	3.04878	17.1951	35.7317
walking	56	-2.25773	-4.69024	-9.27993	NaN	NaN	NaN	-9.5122	18.5366	33.7195
walking	61	-2.34632	-5.18345	-10.1466	2.004	0.069	0.297	-22.6829	20.7927	31.4634
walking	66	-2.3966	-5.82988	-10.817	NaN	NaN	NaN	-35.9146	23.8415	29.2073
walking	71	-2.42772	-6.70376	-11.1642	2.0355	0.0825	0.2805	-48.1707	28.1098	27.3171
walking	76	-2.3942	-7.72369	-11.2934	NaN	NaN	NaN	-57.7439	33.6585	26.0366
walking	81	-2.35829	-8.75081	-11.224	1.992	0.075	0.315	-63.4146	39.4512	25.7927
walking	86	-2.34392	-9.69173	-11.0899	NaN	NaN	NaN	-65.7927	44.7561	26.8902
walking	91	-2.37026	-10.5225	-10.9918	2.0175	0.0825	0.3015	-66.2805	49.1463	28.7195
walking	96	-2.46842	-11.1713	-10.9702	NaN	NaN	NaN	-64.8171	52.5	31.5244
walking	101	-2.66235	-11.5592	-10.963	1.9995	0.0675	0.3285	-61.6463	55.122	34.8171
walking	106	-2.90656	-11.6622	-10.9391	NaN	NaN	NaN	-57.2561	57.5	38.5976
walking	111	-3.16513	-11.4874	-10.9104	2.0175	0.0705	0.3165	-51.7683	60.3049	42.2561
walking	116	-3.40695	-10.9702	-10.8457	NaN	NaN	NaN	-45.6098	63.2927	45.7317

Figure 3-3 Sample of the raw dataset

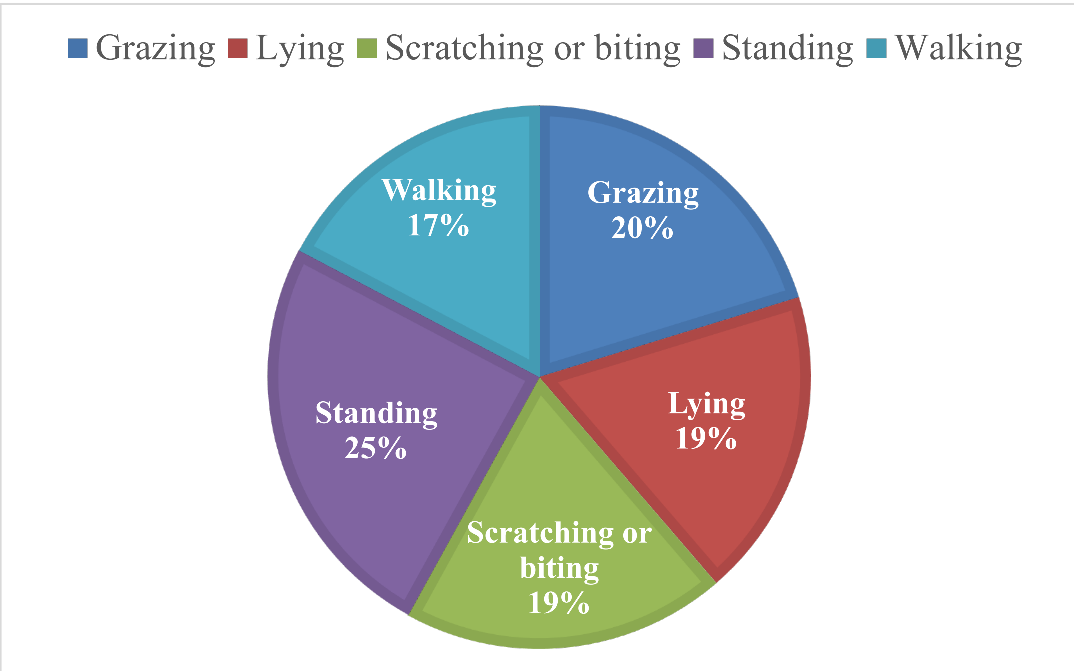


Figure 3-4 Class Distribution

3.1.1.2 Data Pre-processing

The dataset was examined to identify any inconsistencies. Missing values were present in the magnetometer measurements. They were replaced using linear interpolation as it was assumed that there is a linear relationship between the missing and non-missing values. The linear interpolation method is suitable for time-series data [255]. The sensor measurements were segmented into 30-second non-overlapping windows. The selection of the window size was based on previous research which showed that it is big enough to include a representative signal able to describe each activity in the training process. Each data block comprised individual animal behaviours. Consequently, the final data resulted in 1610 randomly selected data blocks containing a relatively even distribution of the five selected activities. The dataset was divided such as 70% was used for training and 30% for final testing.

3.1.1.3 Feature Extraction

A vital phase of the methodology process is to identify a suitable pattern mapping for the ML. To achieve that, and as a first step, a total of 23 features were extracted from all parameters of the dataset (refer to Table 3-1). Therefore, features from various domains (i.e. time and frequency) encompassed 276 (23 features and 12 parameters) newly created features. The choice of the features was based on their performance in previous research concerned with activity recognition problems using acceleration forces, angular momentum and orientation measurements. Features from the time domain are calculated using the raw dataset. However, for frequency features, the raw data had to be translated into the frequency domain using Fourier transform [24, 25]. The final selected features are defined in Table 3-3.

3.1.1.4 Exploratory Analysis

Principal Component Analysis (PCA) mapping was applied and plotted to understand the dataset distribution better. From Figure 3-5, it can be observed that there is a clear distinction between the five classes. Additionally, the PCA plot indicates that the walking and scratching or biting classes display substantial disconnection and may establish a simple decision boundary. Conversely, considerable overlap is demonstrated between the standing and lying classes, thus, the constitution of the class predictions might be more difficult.

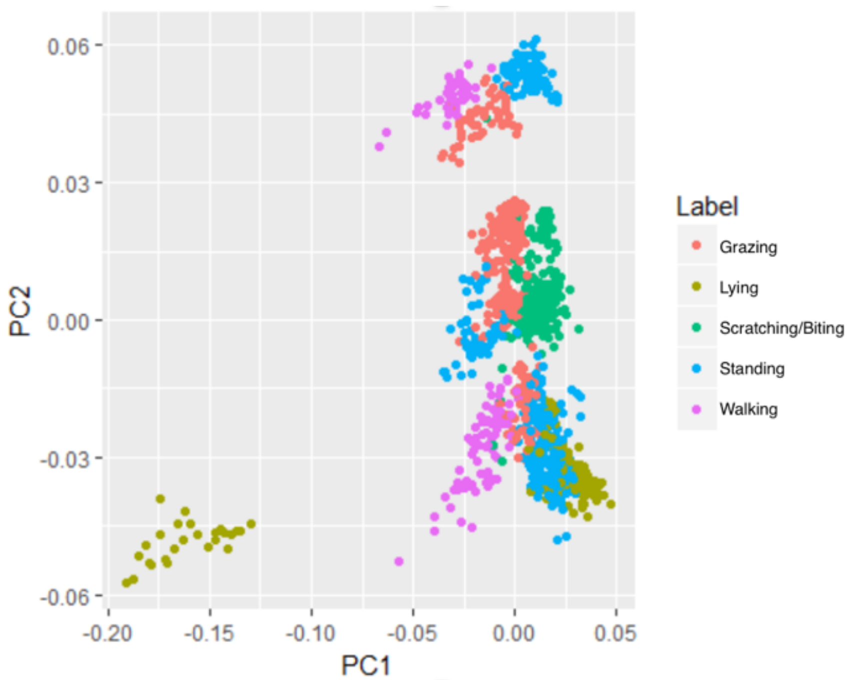


Figure 3-5 . PCA plot: A 2-dimensional illustration of the extracted behavioural data

3.1.1.5 Feature Selection

The feature selection method reduces and optimises the number of features extracted from the raw data. The objective is to understand the dataset better and provide a faster predictor [258]. The features are chosen based on their ranking and best suitability to the classifiers' performance in the feature selection process. The main methods of feature selection are (1) filter, (2) wrapper, and (3) embedded [258].

Filter methods apply statistical measures to allocate ranking scores to each feature. Thus, according to their score, they are either included or excluded from the dataset. Wrapper methods select combinations of feature subsets and evaluate their usability on a given machine learning algorithm. Accordingly, the subsets are scored based on their predictive power. On the other hand, embedded methods learn the best features according to the correctness (accuracy) of the learning model. More information regarding the feature selection methods is provided in Chapter 2.

For this study, the Boruta algorithm was used as the feature selection technique [120]. The Boruta algorithm is a wrapper method that uses a random combination of feature sets and evaluates their importance according to random probes. This method uses the RF classifier for the selection of all relevant features. Consequently, 276 features were used by the Boruta algorithm to identify and accept the most pertinent. This process resulted in a set of features that are ranked in decreasing order. Therefore, the top 15 most important features are selected in this study to train the ML algorithms. Table 3-3 provides the specified features (Chapter 2 provides definitions and formulas for the features). From the sensor measurements, only five parameters are finally used. The x and z-axis of the acceleration, the x and z-axis of the gyroscope, and the x-axis of the magnetometer.

Table 3-3. Final features

Feature	**Parameters**
Mean	x-axis of magnetometer
Standard Deviation	z-axis of gyroscope
Root Mean Square (RMS) velocity	z-axis of gyroscope
Root Mean Square (RMS)	z-axis of gyroscope
Energy	z-axis of gyroscope
Sum of Changes	z-axis of accelerometer
Mean Absolute Change	z-axis of accelerometer
Absolute Integrals	z-axis of gyroscope
Squared Integrals	z-axis of gyroscope, and the x-axis of

	magnetometer
Madogram	x-axis and z-axis of accelerometer, and y-axis of gyroscope
Peak Frequency	x-axis and z-axis of gyroscope

3.1.1.6 Classification and evaluation metrics

A supervised learning approach is used as it is more suitable for the nature of the investigated dataset. In supervised learning, there exists a set of independent variables $\{x_1, x_2 \ldots x_n\}$ and a set of dependent variables $\{y_1, y_2 \ldots y_m\}$. The aim is to fit a model based on the observations of x, and the response y, and predict y based on a new set of x observations. To achieve that, the model must be trained based on the already known data and then test its ability to predict or classify unknown data.

Therefore, a set of examples is created in the form of $\{(x_{(i,1)}, \ldots x_{(i,p)}, y_i)\}$, for $i = 1, \ldots, n$; where p represents the j^{th} feature, i represents the i^{th} observation, and n represents the number of observations. The set may be expressed as the features matrix $X \in R^{nxp}$ and the response $y \in \{1,2,3,4,5\}$, where the numbers characterise indexes to each of the animal's activities (grazing, lying, scratching, standing, and walking), respectively.

The four models selected for the detection of the five activities of the animals in the pasture are Multilayer Perceptron (MLP) [259], Random Forest (RF) [151], Extreme Gradient Boosting (XGB) [131], and k-Nearest Neighbors (KNN) [260].

Sensitivity, specificity, Kappa value, and accuracy are used to evaluate the performance of the models. The classification problem in this study is concerned with a multiclass problem, and the classes are evaluated, using the true prediction of the class against the total false prediction of the other classes [261]. The calculations use the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) metrics. The formulas for each metric are illustrated in Table 3-4.

Table 3-4. List of the evaluation metrics

Metric	**Formula**
Sensitivity	Se = TP/(TP+FN)
Specificity	Sp =TN/(TN+FP)
Accuracy	Ac = (TP+TN)/(TP+TN+FP+FN)
Kappa value	K = (Po – Pe)/(1-Pe), where Po is the proportion of observed agreement and Pe is the proportion of agreement expected by chance.

The classifiers are trained using 70% of the original dataset (training set) and tested its performance on 30%. The model training used 10-fold cross validation technique with one repetition on the training set to tune the model parameters better, as shown in Table 3-5. The overall results by means of accuracy, kappa value, sensitivity, and specificity indicated a high model performance for all with the best results obtained by RF and XGB. For the RF classifier, the final value used for the model was eight variables per level and was selected based on the highest accuracy. The XGB's maximum depth of the tree was three and collected half of the data instances to grow trees to prevent overfitting and decrease the computation time. Using the same criteria as the RF, the optimal model for KNN selected k=5 neighbours, and for MLP the size = 5 hidden layers and the regularisation parameter decay=0.1.

Table 3-5. Summary of performance of the models (10-fold cross validation)

Model	**Accuracy**	**Kappa**	**Sensitivity**	**Specificity**
MLP	95.21%	93.79%	95.36%	98.76%
RF	**96.73%**	**95.74%**	**96.40%**	**99.14%**
XGB	**97.34%**	**96.54%**	**96.96%**	**99.13%**

KNN	92.63%	90.45%	93.19%	98.07%

3.1.2 Results

Following the 10-fold cross validation to train the classifiers, the performance of the models was evaluated on the unseen testing set. Table 3-6 illustrates the results obtained using 30-second windows and 15 features on the test set. It can be observed that the best performance was achieved using RF yielding an accuracy and kappa value of 96.47% and 95.41%, respectively. The second-best results obtained by XGB with 95.85% accuracy and 94.60% kappa value. The grazing, scratching or biting and standing activities identified by RF are higher than the other three models. On the other hand, MLP and XGB obtained the highest sensitivity for walking activity with 98.11%.

All models obtained sensitivity for the grazing activity in the range of 96.09%-97.66%, and specificity in the range of 97.74%-98.31%. Overall, all activities are correctly classified with high results between 91.53% and 99.77%. The KNN classifier obtained the lowest results by means of accuracy and kappa value with 93.57% and 91.66%, respectively. The results are promising for the future implementation of a system to predict the behaviour of animals.

Table 3-6. Model performance – test set

ML Algorithms	**MLP**	**RF**	**XGB**	**KNN**
	Accuracy 94.40%	**Accuracy** **96.47%**	**Accuracy** 95.85%	**Accuracy** 93.57%
	Kappa Value 92.73%	**Kappa Value** **95.41%**	**Kappa Value** 94.60%	**Kappa Value** 91.66%
Grazing				

Sensitivity	96.09%	97.66%	96.09%	96.88%
Specificity	98.02%	97.74%	98.31%	97.74%
Lying				
Sensitivity	91.53%	93.22%	93.22%	93.22%
Specificity	97.87%	99.76%	99.53%	97.16%
Scratching or Biting				
Sensitivity	94.62%	95.70%	94.62%	92.47%
Specificity	99.74%	99.74%	99.49%	99.74%
Standing				
Sensitivity	92.62%	97.32%	96.64%	90.60%
Specificity	97.30%	98.50%	97.60%	97.60%
Walking				
Sensitivity	98.11%	96.23%	98.11%	96.23%
Specificity	99.77%	99.53%	99.53%	99.53%

3.1.3 Discussion

Identifying animal behaviour and calculating circadian rhythms is of immense importance since it can act as an indicator of animal wellbeing [262]. Additionally, grazing activity information offers insights into the food intake and preference of the animals. Knowing where the animals mostly graze allows the farmers to efficiently manage animal distribution, preventing overgrazing and, consequently, soil erosion and soil contamination.

In this study, an online dataset was used from a recent research study featuring five different activities of 4 goats and two sheep: grazing, lying, scratching or biting, standing, and walking. The aim was to identify the most significant features and to select the machine learning algorithm that could be used to classify the five activities with the highest accuracy and kappa value. Various features were extracted from an accelerometer, a gyroscope, and a magnetometer data, resulting in 276 feature mappings. Due to the high dimensionality of the dataset, feature selection was applied using the Boruta wrapper method. The top 15 most important features were then used as the main predictors to train RF, XGB, MLP, and KNN. All four algorithms achieved high accuracy and Kappa values, indicating that the features could discriminate correctly between the classes. The RF classifier obtained the best results having an accuracy of 96.47% and Kappa value of 95.41% for 30 second mutually exclusive behaviours. To the best of our knowledge, those results are higher than previous research concerned with the classification of sheep behaviours, as shown in Table 3-7.

Table 3-7. Comparison with previous studies

Ref	**Animals**	**Classes**	**Signal**	**Window**	**Method**	**Accuracy**
[263]	Sheep	2	Tilt	30-s	LDA	94.4%
[60]	Sheep	5	acc	10-s	DT	91.3%
[66]	Sheep	5	acc	5.3-s	LDA	85.7%
[67]	Sheep	5	acc	-	MLP	76.2%
[68]	Sheep, Goats	5	acc	1-s	DNN	94.00%
[61]	Sheep	3	acc	60-s	DA	93.00%
This work	Sheep, Goats	5	acc, magn, gyr	30-s	RF	96.47%

3.2 Experiment B: Evaluating Random Forest on new data

In the previous experiment, various ML techniques have been used to identify sheep and goat behaviour such as grazing, standing, lying, scratching or biting, and walking. The previous data was sampled at 100Hz and the behaviours were segmented into 30-second windows [64]. In this current experiment, new data were collected from seven Hebridean ewes to categorise six behaviours; grazing, resting, walking, standing, scratching, browsing. The RF performance was evaluated as previous analysis suggested that this algorithm can provide advantages and has been proven to classify the behaviours of sheep and goats adequately. For this reason, new data was collected to re-evaluate the technique. Another aim of this study was to apply RF with lower sampling frequency and smaller window size (i.e. 10 Hz and 10-second windows), and evaluate the resultant performance. This is predicated on previous findings, which indicated that lower frequency rates improve memory and demand less power of the device [180]. It was also previously demonstrated that 16 Hz sample rate and 7-second windows can reduce energy requirements and can be used for real-time sheep activity monitoring [55].

3.2.1 Methodology

To evaluate the performance of the ML technique implemented in previous research [64] using new data , behavioural data was gathered from seven Hebridean ewes from a farm located in Cheshire, Shotwick. In the following sections, the description of the required steps is presented. The experiment's protocol was approved by the Senior Research Officer and LSSU Manager of Liverpool John Moores University (approval AH_NKO/2018-13). Figure 3-6 shows the stages used during the analysis process; this forms the methodology's basis.

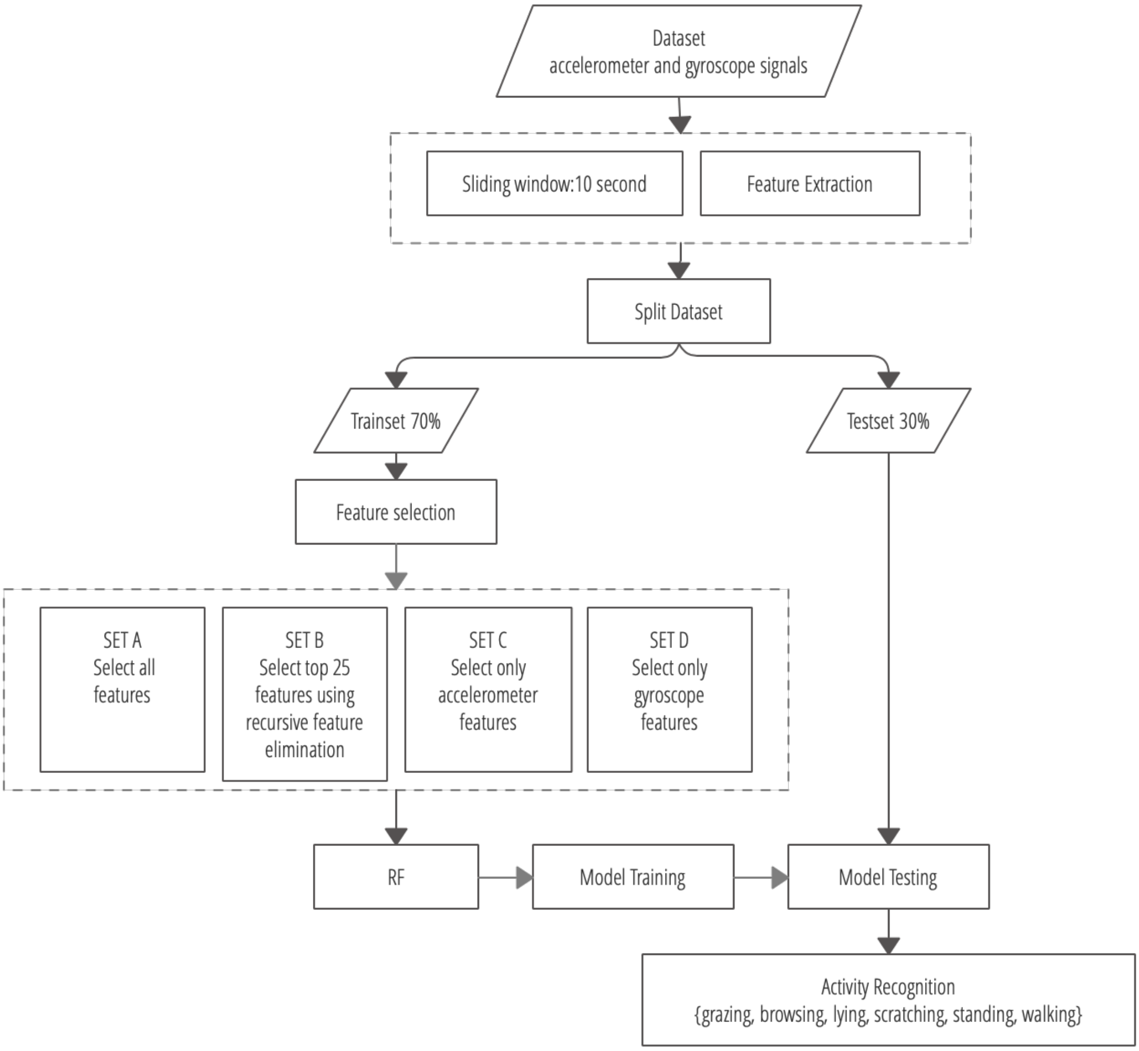


Figure 3-6 A breakdown of the stages used during the analysis process

3.2.1.1 Animals and location

Seven Hebridean ewes are used, located in a paddock in Shotwick OS 53.2391152, -3.0028081, Cheshire with approximate area size and perimeter of 1500 m^2 and 250 m (refer to Figure 3-7), respectively. The behaviour of the ewes (average age 11 years) was observed and recorded using accelerometer and gyroscope measurements daily from 4th July until 18th July 2018 during different times of the day. The animals were free to use the whole area of the paddock and had access to grass and water. The ewes were used in previous research [32] and were habituated to human interaction, so their behaviours were not affected by human presence during the video recordings.

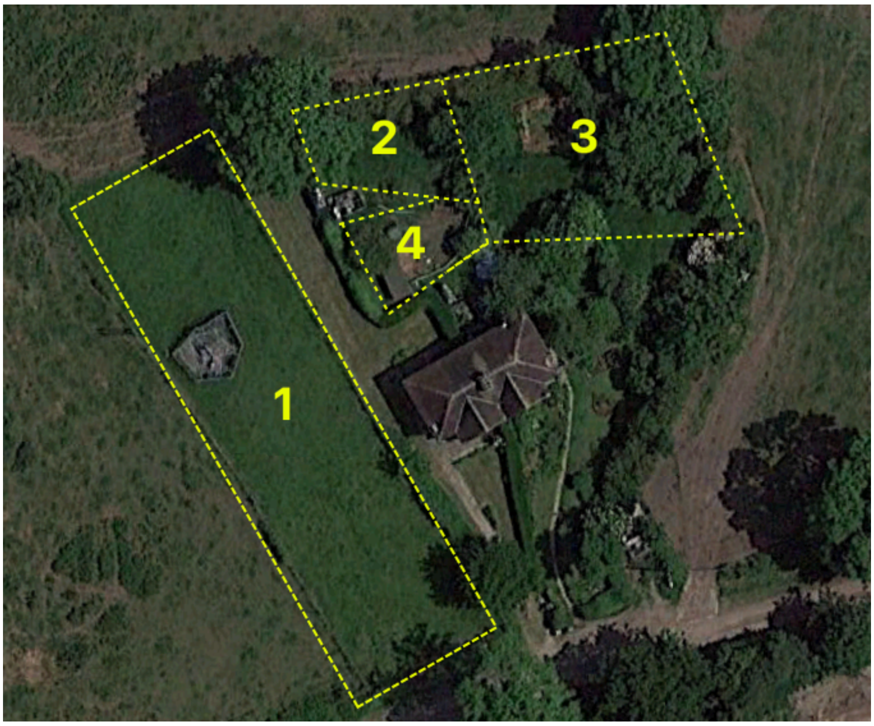


Figure 3-7 Representation of the paddocks (area 1-3) and the pen (area 4). Areas 1 and 2 are linked through a pathway

3.2.1.2 Data Collection and Annotation

During the experimental period, one sheep at a time was fitted with a smartphone device (Samsung Galaxy S5) on the top and side part of the harness. The orientation of the smartphone device was not fixed since the aim is to test whether high accuracy can be achieved similar to previous work [64] without sensor orientation dependency. The smartphone device was set with a data logger application (HyperIMU [264]) which saved the x, y, and z coordinates of accelerometer and gyroscope data at a sample rate of 10Hz as a CSV file for offline data processing. During the data gathering, the animal carrying the smartphone was video recorded using the smartphone camera (Samsung Galaxy S5), and a human observer was at present all time. This information was due to be used as ground truth for activity labelling. In total, 35 hours of recordings were obtained, but only 30 hours were selected for data analysis because they included behaviours with insufficient information.

For the annotation of the animals' behaviour, the ELAN_5.2 annotation tool was used. The CSV files were synchronised with the video recordings to provide an accurate mark of each behaviour. Behaviours with insufficient information were discarded (biting, fighting, running, shaking). Table 3-8 describes the selected animal activities and the duration of each activity.

Table 3-8. Description and duration of the animals' activities

Activity	Description	Time (seconds)	Time in %
Grazing	Grazing while walking and standing.	23150.10	20.95%
Browsing	The animal was reaching upwards for leaves on trees or bushes	593.80	0.54%
Resting	The animal was inactive in a lying posture or ruinating	34453.30	31.18%
Scratching	Scratching with leg movement or pushing against trees or bushes	168.90	0.15%
Standing	Standing idle or ruminating	36640.40	33.15%
Walking	Walking forward, backward or sideward	15506.00	14.03%
Total		110512.50	100%

3.2.1.3 Data Pre-processing

After annotating the data, RStudio open-source IDE for R programming language was used for pre-processing and analysis. The data were tested for missing values, and six behaviours were visualised to understand better each activity (refer to Figure 3-8).

3.2.1.4 Dataset

The behavioural data comprised of a set A={ $t_{i,}$ ax_i, ay_i, $az_{i,}$ gx_i, gy_i, gz_i ,y_i} for i=1,..,n, where n is the number of observations. The t is the timestamp, (ax, ay, az), and (gx, gy, gz) are the accelerometer and gyroscope measurements, respectively, and y is the target vector where y∈{grazing, browsing, resting, scratching, standing, walking}.

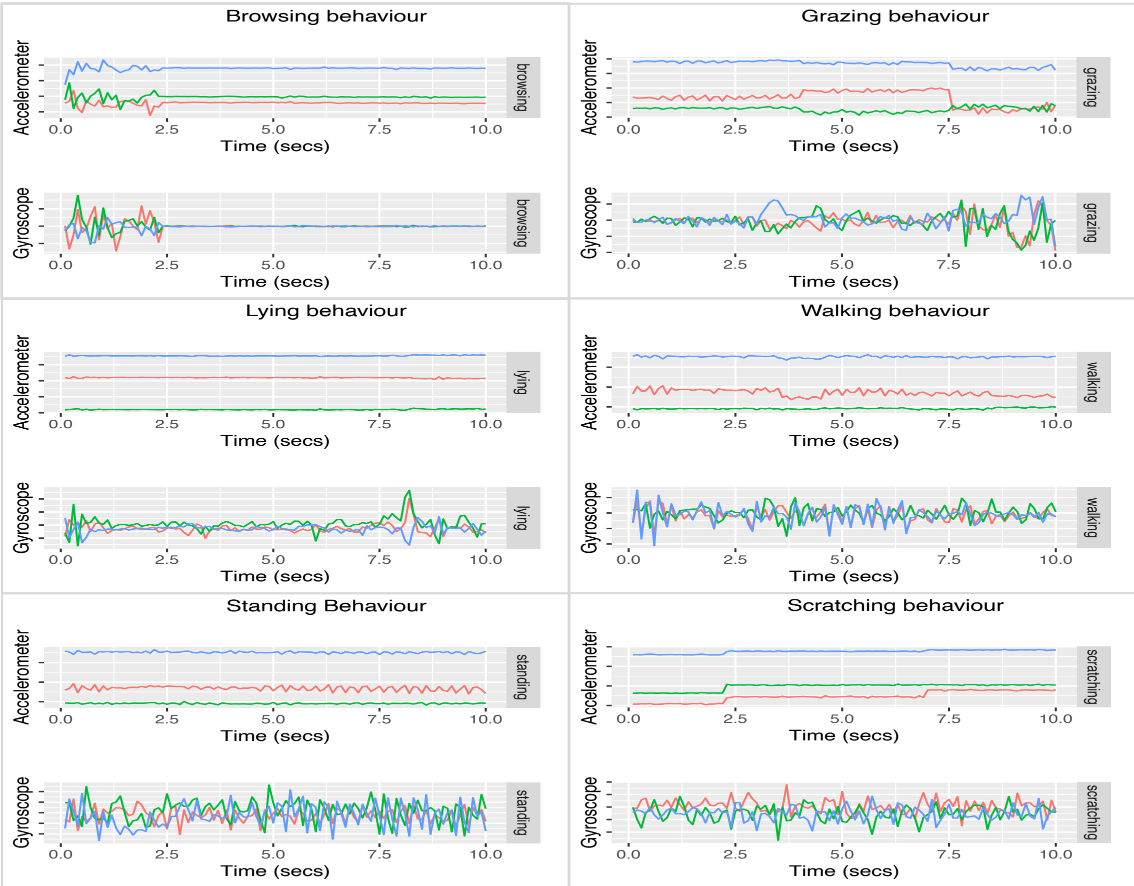


Figure 3-8 Visual representation of the six animal activities. The graphs illustrate the acceleration and gyroscope coordinates of each behaviour in a 10 second time interval

3.2.1.5 Feature extraction

For the analysis of the data and to test the performance of the RF algorithm, firstly, the desired features were extracted from 10-second no-overlapping windows resulting in a total of 98755 windows. The selection of the window size was based on previous finding which noted that smaller window sizes improve memory and battery life and do not succrifice the performance of the ML algorithm [55], [180]. Time-domain features such as the mean, standard deviation, and root mean square are the most usual features for activity recognition according to reference [265]. Additional essential features for activity recognition based on accelerometer measurements are integrals and squared integrals [95]. The features were prechosen and identified by the author of this thesis' previous work [64] and are described in Table 3-9. In this case, 11 features are used from the time and frequency domain resulting in 66 newly created features. Once the features were extracted, the data was split with 70% and 30% for training and testing, respectively.

Table 3-9. Feature names and feature sets

Feature Name	SET A	SET B	SET C	SET D
Mean	acc, gyr	acc, gyr	acc	gyr
Standard Deviation	acc, gyr	ax, az	acc	gyr
Root Mean Square (RMS) velocity	acc, gyr	az, ay	acc	gyr
Root Mean Square (RMS)	acc, gyr	acc	acc	gyr
Energy	acc, gyr	acc	acc	gyr
Sum of Changes	acc, gyr	-	acc	gyr
Mean Absolute Change	acc, gyr	-	acc	gyr
Absolute Integrals	acc, gyr	acc	acc	gyr
Squared Integrals	acc, gyr	acc	acc	gyr
Madogram	acc, gyr	gx	acc	gyr
Peak Frequency	acc, gyr	az, ay	acc	gyr

acc: x,y,z coordinates of the accelerometer, gyr: x,y,z coordinates of the gyroscope

ax: x coordinate of the accelerometer, ay: y coordinate of the accelerometer

az: z coordinate of the accelerometer, gx: x coordinate of the gyroscope

3.2.1.6 Feature selection

The algorithm was tested using four feature sets, Set A, Set B, Set C, and Set D (Table 3-9). Firstly, all features are used from both accelerometer and gyroscope sensors resulting in 66 new features (Set A). The next step was to use only the features from the accelerometer as a new set (Set C), and only features from the gyroscope sensor as Set D. Also, we wanted to isolate the 25 most important features from the whole set. Recursive feature elimination (RFE) with backward selection based on the RF was utilised for this method. At first, the significance level was selected (significance level = 0.05), and the model was fitted using the whole set of predictors. Then, the predictor with the highest p-value (p>significance level) was removed and the model was fitted again. The process was repeated until all the remaining predictors had a p-value lower than the selected significance level. Once the most important features are identified, the top 25 features were selected based on their importance ranking.

3.2.1.7 Classification and evaluation metrics

Table 3-10. Performance evaluation metrics

Metric	Description	Formula
Sensitivity	The correctly Identified the positive class	Se = TP/(TP+FN)
Specificity	The correctly Identified negative classes	Sp =TN/(TN+FP)
Precision	The positive predictive value	Pr = TP/(TP+FP)
Accuracy	The degree of overall correctness	Ac=(TP+TN)/(TP+TN+FP+FN)
Kappa value	The inter-rater agreement measurement	K = (Po – Pe)/(1-Pe)

Based on our previous findings [64], we used RF [151] to train and test the model. The One-vs-All technique was used for the multiclass classification problem [12], constructed from binary classification. Each time, a single class is treated as positive against the rest, which are the negative ones. Sensitivity, specificity, precision, Kappa value, and accuracy were used to evaluate the performance of the ML algorithm and are defined in Table 3-10.

3.2.2 Results

The performance of the proposed algorithm is shown in Table 3-11. The accuracy and kappa values for all four feature sets (Set A, Set B, Set C, Set D; refer to Table 3-9) yielded results higher than 91.69% and 88.40%, respectively. The best results were obtained with the feature set A, which included all features from the accelerometer and gyroscope coordinates with an accuracy of 96.43% and a kappa value of 95.02%. The highest sensitivity was noted in walking behaviour followed by resting. The walking behaviour was at 100%, which is obtained using all feature sets. Specificity was higher than 93.76% for all activities, with walking having the highest at 100%. The lowest accuracy and kappa value were realised using only features extracted from the gyroscope sensor with accuracy and kappa value of 91.69% and 88.40%, respectively. Using the feature set D, browsing activity obtained a very low sensitivity of 36.72%, suggesting that gyroscope coordinates cannot distinguish well the browsing movements of the animal. Yet, browsing activity obtained the lowest sensitivity than the rest of the activities from all feature sets. The algorithm's performance was slightly decreased when using only features from the accelerometer sensor having accuracy and kappa value of 96.03% and 94.46%, respectively.

Table 3-11. RF performance using four different feature sets

	Animal Activities							
	Walking	Browsing	Grazing	Resting	Scratching	Standing	Accuracy	Kappa Value
set A*								
Sens	100.00%	78.91%	94.90%	97.29%	90.91%	95.48%	**96.43%**	**95.02%**
Spec	100.00%	99.99%	98.21%	99.34%	100.00%	97.46%		
Prec	100.00%	98.06%	93.28%	98.73%	100.00%	94.81%		
set B*								
Sens	100.00%	75.00%	94.66%	96.63%	81.82%	93.88%	95.60%	93.88%

Spec	100.00%	99.98%	97.33%	99.18%	100.00%	97.44%		
Prec	100.00%	94.12%	90.31%	98.42%	100.00%	94.69%		
set C*								
Sens	100.00%	79.69%	94.64%	97.14%	90.91%	94.57%	96.03%	94.46%
Spec	100.00%	99.98%	97.85%	99.18%	100.00%	97.45%		
Prec	100.00%	95.33%	92.03%	98.43%	100.00%	94.75%		
set D*								
Sens	100.00%	36.72%	87.10%	93.29%	72.73%	90.77%	91.69%	88.40%
Spec	100.00%	100.00%	96.79%	97.61%	100.00%	93.76%		
Prec	100.00%	100.00%	87.68%	95.38%	100.00%	87.62%		

* Sens = sensitivity, Spec = specificity, Prec = precision

3.2.3 Discussion

Monitoring the activity of livestock and wildlife animals is of great importance because of the highly valuable information that can be gained through such knowledge. For example, the animals' use of pasture can play an essential role in preventing soil erosion and contamination. Additionally, the animals' daily activity could be an indicator of their health status and could help the decision-making process of farm managers. The monitoring of the animals was always a humans' responsibility which is undoubtedly labour intensive and time-consuming. For this reason, automatic classification of animal activity is considered to be a solution to this problem. Also, machine learning can play an important role in a more intense observation of the animals because the machine could identify information that could be more difficult to be noticed through human observation alone.

In this experiment, data from seven Hebridean ewes located in Shotwick, Cheshire, was collected to be used for animal activity classification identifying six behaviours; walking, grazing, resting, browsing, scratching, and standing. Gyroscope and accelerometer measurements were obtained from a smartphone device placed on the animal's harness on the top or side using the HyperIMU application. Similarly with previous studies, we have selected RF to classify the activities as it demonstrated able to detect the activities with high accuracy [57], [64].

Similarities and differences with previous experiment A [64] reported in section 3.1 :

- In previous work, various machine learning algorithms were tested to identify the one that could yield the best results. Therefore, for the data collected in Experiment B, RF was used since it demonstrated the best performance, and hence it was evaluated with the new data.
- In this work, we used only sheep (Hebridean ewes), however, the data used in the previous work consisted of sheep and goats as well. With the outcome, we could be confident that this technique is powerful in classifying sheep and goat behaviour regardless of the breed, and we will test the findings with other breeds in the future.
- Similarly to previous work, the recording device was placed on the animals without fixed orientation. Hence, we have verified that the features used are independent of the sensors' orientation. This result is in agreement with previous work [266].
- Using only accelerometer measurements did not compromise the performance of the algorithm [266]. This conclusion could be used in future work when implementing a solution for real-time monitoring of the animals without conceding the device's battery life.
- The sampling frequency used in this work was at 10Hz in contrast with the previous work, which was at 100Hz. This is because it was noted that higher frequency rates have a negative effect on the memory and power of the device [180]. It was also previously demonstrated that a 16Hz sample rate and 7-second windows could reduce energy needs and be used for real-time sheep activity monitoring [55].

To conclude, RF yielded very high results in the classification of six behaviours of the ewes. Additionally, the kappa value demonstrates that the results are reasonable. Using features from both accelerometer and gyroscope, the algorithm obtained the best results with accuracy and kappa value of 96.43% and 95.02%, respectively. However, using only features extracted from the accelerometer, the accuracy was decreased by only 0.40% and the kappa value by 0.56%. Therefore, in the next experiment, only accelerometer measurements were used. We tested the same features with RF for activity recognition using a different sensor device to improve classification performance further.

3.3 Experiment C: Improving Random Forest algorithm performance using additional features

In this section, the third experiment is explored. This experiment aims to significantly improve the previously tested method [71] by expanding the feature set and decreasing the sliding window from 30 seconds to 5 seconds. In our experiment, we focused on four behaviours; grazing, walking, scratching, and inactive. For the experiment, we used only accelerometer data sampled at 12.5Hz, which was previously demonstrated adequate and did not compromise the battery life of the device [55].

3.3.1 Methods and equipment

This Section describes the materials and methods used to examine the performance of the RF algorithm regarding the classification of four mutually exclusive behaviours of sheep; grazing, walking, scratching, and inactive. Figure 3-9, shows the process followed in conducting the study.

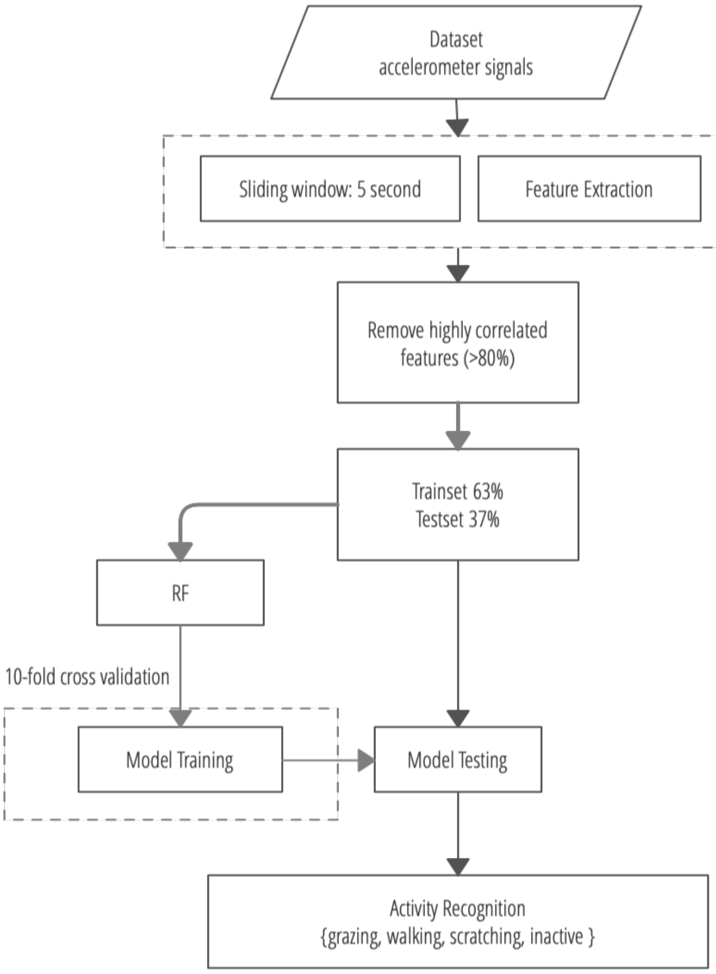


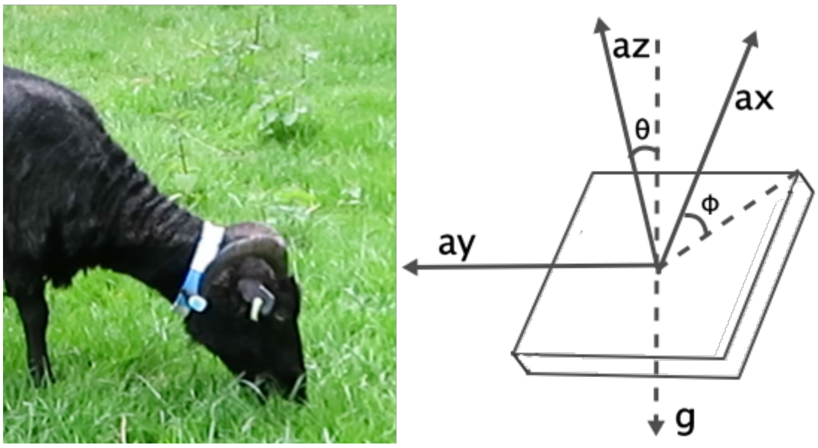
Figure 3-9 Methodology

3.3.1.1 Animals and Location

This study was conducted in July-August 2019 in Cheshire Shotwick (OS location 333781,371970), UK. Eight Hebridean ewes between the ages of 5-12 years were fitted with a sensor device collar. The animals were free to use a paddock of 1500m^2 area size and had access to grass and water all the time. The Senior Research Officer and LSSU Manager of Liverpool John Moores University approved the experiment's protocol (approval AH_NKO/2018-13).

3.3.1.2 Data Collection and Annotation

The MetamorionR® [267] wearable device was used for the current experiment. The sensor device collects motion and environmental data, however for this experiment, only accelerometer measurements were used. The device weighs 0.3oz, and its dimensions are 36mm x 27mm x 10mm with the case. Additionally, a 60mAH MicroUSB rechargeable li-po battery powers it. For this study, accelerometer measurements at a sample rate of 12.5Hz are used. The device logged and saved the data on its offboard memory as a CSV file.



(a) (b)

Figure 3-10 (a) Illustration of the device attached to the collar of the animal; (b) the accelerometer's signals

The animals were fitted with collars, which had the device attached in a nonfixed position to have a more generalised algorithm performance independent of the sensor orientation and position. Figure 3-10(a) illustrates an animal wearing a collar with the sensor device attached. The animals were video recorded during the morning, afternoon or night, and one observer was present each time. At the end of each day, the CSV file was saved for later use. Once all the recordings were completed with 40 hours of recorded behaviours, the accelerometer readings were time synchronised with the video recordings for behavioural annotation. For animal behaviour annotation, the ELAN_5.7_AVFX Freeware tool was used [82] and manually labelled the behaviours as grazing, walking, scratching, and inactive.

3.3.1.3 Data Pre-processing

After the data annotation, all the CSV files were merged and imported into Rstudio® for visualisation and analysis. The behaviours of interest for this study were: grazing, walking, scratching, resting, and standing. Behaviours such as fighting, shaking, and rubbing was not considered for this study. This resulted in utilising 28 out of 40 hours for analysis. Missing values were present in the data, and therefore they were eliminated.

3.3.1.4 Dataset

The behavioural data comprised of a set A={ t_i, ax_i, ay_i, az_i, y_i} for i=1,..,n, where n is the number of observations. The t is the timestamp, (ax, ay, az) is the accelerometer measurements, and y is the target vector where y∈{grazing, walking, scratching, inactive}. Equations (3.1), (3.2), and (3.3) define the acceleration vector (refer to Figure 3-10(b)):

$$ax = 1g*\sin\theta \tag{3.1}$$

$$ay = -1g*\sin\theta*\sin\phi \tag{3.2}$$

$$az = 1g*\cos\theta \tag{3.3}$$

Where θ is the angle between az relative to gravity, ϕ is the angle of ax relative to the ground, and g is the gravitational constant where $1g=9.81m/s^2$.

In this step, the magnitude of the acceleration defined as:

$$\text{Magnitude} = \sqrt{ax^2 + ay^2 + az^2} \tag{3.4}$$

Finally, the magnitude of the acceleration was plotted by activity to gain a visual understanding of the signal (refer to Figure 3-11). The plots show that the inactive state of standing and resting exhibit the same characteristics. Additionally, walking, grazing, and scratching display similar but still distinguishable peaks.

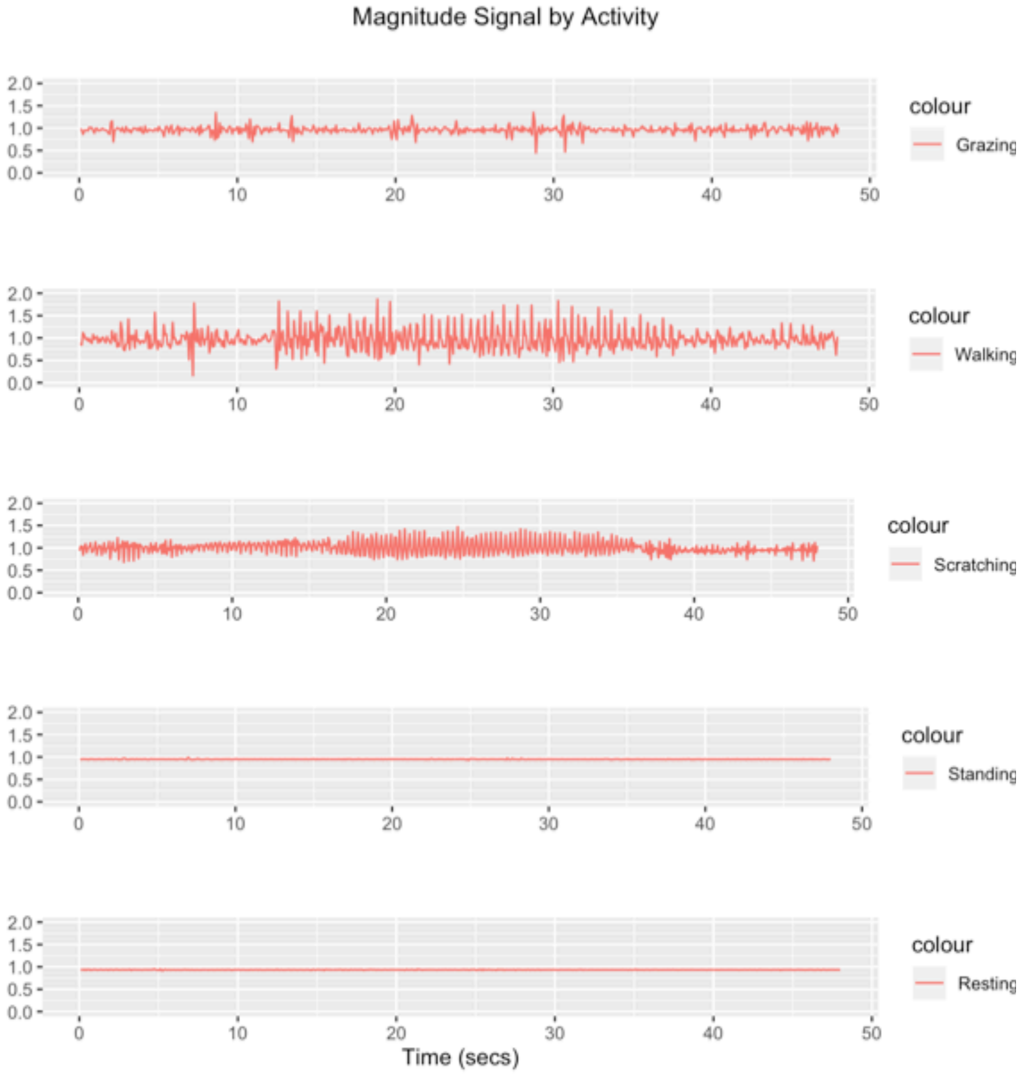


Figure 3-11 Visual Representation of the magnitude plotted by activity; grazing, walking, scratching, standing, resting, respectively

3.3.1.5 Feature Extraction

A total of 17 features were calculated from the new set A={$t_{i,}$ ax_i, ay_i, $az_{i,}$ mag_i, y_i}, resulting in a total of 68 newly created features (i.e. 17 features × four activities). Those features include the mean, standard deviation, root mean square, root mean square velocity, energy, the sum of changes, mean of changes, absolute and squared integrals, madogram

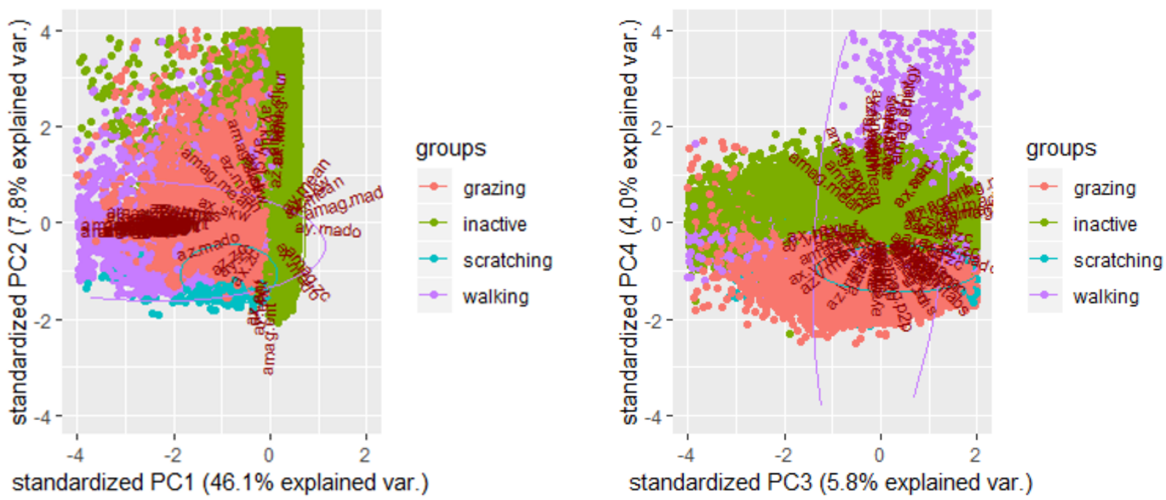
[96], peak frequency, peak to peak value, kurtosis and skewness, zero crossing, crest factor, and signal entropy. The features were extracted using a 5-second sliding window. Having a greater window in a real-time classification could provoke mislabelling because the animal might exhibit more than one behaviour in a short time interval; therefore, a 5-second window is considered sufficient.

3.3.1.6 Feature selection

The distributions for the first four principal components (PCs) concerning target class, original attributes and corresponding impacts of the target classes within the dataset are represented in Figure 3-12(a) and Figure 3-12(b). These figures also indicate the non-linearity of the problem, specifically in terms of the first four PCs covering the highest variances (~65%) within the overall principal components. Though, there is a small degree of overlap between all activities. However, this was expected since the animal's head movements might exhibit similar patterns in some instances. Furthermore, the plots help to understand the corresponding influence of the features within the datasets on the classification of animal behaviours (i.e. four target classes). For instance, in Figure 3-12(a), the Madogram of the magnitude & the Madogram of the y-axis of the accelerometer measurement have a clear impact on class 'inactive' as compared to root mean square velocity, which influences the 'scratching and grazing' classes.

The most commonly used dimensionality reduction technique, PCA, was used [268] to identify the most significant attributes/features within the dataset set and eliminate unnecessary features. In other words, PCA can be used to transform a large dataset containing a large number of features/variables to a lower dimension which still holds most of the information contained in the original high dimensional dataset. One of the important properties of PCA is the attribute loadings on the principal components that can also be used to identify attribute importance within the original dataset.

The correlation coefficient between the dataset attributes is represented by the principal components' loadings (i.e. obtained through PCA). The component rotations provide the maximised sum of variances of the squared loadings. The absolute sum of component rotations gives the degree of importance for the corresponding attributes in the dataset. Figure 3-13 shows the feature significance score within the original dataset, calculated through the PCs loadings. There are variations in the importance measure of features that can be used to identify and remove the dataset's unnecessary features. For instance, the 'madogram' of z and x-axis are indicated as the top-ranked variables compared to magnitude 'integrals' and 'rms' of the ay axis, which are indicated the least important variables within the original dataset.



(a) (b)

Figure 3-12 (a) First two PCA components' distributions; (b) 3rd and 4th components' distributions within the PCA components

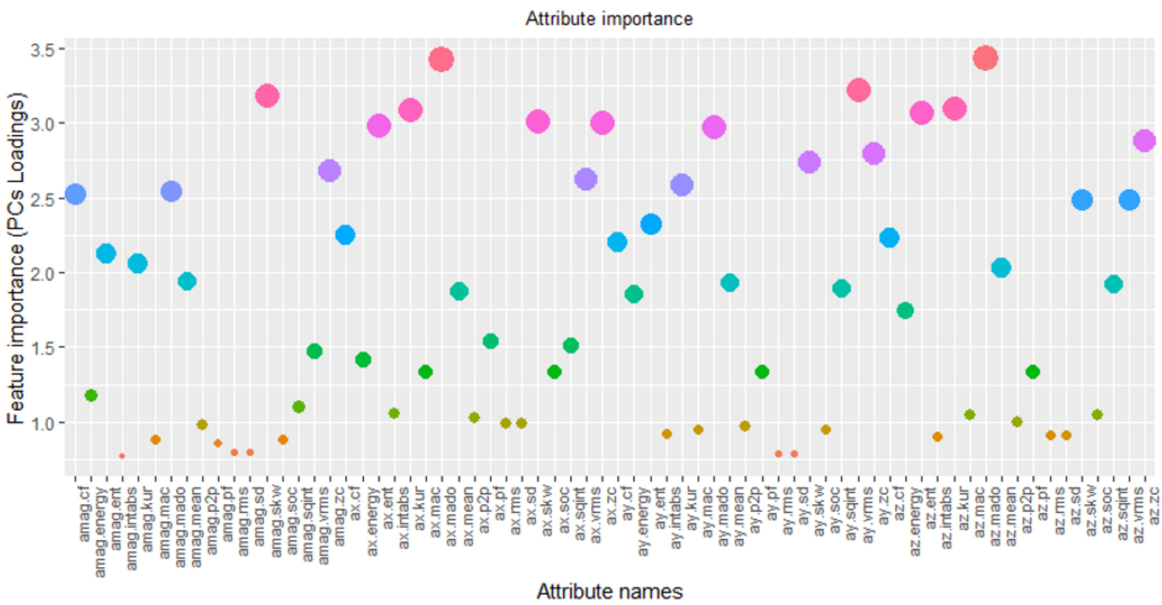


Figure 3-13 Measure of feature importance within the Dataset using principal components loading

To further investigate the features/attributes within the dataset, the correlation coefficients are used. The correlated features with correlation above 80% were removed, and the remaining features agree with the feature importance ranking indicated by the PCA. Therefore, we eliminated our features from 17 to 9. The remaining features are the mean, crest factor, root mean square velocity, skewness, kurtosis, madogram, zero-crossing rate, squared integrals, and signal entropy.

3.3.1.7 Classification and evaluation metrics

The classification algorithm selected to evaluate our dataset and test the activity prediction performance of the animals was the RF algorithm, as it was proved successful in our previous studies and other studies concerned with animal behaviour [9–11, 14, 24]. RF [151] is an ensemble method that consists of a combination of decision trees that are dependent on random values. All trees are sampled independently with the same distribution. The classification decision is then made based on the majority of votes from each tree.

To estimate the algorithm's performance, the model is evaluated with the Out of bag (OOB) accuracy. Decision trees learn from a subset of the dataset (63%), and unseen data (37%) are used for evaluation. This method is a good estimate of the ability of the model to generalise on unseen data [269]. We then recursively evaluated the performance using sensitivity, specificity, accuracy, and kappa value quality measures utilising 10-fold cross validation.

3.3.2 Results

The performance of RF is presented in Table 3-12. The four behaviours are classified correctly at a high rate. The overall accuracy of the algorithm is 99.43%, with a kappa value of 98.66%. Additionally, the f1-score is between 91.53%-99.90%. The lowest F1-score is resulted from scratching, and the highest from inactive behaviour. The sensitivities of all behaviours are between 98.26% to 99.87%. Also, the specificities are between 99.60% to 99.92%. Scratching was misclassified only once with grazing, while grazing was misclassified with scratching and walking in some cases. The same is valid with walking as it was misclassified with grazing and scratching. Only limited cases misclassified inactive behaviour with the other behaviours. However, the misclassification is limited, and consequently, the results showed high accuracy, sensitivity, and specificity in all 4 cases.

Table 3-12. RF Performance on unseen data

	Activities			
	Grazing	**Walking**	**Scratching**	**Inactive**
Sensitivity	98.26%	98.66%	99.87%	99.86%
Specificity	99.91%	99.60%	99.92%	99.84%
F1-score	98.97%	94.64%	91.53%	99.90%
Balanced Accuracy	99.08%	99.13%	99.90%	99.85%
Overall Accuracy: 99.43%, Kohen's Kappa value: 98.66%				

3.3.3 Discussion

In this study, we noted that the animals' movements, while they graze, could sometimes have similarities with the walking and scratching behaviour. Additionally, resting and standing provide a similar pattern of the acceleration signals because of the animals' inactive state, and this was also noted by Barwick et al. [56]. On the other hand, while the animals scratch or bite, the activity is detected easily as the magnitude changes markedly. While the animals are ruminating, the head movements are relatively small, and stationary compared with grazing and it does not interfere with the correct classification of the activity they perform. From the results, 5-second windows can provide a very good representation of activity pattern, and therefore could be suggested that this size is adequate. However, Decandia et al. [62], conducted experiments with various window sizes, such as 5, 10, 30, 60, 120, 180 and 300 seconds, and they identified that the best performance was obtained from a 30-second window having sensitivity of 94.8% for grazing, 80.4% for ruminating, and 92.3% for other behaviours. Though, the two studies cannot be compared because the ML model applied, the selection of features, and the position of the sensor are different. On the other hand, a 5-second window achieved the best performance in a study by Alvarenga et al.[60] when they compared 3, 5, and 10-second windows. The authors achieved an overall accuracy of 85.50% with Decision Trees and 5-second windows, which exhibited higher accuracy than 3 and 10-second windows. However, the variety of feature combinations, ML techniques, sample rate, and window size used in previous and current studies shows that there is still a need for further investigation, and there is no clear indication yet on the more suitable technique to be used for SAR.

This study was focused on detecting four mutually exclusive behaviours of interest to the animal health and production industry. Data were collected from eight Hebridean ewes located in Cheshire Shotwick, UK. Accelerometer signals were collected from a sensor that was attached to the collar of each animal. A total of 28 hours was used to test the performance of RF to detect each behaviour. The behaviours of interest were grazing, walking, scratching, and inactive. To test the algorithm, 17 features were extracted from the x, y, z, and magnitude of the acceleration signal resulting in 68 newly created variables.

Features with higher than 80% correlation are removed and eliminated the features to 9. The evaluation of the RF algorithm was then assessed using an out-of-bag (OBB) estimate, which is empirically proven that is as accurate as using a test set of the same size as the training set [269].

The results were very high for all the activities having accuracies of 99.08% for grazing, 99.13% for walking, 99.90% for scratching, and 99.85% for inactive. The overall accuracy and kappa value were 99.43% and 98.66%, respectively. The results showed that there is an important improvement over the previous method. The technique can be further tested and used for an online activity recognition system and be part of a multifunctional smart device for monitoring and controlling animal behaviour and position.

3.4 Summary

Sensor technologies play an essential part in the agricultural community and many other scientific and commercial communities. AAR is in the interest of the agricultural community, animal behaviourists, and conservationists since it acts as an indicator of the animal's health and nutrition. Accelerometer signals and ML techniques can be used to identify and observe behaviours of animals without the need for an exhaustive human observation which is labour intensive and time-consuming. In this Chapter, we provided information regarding three experiments we conducted concerned with AAR using ML methods. The purpose is to provide suggestions for employing intelligent devices to monitor the activities of free-range animals. Implementing such a device can be used as a smart assistant to provide valuable information regarding the food intake of the animals and their activities during the day, which can improve the decision making of the land managers. Such data can contribute to the animal's welfare, pasture utilisation and overall farm and animal decision management approach. A summary of each experiment is provided.

Experiment A: This experiment proposes a robust machine learning method to classify five activities of livestock. To prove the concept, a dataset was utilised based on the observation of two sheep and four goats. Time and frequency domain features are extracted from accelerometer, gyroscope, and magnetometer signals. Then, a feature selection technique, namely Boruta, was tested with MLP, RF, XGB, and KNN algorithms. The best results were obtained with RF achieving an accuracy of 96.47% and a kappa value of 95.41%. The results showed that the method could classify grazing, lying, scratching or biting, standing, and walking with high sensitivity and specificity.

Experiment B: In this experiment, accelerometer and gyroscope measurements were collected from seven Hebridean ewes located in Cheshire, UK. The animals were video recorded by a human observer. Once the activities of the animals were labelled as grazing, resting, walking, browsing, scratching, and standing, data analysis was conducted. The RF performance was evaluated as previously suggested that this algorithm can provide advantages and has been proven to adequately classify the behaviours of the animals (refer to experiment A). Using features from both the accelerometer and gyroscope, the algorithm obtained the best results with accuracy and kappa value of 96.43% and 95.02%, respectively. However, using data from the accelerometer exclusively decreased accuracy by 0.40% and kappa value by 0.56%. Therefore, in the following experiment, only accelerometer sensor data are used.

Experiment C: In this experiment, we employed an RF algorithm to identify grazing, walking, scratching, and inactivity (standing, resting) of 8 Hebridean ewes located in Cheshire, Shotwick, in the UK. We gathered accelerometer data from a sensor device that was fitted on the collar of the animals. The algorithm selection was based on previous research by which RF achieved the best results among other benchmark techniques. Therefore, in this study, more focus was given to feature engineering to improve prediction performance. Seventeen features from the time and frequency domain were calculated from the accelerometer measurements and the magnitude of the acceleration. Feature elimination was utilised in which highly correlated ones were removed, and only nine out of seventeen features were selected. The algorithm achieved an overall accuracy of 99.43% and a kappa value of 98.66%. The accuracy for grazing, walking, scratching, and inactive was 99.08%, 99.13%, 99.90%, and 99.85%, respectively. The overall results showed a significant improvement over previous methods and studies for all mutually exclusive behaviours. Those results are promising, and the technique could be further tested for future real-time activity recognition.

The next chapter will discuss the use of CNN Transfer learning for SAR.

Chapter 4 Deep Transfer Learning for Sheep Activity Recognition

In this Chapter, a system for monitoring sheep activity using accelerometer data with robustness to accelerometer specifications, position, and orientation is proposed. To solve the problem of heterogeneity of accelerometer sensors, Transfer Learning (TL) based on Convolutional Neural Networks (CNN) in animal behaviour classification is suggested. The aim is to use the pre-trained model obtained from the source dataset (sensor 1), and test it on the target dataset (sensor 2), using the learned features and weights. This will allow testing of the hypothesis of generalisation in animal behaviour classification, independent of the type and orientation of the accelerometer device, by leveraging the pre-trained model obtained on a larger dataset (sensor 1), to classify behaviour on a smaller dataset (sensor 2).

The contributions of this part of the thesis are as follows:

- A real-time data-driven approach for animal activity recognition is proposed comprising: a) grazing, b) active, and c) inactive states using a composite of the CNN and hand-crafted features, which significantly improves classification performance.
- Two primary datasets are acquired through different sensors, which are made publicly available to the research community.
- The use of deep Transfer Learning on data gathered from two types of sensors located on the collar of sheep in a non-fixed orientation, to introduce variability in the dataset, hence evaluating the generalisation properties of the proposed sheep activity recognition

approach. To the best of the researcher knowledge, the proposed Transfer Learning approach is used for the first time in animal behaviour recognition.

The remainder of this Chapter is organised as follows. Section 4.1 provides an overview of transfer learning works. Section 4.2 describes the proposed activity recognition algorithm and presents the description of the datasets and the system methodology. Section 4.3 presents the results and discussion, followed by Section 4.4, which summarises Chapter 4.

4.1 Introduction

Various studies explored the identification of animal activities based on accelerometer data [9], [34-36]. Wearable devices using accelerometers have been commonly used with ML techniques to recognise cattle behaviour [34], [37-45]. Furthermore, ML was used to identify the activities of horses [279], sharks [180], seals [280], goats [49, 50] and other domesticated or wild animals. While many research studies addressed the classification of animal behaviour, specifically in sheep, there is still a need for further investigation on the optimisation of the devices and techniques used [46]. Different studies propose diverse models, setups and devices using ML[263][68].

In the context of Transfer Learning, Oquab et al. [282] proposed transfer learning to extract mid-level features from the ImageNet dataset [283] and reuse the representations on smaller datasets. Xia et al. [284] proposed ensemble concepts of multiple Transfer CNNs (TCNN) to improve model generalisation, by introducing three ensemble TCNNs. The authors used several datasets and reported enhanced accuracy. They showed better generalisation over CNNs and a single TCNN. Transfer learning has been successfully used in object recognition [285] and human activity recognition [286].

Those research works demonstrate the importance of using TL in utilising knowledge acquired from one domain to another, which saves time and supports the application of robust models, specifically in limited size datasets. Additionally, in the field of AAR and wearable devices in general, it provides opportunities to explore various sensor devices and quickly adapt to new sensor configurations, building upon the time and effort spent on the original work. In summary, transfer learning supports adapting new sensors in the problem domain without thorough training in developing new predictive models. This research is the first of its kind to propose CNN transfer learning for animal activity recognition, specifically in sheep, and highlight the benefits of such an approach in the context of sensor devices, i.e., sensor heterogeneity, variations in sensor orientation and data gathering.

4.2 Methodology

Two types of accelerometer sensors (metamotionR and SenseHat) are used in this research placed on the sheep collar to capture the primary data with a sample rate of 12.5Hz. Let D_S represents the data captured from metamotionR, which is considered as the source data, whereas D_T represents the data acquired through RaspberryPi (with the SenseHat board attached) and will be used to validate the reusability of transfer learning, i.e., the target data. Both datasets were labelled manually and normalised using the z-score. CNN was used to identify the activities of animals using supplementary time and frequency domain features. Temporal and spectral features were extracted using a sliding window of 2s with 50% overlap, resulting in two additional datasets from the metamotionR and SenseHat sensors referred to as D_{S+}, D_{T+}, respectively. It should be noted that D_{S+}, D_{T+} relate to the augmented datasets, which include the hand-crafted features, whereas D_S and D_T consist only of the x, y, and z accelerometer values and their magnitude.

Extensive simulation experiments were carried out on six CNN models, which were trained using D_S and D_{S+}. The outcome of the simulations was the selection of the top-performing CNN configuration based on the accuracy obtained on the test sets. These models were then stored, and transfer learning was used on D_T and D_{T+}. Figure 4-1 illustrates the overall procedure, including the application of transfer learning.

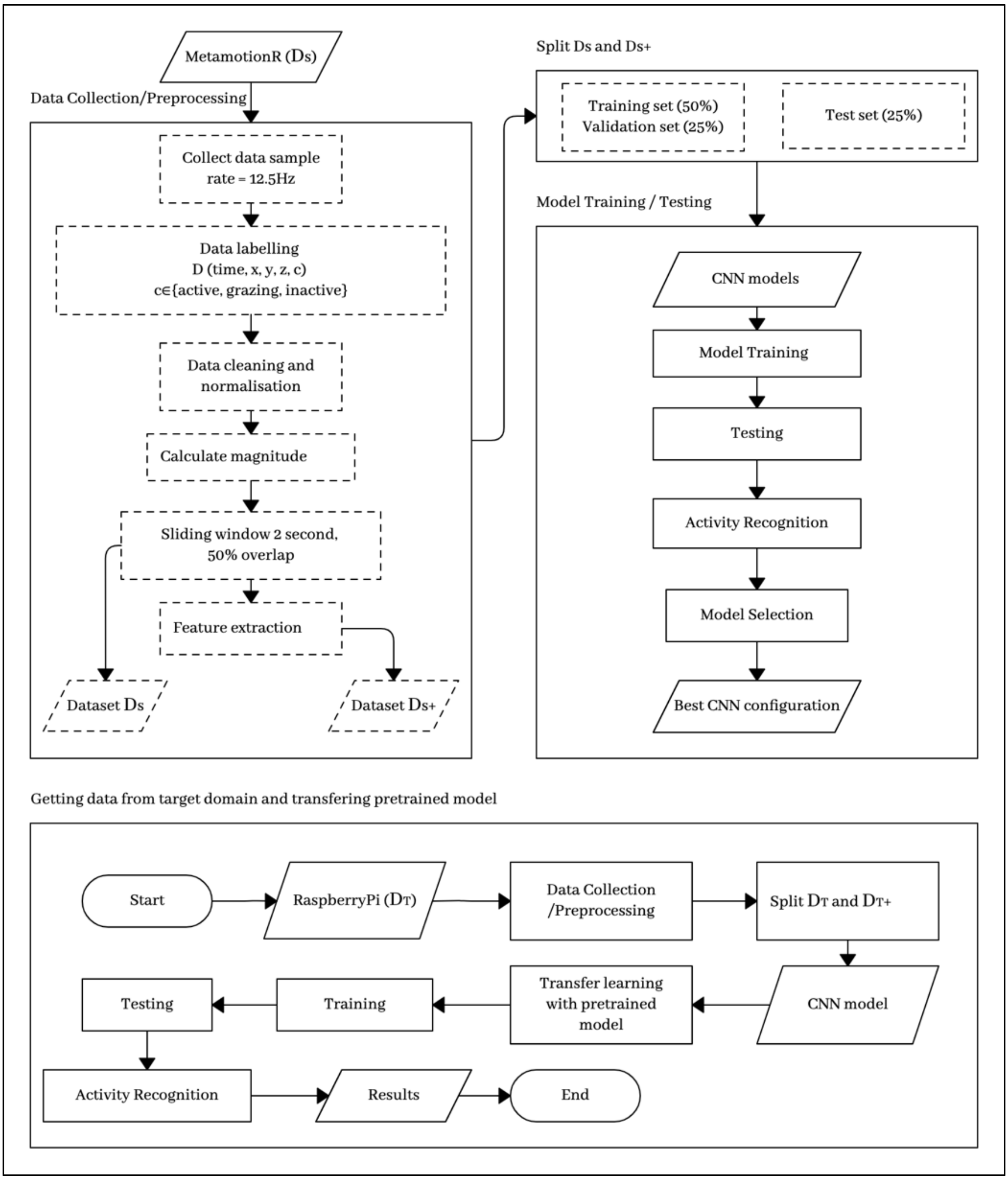


Figure 4-1 System methodology

4.2.1 Datasets

In this work, two primary datasets comprising the accelerometer measurements from a flock of 9 Hebridean ewes 35±5 kg, 9±5 years old at a farm located in Cheshire Shotwick (OS location 333781,371970) were collected. Ethical approval was obtained from Liverpool John Moores University to collect the datasets and conduct the experiments (Ref: AH_NKO/2018-13). To acquire the data, two types of devices were used, mounted on the collar of the animals. The first device was mounted on a fixed position (at 270° degrees) and orientation; however, the second device was mounted in a non-fixed manner (at 0°, 90°, or 180° degrees) to test the performance of the proposed methodology, independent of sensor orientation and position. MetamotionR [267] was used to collect motion and environmental data, while RaspberryPi with the SenseHat board collected temperature, humidity, pressure and motion measurements. Both sensor outputs were logged, and the data was stored with a sampling rate of 12.5Hz. In total, recordings of over 65 hours of activity were obtained, which resulted in a dataset of 2,925,000 samples.

The datasets collected from the ewes were loaded into the ELAN_5.7_AVFX Freeware tool [82]. Behaviours such as running, fighting, and shaking were limited and were excluded from the datasets. Walking and scratching behaviours were merged and considered as a unified behaviour, labelled as 'active'. Likewise, standing and resting were joined together and labelled as 'inactive'. Chewing with the animal head down while walking or standing was labelled as 'grazing'. The dataset acquired with MetamotionR (D_S) contains 1,048,575 samples, whereas the dataset captured via RaspberryPi (D_T) comprises 762,860 samples. We refer to these datasets as the source and target datasets, respectively. Let D_k represents the joint dataset, where $D_k=\{t_i, x_i, y_i, z_i, c_i\}$, i=1,..,n, n is the number of observations, and the datasets $k=\{S,T\}$. Parameter t relates to the timestamp, while x, y, z are the accelerometer measurements in the x, y, and z axes, respectively, and c is the label variable, where $c \in \{active, grazing, inactive\}$. Figure 4-2 presents the distribution of the three activities within our datasets. The charts clearly indicate the imbalanced distribution of activities within both datasets, as expected, due to the nature of the study and the activities considered.

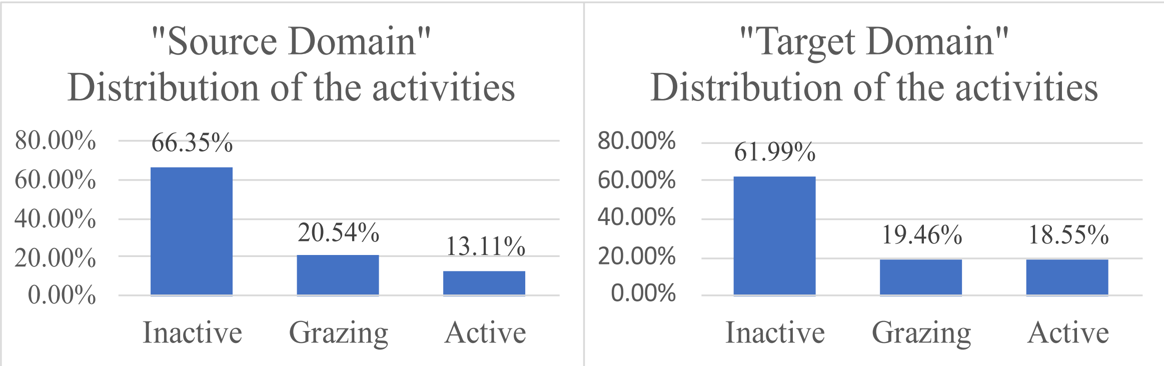


Figure 4-2 Duration of the activities for source and target domains

4.2.2 Data pre-processing

The magnitude of the accelerometer measurements was calculated, resulting in an additional feature in $\mathbf{D_k}$={ t_i, x_i, y_i, z_i, mag_i, c_i} , calculated as shown in Equation (4.1):

$$\text{mag} = \sqrt{x^2 + y^2 + z^2} \tag{4.1}$$

where, x, y and z represent the 3-dimensional accelerometer data. The dataset was normalised for zero mean (μ) and a standard deviation (σ) of 1. The normalised datasets were then partitioned into training, validation, and testing, with a ratio of 50%, 25%, and 25%, respectively. Overlapping was used to enable real-time classification with a ratio of 50%, which has been shown to be effective in previous activity recognition studies [97]. Using a larger window size in real-life classification could lead to mislabelling since animals may exhibit more than one behaviour in a short time interval. Thus, a 2s window was considered sufficient while not compromising the device battery life. When animals showed more than one behaviour within the same time window, this was labelled with the most frequently occurring activity.

4.2.3 Feature extraction

Temporal and spectral analysis of the overlapping window data to extract meaningful information from the datasets was performed, supporting behaviour recognition. The features were selected based on previous findings [64], [71], as well as state-of-the-art research in terms of feature importance using gait information in human and animal activity recognition [95], [96], [287]. A total of 13 features were calculated for each of the x, y, z accelerometer data and the magnitude of the acceleration signal for each activity, resulting in a 52-dimensional feature set as illustrated in Table 4-1.

Boruta algorithm [119] was used to explore and rank the importance of the extracted feature set and confirm that all features contribute meaningful information, as shown in Figure 4-3. It can be deciphered that the mean of the x-axis acceleration, skewness and kurtosis of the acceleration magnitude, fractal dimension of the z-axis acceleration, and skewness of the y axis acceleration are the most significant features. On the contrary, energy, integrals, RMS, peak frequency and squared integrals of the y axis acceleration are low ranked features. However, all features contribute to the discrimination of the three activities, confirmed by the Boruta algorithm. Thus, no feature elimination was performed.

Table 4-1 List of extracted features

#	Feature name	Equation
1	Mean	$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$
2	Standard Deviation	$s = \sqrt{\sum_{i=1}^{n} \frac{(a_i - \bar{a})^2}{n}}$
3	Skewness	$skewness = \frac{1}{n}\sum_{i=1}^{n} \frac{(a_i - \bar{a})^3}{s^3}$
4	Kurtosis	$kurtosis = \frac{1}{n}\sum_{i=1}^{n} \frac{(a_i - \bar{a})^4}{s^4}$

5	RMS	$rms = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {a_i}^2}$
6	RMS velocity	$rmsV = \sqrt{\frac{1}{n}\sum_{i=1}^{n} diffinv(a_i)^2}$, *diffinv()* is the inverse function of the *diff()*
7	Sum of Changes	$soc = \sum_{i=1}^{n} diff(a_i)$, where *diff()* computes the consecutive differences of the vector
8	Mean of Changes	$moc = \frac{1}{n}\sum_{i=1}^{n} \lvert diff(a_i)\rvert$
9	Integrals	$Integrals = \int_{t=0}^{T} \lvert a_1\rvert\, dt + \int_{t=0}^{T} \lvert a_2\rvert\, dt + \int_{t=0}^{T} \lvert a_3\rvert\, dt$
10	Squared Integrals	$Integrals^2 = \left(\int_{t=0}^{T} \lvert a_1\rvert\, dt\right)^2 + \left(\int_{t=0}^{T} \lvert a_2\rvert\, dt\right)^2 + \left(\int_{t=0}^{T} \lvert a_3\rvert\, dt\right)^2$
11	Madogram	$\gamma_p(t) = \frac{1}{2}\mathrm{E}[a_i - a_{i+u}]$, where t=lag, E[.]=expectation
12	Energy	$energy = \sum_{i=1}^{n} {a_i}^2$
13	Peak Frequency [52]	$pf = fmax = \arg\left(\frac{fs}{n} max_{i=0}^{n-1} P(i)\right)$ fs = sampling frequency, P(i) = power of the spectrum
a=accelerometer signal x, y, and z where a_1=x, a_2=y, a_3=z n is the number of rows in the signal window T is the number of data points in each window		

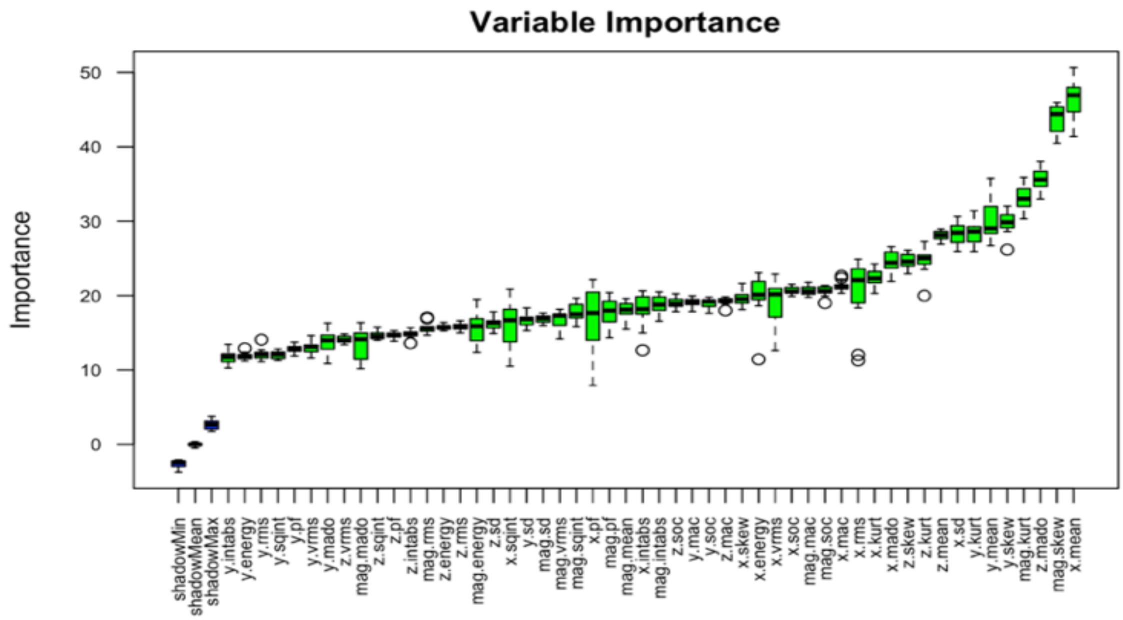


Figure 4-3 Feature Importance illustration from the Boruta Algorithm for $\mathbf{D_{S+}}$. Note that features shadowMin, shadowMean, shadowMax are created by the Boruta algorithm in order to rank the original features. More information can be found in [119]

4.2.4 CNN and Transfer Learning

The recorded measurements and associated features were used in the proposed Deep Learning models to classify the target activities and also during transfer learning.

CNN is a hierarchical feed-forward neural network, widely used in image classification tasks because of its ability to perform well on extensive and complex datasets [64, 65]. In recent years, CNN has also become popular in activity recognition problems [290], [291]. Further details about CNN can be found in [289], [292].

The main idea behind transfer learning is to gain knowledge from a dataset (source domain D_S) and then transfer knowledge to a new dataset (target domain D_T) to improve learning in the target domain [293]. Thus, a source domain D_S: $X_S \rightarrow Y_S$, is defined with feature space X_S, and a label set Y_S, such that $D_S=\{(\mathbf{x_i}, y_i)\ldots,(\mathbf{x_n}, y_n)\}$, for i=1,…,n, where n is the number of observations in the dataset, and $\mathbf{x_i} \in X_S$, $y_i \in Y_S$. Additionally, we have a target domain D_T: $X_T \rightarrow Y_T$, with feature space X_T, and a label space Y_T, such that $D_T=\{(\mathbf{x_j}, y_j),\ldots,(\mathbf{x_m}, y_m)\}$, where $\mathbf{x_j} \in X_T$, and $y_j \in Y_T$, for j=1,…,m, where m is the number of observations in D_T. Furthermore, a "task" (T) consists of the label Y and the predictive function $f(\cdot)$, and is denoted as $T = \{Y, f(\cdot)\}$, which can be learned from the training data, and used to predict the labels of unseen vectors.

In the case where $X_S \neq X_T$, i.e., source and target datasets, may come from different domains, including different marginal/predictive distributions and feature/label spaces, transfer learning is described as heterogeneous. Alternatively, when $X_S = X_T$, transfer learning is defined as homogeneous. In the current study, homogeneous transfer learning is used because both source and target domains' feature space and domain characteristics are the same. The difference between the source and target datasets in this study is the accelerometer sensors used and the orientation and position of the sensor. Additionally, the motion measurements of the second device exhibit some noise and the size of the second dataset is smaller.

4.2.5 Experimental Design

To set the baseline for the experiments, several classification trials were conducted to investigate the proposed methodology's performance and configure the deep learning models. For the CNN, datasets D_S and D_{S+} are used, including the original measurements and the hand-crafted features, with 50%, 25%, and 25% ratios for training, validation, and testing, respectively, for both datasets. The feature extraction process accommodates for the extraction of temporal and spectral features that represent animal behaviour in a very succinct fashion. Therefore, it was not necessary to incorporate dynamics in the actual ML system as per the application of LSTM. For this reason, CNN model was selected for the purpose of this experiment.

Six CNN configurations (models A to F) were utilised, where the number of convolutional layers and the dropout rate (in the range of [0.1, 0.5]) are varied. The common configuration for all models includes: a) for each layer, the rectified linear activation function

(ReLu) was applied; b) smaller dropouts were applied in the convolution layers, whereas larger dropouts were applied in the fully connected layers (i.e., 0.5); c) for all six CNN models, early stopping is used, modified to monitor the minimum validation loss with patience of 20. Early stopping was selected to avoid overfitting during the training of the data, as well as to reduce the number of epochs needed. During the training, a record is kept of the loss function on the validation data and once there is no improvement on the performance, the model stops the training without going through all the epochs using the "patience" parameter. Therefore, the epochs are reduced; d) for model optimisation, the Adam optimiser was applied with a learning rate lr=0.001, while the loss function was set to categorical cross-entropy [289]; e) a SoftMax layer was used for the classification of target activities. The weight regularisation with an L_2 vector norm of [0.1, 0.01, 0.001, 0.0001, 0.00001] is used to identify the best configuration during training. The vector norm was only applied to the fully connected layers of the models.

The performance of the models for each configuration is shown in Table 4-2. The graphical representation of the accuracy obtained from all models is illustrated in Figure 4-4. Model A, which has the highest accuracy, was selected with L_2=0.00001 to be further used for transfer learning. Model A consists of two convolutional layers, one fully connected layer and an output layer. The first convolutional layer uses 16x16 convolutional filters followed by a 10% dropout layer, and the second convolutional layer uses 32x32 convolutional filters followed by a 20% dropout. The fully connected layer has 64 filters, followed by a 50% dropout, and the output layer uses SoftMax (refer to Figure 4-5)

Table 4-2 CNN Performance on DS and DS+ for each model using weight decay with the l2 norm

		Accuracy			
		D_S		D_{S+}	
Model	L2	Training set	Test set	Training set	Test set
A	1.00E-01	0.9600	0.9574	0.9564	0.9576
A	1.00E-02	0.9701	0.9652	0.9622	0.9624
A	1.00E-03	0.9753	0.9700	0.9716	0.9690

A	1.00E-04	0.9776	0.9700	0.9787	0.9768
A	1.00E-05	0.9759	0.9746	0.9892	0.9855
B	1.00E-01	0.975	0.9670	0.9799	0.9792
B	1.00E-02	0.9807	0.9715	0.9850	0.9821
B	1.00E-03	0.9813	0.9684	0.9853	0.9816
B	1.00E-04	0.9800	0.9685	0.9832	0.9809
B	1.00E-05	0.9778	0.9677	0.9847	0.9806
C	1.00E-01	0.9403	0.9365	0.8696	0.8697
C	1.00E-02	0.9606	0.9566	0.9578	0.9608
C	1.00E-03	0.9641	0.9607	0.9691	0.9700
C	1.00E-04	0.9713	0.9659	0.9689	0.9702
C	1.00E-05	0.9744	0.9658	0.9796	0.9796
D	1.00E-01	0.9535	0.9548	0.9622	0.9605
D	1.00E-02	0.9498	0.9479	0.9752	0.9733
D	1.00E-03	0.9576	0.9545	0.9804	0.9781
D	1.00E-04	0.9552	0.9501	0.9863	0.9824
D	1.00E-05	0.9664	0.9580	0.9897	0.9841
E	1.00E-01	0.7904	0.7932	0.9493	0.9485
E	1.00E-02	0.9193	0.9210	0.9683	0.9675
E	1.00E-03	0.9449	0.9453	0.9774	0.9743
E	1.00E-04	0.9183	0.9187	0.9830	0.9798
E	1.00E-05	0.9584	0.9532	0.9873	0.9821
F	1.00E-01	0.9739	0.9710	0.9826	0.9789
F	1.00E-02	0.9790	0.9716	0.9877	0.9822
F	1.00E-03	0.9795	0.9726	0.9871	0.9815
F	1.00E-04	0.9734	0.9721	0.9833	0.9810

F	1.00E-05	0.9753	0.9730	0.9851	0.9829

Figure 4-4 Graphical representation of accuracy obtained from the CNN models trained on D_S^{tr}, D_{S+}^{tr}, and tested D_S^{ts}, D_{S+}^{ts}.

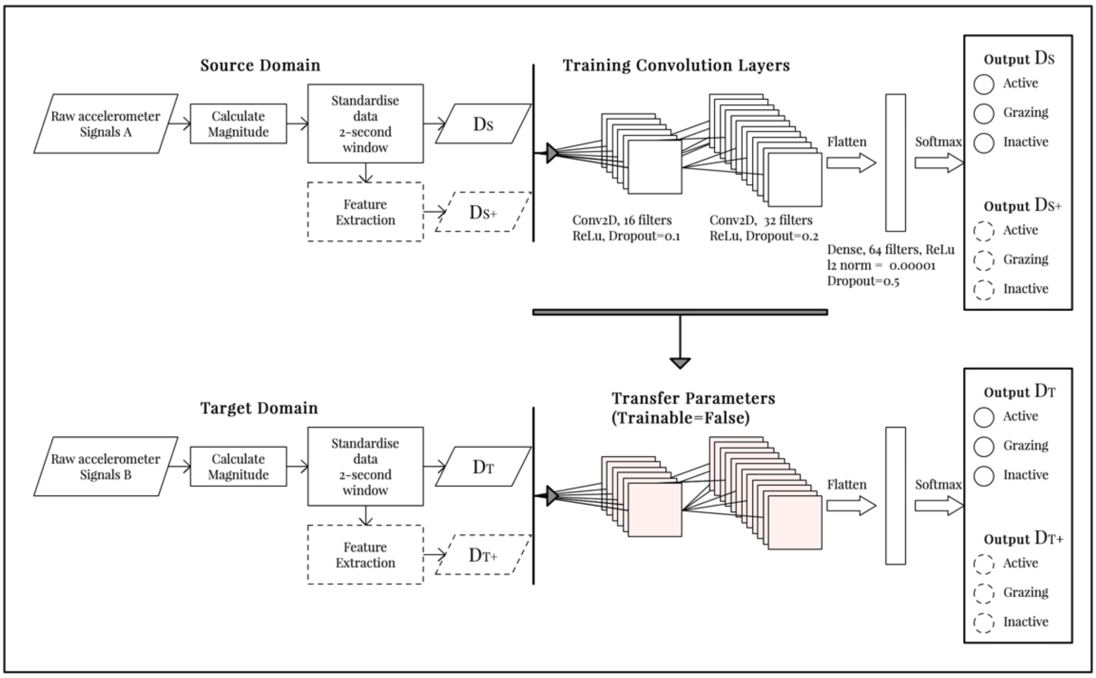


Figure 4-5 CNN Architecture for the proposed model

4.3 Results and Discussions

In this work, two experiments using CNN model A were conducted since the model achieved the best results while partitioning the datasets into 50%, 25%, and 25% for the training, validation, and testing purposes, respectively. The first experiment used the original source and target datasets. In contrast, the second experiment followed the same procedure as the first, but this time the datasets which included the hand-crafted features are used. The purpose is to investigate whether the addition of the time and frequency domain features affect the performance of the model and the generalisation of the algorithm since the direction and placement of the accelerometers differ between the two sensor setups. For each experiment, several statistical metrics, including precision, recall, F_1 score, and accuracy, were used to evaluate the performance of deep learning and transfer learning over different combinations of training and testing datasets. Figure 4-5 illustrates the architecture of the proposed model and transfer learning procedure.

4.3.1 Experiment A: Transfer learning from D_S to D_T

Model A was trained on the training set D_S^{tr}, validated on D_S^{val}, and then tested on D_S^{ts}. For transfer learning, the trained model was stored to be later reused on the target domain, i.e., DT. The only trainable layers during transfer learning on D_T^{tr} were the fully connected layers, which are responsible for the classification of the target activities. The results obtained from both experiments are shown in Table 4-3. Figure 4-6 and Figure 4-7 illustrate the accuracy and loss of the model per epoch, respectively.

Table 4-3 CNN model A classification results on D_S^{ts}, and using Transfer Learning on D_T^{ts}

	D_S^{ts}			D_T^{ts}		
Activities	Accuracy: 0.9746			Accuracy: 0.9479		
	Precision	Recall	F_1 score	Precision	Recall	F_1 score
Active	0.9505	0.8942	0.9215	0.9322	0.8898	0.9105
Grazing	0.9332	0.9745	0.9534	0.8986	0.9417	0.9196
Inactive	0.9963	0.9935	0.9949	0.9982	0.9943	0.9963
Average	0.9600	0.9541	0.9566	0.9430	0.9419	0.9421

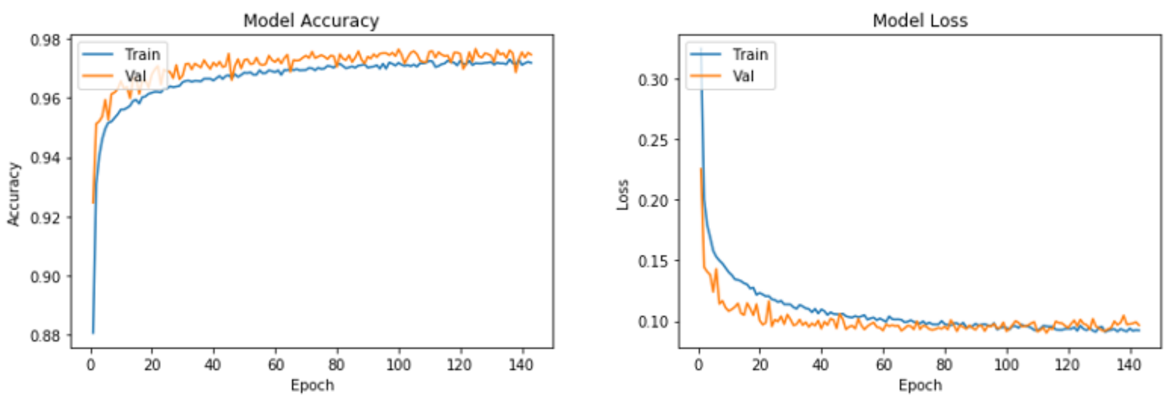


Figure 4-6 Experiment A: CNN model A training on source domain

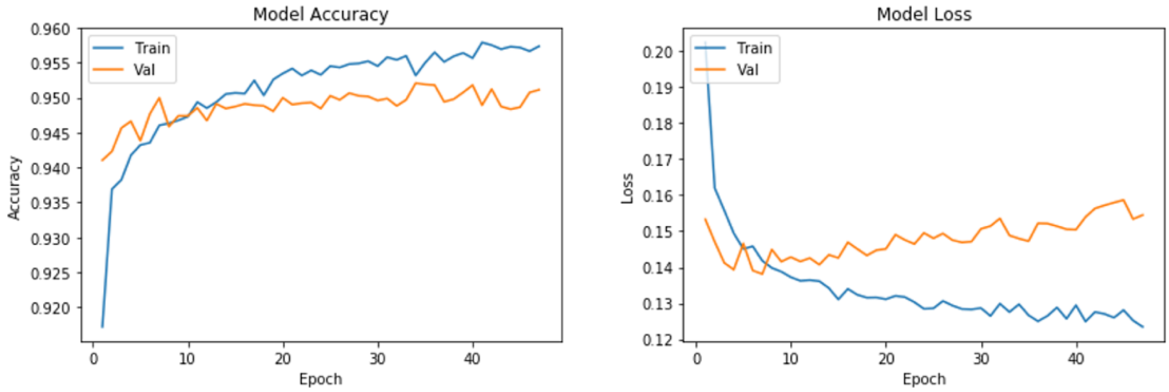


Figure 4-7 Experiment A: CNN model A transfer learning on target domain D_T^{tr}

Table 4-3 indicates the overall accuracy obtained is 97.46% on D_S^{ts} and 94.79% on D_T^{ts}. The highest precision, recall, and F_1 score on both test sets is noticed on the 'inactive' behaviour, having scores above 99.35%. The lowest recall was obtained in D_T^{ts} having 88.98% on the 'active' behaviour, which is similar for D_S^{ts}, indicating that recall is 89.42%. On the other hand, the precision of active behaviour on both test sets is higher than that of the grazing behaviour with 95.05% on D_S^{ts} and 93.22% on D_T^{ts}. Recall results for grazing behaviour are 97.45% on D_S^{ts} and 94.17% on D_T^{ts}, respectively. The best predictive rate was achieved in 'inactive' behaviour. The F_1 scores for D_S^{ts} are 92.15%, 95.34%, and 99.49% for active, grazing, and inactive behaviour, respectively. For D_T^{ts}, the F_1 scores for active and grazing behaviour are at 91.05% and 91.96%, respectively, which is lower than the associated scores on D_S^{ts}. However, inactive behaviour on D_T^{ts} has an F_1 score of 99.63%. Overall, it can be observed that the model performed better on the source dataset in all cases, and the accuracy decreased when the model was transferred to the target dataset. A reason for this decrease may be that the orientation of the sensor on the second device was not fixed, showing different patterns.

On the other hand, multiple factors may contribute to the model's biased performance in the case of inactive behaviour (i.e., higher precision, recall and F_1 score than other classes). Firstly, grazing behaviour can be easily misclassified as walking or scratching since these behaviours exhibit similar movements in some cases. Likewise, inactive behaviour can be easily classified, contrary to active and grazing, since the pattern does not indicate changes and remains stable due to motionless behaviour. Finally, the distribution of data samples for inactive behaviour is comparatively more prominent than the other two classes. Thus, class imbalance may be one of the major causes that the model performs better in identifying inactive behaviour.

4.3.2 Experiment B: Transfer learning from D_{S+} to D_{T+}

Experiment B is identical to experiment A, except for the datasets, including hand-crafted features (i.e., D_{S+}, D_{T+}). Model A was trained on D_{S+}^{tr}, validated on D_{S+}^{val}, and then transfer learning was performed on the target data, i.e., D_{T+}^{tr}. The final model was then tested on the unseen data from D_{S+}^{ts} and D_{T+}^{tr}. Similar to experiment A, only the fully connected layer was allowed to be trained. Results obtained from both tests are presented in Table 4-4. Figure 4-8 and Figure 4-9 illustrate the accuracy and loss of the model per epoch.

Table 4-4 CNN model A classification results on D_{S+}^{ts}, and using Transfer Learning on D_{T+}^{ts}

	D_{S+}^{ts}			D_{T+}^{ts}		
Activities	Accuracy: 0.9855			Accuracy: 0.9659		
	Precision	Recall	F1 score	Precision	Recall	F1 score
Active	0.9498	0.9422	0.9460	0.9309	0.8712	0.9000
Grazing	0.9646	0.9669	0.9657	0.9248	0.9551	0.9397
Inactive	0.9987	0.9994	0.9991	0.9917	0.9949	0.9933
Average	0.9710	0.9695	0.9703	0.9491	0.9404	0.9443

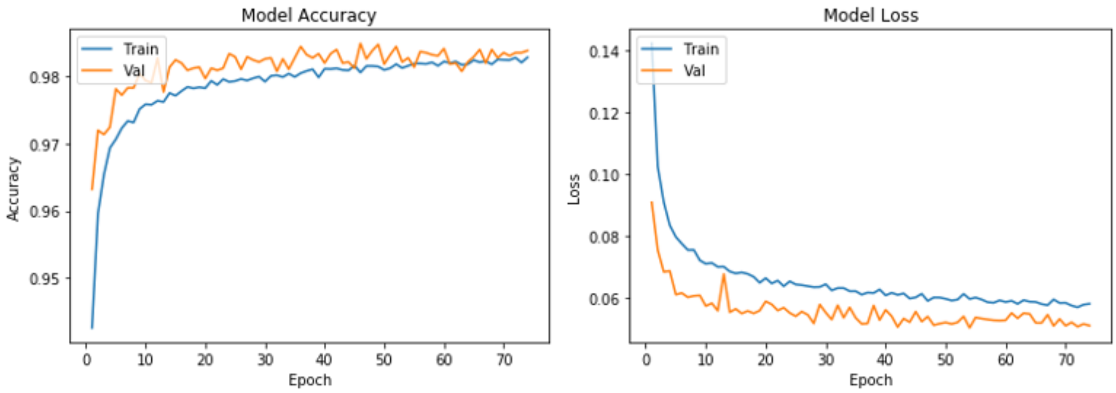


Figure 4-8 Experiment B: CNN model A training on source domain D_{S+}^{tr}

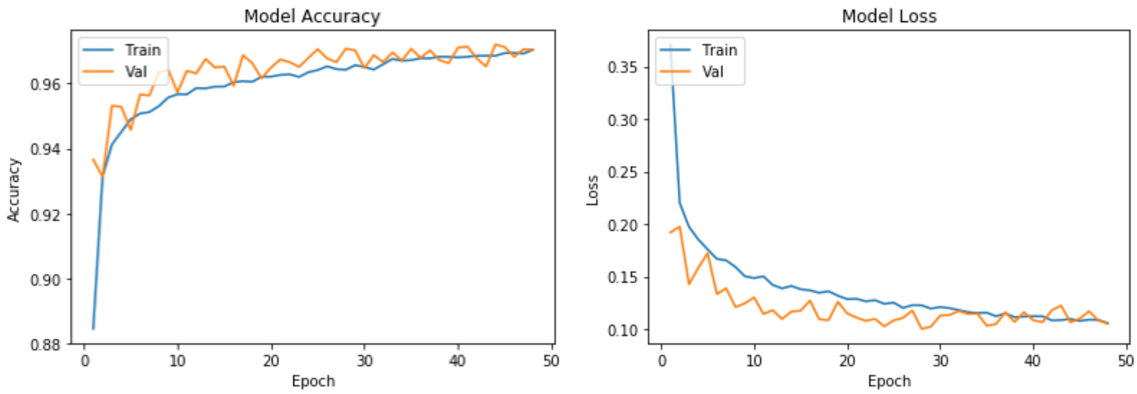


Figure 4-9 Experiment B: CNN model A transfer learning on target domain D_{T+}^{tr}.

Table 4-4 presents the model performance over the datasets with the hand-crafted features, which indicate overall accuracies of 98.55% and 96.59% for D_{S+}^{ts} and D_{T+}^{ts}, respectively. These outcomes also align with the results from experiment A, where accuracy decreased when the model was transferred to the target domain. Likewise, the classification accuracy of inactive behaviour is comparatively better with precision and recall of 99.87% and 99.94%, respectively, on D_{S+}^{ts}, and 99.17% and 99.49% on D_{T+}^{ts}, respectively. Precision and recall for the source domain's active behaviour achieved 94.98% and 94.22%, respectively. There is a noticeable increase of 4.8% in the recall result on active behaviour in experiment

B, compared with experiment A, which achieved a recall of 89.42% on the source domain. The results from all behaviours for experiment B on the source domain are comparatively balanced. On the other hand, recall on active behaviour in the target domain is 87.12%, which is slightly decreased (i.e., by 1.8%), compared to the recall on the target domain from experiment A. The F_1 scores obtained on the source domain are 94.60%, 96.57%, and 99.91% for active, grazing, and inactive behaviour, respectively, which is also slightly higher when compared to the F_1 scores obtained in the experiment A over the source domain. On the other hand, testing the model on the target domain shows a slight decrease in the F_1 score (i.e., 1.05%) on active behaviour but an increase of 2.01% on the grazing behaviour. In both experiments, the F_1 scores for inactive behaviour are above 99.33% on the target domain. Regarding the grazing behaviour, precision and recall on the source domain is 96.46% and 96.69%, respectively. Grazing behaviour precision on the target dataset is 92.48%, while recall is 95.51%.

A summary of the obtained results is presented in Table 4-5 and Table 4-6. From the overall results, it is observed that the CNN achieved the best accuracy on both source and target domains when using the datasets with the hand-crafted features. The experiments also indicate high F_1 scores for the inactive behaviour, in the range of 99.33%-99.91%. The lowest F_1 score (90.00%) is achieved on the grazing behaviour when the TL is applied to D_{T+}^{ts}. However, when conducting experiment B on D_{S+}^{ts}, F_1 scores of 96.57% on grazing and 94.60% for active behaviours are obtained. These outcomes indicate the superiority of the proposed model, when compared to previously published studies. For instance, [21] showed 69.8% and 45.2% precision performance for the lying and standing behaviours, respectively, compared to 99.87% and 99.17% in the proposed study for the source and target domain to classify inactivity (lying, standing). Likewise, [58] and [62] reported 85.18% and 89.7% overall accuracy, which is significantly lower than the proposed model, which achieved accuracies in the range of 96.59%-98.55%. However, it is important to note that the number of activities was different in these studies compared to the current work.

Table 4-5 Overall Results on D_S and D_{S+} for CNN

Activity	Experiment A	Experiment B
	D_S^{ts}	D_{S+}^{ts}
Active F1 score	92.15%	94.60%
Grazing F1 score	95.34%	96.57%
Inactive F1 score	99.49%	99.91%
Accuracy	97.46%	98.55%

Table 4-6 Results from Transfer Learning on DT and DT+ for CNN

Activity	Experiment A	Experiment B
	D_T^{ts}	D_{T+}^{ts}
Active F1 score	91.05%	90.00%
Grazing F1 score	91.96%	93.97%
Inactive F1 score	99.63%	99.33%
Accuracy	94.79%	96.59%

As mentioned, active behaviour is comparatively complex, comprising overlapping behaviours, such as running, shaking, scratching and walking, which exhibit more complex movements. For instance, Barwick reported poor classification (54%) for the standing activity using collar data [56]. Likewise, [66] indicated limited performance with accuracies of 69.8%, 45.2%, and 25.1% for lying, standing and walking behaviours, respectively. Based on these findings and expert recommendations, the proposed model integrated the lying and standing activities into a single behaviour (i.e., inactive), which significantly improved accuracy (over 99%). A similar work proposed by Umstater et al. [263] also indicated that there are instances, where walking and grazing show overlapping patterns because sheep may graze while walking, which makes it more difficult for ML models to distinguish between these two behaviours.

Studies also indicated performance variations with respect to sensor attachments for data collection. For instance, [46] reported accuracy of 67%-88% for collar-based data compared to 86%–95% for ear-tag sensors. Barwick [56] obtained better results when the sensor was attached to the ear-tag of the animal with prediction accuracies of 94%, 96% and 99% for the grazing, standing and walking behaviours, respectively.

The statistical results indicate the robustness of CNN in terms of generalisation on unseen data, which support its use in real-life applications. Concerning real-life applications and real-time decision making, the CNN model is beneficial, specifically, because it can automatically extract features from the raw sensor data while producing robust results, as demonstrated in our experiments. When using CNN with the hand-crafted features in experiment B, higher results were achieved. It was shown that the transfer learning application is more robust when compared to experiment A. In other words, the use of CNN with hand-crafted features supports real-time operation in real-life scenarios of animal monitoring and warning generation.

In addition to the reliable and efficient performance, the use of transfer learning is advantageous because of the reusability and generalisation of pre-trained models from other applications within similar domains on unseen datasets. Furthermore, as CNN performs better with larger datasets, transfer learning can be leveraged to provide a cost-effective solution (regarding time and resources) while reusing it for limited size datasets. In this way, the new dataset can be used in transfer learning while using the knowledge acquired from the model trained in relevant larger datasets.

4.4 Summary

In this chapter, the problem of SAR was considered in the context of two sensor types and configurations. The first sensor (MetamotionR) was placed at a fixed orientation and was characterised by lower noise density, while the second sensor had varying orientation and higher noise density. The research investigated the use of CNN and TL in two problem settings. Firstly, CNN and transfer learning on the original datasets were applied. Next, a large number of temporal and spectral features were extracted, which resulted in two augmented datasets. The simulation studies and associated analysis indicated that CNN and transfer learning could generate high accuracy in classifying three sets of activities, including active, grazing and inactive behaviours. High-quality classification results were achieved in terms of accuracy, F1 score, precision and recall quality measures when benchmarked with other results in the literature. It was observed that the inclusion of hand-crafted features improved the performance of both the employed CNN and transfer learning solutions. Furthermore, the simulation results showed the advantage of using deep learning in terms of generalisation, indicating its reusability when datasets are limited in animal behaviour recognition.

In the next Chapter, trials regarding training sheep using audio are presented, analysing the animals' response based on various audio signals and evaluating the possibility of using audio cues only as an alternative solution for a VF system.

Chapter 5 Training Sheep to Change Direction Using Audio for a Virtual Fencing System

This chapter provides information regarding three trials conducted to identify whether audio stimulus can replace electric shocks in a virtual fence scenario (information regarding virtual fences and findings of auditory awareness in sheep is provided in Chapter 2). This chapter also gives a detailed insight into the process followed in the experiments. The experiments performed in this study aim to answer the following questions:

- Can the position of the Sheep be manipulated based solely on acoustic cues without the need for a stressor?
- Which frequencies are more effective in restricting sheep access to an attractant or a restricted area?
- Are there correlations between the time to respond to the sound and i) the attractant, ii) animal personality, and iii) frequency?
- Is the time of the animal to respond to the emitted sounds different based on the personality type of the Sheep?
- Are there any differences and/or commonalities between the animal responses to the frequencies based on the breed?

5.1 Introduction

Virtual fencing systems have been suggested as new means of keeping animals in specified areas. Over the past years, research has been conducted to test a virtual fence's ability to control and manage animals' spatial distribution, using sounds followed by electric shocks to deter animals' access to restricted areas. Electric shocks are banned in various countries due to the stress that they can cause to the animals. Therefore, this study aimed to test only sounds in the range of 125Hz-17kHz and white noise to determine whether it is possible to discourage a small flock of seven Hebridean ewes from entering a restricted area without using electric stimulus. The selection of those frequencies was based on previous suggestions that the sheep's hearing range is between 125Hz and 42kHz [248]. A common highest frequency of commercial speakers is around 20kHz, therefore, a maximum of 17kHz was selected in the trials for cost-efficient and convenient reasons to make sure that the frequency can be audible to the observer as well, as the highest frequency a human can hear is around 17,6kHz [248]. In the trials, two means of testing are used; 1) sounds were emitted when the animal approached an attractant (food bowl filled with pellets), and 2) sounds were emitted when the animal approached a specified restricted area.

Trials are conducted using two breeds: Hebridean and Greyface Dartmoor ewes. The results on Hebrideans showed that white noise, 125Hz-440Hz, 10kHz-14kHz, and 15kHz-17kHz could successfully discourage animals from reaching a specific area with an overall success of 89.88%. The successful responses included 78.53% that turned and walked away, 9.46% turned and run, and 12.01 stopped. Similarly, Greyface Dartmoor ewes were discouraged from entering a restricted area with an overall of 95.93% (correct responses: turned and walked away: 68.64%, turned and run: 19.49%, and stopped: 11.86%). Additionally, results from the three trials revealed that the use of attractant, the sheep temperament, and the type of frequency, have a statistically significant effect on the time needed for the animal to respond. This study considered the "turn and walk away" response to be the most desirable as it indicates that the animal is not experiencing further stress. The study demonstrated the potential of replacing electric shocks with sounds to manage the animals' spatial distribution on the pasture, but there is still a need for further investigation. The method can be considered when 'leaky' boundaries are acceptable, as it is not 100% stock-proof and cannot keep predators away. Therefore, this method should be tested further using larger flocks of animals and other breeds as well. Additionally, research is needed to investigate how to keep predators out before it is considered for commercialization. On the other hand, the method can safely be used to manage the spatial distribution of the animals on the land they graze as it can be beneficial to land utilization and prevent overgrazing and soil erosion.

5.2 Materials and Methods

5.2.1 Ethical statement, animals and location

The Senior Research Officer and LSSU Manager of Liverpool John Moores University approved the experimental protocol (approval AH_NKO/2018-13). Three trials have been conducted. The first two trials were conducted in Shotwick Village (Chester) in a paddock with an area of approximately 20m x 60m. Seven Hebridean ewes aged 5-14 years were used. The third trial was conducted in Wales in a paddock comparable in scale to the Shotwick paddock. In the third trial, 8 Greyface Dartmoor ewes are used, aged 2-8 years old. The animals in all three trials were habituated to the presence of people, therefore their behaviour was not influenced by the presence of the observers. Throughout the experiments, the animals were free to use the paddock's whole area and access grass and water.

5.2.2 Equipment

A commercial Bluetooth speaker (EWA A106 Pro Wireless Mini Bluetooth Speaker) was attached to the animal's collar using a small carrying case. The speaker weighs 175.77 grams, with dimensions of 4.8 x 4.8 x 3.84 cm. Additionally, to test the animals' response to various audio signals, a custom sound system in Cycling '74 MAX/MSP visual programming language was developed (refer to Figure 5-1). The system was paired with the collar to manually send audio cues to the animal for the testing. The system generates white noise and sounds between 100Hz-20kHz, within the Sheep's hearing range [248]. The sounds were intermittent or continuous, and the volume could be manipulated through the system manually. Additionally, once the sound was emitted, a log file is generated reporting the start time of the sound, the type of the sound (sine wave or white noise), the volume level, and stop time. To record the behaviour of the animals, a Canon SX720HS video camera was used.

5.2.2.1 Custom sound system

The created sound system was used in the experiments by connecting the Bluetooth speakers with the laptop. One of the observers was responsible for choosing and emitting the sounds during the experiment. We did not use all the functionalities of the system, however, the reason we added them was that we will consider using the full functionality of the system in future work. The system consists of the following controllers.

- SYNTHESIZER: The synthesizer part of the system allows the user to choose between either white noise or sine waves (with adjustable frequency). Additionally, there is an option to emit the desired sound as continued or intermittent.
- ORIENTATION: A slide bar with the capability to control the localization of the sound (for example, the orientation manages the volume of the left and right speaker)
- POSITION: The sheep icon can be positioned in five different volume levels between 0 to 5, having 0%, 25%, 50%, 75%, and 100% volume level, respectively
- PUNISH: Once this button is pressed, the sound changes to white noise with 100% volume level.
- STATUS: This section consists of visual information of the current state of the settings at a particular time. It also shows how many seconds the system has been active once the volume is above 0%. Consequently, this information is logged in text files.
- EVENT LOG: It allows the user to view, save, or clear the activity logs created when the program is active
- MASTER: It controls the master volume and shows the total output level in dB.
- AUTOMATOR: Allows the user to define the levels of volume, the frequency type, and time duration of the sounds, and then it can be played automatically with the press of the start button

Figure 5-1 Custom sound system developed using the Cycling '74 MAX/MSP media tool

5.2.3 Experimental Protocol

Overview of the experimental setting

Seven Hebridean ewes were used age range 4-12 years, and, 8 Greyface Dartmoor ewes aged 2-8 years old. The ewes were grazed in two contiguous paddocks of 20 x 60 meters making up and area of ¾ acre in total. The experiments took place in one of these two paddocks. The acoustic stimuli were applied to sheep on an individual basis while they remained in their flock, using food of high calorific value (pellets and sugar beet shreds) as a strong attractant.

The strength of attraction to bait in experiments involving small groups of sheep has been positively correlated with calorie value of the bait [294]. Rare breeds of sheep are reported to be significantly more averse to an experimental deterrent than heritage or commercial breeds (e.g. withdrawal on approach of a human with a device for administering a paint mark [294] or the presence of a dog outside the bait pen [295]. Hebridean sheep are classified as a heritage breed and are often used in conservation grazing management schemes in the UK. Hebridean sheep are therefore comparable to commercial breeds in terms of grazing management requirements that might be benefited by acoustic fencing. Isolation is extremely stressful for any breed of sheep [295], [296] and provides an unacceptable source of variation in response to experimental treatments [296] This is because the emotional state of a sheep affects its concentration, decision making powers and memory [297]. Visual cues from flock members are of paramount importance in influencing behaviour via position and body language [240].

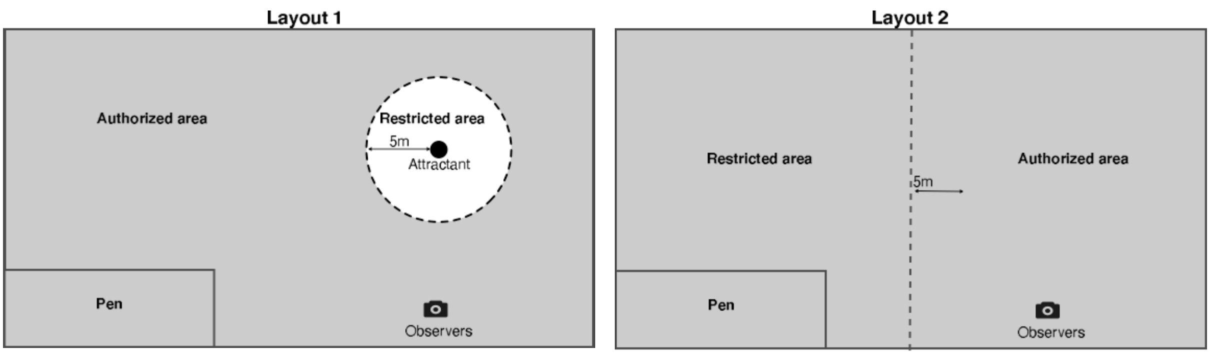
Taking the above information into account, to eliminate unethical levels of stress on sheep and a source of uncontrolled variation in the behavioural data, the experiments with an acoustic deterrent were performed in the normal small flock situation. The human observers were people that the flock were very familiar with. The space available for a given reaction to a sound stimulus was ample, so that any stress associated with a sensation of undue confinement would not affect behavioural responses. This flock-level testing protocol was also designed to emulate a commercial setting, where several sheep would be competing for an attractive food resource. This competitive situation created the strongest drive for a given sheep to ignore any acoustic stimuli. The impact of a sheep's reaction to acoustic stimuli on other members of the flock could also be assessed in terms of how efficient the system was at influencing more than one animal. One animal therefore wore the collar containing the acoustic stimulus at any given time. The experiments were video recorded .The observers were very familiar to the sheep. One observer operated the video camera and the other observer operated the software for controlling the sound stimulus.

Experimental procedure

The experimental procedure was common to trials 1, 2 & 3 and was as follows. The ewes were penned and one ewe was randomly selected and fitted with the collar. The collar was easy to fasten and unfasten to minimize the time an animal had to be restrained. The animals had approximately 20-30 minutes to settle down before each experimental session. At the end of each experimental session, the animal wearing the collar was returned to the pen to remove the collar, and the collar was fitted to the next animal. Once testing was completed on all animals, all of the animals were released in their normal area of pasture. The video recordings were time synchronized with the log file generated from the sound system to label the response on each sound. The sounds used were in the range of 125Hz-17kHz and white noise with random volume levels between 25%-100%. The number of repetitions was selected based on the expert's suggestion (Dr. Jennifer Sneddon), which was also present during the trial. For example, if the animal was distressed, the repetitions of the sounds were either reduced or stopped and another animal was selected for the trial. Therefore, there was no fixed number of repetitions for each sound. These acoustic stimuli were randomly applied to the sheep with the collar, in series, for indicative amounts of time up to 15s, until a given behaviour was observed. If the stimulus was ignored the sound was switched off after 15s. The sheep was rewarded for a favourable behavioural response by immediate observer-controlled cessation of the sound. The mutually exclusive behavioural responses of the animals exposed to the acoustic stimuli consisted of 1) turn and walk away, 2) turn and run, 3) stop, 4) no response (no response was reported if the animal continued to walk towards the restricted area, or if the animal responded by moving forward instead of backward from the area). The response that was considered most desirable was for the animal to turn and walk away calmly from the bait bowl. This behavioural response indicated that the animal did not experience undue but reacted calmly with time to decide how to respond to attain the reward of immediate sound cessation. Turning and running was a sign of a higher level of stress and was to be preferably avoided. Stopping with or without further exposure to sound to move the animal away from the area was considered an acceptable level of stress.

5.2.4 Trial 1: Exploring frequency bands (vs animal behaviour)

The purpose of this experiment was to answer the experimental aims 1) Can the position of the sheep be manipulated based solely on acoustic cues?, and 2) Which frequencies are more effective in restricting access of the sheep to a feed bowl with sugar beet shreds?. For the trial, the layout 1 in Figure 5-2 was used. The purpose of the experiment was to qualify which frequencies were successful at restricting the animal's access to a food bowl with sugar beet shreds, and quantify the strength of the reaction in seconds. When the animal approached the bowl, a sound was emitted to describe and time which reaction occurred. As mentioned above, the recorded reactions were: 1 turn and walk away, 2 turn and run, 3 stop, or 4 no response (ignored the acoustic stimulus). The first operator observed the animal, while the second operator sent acoustic signals through the customised sound system to the Bluetooth speaker attached to the collar on the sheep. In order to obtain a consistent approach, stones were placed 5 m from the bait bowl in order to know when to emit the sound and give the sheep sufficient time to make a decision and respond. The observer video recording the animal would inform the observer in control of the sound-emitting software when to emit a sound. As shown in Figure 5-2 (a), in layout 1, the animals could use the whole pasture, with the exception of the pen, and pasture surrounding the food bowl. The presence of the observers did not cause the animals to feel stressed or change their behaviour because they are habituated to the presence of people.



(a) (b)

Figure 5-2 Both layouts 1 and 2 were surrounded with physical fences. (a) Layout 1: Experimental layout used in both trials. The attractant was a bowl filled with pellets in order to attract the attention of the Sheep. The sound was emitted as soon as the animals were 5 meters away from the bowl. (b) Layout 2: Experimental layout used only in trial 2. The pasture was divided into two areas (authorized and restricted). The dash line represents the end of the authorized area. The sound was emitted as soon as the animals were 5 meters away from the dashed line.

During the trial, sounds were tested in the range of 125Hz-17kHz and white noise, to determine whether any sounds could restrict animal access to the food bowl. The response of seven Hebridean ewes was examined. The animal wearing the collar was encouraged to approach the food and then an intermittent sound was emitted for a maximum of 15s and then stopped, if the animal did not display any response. Once the animal started to approach the bowl, a sound was emitted to test if a reaction could be recorded in terms of change in animal direction. If the animal responded with one of the three actions: (1) turn and walk away, (2) turn and run, or (3) stop, then the audio stimulus stopped. If the animal approached the food bowl again, then another audio signal was emitted until the desired response was observed. Intermittent white noise and sine wave sounds were used to restrict the access of the animal to the food. Intermittent sound was used in the experiment because it has been previously observed that discontinuous sound, generates better responses in animals [249] [250]. We used four[250]. Four volume levels are utilised; 25%, 50%, 75%, and 100% with starting dB level being at 0.0 (volume level at 0%) to -3.0 full scale (volume level at 100%). The volume was randomly selected and it was determined that volume did not have a concrete effect on the response of the animals, hence the focus was on the effect of the frequency. Once all the desired audio sounds were tested on each animal, the trial was completed. The sound frequencies, repetitions used on the seven sheep, and response during the experiment are presented in Table 5-1. Table 5-1 shows that frequency band between 1kHz and 9kHz was not able to keep the animals away most of the times, having success rate less than 50%, therefore those bands were excluded from the second trial.

Table 5-1 Frequency bands of sound signals, number of repetitions and response rate (trial 1)

Sound	**Repetitions**	**Response %**	
		Yes	**No**
White noise	52	**100.00%**	**0.00%**
125Hz-440Hz	35	**74.29%**	**25.71%**
1kHz-5kHz	19	47.37%	52.63%
6kHz-9kHz	17	35.29%	64.71%
10kHz-14kHz	19	**78.95%**	**21.05%**

15kHz-17kHz	38	**84.21%**	**15.79%**

The frequency bands that are selected for the second trial are presented in bold

5.2.5 Trial 2: Exploring sound duration vs animal response

The second trial was conducted during the summer of 2019 (June-August) at the same location as trial 1. For the second trial, only the frequency bands that were successful in the previous trial were used (refer to Table 5-1). In the first setting (layout 1, Figure 5-2 (a)), a bowl filled with pellets or sugar beets was used and the sounds were emitted in order to test if the animal can be stopped from either eating or even approaching the bowl. In this setting, the same procedure as in trial 1 was followed. In a different setting (layout 2, Figure 5-2 (b)), there was no food bowl. In these settings, we were emitted when the animal was standing or walking towards the restricted area to test whether it is possible to strict restrict access to that area. Once the animal was 5 meters away from the zone of the restricted area, an audio signal was emitted for a maximum of 15 seconds. If the animal turned and walked away, turned and run, or stopped, then the audio signal was ceased. Only one animal at a time was tested. Table 5-2 presents the selected audio signals used for the trial 2, and the number of repetitions for each band.

Table 5-2 Selected audio signal frequencies and number of repetitions used in trial 2

Sound	Repetitions
White noise	207
125Hz-440Hz	112
10kHz-14kHz	73
15kHz-17kHz	205
Total	597

5.2.6 Trial 3: Frequency vs animal behaviour on a new breed (Greyface Dartmoor)

In order to test the selected sounds on a different breed and answer the question whether there are any differences and/or commonalities between the animal response to the frequencies based on the sheep breed, a farm in Wales on March 2020 was visited. The sound experiment was conducted on eight Greyface Dartmoor ewes. Same protocol was followed as in the previous trials which is described in section 5.2.3. In this trial layout 3 was utilized (refer to Figure 5-3). The paddock was comparable in scale to the Shotwick paddock. The animals were kept in the pen in order to fit the collar and then they were released in the paddock. In this experiment sounds were emitted when the animal was 5 meters away the restricted area and approaching. For this experiment no food bowl was used as an attractant. Similarly, with previous experiments, all animals were tested and by the end of the experiments the time was synchronized with video recordings and the log file of the sound system to report the behaviour of the animals on each repetition.

Table 5-3 shows the number of repetitions used in the trial based on the frequency bands.

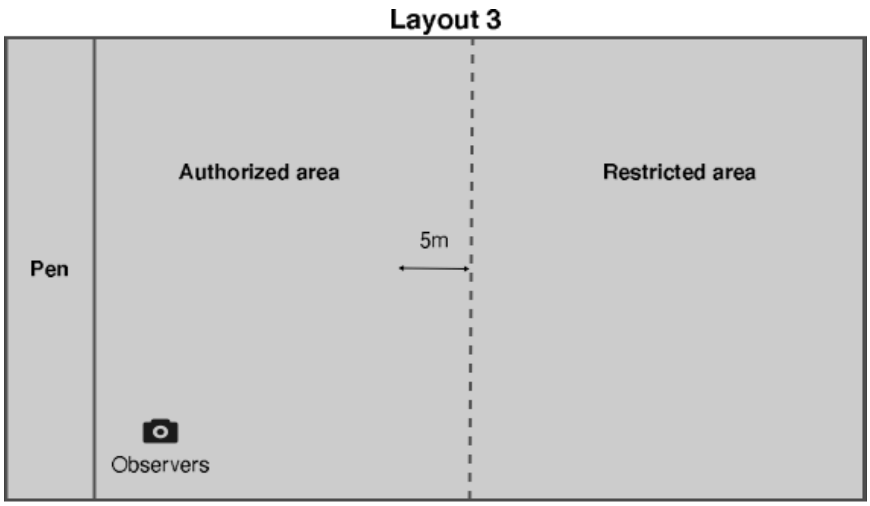


Figure 5-3 Experimental layout for trial 3. The pasture was divided into two areas (authorized and restricted). The outside pasture was surrounded with physical fences. The dash line represents the end of the authorized area. The sound was emitted as soon as the animals were 5 meters away from the dashed line

Table 5-3 Selected audio signal frequencies and number of repetitions used in trial 3

Sound	Repetitions
White noise	**33**
125-440	**29**
10000-14000	**39**
15000-17000	**22**
Total	**123**

5.2.7 Data analysis

To identify the proportion of desired responses to the sound cues against no responses, frequency tables have been generated that present the sound band used, repetitions, and whether there was a response, or not for all three trials. The response was divided into three categories; 1) Turn+walk away, 2) Turn+run, and 3) stop. These tables provided a clear understanding of the success rate of each sound. The tables were developed to answer the first two questions of our study i.e., whether the position of the sheep can be manipulated based solely on acoustic cues and to identify which frequencies are more effective in restricting access of the Sheep to an attractant or to a restricted area.

Furthermore, we used statistical software SPSS Statistics V26 for the analysis. The data analysis used "duration" as the dependent variable, which was tested against "sheep_type", "frequency", "attractant", and "response". The analysis was conducted to answer the question whether there is any effect on duration based on the personality type of the Sheep, and applied audio stimulus. For the analysis, generalized linear model (GLM) was utilized [298] under Tweedie distribution [299] with Log Link function. The GLM uses a training method for various sets of regression models. Even if the relationship between the independent variables X, and the dependent variable Y may not be additive or linear, the GLM allows additive (additive assumption means the effect of changes in a predictor on a response is independent of the effects of changes in other predictors), and linear way. A GLM consists of a linear predictor (η_ι), a link function (g(.)), and a variance function var(Y_i) and are presented in Equations (5.1), (5.2), and (5.3), respectively:

$$\eta_\iota = \beta 0 + \beta 1 x 1_\iota + \ldots + \beta \varrho x_\varrho \tag{5.1}$$

where β are the coefficients to be estimated, for i=1..,ϱ where ϱ is the number of independent variables

$$g(\mu_i) = \eta_\iota \tag{5.2}$$

the link function g(.) describes the dependen0ce of the mean μi on the linear predictor

$$\mathrm{Var}(Y_i) = \Phi v(\mu_i) \tag{5.3}$$

the variance function describes the variance on the mean , where Φ is the dispersion parameter. Further information for the generalized linear model can be found in Baxter et al., 1990 and Dunteman & Ho., 2011.

For trials 1 and 2, the main effects and interactions of frequency, attractant, sheep personality type and response on the duration were tested using a Wald chi-square test [301], whereas $p < 0.05$ was considered to be statistically significant. The results provided information regarding the third question considered in these investigations, i.e., whether there is a significant effect between the time to respond to the sound and i) the attractant, ii) personality type of the animal, and iii) frequency. The investigation aimed to understand the time needed for the animals to respond when they are exposed to the sounds as it will play an important role in the design of the virtual fence system. Additionally, the main effect of frequency vs duration is tested on a new breed (trial 3) to compare with the findings from trial 1 and 2 and identify if there are any differences or similarities between the behaviour of the two breeds.

5.3 Results

5.3.1 Trial 1: Exploring frequency bands (vs animal behaviour)

The aim of the 1st trial was to identify if only sound cues without electric shocks can generate a response (turn and walk away, turn and run, stop) of the animals on the pasture. For the 1st trial, sounds are emitted every time the animal attempted to approach a food bowl filled with pellets. If the animal responded, the sound was ceased. The sound was not emitted for more than 15 seconds in order to avoid causing distress to the animal. Therefore, if the animal continued to move forward towards the food bowl, the outcome was recorded as no response. To test this, white noise was utilized, and also randomly selected sounds in the range of 125Hz-17kHz divided in 5 bands: 125Hz-440Hz, 1kHz-5kHz, 6kHz-9kHz, 10kHz-14kHz, and 15kHz-17kHz. Figure 5-4 illustrates the overall response of the animals to the sounds. From Figure 5-4, it can be noticed that the frequency range of 1kHz-9kHz did not manage to alert the animals with a reasonable success rate (47.37% for 1kHz-5kHz, and 35.29% for 6kHz-9kHz) and therefore the bands were excluded from trial 2.

Table 5-4 Selected frequency bands and animal response type in trial 1

Sound	**No Response**	**Total Response**	**Turn+walk away**	**Turn+Run**	**Stop**
White noise	0.00%	100.00%	42.31%	48.08%	9.62%
125Hz-440Hz	25.71%	74.29%	69.23%	19.23%	11.54%
10kHz-14kHz	21.05%	78.95%	66.67%	26.67%	6.67%
15kHz-17kHz	15.79%	84.21%	87.50%	0.00%	12.50%
Average	**13.19%**	**86.81%**	**62.40%**	**27.20%**	**10.40%**

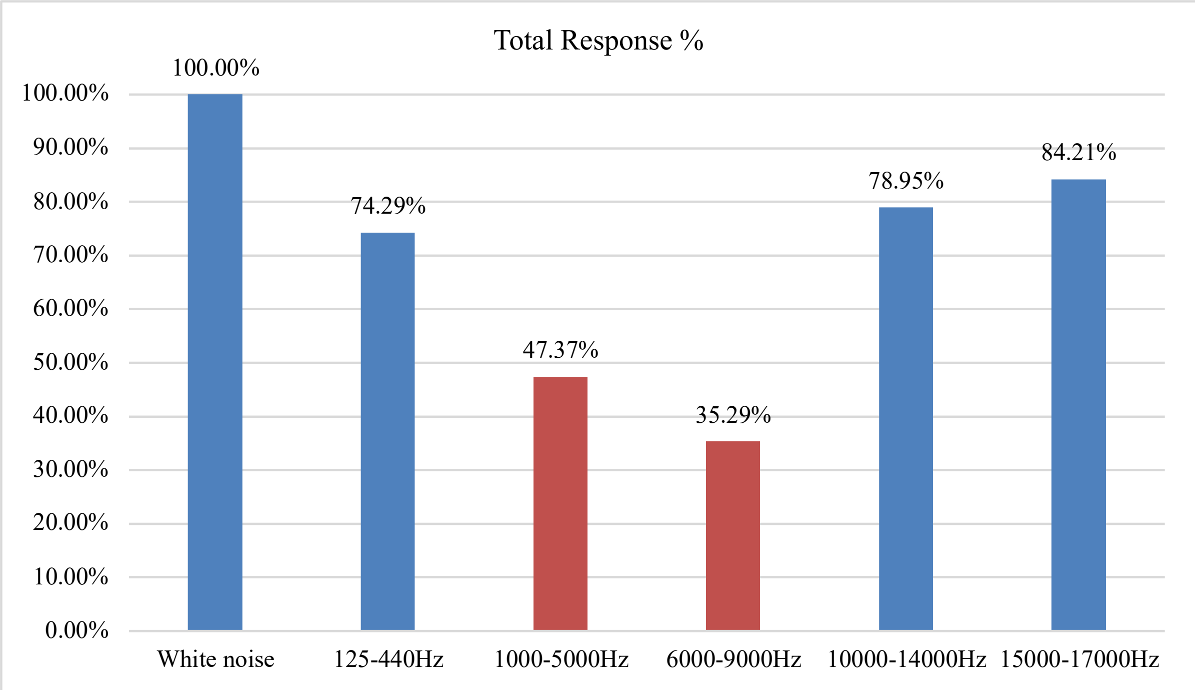


Figure 5-4 Average score of the animals' responses to each sound in trial 1. The blue bars present the sounds which are considered successful in the trial 1. The orange bars present the sounds which were excluded in the next trial (trial 2)

In

Table 5-4**,** the results from the trial in more detail are shown. As it can be noticed white noise was 100% successful in restricting animal access to the food bowl, followed by the 15kHz-17kHz frequency range with 84.21% response. The frequency bands of 125-440Hz and 10kHz-14kHz restricted the animal with 74.29%, and 78.95% success, respectively. The highest score where the animals turned and walk away from the food bowl was achieved using frequencies between 15kHz-17kHz in contrast with the white noise where the animals turned and walk away only 42.31% of the time. Emitting 125-440Hz and 10kHz-14kHz, the animals turned and walk away from the bowl 69.23% and 66.67%, respectively.

The animals turned and run away from the food bowl 48.08% of the time when white noise was emitted, in contrast with the frequency band 15000-17kHz where the animals did not run.

5.3.2 Trial 2: Testing the selected sounds vs category of animal response

The results from the second trial are presented in Table 5-5. The response rate of 90.62% was achieved. The highest response for a specific type of audio stimulus was 92.86%, achieved using low frequency sounds between 125-440Hz. The lowest response rate, i.e., 89.27%, was recorded when using signals in the 15kHz-17kHz range. The average rate of response for animals turning and walking away was 82.26%, with the highest rate of this response (90.16%) recorded at the 15kHz-17kHz range. In regard to the behaviour where animals turned and ran, this occurred 12.83% of the time, when white noise was emitted. The animals did not run when the sound frequency was between 10kHz-17kHz. Based on all selected frequency sounds, the animals stopped while walking towards the bowl between 9.84%-14.93%. An overall response using all selected sounds was that only 9.38% of the time the animals did not respond and proceeded towards the bowl.

Table 5-5 Selected frequency bands and animal response in trial 2

Sound	No Response	Total Response	Turn+walk away	Turn+Run	Stop
White noise	**9.66%**	**90.34%**	**74.33%**	**12.83%**	**12.83%**
125Hz-440Hz	**7.14%**	**92.86%**	**80.77%**	**4.81%**	**14.42%**
10kHz-14kHz	**8.22%**	**91.78%**	**85.07%**	**0.00%**	**14.93%**
15kHz-17kHz	**10.73%**	**89.27%**	**90.16%**	**0.00%**	**9.84%**
Average	**9.38%**	**90.62%**	**82.26%**	**5.36%**	**12.38%**

5.3.3 Trial 3: Testing the selected sounds vs animal response on new breed

The results from the third trial are presented in Table 5-6. During this trial, an average response rate of 95.93% was achieved. Using White noise and frequencies between 15kHz-17kHz were successful at 100%, the animals responded as desired in all cases. The lowest response rate of 89.74%, was achieved when the frequencies were between 10kHz and 14kHz. The average rate of response for animals turning and walking away was 68.64%. The animals turned and walked away at 57.58%, 78.57%, 68.57%, and 72.73% rate, with white noise, 125-440Hz, 10000-14kHz, and 15000-17kHz, respectively. The animals turned and run at the rate of 42.42%, and 27.27%, when white noise, and 15000-17kHz were emitted, respectively. When 125-440hz was emitted, the animals run at the rate of 3.57%, and at the rate of 5.71% when 10000-14kHz was played. Regarding the behaviour when the animals stopped to the sound, 17.86% rate was recorded at 125-440Hz, and 25.71% at 10000-14kHz. There was not any recorded “stopped” with white noise and 15000-17kHz.

Table 5-6 Selected frequency bands and animal response in trial 2

Sound	No Response	Total Response	Turn+walk away	Turn+Run	Stop
White noise	0.00%	100.00%	57.58%	42.42%	0.00%
125Hz-440Hz	3.45%	96.55%	78.57%	3.57%	17.86%
10kHz-14kHz	10.26%	89.74%	68.57%	5.71%	25.71%
15kHz-17kHz	0.00%	100.00%	72.73%	27.27%	0.00%
Average	**4.07%9**	**95.93%**	**68.64%**	**19.49%**	**11.86%**

5.3.4 Merging Trial 1 and 2: testing Hebridean ewes

Results from Trial 1 and 2 are presented to illustrate the effects of personality type, attractant, and frequency, on duration. Both trials used Hebridean ewes.

5.3.4.1 Duration statistics and effect of sheep personality type, attractant, frequency, and response on duration

From the data of trials 1 and 2, the statistics of the duration, the main effects and interactions of the sheep personality type, attractant, frequency, and response, on the duration needed from the animal to show alerting behaviour are investigated. The minimum, maximum, mean, and standard deviation of the dependent variable (duration) is presented in Table 5-7. Table 5-8 provides information of the results from the overall model. The omnibus test is a likelihood-ratio chi-square test of the model versus the null model. The significance value of less than 0.05 indicates that the current model outperforms the null model. From the likelihood ratio test of all the independent variables, a *p*-value of 0.000 was demonstrated, indicating a statistically significant overall model as shown in Table 5-8.

The model is then tested to identify which of the independent variables have a significant effect on the depended variable. Main effects are tested using one variable at a time versus the depended variable. Variables with significance value less than 0.05 show that they have some apparent effect. Additionally, the interactions of the variables are considered and tested to identify if they have any significant effect on the duration (depended variable). Interactions test whether two or more variables influences the relationship between the independent variable and the depended variable. The results of main effects and interactions are shown in the following subsections. In the sections below, the results of the main effects of attractant, sheep personality type, and frequency on the duration are presented. Additionally, the interaction effects on the duration of the variables are tested. Interactions of Sheep type vs Frequency on duration are tested to investigate further whether the frequency bands play a role in the duration the animals respond based on the sheep personality type (i.e., bold vs shy).

Table 5-7 Depended variable statistics

		Minimum	Maximum	Mean	Std. Deviation
Dependent Variable	Duration (in seconds)	1	14	4.92	3.316

Table 5-8 Results from the overall model (Omnibus test)

Likelihood Ratio Chi-Square	df	Sig.

132.635	22	.000

5.3.4.2 Exploring main effects between the independent variables and dependent variable

Attractant vs duration

The mean duration in seconds of the animal to respond once it is exposed to the sounds is presented in Table 5-9. It can be noted that when there is no food bowl involved, the mean time of the animal to respond is 1.41 seconds faster. From the pairwise comparisons of the estimated marginal means on the duration it is indicated that the mean difference is statistically significant having p-value of 0.016. The boxplot in Figure 5-5 shows that the maximum time to respond when there is no attractant is approximately 1 second faster than when there is no food bowl involved. Median time in both situations is 4 seconds, and minimum time is around 1 second. Table 5-9 presents the estimates of duration by the attractant variable, and Table 5-10 shows the results from pairwise comparison of food bowl vs no food bowl on the duration.

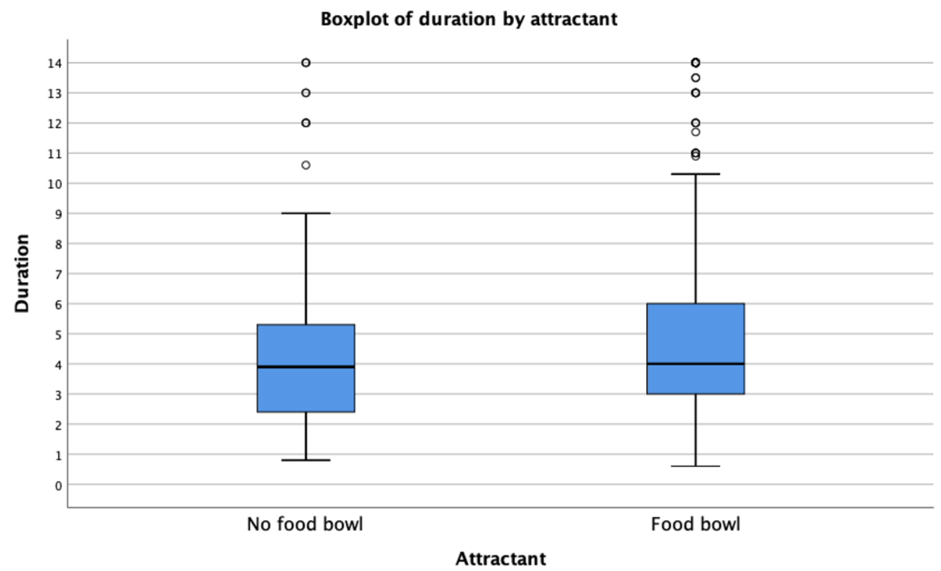


Figure 5-5 Proportion of time in seconds needed from the animals to respond to sound cues while having or not having an attractant; i.e., Food bowl filled with pellets

Table 5-9 Estimates of duration (in seconds) by attractant

Attractant	Mean	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper
No_food_bowl	3.19	.505	2.34	4.35
Food_bowl	4.60	.342	3.97	5.32

Table 5-10 Pairwise comparison of food bowl vs no food bowl on the duration

(I) Attractant	(J) Attractant	Mean Difference (I-J)	Std. Error	Sig.	95% Wald Confidence Interval for Difference Lower	95% Wald Confidence Interval for Difference Upper

No_food_bowl	Food_bowl	-1.41a	.581	.016	-2.54	-.27
Food_bowl	No_food_bowl	1.41a	.581	.016	.27	2.54

Sheep personality type vs duration

From Table 5-11 it can be seen that bold animals respond to the sounds with a mean duration of 5.17 seconds, in contrast with shy animals who responded with a mean of 3.05 seconds. The pairwise comparisons showed that the sheep personality type has a statistically significant effect on the duration to respond with p-value of 0.0001 having a mean difference of 2.12 seconds faster response when the animal is considered shy (refer to

Table 5-12). Figure 5-6 shows that shy animals need less time to respond to the sounds than bold animals. The maximum time of shy sheep to respond was 7 seconds, and the maximum time of bold animals to respond was 12 seconds. On the other hand, the minimum time both personality types needed to respond was approximately 1 second with median of 4 seconds and 3 seconds for bold and shy sheep, respectively.

Table 5-11 Estimates of duration (in seconds) by sheep personality type

			95% Wald Confidence Interval	
Sheep_type	Mean	Std. Error	Lower	Upper
bold	5.17	.365	4.50	5.94
shy	3.05	.374	2.40	3.88

Table 5-12 Pairwise comparison of shy vs bold sheep on the duration

(I) Sheep_type	(J) Sheep_type	Mean Difference (I-J)	Std. Error	Sig.	95% Wald Confidence Interval for Difference	
					Lower	Upper
bold	shy	2.12a	.417	.000	1.30	2.93
shy	bold	-2.12a	.417	.000	-2.93	-1.30

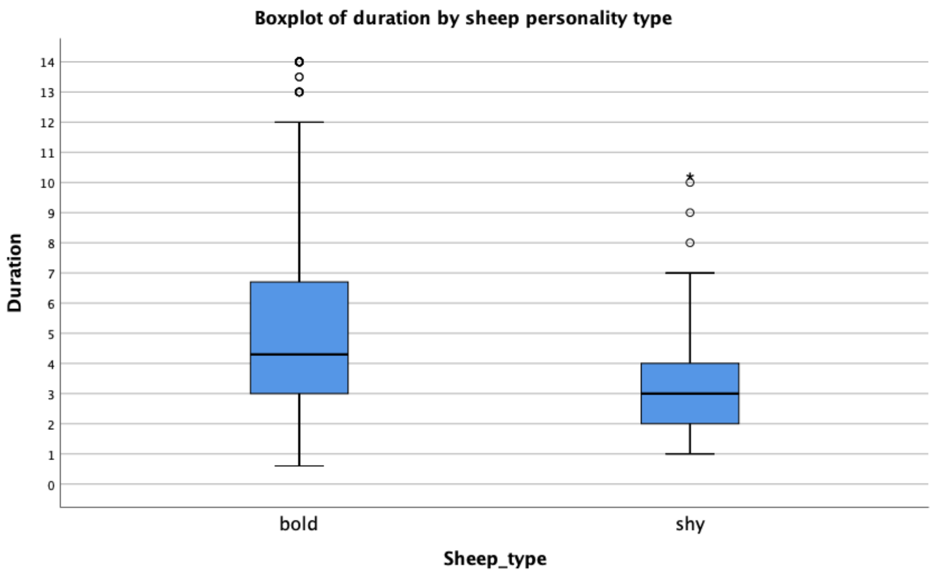


Figure 5-6 Proportion of time in seconds needed from the animals to respond to sound cues based on the personality type of the animal. We have divided the animals' in bold and shy.

Frequency vs duration

The results indicated that there is no obvious difference in the means of duration of behavioral response between the four frequency bands. Table 5-13 shows that only the low frequency band (125-440Hz) has a faster response with mean of 3.12 seconds, in comparison with the rest three (white noise, 10kHz-14kHz, and 15kHz-17kHz). An investigation of each frequency band and pairwise comparison using Wald chi-squared test showed a p-value of 0.023 which is considered statistically significant (refer to Table 5-14). Looking further into the frequencies it can be identified that there is a statistically significant effect between white noise and 125-440Hz with mean difference in time to respond of 1.35 seconds. Frequency bands of 10kHz-14kHz, and 15kHz-17kHz was found not to be significantly different from each other in terms of mean duration of behavioral response data and therefore not included in Table 5-14. Completed table which includes all pairwise comparisons can be found in Appendix A.1.

Table 5-15 presents an overall estimate of the main effect between frequency and duration and show that there is a significant effect of p=0.033.

Table 5-13 Estimates of duration (in seconds) by frequency

			95% Wald Confidence Interval	
Frequency	Mean	Std. Error	Lower	Upper
White noise	4.48	.623	3.41	5.88
125Hz-440Hz	3.12	.346	2.52	3.88
10kHz-14kHz	4.25	.497	3.38	5.35
15kHz-17kHz	4.19	.630	3.12	5.62

Table 5-14 Pairwise comparison of frequencies on the duration

(I) Frequency	(J) Frequency	Mean Difference (I-J)	Std. Error	Sig.	95% Wald Confidence Interval for Difference	
					Lower	Upper
white_noise	125-440Hz	1.35[a]	.595	.023	.19	2.52
125-440Hz	white_noise	-1.35[a]	.595	.023	-2.52	-.19

a. The mean difference is significant at the .05 level.

Table 5-15 Overall Test Results: frequency vs duration

Wald Chi-Square	df	Sig.
8.761	3	.033

5.3.4.3 Exploring interaction effects of sheep type and frequency on duration

Table 5-16 presents the model-estimated marginal mean, standard error, and confidence interval of the duration at sheep type and frequency category. From Table 5-16 it can be observe that the mean duration ranges from a low of 2.39 seconds for shy Sheep exposed to 125-440Hz sounds, to a high of 6.07 seconds for bold Sheep exposed to a high frequency band of 10000-14kHz. Figure 5-7 illustrates the time in seconds needed from the animals to respond to sound cues based on the personality type of the animal and frequency. It can noted that there is a difference in duration between bold and shy sheep. however it cannot clearly indicate that the differences are significant.

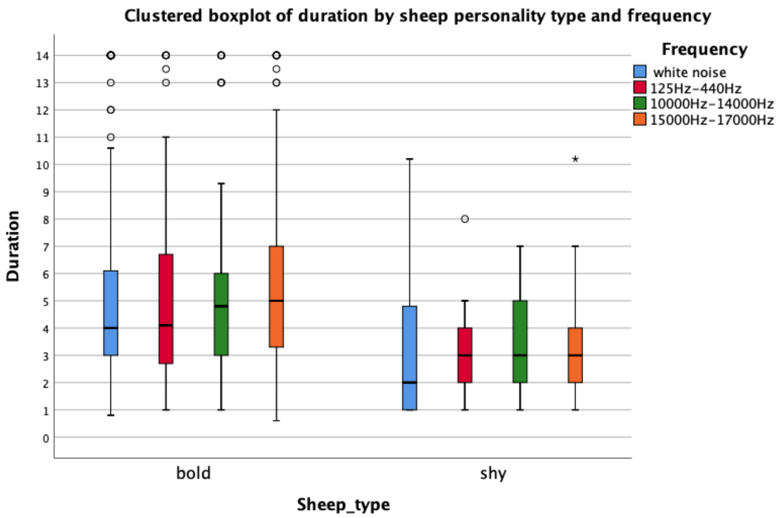


Figure 5-7 Proportion of time in seconds needed from the animals to respond to sound cues based on the personality type of the animal and frequency.

Table 5-16 Estimates of duration (in seconds) by attractant*frequency

Sheep_type	Frequency	Mean	Std. Error	95% Wald Confidence Interval	
				Lower	Upper
bold	white_noise	4.77	.497	3.89	5.85
	125Hz-440Hz	4.09	.389	3.39	4.92
	10kHz-14kHz	6.07	.801	4.69	7.86
	15kHz-17kHz	6.03	.893	4.51	8.06
shy	white_noise	4.20	.904	2.76	6.41
	125Hz-440Hz	2.39	.413	1.70	3.35
	10kHz-14kHz	2.98	.573	2.04	4.35

	15kHz-17kHz	2.91	.532	2.03	4.16

Table 5-17 Pairwise comparison of sheep type*frequencies on the duration

(I) Sheep_type*Frequency	(J) Sheep_type*Frequency	Mean Difference (I-J)	Std. Error	Sig.	95% Wald Confidence Interval for Difference	
					Lower	Upper
bold*125Hz-440Hz	bold*10kHz -14kHz	-1.99[a]	.895	.026	-3.74	-.23
	bold*15kHz-17kHz	-1.94[a]	.885	.028	-3.68	-.21
bold*10kHz-14kHz	bold*125Hz-440Hz	1.99[a]	.895	.026	.23	3.74
bold*15kHz-17kHz	bold*125Hz-440Hz	1.94[a]	.885	.028	.21	3.68
	shy*15kHz-17kHz	3.13[a]	.683	.000	1.79	4.46
shy*white_noise	shy*125Hz-440Hz	1.81[a]	.850	.033	.15	3.48
shy*125Hz-440Hz	shy*white_noise	-1.81[a]	.850	.033	-3.48	-.15

Table 5-18 Overall Test Results: sheep type*frequency on duration

Wald Chi-Square	df	Sig.
39.165	7	.000

The Wald chi-square tests the effect of Sheep_type*Frequency. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Table 5-17, shows whether the values shown in Table 5-16 represent significant changes. Results in Table 5-17 indicate that the interaction effect of bold*frequency on duration is statistically significant between frequency band 125Hz-440Hz vs 10kHz-14kHz and 15kHz-17kHz with p values of p=0.026 and p=0.028 respectively. Bold animals react 1.99 seconds faster at 10kHz-14kHz compared to 125Hz-440Hz, and 1.94 seconds faster when 15kHz-17kHz are emitted compared to 125Hz-440Hz. On the other hand, the pairwise comparison between frequency bands and shy Sheep showed significant difference with p value of p=0.033 between 125Hz-440Hz band vs white noise where in this situation, shy sheep react 1.81 seconds faster with 125Hz-440Hz. The Wald chi-square test presented in Table 5-18 shows an overall significance value of 0.0001 on the sheep type * frequency effect.

5.3.5 Trial 3: testing Greyface Dartmoor ewes Duration statistics and main effect of frequency bands on duration

In this section we present results from trial 3 concerned with the effects of frequency sound on duration when the sounds were used on Greyface Dartmoor ewes. In the trial the animals are exposed to sounds and restrict them from crossing a restricted area. As information regarding the personality type of the ewes were not provided, only the main effect of the frequency bands on the duration was tested. Results are presented in the following subsection.

Duration statistics and main effect of frequency bands on duration

The Greyface Dartmoor ewes responded to the sounds with a minimum of 1.40 seconds, maximum of 14.00 seconds, and mean of 4.41 seconds, and standard deviation of sd=2.46 seconds (refer to Table 5-19) . In Table 5-20, a detailed duration mean by frequency band is presented. The animals reacted to white noise with a mean of 2.85 seconds which is less than the mean time of the other three bands. The mean duration to respond to 125Hz-440Hz was 4.28 seconds. The other two bands, 10kHz-14kHz, and 15kHz-17kHz, had a similar mean duration of 3.76 seconds, and 3.39 seconds respectively.

The box plot in Figure 5-8 illustrates the proportion of time in seconds needed from the animals to respond to sound based on each the frequency. From the figure, there is a clear difference between 125Hz-440Hz versus the other three bands. The analysis of the frequencies vs duration showed that there is a statistically significant effect of the frequency on duration with a p-value of 0.000, and is presented in Table 5-21.

Table 5-19 Depended variable statistics

		Minimum	Maximum	Mean	Std. Deviation
Dependent Variable	Duration (in seconds)	1.40	14.00	4.4144	2.45570

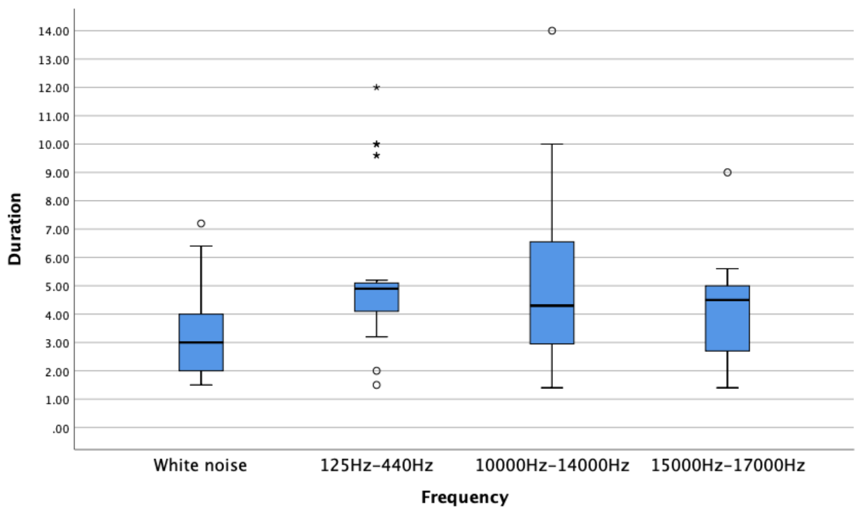


Figure 5-8 Proportion of time in seconds needed from the animals to respond to sound cues based on each frequency band.

Table 5-20 Estimates of duration (in seconds) by frequency

Frequency	Mean	Std. Error	95% Wald Confidence Interval	
			Lower	Upper

White noise	2.8587	.17794	2.5304	3.2296
125Hz-440Hz	4.2769	.26918	3.7806	4.8384
10kHz-14kHz	3.7577	.36060	3.1135	4.5353
15kHz-17kHz	3.3908	.20094	3.0190	3.8084

Table 5-21 Results from the overall model (Omnibus test)

Wald Chi-Square	df	Sig.
18.933	3	.000

The Wald chi-square tests the effect of Frequency. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

The pairwise comparisons of each frequency band on the duration are presented in Table 5-22. In this table, the comparisons which have no significant are excluded, however all values for the comparisons can be found in Appendix A.2. From Table 5-22 it can be noted that white noise cause faster response in comparison with 125Hz-440Hz, 10kHz-14kHz, and 15kHz-17kHz, and the effect is statistically significant with p-values of p=0.000, p=0.025, and p=0.047, respectively. Additionally, the table reveals that 125Hz-440Hz have a statistically significant mean difference with the 15kHz-17kHz, on the duration, finding that the reaction caused on the animals from the latter is faster by approximately 1 second.

Table 5-22 Pairwise comparison of frequencies on the duration

a. The mean difference is significant at the .05 level

(I) Frequency	(J) Frequency	Mean Difference (I-J)	Std. Error	Sig.	95% Wald Confidence Interval for Difference

					Lower	Upper
White noise	125Hz-440Hz	-1.4182[a]	.32267	.000	-2.0507	-.7858
	10kHz-14kHz	-.8990[a]	.40211	.025	-1.6872	-.1109
	15kHz-17kHz	-.5321[a]	.26841	.047	-1.0582	-.0060
125-440Hz	15kHz-17kHz	.8861[a]	.33591	.008	.2278	1.5445

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Duration

.

5.4 Discussion

The technology of virtual fencing systems using only sounds is not novel as it has been established in cattle [187], [197], [302]. Studies for a virtual fence system for sheep that use only sounds are limited and most of them used auditory warning and electric stimuli to train the animals [200], [207], [208]. However, it has been reported that training animals using electric stimulus could have a negative impact on animal welfare [303]. Thus, the aim of this study was to assess whether sound cues could replace electric stimulus in a virtual fence scenario using Hebridean ewes located in Shotwick. Furthermore, we aimed to test the sounds on a different sheep breed (Greyface Dartmoor). We have conducted three trials. In the first trial, we randomly tested various frequency bands to identify if and which sounds cause alerting responses to the animals and restrict them access to a food bowl filled with pellets. We exposed the animals to white noise, 125Hz-440Hz, 1kHz-5kHz, 6kHz-9kHz, 10kHz-14kHz, and 15kHz-17kHz. From the results obtained, we identified that the frequencies between 1kHz-5kHz, 6kHz-9kHz were not successful in causing satisfactory responses from the animals and therefore were excluded. From the results of the first trial, we concluded that very low (125Hz-440Hz), and high frequency sounds greater than 10kHz are promising and the animals exposed to them react as desired. This might be due to the sensitivity of the animals to sounds with frequencies above 10kHz [248]. Additionally, low frequencies might be more disturbing as they cause an effect on the nervous system of humans and animals and might cause the animals a flight reaction in order to avoid it [304]. In the second and third trial, we used only the selected sounds from trial 1 to look further in the animals' responses and investigate the effects of the variables on the duration. The technology of virtual fencing systems using only sounds is not novel as it is applied mostly on cattle [187], [197]. Studies for a virtual fence system for sheep that use only sounds are limited and most of them used auditory warning and electric stimuli to train the animals [200], [207], [208]. However, it has been reported that training animals using electric stimulus could have a negative impact on animal welfare [303]. Thus, the aim of this study was to assess whether sound cues could replace electric stimulus in a virtual fence scenario using Hebridean ewes. Furthermore, we tested the sounds on a different sheep breed (Greyface Dartmoor). Three trials have been conducted. In the first trial, various frequency bands are randomly tested to identify if and which sounds cause alerting responses to the animals and restrict their access to a food bowl filled with pellets. The animals are exposed to white noise, 125Hz-440Hz, 1kHz-5kHz, 6kHz-9kHz, 10kHz-14kHz, and 15kHz-17kHz. From the results obtained, we identified that the frequencies between 1kHz-5kHz, 6kHz-9kHz were not successful in causing satisfactory responses from the animals and therefore were excluded. From the results of the first trial, it was concluded that very low (125Hz-440Hz), and high frequency sounds greater than 10kHz are promising and the animals exposed to them react as desired. Then, in the second and third trial, only the selected sounds from trial 1 are selected to look further in the animals' responses and investigate the effects of the variables on the duration.

In the following subsections 5.4.1, and 5.4.2 the results obtained from trial 1 and trial 2 are discussed. The focus is on the frequency bands used to investigate whether the behaviour of the animal can be controlled by restricting access to an attractant or to a specific area (Section 5.4.1). Then, in Section 5.4.2, the findings from the tests performed are used to investigate main effects and interactions between the independent variables: frequency, attractant, and sheep personality type, on the dependent variable: duration. In Section 5.4.3 the results from the 3rd trial are discussed conducted on a different breed. Lastly in Section

5.4.4, the results of the two different breeds are compared: Hebridean ewes vs Greyface Dartmoor ewes.

5.4.1 Frequency bands vs response vs sheep personality type

Overall, the animals reacted with satisfactory levels above 88.48% on all four selected sounds. A total of 89.88% of the response times indicated that it is possible to monitor the animals' location on the land they graze. The most desired response of the animals was to turn and walk away calmly, indicating that the sound may not cause any stress on the animal [200], [207], [208]. From Table 5-23 we can make the following observations . White noise appeared to be more alarming to animals than the rest of the frequency bands, since 20.50% of times, the recorded response was to turn and run away. On the other hand, when sounds in the high frequency band of 15-17 kHz were emitted, the animals either turned and walk away with a rate of 89.77%, or stopped with a rate of 10.23%. Using this band, the animals did not run away and this indicated that this band was successful in manipulating animal behaviour while not causing unnecessary stress. Moreover, when sounds in the frequency band of 10-14 kHz were emitted, the animals turned and walked away 81.71% of the times, turned and run 4.88% of the times, and stopped with a rate of 13.23%. This band achieved the second highest response with 89.13% of the desired response. The literature reports that sheep can hear best at 10 kHz [248], and this might be the reason for the high rate of occurrence of desired reactions.

Differences in the reaction of the frequencies between bold and shy sheep were investigated. Table 5-24 presents the percentage of the desired response on each frequency band based on the personality type of the animal. Shy animals reacted as desired with a total response of 98.40%. Using white noise, and sounds in the 125Hz -440Hz, and 10kHz-14kHz bands, the animals responded each time (100% response). When using the high frequency band of 15kHz-17kHz, we had a 1.60% of no response. Yet again, the response of shy animals on all four bands was very encouraging and this needs to be further investigated. For bold animals, it was noticed that a significant decrease of approximately 10% in terms of responses in comparison to shy animals. Bold animals reacted with a rate of 88.11%, having the highest percentage of response with white noise (91.56%), and the lowest at 83.05%

using very low frequencies between 125Hz-440Hz. Using the sounds between 10kHz-14kHz led to the second most successful rate of responses, i.e., 87.34%.

Table 5-23 Response results from both trials by frequency and response type

Sound	No Response	Total Response	Turn+walk away	Turn+Run	Stop
White noise	7.72%	92.28%	67.36%	20.50%	12.13%
125Hz-440Hz	11.56%	88.44%	78.46%	7.69%	13.85%
10kHz-14kHz	10.87%	89.13%	81.71%	4.88%	13.41%
15kHz-17kHz	11.52%	88.48%	89.77%	0.00%	10.23%
Total	10.12%	89.88%	78.53%	9.46%	12.01%

Table 5-24 Response results from both trials by frequency and personality type

Response vs No response			Response per frequency			
Personality	No Response	Response	White noise	125Hz-440Hz	10kHz-14kHz	15kHz-17kHz
Bold	11.89%	88.11%	91.56%	83.05%	87.34%	85.79%
Shy	1.60%	98.40%	100.00%	100.00%	100.00%	98.40%

From the results of Table 5-23 and Table 5-24, it is suggested that all four sounds (i.e., white noise and selected frequency bands) could be used in a concept of a virtual fencing system on shy animals. However, some sounds may be subject to habituation. The results suggest that in a virtual fence system, three bands of 125Hz-440Hz, 10kHz-14kHz, and 15kHz-17kHz could be randomly played to achieve one of the desired behavioural reactions. It was observed that white noise caused more stress/irritation to the animals based on their reaction (i.e., turned and run away from the area), and thus it may not be subject to habituation. Our suggestion is in agreement with Umstatter et al., that the development of a smart virtual fence, should trigger different sounds in a random pattern to avoid habituation, and white noise could be used as a last resort [187]. Based on these results, it could be concluded that the temperament of an animal plays an important role in terms of its behavioural response to the use of a virtual fence. Shy animals reacted as desired with a rate of 98.40% and additionally often did so when bold animals were wearing the collar, and reacted as desired. Therefore, it could be suggested that audio cues were successful at restricting sheep in their access to a restricted area and that they have potential as a replacement to an electric stimulus in a VF system. The ability of the animals to learn a virtual fence system based on their temperament needs to be further investigated. Other studies have found no association between temperament and learning [208]. Sheep have excellent learning and memory abilities and can follow sophisticated rules including reversal learning [241], [243], [246], [253] and response inhibition [247]. Further, experimental time required for training in sheep is markedly shorter than has been reported in primates, where training and testing typically takes many months [246]. In this study we have demonstrated experimental evidence that that response inhibition capability in sheep individually or in small flocks could be linked to control of position via an acoustic stimulus to instruct animals to stop moving in one direction and take another away from a virtual boundary.

5.4.2 Main effects and interactions of sheep personality type, attractant, and frequency, on duration

In this subsection the main effects and interactions of the independent variables on the duration will be discussed. With this analysis the aim is to understand further the response time of the animal when exposed to the sounds and investigate whether there is any statistically significant difference on the response time based on the sheep personality type, attractant, and frequency.

Attractant vs duration and sheep personality type vs duration

From the results obtained while analyzing the main effect of the attractant on duration needed for the animal to respond, it was confirmed that the mean difference of the estimations is statistically significant ($p = 0.016$). The animals reacted faster to the exposure of the sounds, when there is no food bowl involved. This suggested that the sound irritated the animals and thus, they reacted faster when there was no motivation/reward [305]. In a real-world scenario, this could indicate that if the virtual fence is used in a pasture, where taller and better grass is available in the restricted area, the animals may be willing to attempt to cross over more times. On the other hand, they could be easily manipulated with the emission of sounds, if the restricted area is not as attractive to them. Table 5-25 shows the percentages of the total proportional behavioural response data pooled for all animals across both trials. It is clear that the emitted sounds were successful at eliciting a response with a rate of 97.73% when there was no food incentive, making the animals turn and walk away from the restricted area at 89.15%. It should be noted that the 2.27% failure rate of the sounds is related to only bold animals. Shy animals responded 100% of times to auditory stimuli. When access to a food bowl with sugar beet shreds was restricted, the success rate for pooled data for all animals across both trials was 88.18%. From Table 20 it can be seen that 11.82% of the tests failed to restrict access to the animals. In this situation, there were only two occasions where the emission of sounds was unsuccessful, involving shy animals (0.32%).

It was observed that duration (s) needed to respond to an acoustic stimulus based on sheep personality type, was shorter in shy animals, having mean duration of 3.05 seconds, which is approximately 2 seconds faster than the corresponding mean duration for bold animals. This was proved to be statistically significant, ($p = 0.0001$). Therefore, we concluded that the temperament of the animal is a factor that influences the duration and nature of the behavioural response to an acoustic stimulus. This was also stated by Umstatter et al. [197], who said that the temperament of the animals might be a factor that needs to be further explored.

Table 5-25 Total response to sounds: food bowl vs no food bowl

Attractant	No response	Response	Turn+walk away	Turn+run	Stop
No	2.27%	97.73%	89.15%	0.00%	10.85%
Yes	11.82%	88.18%	75.98%	11.73%	12.29%

Frequency vs duration effect and frequency by sheep personality type vs duration

The overall estimate of the main effect between frequency and duration on behavioural response showed a statistically significant value of p=0.033. Looking into the effects of the frequency bands of the emitted sounds, we it was observed that the low frequency (125-440 Hz) caused a faster reaction on the animals compared to the other three frequency bands. This low-frequency band though, needs to be further investigated as based on the strength and rapidity of behavioural reactions it might be comparatively too alarming for sheep for use in a VF system.

Bold vs Shy animals

As previously discussed, bold animals reacted significantly more slowly to sounds than shy animals. It was identified that bold animals reacted with a mean duration of 4.77, and 4.09 seconds, when white noise and sounds in the range of 125-440 Hz are emitted, respectively. The mean time was faster than the duration needed when sounds of high frequencies are emitted having a mean of 6.07 seconds in the range of 10-14 kHz, and 6.03 seconds in the range of 15-17 kHz. Bold animals reacted 1.99 seconds (p= 0.26) faster to sounds in the range of 10-14 kHz, compared to sounds in the range of 125-440 Hz, and 1.94 seconds (p=0.028) faster, when sounds in the range of 15-17 kHz are emitted, compared to the range of 125-440 Hz. Yet again, the band of 125Hz-440 Hz showed a dominant response and needs further investigation. The same was observed with shy animals for the same frequency range. In this situation, low-frequency sounds caused a faster and statistically significant response of approximately 2 seconds (p=0.033) in comparison to white noise. The reaction to the low frequencies is interesting and it is worth investigated further with larger flocks and different breeds . This frequency band might be more alarming, or the observed behavioral responses may have been due to an unknown quality about the hearing sensitivity of sheep that could be associated with age, [33].

5.4.3 Trial 3: Frequency bands vs animal response on new breed

The experiment proved successful when the sounds were tested on Greyface Dartmoor ewes with an overall rate of 95.93%. It was noted that white noise and frequencies between 15kHz-17kHz were successful at 100%, leading to a conclusion that those bands can keep animals away from a restricted area. However, those two bands cause further stress to the animals as it was reported that the animals ran away at the rate of 42.42%, and 27.27%, with white noise, and 15kHz-17kHz, respectively. From the results it can be assumed that white noise is very irritating for the animals, as it was also found when white noise was tested on Hebrideans. In contrast, when 15kHz-17kHz was tested in Hebrideans, the animals never ran away. Factors that may be influencing the response might be the hearing ability of the animals, as the Greyface Dartmoor are younger animals and might conceive the high frequencies as more irritating, or the surroundings of each flock. Greyface Dartmoor might be not be used to high frequency sounds.

5.4.4 Hebrideans vs Greyface Dartmoor: frequency vs duration

From both breeds, similar reactions were observed when the animals are exposed to white noise. The sound was 100% successful on Greyface Dartmoor, and 92.28% on Hebrideans. Additionally, both breeds ran away with higher rates in comparison with the other three frequency bands. Another observation is that when 125Hz-440Hz, and 10kHz-14kHz were emitted, the animals mostly turned and walked away, or stopped. The rate of stopped behaviour was higher than the ran in both situations

Regarding the response time of the animals, in Table 5-26 it can be observe that Hebridean ewes respond with an overall mean of 4.01 seconds, in comparison with Greyface Dartmoor responds with an average of 3.57 seconds. However, the comparison is not clear as the trials followed different settings. Hebridean ewes were tested under two situations; i) having a food bowl filled with pellets in a restricted area, and ii) not having a bowl. Previously, it was noted that when the animals are exposed to the sound when no attractant is involved, the reaction time is with an average of 3.19, which agrees with the findings from the 3rd trial, as the sounds was tested on the new breed with no food bowl involved.

Table 5-26 Mean response duration: Hebrideans vs Greyface Dartmoor

	Mean Duration (seconds)	
Frequency	Hebridean ewes	Greyface Dartmoor ewes
white_noise	4.48	2.86
125Hz-440Hz	3.12	4.28
10kHz-14kHz	4.25	3.76
15kHz-17kHz	4.19	3.39
Average	**4.01**	**3.5725**

Looking at the main effect of the frequencies on the duration on both breeds the following observations are obtained. In the case of Hebrideans, a statistically significant effect between 125Hz-440Hz and white noise can be found, having the latter to cause faster response. On the contrary, Greyface Dartmoor animals showed an opposite effect, having white noise to cause faster reactions than all the other bands. As mentioned in a previous section, the reaction might be influenced by either the hearing ability of the animals, or by the habituation of the animals with their surroundings. Nevertheless, the findings from all trials indicate that sounds are possible alternative of an electric shock, and they might offer ethically acceptable solutions in a virtual fencing system. Virtual fence systems concerned with sheep need to be further investigated having large flocks involved to test the hypothesis further. The suggestion is that white noise is the most irritating sound, and it might be considered in the training phase of the flocks.

5.5 Summary

This chapter showed that there is potential for replacing electric shocks with audio stimuli within the context of a virtual fence scenario, and there is considerable evidence for this to be further explored in a commercial application. In the study, we identified four frequency bands, which caused a desired influence on behavioural responses relating to the spatial control of sheep, specifically, restricting access to a food bowl, or a specified area. Two different breeds were tested: 1) Hebridean ewes, and 2) Greyface Dartmoor. Results from Hebrideans and Greyface Dartmoor yielded an overall success of 89.88%, and 95.93%, respectively. White noise caused more overt stress to the animals than the frequency bands of 125Hz-440Hz, 10kHz-14kHz, and 15kHz-17kHz as the sheep turned and ran away from the area with higher rates compared to the rest of the sounds. This study also considered the main effects of sheep personality type, presence of attractant, and sound frequency on the duration in seconds that the sheep took to respond. This could be an indicator of the time needed for a VF system to play a sound in a commercial farming scenario. It was found that the personality of the animals, i.e., bold vs shy, had an effect on duration, with shy animals reacting faster to avoid an acoustic stimulus. Additionally, the presence of a food bowl motivated sheep to approach the restricted area and caused them to react significantly more slowly to the sounds. This information is vital to understand the potential for using acoustic stimuli in virtual fence systems, should restricted areas have better quality and taller grass. In this scenario, using a virtual fence system may be more challenging as the drive for feeding may overcome the fear or annoyance of an acoustic signal for directing spatial position. Lastly, it was observed that very low frequencies cause significantly faster reaction, s and this frequency band needs to be further investigated. As a final remark, this research used only a small number of animals and the findings need to be further evaluated in the context of larger flocks and different animal breeds.

In the next Chapter, a multifunctional system is proposed using CNN transfer learning based on the findings of Chapter 4, and the findings from Chapter 5 regarding the sound use to train animals to learn a VF system.

Chapter 6 Intelligest system proposal

In this Chapter, an algorithm for monitoring sheep activity using accelerometer data with robustness to accelerometer specifications, position, and orientation is proposed. In Chapter 4, Transfer Learning (TL) based on Convolutional Neural Networks (CNN) was described for animal behaviour classification to solve the problem of heterogeneity of accelerometer sensors. The findings are used in this study to identify the active and inactive behaviour of the animals using the selected CNN TL model from Chapter 4. The proposed algorithm detects whether the animal is active or inactive, the position of the animal on the pasture, and the virtual boundaries. The suggested algorithm will control the animals' position by calculating animals' location and checking whether the animal is within the predefined virtual boundary.

6.1 CNN Transfer Learning for SAR

In this section, the methodology for sheep activity recognition is shown. In this experiment, the findings and the datasets from the previous investigation (refer to Chapter 4) is used and applied CNN TL for the classification of the "active" and "inactive" state of the animals.

6.1.1 Methodology

The accelerometer devices metamotionR, and SenseHat, are used in this research. The sensors were attached to the collar of the animals to collect accelerometer measurements at a sample rate of 12.5Hz. Similar to Chapter 4, let D_S represents the data captured from metamotionR, whereas D_T represents the data acquired through RaspberryPi that will be used to validate the reusability of TL, i.e., the target data. Both datasets were labelled manually and normalised using the z-score.

The CNN model from Chapter 4 was used to train D_S (source domain). Then, the saved model was used and transferred the knowledge on the target domain D_T. Figure 4-1 illustrates the overall methodology.

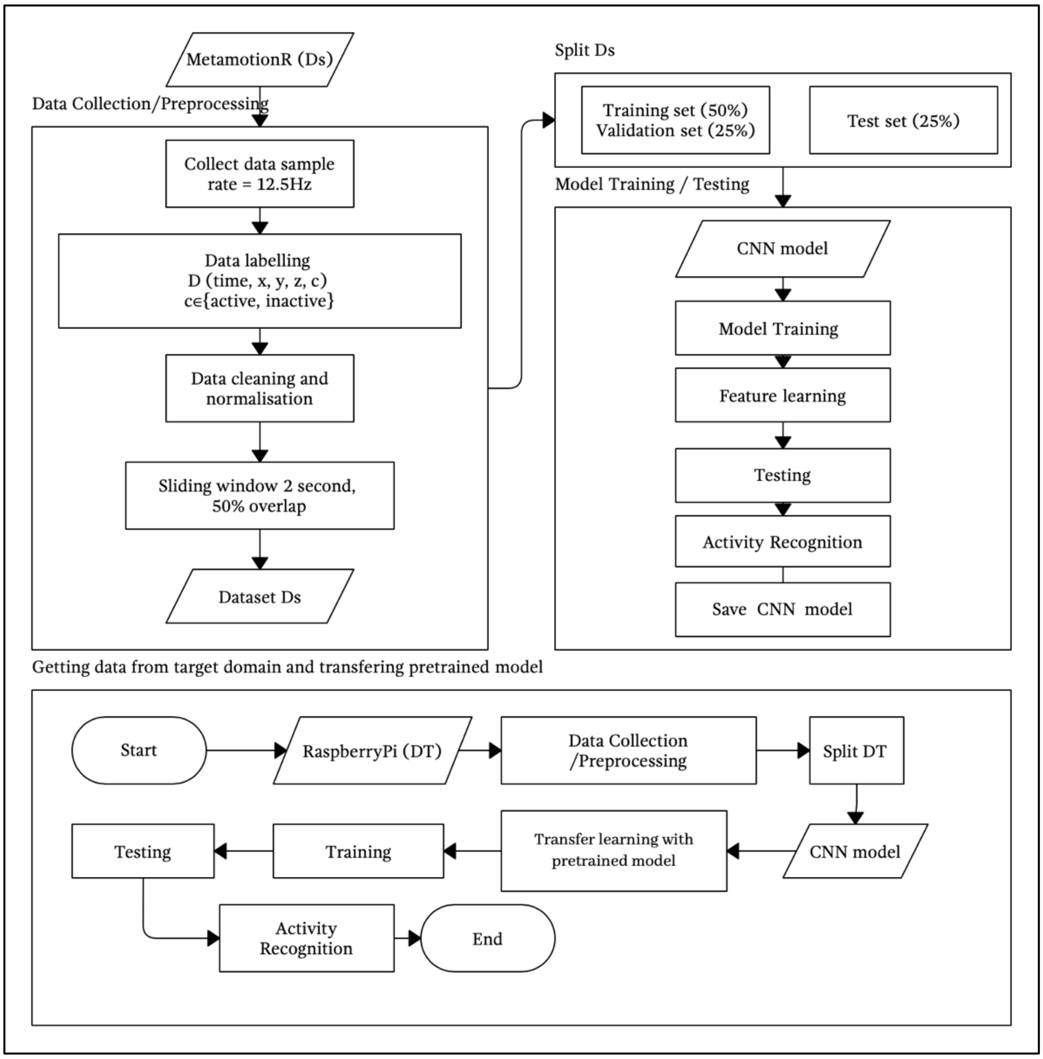


Figure 6-1 System methodology

6.1.1.1 Datasets

In this work, two primary datasets comprising the accelerometer measurements from a flock of 9 Hebridean ewes 35±5 kg, 9±5 years old at a farm located in Cheshire Shotwick (OS location 333781,371970) are used. Two types of devices were used, attached to the collar of the sheep. The first device was fitted on a fixed position (right side of the collar) and orientation (at 270° degrees). The second device was placed in a non-fixed way (at 0°, 90°, or 180° degrees). This was to test the performance of the model transferring learning from one domain to another by which the measurements differ because of sensor orientation and position. Figure 6-2 illustrates an example of the sensor orientation MetamotionR and RaspberryPi collected accelerometer measurements at a sampling rate of 12.5Hz. The data from both sensors were stored as timestamped CSV files. Recordings of 40 hours of activity were obtained; 23 hours comprised D_S , and 17 hours comprised D_T. Specifications of the two devices are presented in Table 6-1.

Table 6-1 Specifications of the sensor devices

Specifications	**MetamotionR device**	**RaspberryPi+SenseHat**
Size	26mm x 17mm x 2.5mm	85mm x 56mm x 17mm
CPU	ARM® Cortex-M4F, 32 bit	4× ARM Cortex-A53, 1.2GHz
RAM	64 kB	1Gb
Bluetooth	Bluetooth LE 4.2	Bluetooth LE 4.2
Wifi	No	Dual-band 802.11ac wireless LAN (2.4GHz and 5GHz)
Battery	100mAH micro-USB rechargeable Lipo	External battery
Accelerometer	Bosch®MI160 3-Axis Accelerometer	ICM20948- 3-Axis accelerometer
Set Sample Rate	12.5Hz	12.5Hz
Noise density	**180 µg/√Hz**	**230 µg/√Hz**
Resolution	16 bit	16 bit
Position on collar	**right side of the collar**	**bottom of the collar**
Orientation	**Fixed: 270° degrees**	**Non-fixed: 0°, 90°, or 180°**

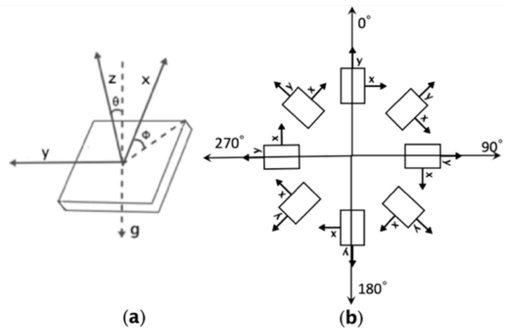
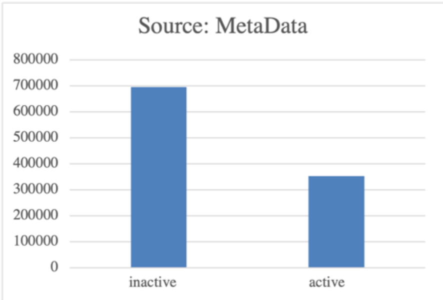


Figure 6-2 (a) Accelerometer coordinates x, y, and z, where θ is the angle of z relative to gravity, ϕ is the angle of x relative to the ground, and g is the gravitational acceleration constant g=9.81m/s2. (b) Illustration of sensor orientation.

Both datasets were labelled using the annotating tool ELAN_5.7_AVFX Freeware tool [82]. The behaviours were labelled as active and inactive. Behaviours such as running, fighting, shaking, walking, scratching, and grazing were labelled as 'active'. Likewise, standing and resting were labelled as 'inactive'. The source domain (D_S) contains 1,048,575 samples, whereas the target domain (D_T) comprises 762,860 samples. Let D_k represents the joint dataset, where $D_k=\{t_i, x_i, y_i, z_i, c_i\}$, i=1,..,n, where n is the number of observations, and k={S,T} being the source and target dataset. Variable t relates to the timestamp, while x, y, z are the accelerometer measurements in the x, y, and z axes, respectively, and c is the label variable, where c ∈ {active, inactive}. Figure 4-2 presents the distribution of the two activities within the datasets. The charts indicate the imbalanced distribution of activities within both datasets, as expected, due to the nature of the study and the activities considered. The graphs in Figure 6-4 and Figure 6-5 illustrate the acceleration coordinates of active and inactive state captured from D_S, and D_T, respectively.



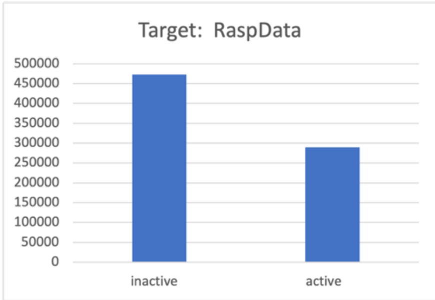
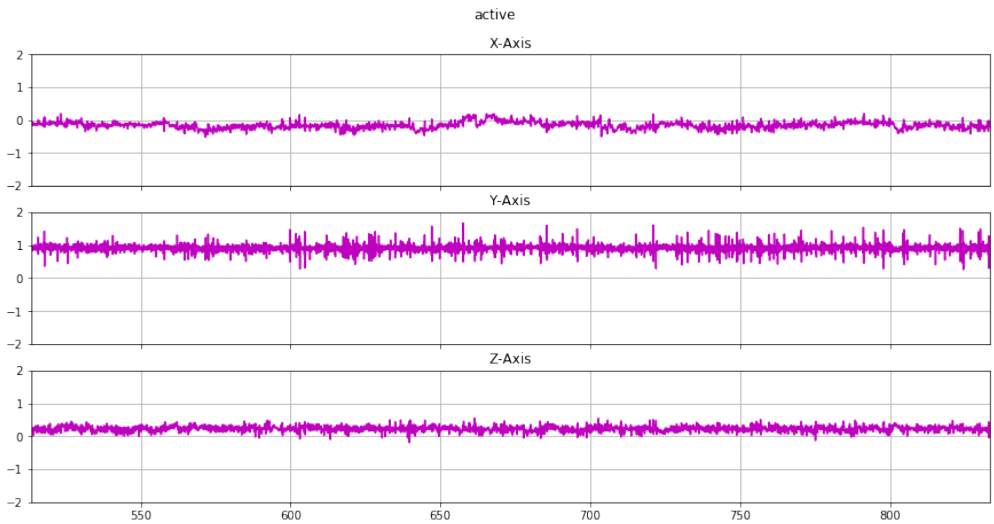
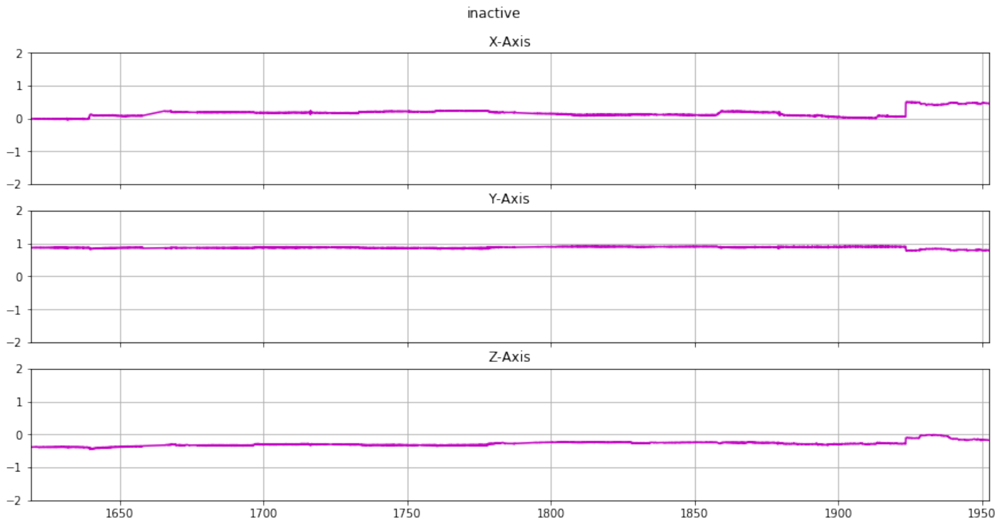


Figure 6-3 Duration of the activities for source and target domains

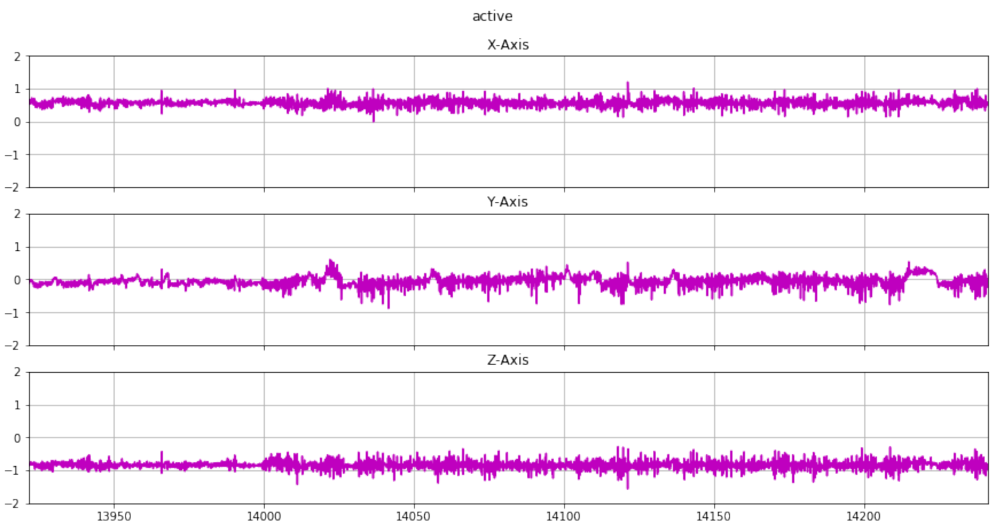


(a)

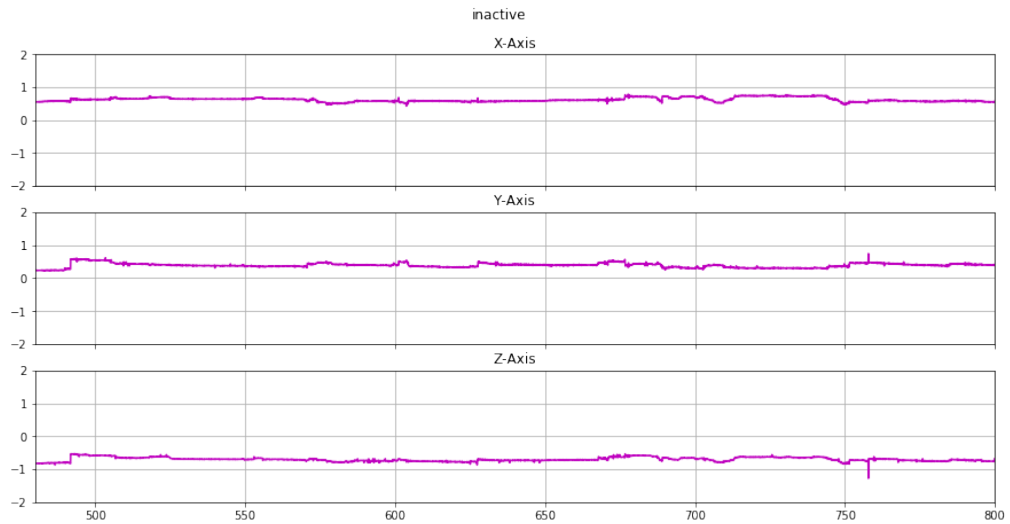


(b)

Figure 6-4 Acceleration coordinates of active (a) and inactive (b) state captured from D_S



(a)



(b)

Figure 6-5 Acceleration coordinates of active (a) and inactive (b) state captured from D_T

6.1.1.2 Data pre-processing

The datasets were normalised with zero mean (μ) and a standard deviation (σ) of 1. Then, the datasets were partitioned into training, validation, and testing, with a ratio of 50%, 25%, and 25%, respectively. Overlapping was used to enable real-time classification with a ratio of 50%, which has been shown to be effective in previous activity recognition studies [97]. A larger window size could lead to misclassification since animals may display more than one behaviour in a short time interval. Thus, a 2s window was considered sufficient and served the purpose of the virtual fence algorithm, which will check the animal's status every two seconds. When the animals presented more than one behaviour in the 2-second interval, (i.e. walking to standing, resting to walking), the labelling followed the most frequently occurring activity.

6.1.1.3 CNN and Transfer Learning

The recorded measurements were used in the proposed Deep Learning model to classify the target activities and also during TL. In this study, homogeneous TL is used because both source and target domains' feature space and domain characteristics are the same. The difference between the source and target datasets in this study is the accelerometer sensors used and the orientation and position of the sensor. Additionally, the motion measurements of the second device exhibit some noise and the size of the second dataset is smaller.

6.1.1.4 Experimental Design

The CNN model consists of a convolutional layer that uses 16x16 filters with a 10% dropout. The second convolutional layer uses 32x32 filters, with a 10% dropout. After the two convolutional layers, a fully connected layer is added with 64 filters, l2-norm of 0.00001, and 50% dropout, followed by an output layer that uses SoftMax. The CNN architecture is illustrated in Figure 6-6, and the TL proposed method is illustrated in Figure 6-7.

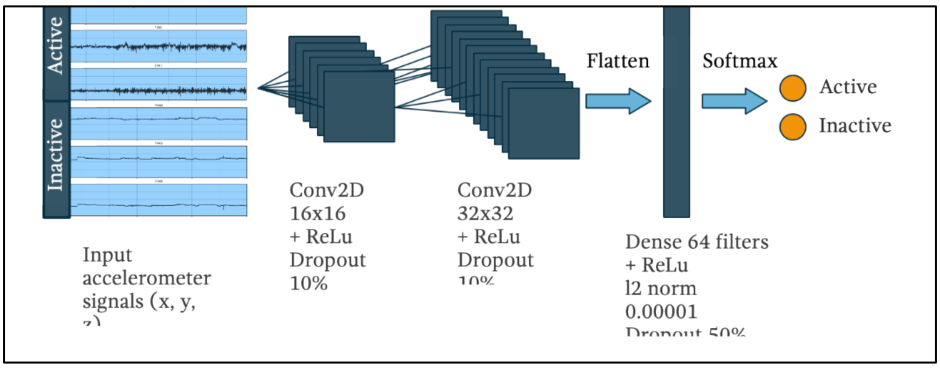


Figure 6-6 CNN model architecture

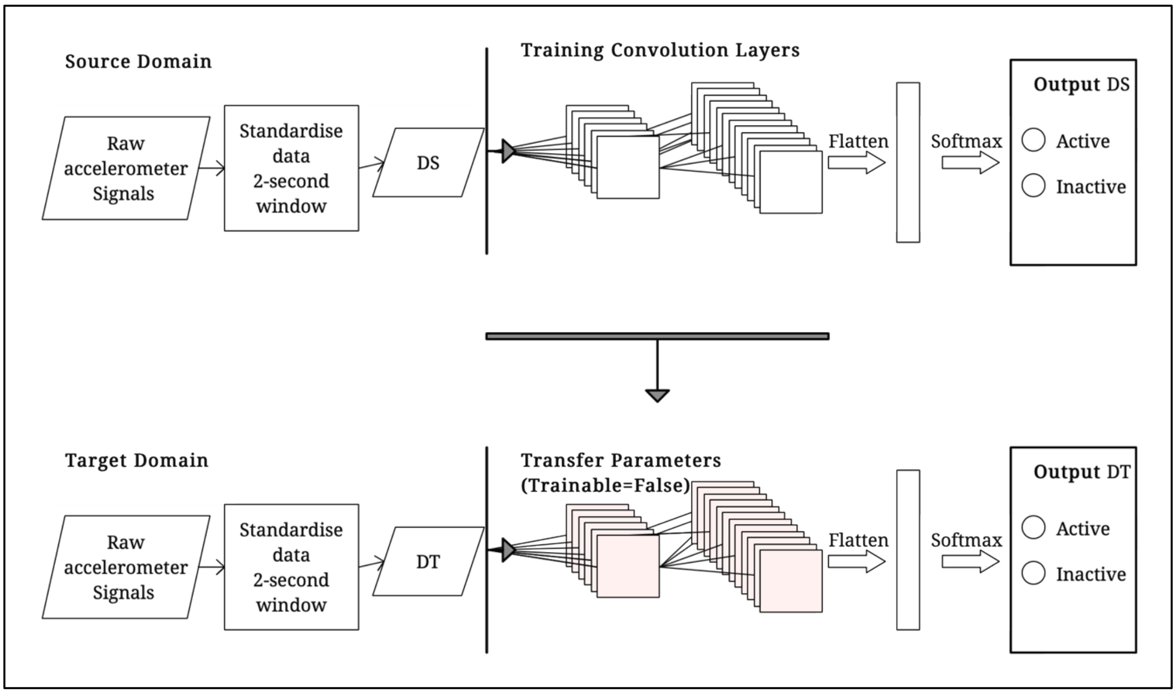


Figure 6-7 CNN proposed TL method

6.1.2 Results and Discussion

A binary classification using CNN on D_S and D_T is conducted while dividing the datasets into 50%, 25%, and 25% for the training, validation, and testing purposes, respectively. The performance of the model was evaluated using precision, recall, F_1-score, and accuracy.

The CNN model was trained on the training set D_S^{tr}, validated on D_S^{val}, and then tested on D_S^{ts}. The model was then saved and reused on the target domain, i.e., D_T. The only trainable layers during TL on D_T^{tr} were the fully connected layers. The accuracy and loss per epoch during model training are illustrated in Figure 6-8 and Figure 6-9. The results from the evaluation on the tests sets D_S^{ts}, and D_T^{ts} are presented in Table 6-2. The confusion matrix from both tests is illustrated in Figure 6-10.

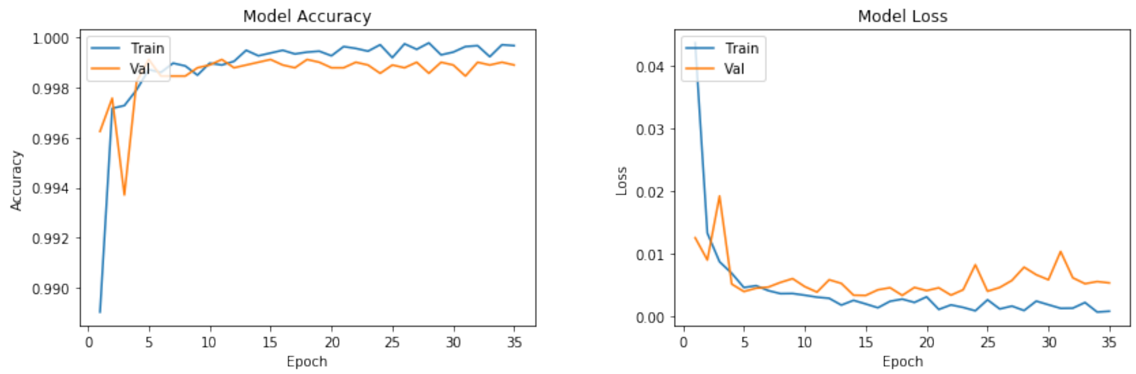


Figure 6-8 CNN model training on source domain D_S^{tr}

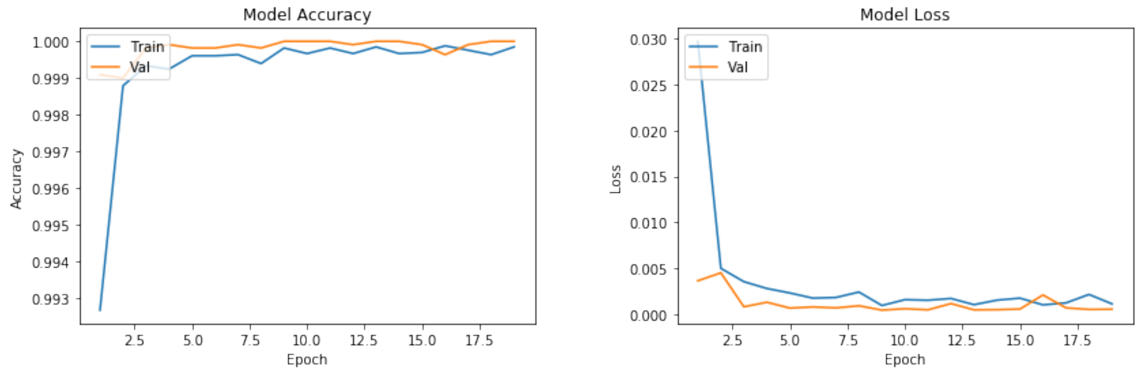
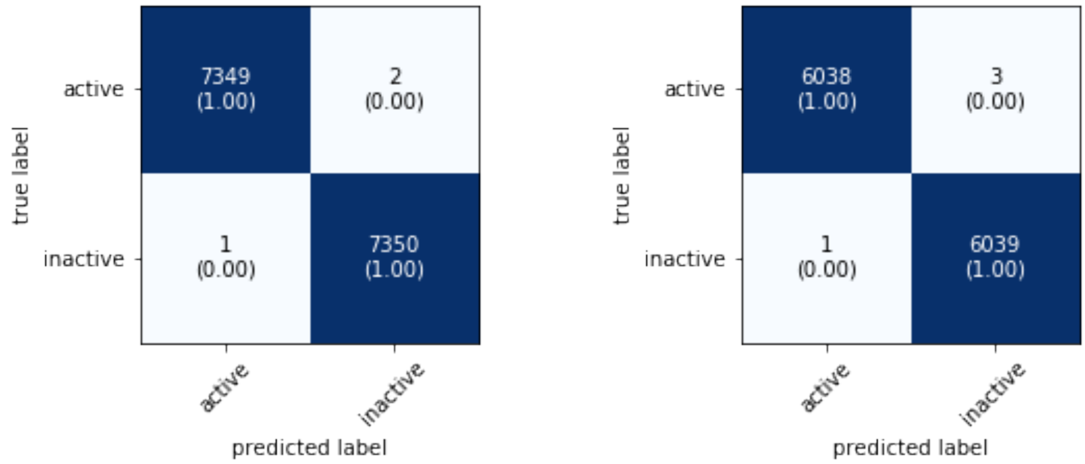


Figure 6-9 CNN model training on target domain D_T^{tr}

Table 6-2 CNN model classification results on D_S^{ts}, and using Transfer Learning on D_T^{ts}

Activities	D_S^{ts} **Accuracy: 99.99%**			D_T^{ts} **Accuracy: 99.97%**		
	Precision	Recall	F_1 score	Precision	Recall	F_1 score
Active	99.99%	99.99%	99.99%	99.98%	99.95%	99.97%
Inactive	99.99%	99.99%	99.99%	99.95%	99.98%	99.97%



(a) (b)

Figure 6-10 Confusion matrix of the classification results on D_S^{ts} and D_T^{ts} presented in (a), and (b), respectively

Table 4-3 shows that the overall accuracy achieved is 99.99% on D_S^{ts} and 99.97% on D_T^{ts}. The precision, recall, and F_1 score on the source domain is 99.99% for both active and inactive state. Also, very high scores are yielded on the target domain with precision, recall, and F_1 scores above 99.95%. An important factor that may contributed towards the high performance of the model is the differences that are present in the patterns of the two classes. For example, in inactive behaviour, the accelerometer signal patterns do not exhibit changes and remain stable due to motionless behaviour, in contrast with active behaviour, which displays more complex patterns.

These outcomes indicate the superiority of the proposed model when benchmarked with other ML techniques, as the model can be reused without the need for exhaustive data pre-processing, data labelling, and evaluation of various ML techniques. For example, the statistical results in this method indicate the robustness of CNN in terms of generalisation on unseen data, which support its use in real-life applications. In relation to real-life applications and real-time decision making, the CNN transfer model is powerful, specifically, because it can automatically extract deep learning features from the raw sensor data while producing robust results, as demonstrated. In other words, the use of CNN TL supports real-time operation in real-life scenarios of animal monitoring. As CNN performs better with larger datasets, TL can be leveraged to provide a beneficial solution (regarding time and resources) while reusing it for limited size datasets. In this way, the new dataset can be used in TL while using the knowledge acquired from the model trained in relevant larger datasets.

In this section, a previously trained CNN model was reused and successfully achieved high classification results. This method will be part of the virtual fence algorithm proposed in Section 6.2.

6.2 Virtual Fence design

The proposed virtual fence system is described in the following sections. The proposal VF is a multifunctional virtual fencing system able to monitor whether the animal is active or inactive and control as well as manipulate the position of the animals.

6.2.1 Monitoring and controlling sheep

The system's whole purpose is to monitor sheep behaviour (Section 6.1 experiment) and sheep position on the land they graze by fitting them with a smart collar. To monitor the position of the animals in real-time, this work proposes the use of GPS. Finally, to control the position and keep the animals within the virtual fence boundaries, sounds are suggested based on the sounds experiments conducted in Chapter 5.

6.2.2 Collar functionality

The collar will collect accelerometer signals and GPS coordinates with a set frequency (typically 0.5 Hz). The collar will check iteratively (i.e. every two seconds) if the animal is active or inactive and will also check whether the animal is within the VF boundaries. If the animal is approaching the restricted area, a sound warning is emitted for a fixed time (typically, lasting 15s). If the animal stops and turn back, the sound stops, otherwise, it continues for the entire duration of 15s (or a suitable custom duration setting). If the animal does not respond to the sound after the specified period, an acute audio signal (i.e. white noise) will be omitted to move sheep away from the boundary. The collar will be able to communicate with a farmer via a text message to let them know the animal's location. If the animal returns to the specified area, the system sends a text message to inform the interested parties that everything is back to normal. The sounds randomly change every two weeks to avoid habituation. The time periods were chosen based on observations during previous trials and recommendations from domain experts, i.e., face to face conversations with farmers and animal behaviourists.

6.2.3 Multifunctional VF system algorithms

Algorithm 1 and 2 represent the proposed activity recognition and virtual fence system. The system's primary functions are to identify the animal's activity and prevent the animal from crossing the VF boundary. Algorithm 1 is responsible for the activity recognition task. It presents the acquisition of accelerometer data at a sample rate of 12.5 Hz. Once the accelerometer signals are collected, the CNN model with the obtained measurements is responsible for making predictions of the activity at the specific time ("active" or "inactive"). Model predictions are stored in a new dataset, which will be sent from the collar device to the gateway. The gateway will be the means of communication between the collar and the web application hosting the information gathered from each collar.

Algorithm 1: Data Processing and Activity Recognition using CNN

Input:

Accelerometer data Acc$\{x_i, y_i, z_i\}$ // Data captured through the collar sensor

Output:

1. Matrix M $\{(f_{ij,..,}f_{tm})\}$, for i = 1,..,t where t = windowSize, and j=1,..,m where m is the accelerometer measurements

2. Classified Activity Label a, a ∈ {"active", "inactive"}

Procedure:

Let:

sampleRate ← 12.5 // The sample rate of the accelerometer is set to 12.5Hz)

```
window (data segment) ← 2 seconds      // The window size is set to 2 seconds
windowSize ←window x sampleRate
t ← currentTime
prediction ← CNNmodel                  // The CNN model
REPEAT
getActivity:
  t = i + windowSize
  While i in i ≤ t                     // get the 2 second segment
    accx = Acc[xi]
    accy = Acc[yi]
    accz = Acc[zi]
    Generate Data segment              D{( accx_i , accy_i, accz_i),..( accx_t , accy_t, accz_t )}
  End While
M ← D{(fij,..,ftm)}
a ← prediction(M)
UNTIL system stops by user input
```

Algorithm 2 presents the sequential procedure used within the proposed virtual fence system. First, the geopoints which are responsible for updating the position of the smart collar are defined. Additionally, virtual boundaries, including the desired virtual fenced area and the restricted area, A and B, respectively, are defined. In the algorithm, an important input is the personality type of each animal, as it will play an essential role in the sounds to be emitted. From the experiment in Chapter 5, it was identified that bold and shy animals react differently regarding the time needed to respond; therefore, the collars will emit different sounds based on the sheep's personality type. The sounds to be emitted are also defined, having three different categories; low frequency between 125Hz-440Hz, high frequency between 10kHz-17kHz, and white noise. The low frequency will be emitted to shy animals, and high frequency will be emitted to bold animals when the animals approach the restricted area. White noise will be used for both types of animals as the last sound to be emitted and will be used only if the animal goes outside the specified area. An alert message is defined in the algorithm, which will be sent to the user when the animal approaches the restricted zone, and when the animal is back to the desired area. The last definition includes activity monitoring, which will detect if the animal is active or inactive. The system will be updated every 2 seconds as it is a reasonable timeframe for emitting sounds and detecting the activity of the animals.

The procedure defined in Algorithm 2 identifies the personality type of the animal, then determines the sounds to be used for each case. Additionally, it gets the two predefined virtual boundaries, the animal's location at the current state, calculates the position compared to the defined virtual fence boundaries and recommends action. If the animal is inside the predefined safe area (refer to Figure 6-11, the area within polygon B), the system continues monitoring without warning. Otherwise, if the animal is within the warning area (i.e., outside polygon B, but inside polygon A), the system automatically emits sounds to restrict animals within the safe zone boundary while monitoring their location until the animal returns to the desired area. Moreover, the user receives an alert regarding the status of the position of the animals. Two alerts are generated; the first is to inform the interested party that the animal with a specified ID is outside the boundary, and the second alert informs that the animal is back to the desired area. A brief overview of the system is illustrated in Figure **6-12**.

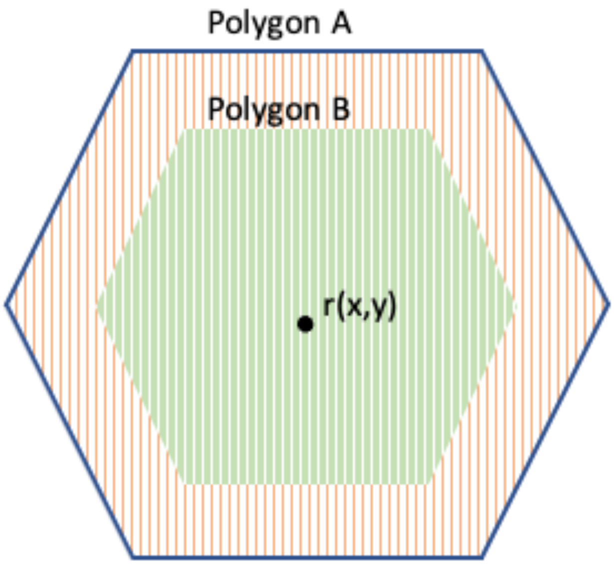


Figure 6-11 Virtual fence areas defined by the user. Point r(x,y) is the animal's position at a specific time, regularly updated from the collar sensor where x = latitude and y= longitude. Polygon A defines the restricted boundaries, and the polygon B defines the safe zone. The purpose of the system is to emit warning sounds once the animal is between polygon A and B. if the animal continues and walks toward outside polygon A, the sound cues become stronger and continue for a maximum of 15 seconds.

Algorithm 2: Virtual fence algorithm

Definitions for the virtual fence algorithm

Defining the geopoints

R {x, y} is a set of the sensor's geopoints where r_{xy} represents the sheep position

Where x = latitude , and y= longitude

Defining the personality of the sheep

P {p_s, p_b} is the set of the animal's personalities

Where p_S = "shy", and p_b="bold"

Defining the geopoints to create the virtual boundaries

A $\{a_1,..,a_n\}$ is a set of geopoints for the polygon A, Need minimum number of points

B $\{b_1,..,b_n\}$ is a set of geopoints for the polygon B, in which B subset of A n is the number of points set by the user

polyA = the polygon created from A

polyB = the polygon created from B

Defining the sounds to be used

S $\{s_{high}, s_{low}, s_{wn}\}$ is the set of the sounds to be emitted

for s_{high} = high frequency (10kHz-17kHz), s_{low} = low frequency (125Hz-440Hz), $s_{wn.}$ = white noise

Defining the sounds to be sent when the animal is outside Polygon B and inside Polygon A, or the animal is outside Polygon A

$sound_{in}$ is the sound to be sent when the animal is within Polygon A and outside Polygon B

$sound_{out}$ is the sound to be sent when the animal is outside Polygon A

Two different messages will be sent to the interested parties

alert_1 = "Animal with "ID" outside the boundary"

alert_2 = "Everything is ok now!"

Defining the activity of the animal

activity is {active, inactive}

PROCEDURE

GET [polyA, polyB] — # We get the polygons defined by the user

GET p — # Get the personality type of the animal and choose the sound to be emitted according to the type. If the animal is shy we select low frequency, if it is bold we select high frequency

IF ($p = p_s$):

    $sound_{in} = s_{low}$

ELSE

    $sound_{in} = s_{high}$

$sound_{out} = s_{wn}$ — #Defining the white noise. White noise will be emitted only if the animal goes outside the polyA

LOOP:

```
GET rxy                                  #Get latitude and longitude of the smart collar every
IF (rxy within polyB):                   2 seconds
        exit
ELSE IF (rxy within polyA):
        sound = soundin
        play(sound)
        GET activity                     # refer to Algorithm 2
                IF (activity = “active”):
                        play(sound)
                ELSE:
                        stop(sound)
                        exit
ELSE:
        sound = soundout
        FOR i=0 in i<=x:
                play(sound)
                IF (activity = “active”):
                        play(sound)
                        alert = alert_1
                        send(alert)
                ELSE:                    # x is the maximum number of seconds where x
                        stop(sound)      <=15
                        alert = alert_2
                        send(alert)
                        exit
                End
End
```

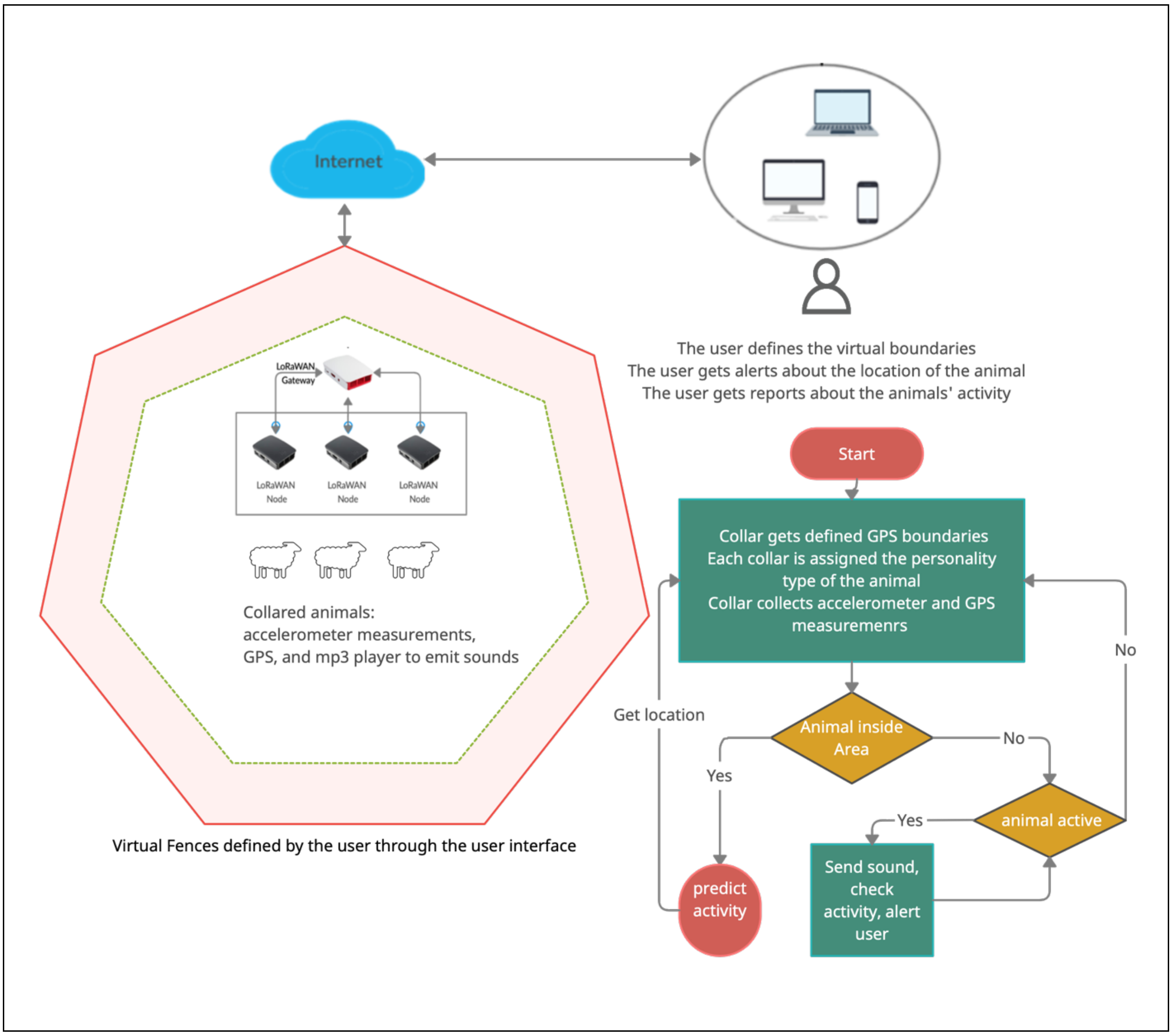


Figure 6-12 System overview

6.2.4 Discussion

In this section, the concept of VF is introduced, which manipulated the position of sheep using sounds and monitor if the animals are active or inactive. The activity of the animals is vital as it can provide valuable information on the animals' health. In this case, binary classification is proposed and differentiated only the active and inactive state of the animals. However, the activity recognition can be of any behaviour and can be easily applied in the algorithm. For example, the algorithm may be needed to identify other activities such as grazing, walking, running, scratching, fighting, or resting. The decision depends on the farmer and the required information to be analysed.

Additionally, the virtual boundaries defined in the proposed algorithm can be of any geometric size and shape. The users will be able to select in advance based on their needs. In the algorithm, the personality type of the animals is considered before the sounds are selected. This is because based on the findings in the experiment described in Chapter 5, it was identified that shy sheep are more easily manipulated as they have a strong feeling of flight response to external factors, such as a sound in this case. Based on the experiments, it was noted that shy animals respond faster (mean of 3.05 seconds) than bold animals (mean of 5.17 seconds). It was identified that shy sheep reacted faster when the low frequency was applied, and bold sheep react faster when the high frequency was used. Therefore, the decision in the virtual fence algorithm is based according to the findings. To be noted that, when testing sounds on animals without having information on their personality type, it was identified that high frequency had a statistically significant effect on the response time, having faster response on the high frequency. Therefore, it is suggested to use high frequencies when the personality type of the animal is unknown. Finally, the selection of white noise as the "punishment" sound for all animals allows a fast reaction, as it causes more stress, and hence, this sound will be emitted only if the animal goes outside the desired area.

6.3 Summary

In this Chapter, the SAR problem was evaluated in the context of two sensor types and configurations. We investigated the use of a CNN pre-trained model to TL on two datasets to classify the active and inactive state of sheep. In this research, it was proved that the suggested method (presented in Chapter 4) successfully classified the activities with accuracies above 99.95% and highlighted the importance of the use of TL in SAR. To the best of the researcher knowledge, the CNN TL for animal activity recognition was not used before. High quality classification results were achieved in terms of F1 score, precision and recall quality measures when benchmarked with other works in the literature.

Furthermore, it is suggested to use CNN TL model in the proposed virtual fence algorithm. The suggested smart system can be used for land utilisation, preventing overgrazing, and controlling if any anomalies are identified with the animals' state. Meaning, if the animal is inactive for a significant amount of time, this must alert the farmer that something might be wrong with the animals. As stated, the sheep activity recognition can include various behaviour that may be helpful. For example, grazing behaviour can be monitored and provide insight to the farmer of the animal's food intake. Also, grazing will provide information about the location the animals graze mostly and help the farmer control where the animals graze to avoid overgrazing.

In the next chapter, the summary of the whole thesis and the conclusions derived from the experiments will be presented. Additionally, future research directions will be provided.

Chapter 7 Conclusion and Future Directions

This chapter presents the summary of each chapter and the conclusions that originated from the obtained results regarding SAR problems. Additionally, this chapter presents the findings derived from the experiments conducted to test the ability of sounds to alert and control the position of sheep on the land they graze. Moreover, it reviews the contributions of the whole thesis and highlights the importance of the results from the agricultural point of view and their practical implementations in a farm setting with the help of transfer learning and acoustic cues. Possible future directions for generic AAR and VF solutions are provided at the end of this chapter that benefit from the results of this research work.

7.1 Research Summary

This thesis focuses on SAR problems using accelerometer measurements fitted on the collars of the animals using ML and CNN transfer learning. As the performance of SAR depends on various factors such as sensor orientation and position, heterogeneity of animals and accelerometers, extracted features and ML methods, these elements are considered through the following question:

Which methods are optimal in recognising sheep activities while considering orientation-independence of the sensors and heterogeneity of the animals and the sensor device to be used across various accelerometer sensors and reused when new animals and new devices are introduced?

Additionally, opportunities to use alternative solutions for a virtual fence system using acoustic cues to replace electric shocks are investigated. The main question was as follows:

How to control the position of the sheep on the land they graze without the use of electric shocks? Is there any potential to control it using solely acoustic cues?

In the thesis, we conducted an exhaustive investigation to meet the aims through the following contributions.

Contribution 1: State of the art for SAR and VF

The main stages for the SAR problem are presented in this thesis as it offers considerable potential for an efficient decision-making approach considering the wellbeing of animals and proficient land utilisation. The essential techniques to solve a SAR problem and identified opportunities and challenges are identified. This extensive review suggested that each SAR problem can be solved differently, depending on the research aim and data availability. A vital characteristic in the system design is the selection of sheep activities, as this will form the basis of the methodology. The literature review identified that lameness in sheep using accelerometer signals is a problem that has not been adequately studied, thus offering opportunities for novel contributions in this field. The majority of research works focus on using collar-borne and ear-borne accelerometers. Animal activity prediction results indicated that these two sensor positions result in higher accuracy when compared to sensors mounted on the leg and under the jaw. However, limited studies exist regarding these two latter sensors, and therefore, more research should be conducted to further explore the advantages of using leg and under the jaw based sensors. Indeed, a leg-borne sensor can provide valuable information regarding movement activities.

In contrast, the signals collected from the animal's jaw could provide more information regarding feeding activities, such as grazing, biting, chewing, and ruminating, which are critical behaviours for the sheep industry and conservation purposes. Limited studies exist regarding the use of deep learning for sheep activity recognition. Therefore, in this research work, the opportunity to this avenue is investigated.

Contribution 2: Datasets

In this research, four datasets are presented as it was identified that there is a need for open access datasets that contain sheep behaviour from motion measurements; the first dataset includes accelerometer and gyroscope measurements at a sample rate of 10Hz, obtained from a smartphone device placed on the animal's collar demonstrating six activities namely, walking, grazing, resting, browsing, scratching, and standing; the second dataset includes accelerometer measurements at a sample rate of 12.5Hz collected from a sensor device attached to the collar of sheep using MetamorionR® commercial device featuring four behaviours namely grazing, walking, scratching and inactive; the third and fourth datasets contain accelerometer measurements collected from MetamorionR®, and RaspberryPi device, respectively, at a sample rate of 12hz including active, grazing and inactive behaviours. In general, the datasets will offer the opportunity to researchers to investigate further heterogeneity and generalisation performance of ML algorithms as they contain various measurements from diverse devices, featuring a combination of sheep behaviours.

Contribution 3: Identifying optimal feature sets and ML techniques for SAR

Three different approaches using various features sets are presented to identify animal behaviours. The sensor used and window sizes are investigated to guide the research community on how to approach SAR problems based on the researchers' requirements, as the solution is not a one-fits-all. In the experiments, we also focused on the device's energy efficiency; therefore, we aimed to limit the sensor measurements, window size, and sample rate to a minimum.

Three experiments concerned with feature extraction and ML methods for AAR are presented. The purpose of the experiments was to investigate and suggest solutions for implementing an intelligent monitoring system. Various combinations of feature sets were also presented and suggested. The first experiment used an online dataset featuring five different activities of four goats and two sheep, including grazing, lying, scratching or biting, standing, and walking. The aim was to identify the most significant features and select the machine learning algorithm that could be used to classify the five activities with the highest accuracy and kappa value considering the heterogeneity of the animals. Time-domain and frequency-domain features were extracted from accelerometer, gyroscope, and magnetometer measurements. While considering the energy efficiency of a device, 15 most important features are used and trained the data using RF, XGB, MLP, and KNN to identify the most suited ML algorithm for the specific problem. The results showed that all four algorithms achieved high accuracy and Kappa values, indicating that the features could discriminate correctly between the five activities. The random forest classifier obtained the best results with an accuracy of 96.47% and a Kappa value of 95.41% for 30 second mutually exclusive behaviours, which showed a significant improvement in the results compared to previous studies.

To find optimal feature sets with RF, the second experiment conducted to evaluate the findings and test the ability of RF to discriminate six mutually exclusive behaviours successfully; walking, grazing, resting, browsing, scratching, and standing. Therefore, the collected data from a small flock comprising seven Hebridean ewes located in Shotwick, Cheshire. Gyroscope and accelerometer measurements were obtained from a smartphone device placed on the animal's harness on the top or side using the HyperIMU application. In this work, four different feature groupings are used, including a mixture of features using accelerometer and gyroscope, and identified that using only accelerometer measurements did not compromise the performance of the algorithm; therefore, the use of accelerometer measurements only is further investigated to reduce the amount of energy needed from a device when more than one sensor is used. The algorithm obtained the best results with accuracy and kappa value of 96.43% and 95.02% when using both accelerometer and gyroscope features. However, using only features extracted from the accelerometer decreased the accuracy by 0.40% and the kappa value by 0.56%. From this work, it was shown that RF could obtain high results using two different datasets, including gait movements from two species (goats and sheep), thus introducing the power of this technique for generalisation.

Finally, the third experiment used data collected from MetamotionR sensors comprised only of accelerometer measurements. The aim was to test the performance of random forest to detect grazing, walking, scratching, and inactive behaviour of sheep using 5-second windows. To test the algorithm, 17 features were extracted from the x, y, z, and magnitude of the acceleration signal resulting in 68 newly created variables. The features were then eliminated to 9, as features with higher than 80% correlation were removed. The RF results were very high for all the activities with an accuracy of 99.08% for grazing, 99.13% for walking, 99.90% for scratching, and 99.85% for inactive. The overall accuracy and kappa values were 99.43% and 98.66%, respectively. The third experiment concluded that only an accelerometer sensor can suffice excellent performance and can be used for SAR and AAR problems. Additionally, adding magnitude values of the accelerometer and reducing the window size to 5 seconds is sufficient for SAR and saves energy for the device as the computational requirements can be kept to a minimum.

Contribution 4: Transfer Learning for SAR

The use of two sensor types and configurations for the SAR problem is considered to introduce variability in the dataset, thus evaluating the generalisation properties of the CNN Transfer learning approach to identify active, grazing and inactive behaviours, which is used for the first time in a SAR problem. Two approaches were used. The first approach relied only on DL features using CNN and obtained accuracy and F1 score of 97.46% and 95.66%, respectively, on the source data. When transfer learning was applied, accuracy and F1 score of 94.79% and 94.21% were obtained. In the second approach, the extraction of handcrafted features is introduced to test whether the performance can be improved. Using CNN with the handcrafted features, an accuracy of 98.55% was achieved on the source domain. Applying transfer learning to the target data, the accuracy was 96.59%. The studies and associated analysis indicated that CNN and transfer learning could generate high accuracy in classifying the three activities. The results showed that the inclusion of handcrafted features improved the performance of both the employed CNN and transfer learning solutions. Furthermore, the simulation results showed the advantage of using deep learning in terms of generalisation, indicating its reusability when datasets are limited in animal behaviour recognition.

Contribution 5: Using audio sounds to train sheep to learn a VF system

The prospects to develop a VF system that can be ethically acceptable by identifying animal training means discarding electric shocks was investigated. Over the past years, research has been conducted to test a virtual fence's ability to control and manage animals' spatial distribution using sounds followed by electric shocks to deter animals' access to restricted areas. As electric shocks are banned in various countries because they may cause distress to the animals, alternative means to shocks are needed. Therefore, the aim is to test whether seven Hebridean ewes can be discouraged from entering restricted areas, or block their access to an attractant (food bowl filled with pellets), by using sound frequencies between 125Hz and 17000Hz, and white noise. The trials were conducted using two breeds; Hebridean ewes and Greyface Dartmoor, to identify similarities and differences between the two breeds. Bluetooth speakers attached to the animals' collars triggered sounds from a custom sound system that an observer manually controlled once the animals approached a restricted area. The results on Hebrideans showed that white noise, 125Hz-440Hz, 10000Hz-14000Hz, and 15000Hz-17000Hz could successfully discourage the Hebrideans from reaching a specific area with an overall success of 89.88%. Comparably, Greyface Dartmoor ewes were discouraged from entering a restricted area with an overall of 95.93%. The results revealed that the use of attractant, the sheep temperament, and the type of frequency have a statistically significant effect on the time needed for the animal to respond. The trials showed a potential to replace electric shocks with sounds to manage the animals' position; however, more research is needed. Therefore, this method should be tested further using larger flocks of animals and other breeds as well in order to establish a more profound understanding of whether a VF system could replace traditional fences on the land they graze using sounds only. The method can be applied to manage the spatial distribution of the animals in a fenced area that can benefit land utilisation and prevent overgrazing and soil erosion.

Contribution 6: Smart VF system proposal

An algorithm for monitoring sheep activity using accelerometer data with robustness to accelerometer specifications, position, and orientation was proposed.

Firstly, using a CNN pre-trained model was used to transfer learning on two datasets to classify sheep's active and inactive state. The suggested method successfully classifies the active and inactive with accuracy above 99.95%. The research highlights the importance of the use of Transfer learning in animal activity recognition, as, to the best of the researcher knowledge, it was not used before. Additionally, the suggested CNN transfer learning model is used to use for the proposed virtual fence algorithm. Additionally, the proposed algorithm can be used for alternative solutions such as detecting grazing, running, walking, lameness if the researchers desire to test more activities using transfer learning in combination with a virtual fence solution using only sounds.

7.2 Future Work

Future works for this research can involve the following

1. Limited studies exist regarding the use of deep learning for sheep activity recognition, and indeed, this is an avenue that needs to be further explored. Classifying sheep behaviour using DL could overcome limitations that arise from conventional approaches. For example, when using traditional ML methods, there is a requirement to develop and investigate the appropriateness of feature extraction, which takes valuable time and effort. This can be resolved by utilising DL models, which automatically learn to extract relevant features during the learning process. In this research work, the use of DL to solve SAR problems with success was explored, however, this avenue needs to be further tested using more extensive datasets and more activities such as fighting, running, browsing, foraging, lame walking and more. Identifying all those activities could be beneficial to the farmer managers as they could better understand the environment the animals live in. For example, identifying fighting might also be used with sounds to alert the animals to stop. Additionally, running might indicate that a predator or a thief is present on the farm, and the farmer will act. Therefore, creating a system able to detect a large number of sheep activities will act as a "virtual" observer for the everyday activities of the animals and comprise a complete solution for decision making on animal and land management.

2. Another future direction will be to test the proposed methods in zoos or safari parks. The directions are to control and monitor the animals online and manage and decide for their wellbeing and food intake, using an efficient system that can work with the managers for more effective animal management. The use of CNN Transfer learning might be further applied to other animals and improve the generalisation of the predictive models. Having an intelligent monitoring system in safari parks could provide a more efficient environment for the animals, as their behaviour will be monitored on a 24/7 basis. Additionally, sounds can be tested on wild animals and whether their behaviour can be altered. Suppose there is a potential use of the sounds to wild animals. In that case, the

monitoring system can also prevent aggressive behaviours of wild animals towards visitors of the park.

3. Future research will consider the use of the selected sounds to a larger flock of sheep. Additionally, looking further into the hearing ability and sheep personality type in order to evaluate the sounds used and investigate whether the selected sounds form a strong basis of successful use for a VF system.

4. Proposing the development and further testing of real-time VF system to control and monitor the position and behaviour of the animals. The system will comprise of collared devices, a gateway device, and a web application. The purpose is to allow farming managers to have complete control of the farm setting by retrieving reports regarding animal behaviour, creating virtual fences to the land the animals graze and be aware of any issues regarding escaping animals, or anomalies in the animals' behaviours. The collar device will be able to gather accelerometer measurements and GPS location in real-time. Additionally, the collar will emit sounds once the animal approaches a virtual boundary defined by the user through the gateway device.

Appendix A Pairwise comparisons tables

A.1 Pairwise comparison of frequencies on the duration: trials on Hebridean ewes

(I) Frequency	(J) Frequency	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
white_noise	125-440Hz	1.35[a]	.595	1	.023	.19	2.52
	10000-14kHz	.22	.757	1	.769	-1.26	1.71
	15000-17kHz	.29	.760	1	.701	-1.20	1.78
125-440Hz	white_noise	-1.35[a]	.595	1	.023	-2.52	-.19
	10000-14kHz	-1.13	.583	1	.053	-2.27	.01
	15000-17kHz	-1.06	.626	1	.090	-2.29	.17
10000-14kHz	white_noise	-.22	.757	1	.769	-1.71	1.26
	125-440Hz	1.13	.583	1	.053	-.01	2.27
	15000-17kHz	.07	.798	1	.931	-1.50	1.63
15000-17kHz	white_noise	-.29	.760	1	.701	-1.78	1.20

	125-440Hz	1.06	.626	1	.090	-.17	2.29
	10000-14kHz	-.07	.798	1	.931	-1.63	1.50

A.2 Pairwise comparison of frequencies on the duration: trial on Greyface Dartmoor ewes

(I) Frequency	(J) Frequency	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
white_noise	125-440Hz	-1.4182[a]	.32267	1	.000	-2.0507	-.7858
	10000-14kHz	-.8990[a]	.40211	1	.025	-1.6872	-.1109
	15000-17kHz	-.5321[a]	.26841	1	.047	-1.0582	-.0060
125-440Hz	white_noise	1.4182[a]	.32267	1	.000	.7858	2.0507
	10000-14kHz	.5192	.44999	1	.249	-.3628	1.4012
	15000-17kHz	.8861[a]	.33591	1	.008	.2278	1.5445
10000-14kHz	white_noise	.8990[a]	.40211	1	.025	.1109	1.6872
	125-440Hz	-.5192	.44999	1	.249	-1.4012	.3628
	15000-17kHz	.3669	.41281	1	.374	-.4421	1.1760
15000-17kHz	white_noise	.5321[a]	.26841	1	.047	.0060	1.0582
	125-440Hz	-.8861[a]	.33591	1	.008	-1.5445	-.2278
	10000-14kHz	-.3669	.41281	1	.374	-1.1760	.4421

References

[1] D. M. Weary, J. M. Huzzey, and M. A. G. Von Keyserlingk, “Board-invited Review: Using behavior to predict and identify ill health in animals,” *Journal of Animal Science*, vol. 87, no. 2, pp. 770–777, Feb. 2009.

[2] K. M. McLennan *et al.*, “Technical note: Validation of an automatic recording system to assess behavioural activity level in sheep (Ovis aries),” *Small Ruminant Research*, vol. 127, pp. 92–96, Jan. 2015.

[3] A. . Frost, C. . Schofield, S. . Beaulah, T. . Mottram, J. . Lines, and C. . Wathes, “A review of livestock monitoring and the needs for integrated systems,” *Comput. Electron. Agric.*, vol. 17, pp. 139–159, 1997.

[4] J. P. Bikker *et al.*, “Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity,” *Journal of Dairy Science*, vol. 97, no. 5, pp. 2974–2979, May 2014.

[5] J. Barwick, D. Lamb, R. Dobos, D. Schneider, M. Welch, and M. Trotter, “Predicting lameness in sheep activity using tri-axial acceleration signals,” *Animals*, vol. 8, no. 1, pp. 1–16, Apr. 2018.

[6] J. Kaler, J. Mitsch, J. A. Vázquez-Diosdado, N. Bollard, T. Dottorini, and K. A. Ellis, “Automated detection of lameness in sheep using machine learning approaches: novel insights into behavioural differences among lame and non-lame sheep,” *Royal Society Open Science*, vol. 7, no. 1, p. 190824, Jan. 2020.

[7] D. A. Gougoulis, I. Kyriazakis, and G. C. Fthenakis, “Diagnostic significance of behaviour changes of sheep: A selected review,” *Small Ruminant Research*, vol. 92, no. 1, pp. 52–56, Jan. 2010.

[8] Z. Al-Rubaye, A. Al-Sherbaz, W. D. McCormick, and S. J. Turner, “The use of multivariable wireless sensor data to early detect lameness in sheep,” 2016.

[9] A. C. Winter, “Lameness in sheep,” *Small Ruminant Research*, vol. 76, no. 1–2, pp. 149–153, 2008.

[10] N. Chapinal, A. M. de Passillé, M. Pastell, L. Hänninen, L. Munksgaard, and J. Rushen, “Measurement of acceleration while walking as an automated method for gait assessment in dairy cattle,” *Journal of Dairy Science*, vol. 94, no. 6, pp. 2895–2901, Jun. 2011.

[11] S. Neethirajan, “The role of sensors, big data and machine learning in modern animal

farming," *Sensing and Bio-Sensing Research*, vol. 29. Elsevier, p. 100367, 01-Aug-2020.

[12] P. D. PENNING, "A technique to record automatically some aspects of grazing and ruminating behaviour in sheep," *Grass and Forage Science*, vol. 38, no. 2, pp. 89–96, Jun. 1983.

[13] S. M. M. Rutter, R. A. A. Champion, and P. D. D. Penning, "An automatic system to record foraging behaviour in free-ranging ruminants," *Applied Animal Behaviour Science*, vol. 54, no. 2, pp. 185–195, Jan. 1997.

[14] P. P. Nielsen, "Automatic registration of grazing behaviour in dairy cows using 3D activity loggers," *Applied Animal Behaviour Science*, vol. 148, no. 3–4, pp. 179–184, Oct. 2013.

[15] P. Martin and P. Bateson, "Measuring Behaviour: An Introductory Guide," 1986.

[16] C. Carslake, J. A. Vázquez-Diosdado, and J. Kaler, "Machine learning algorithms to classify and quantify multiple behaviours in dairy calves using a sensor–moving beyond classification in precision livestock," *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–14, Dec. 2021.

[17] A. Jukan, X. Masip-Bruin, and N. Amla, "Smart computing and sensing technologies for animal welfare: A systematic review," *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–15, Nov. 2017.

[18] E. E. L. C. C. Shepard *et al.*, "Identification of animal movement patterns using tri-axial accelerometry," *Endangered Species Research*, vol. 10, no. 1, pp. 47–60, Mar. 2010.

[19] S. Neethirajan, "Recent advances in wearable sensors for animal health management," *Sensing and Bio-Sensing Research*, vol. 12. Elsevier, pp. 15–29, 01-Feb-2017.

[20] D. N. Tran, T. N. Nguyen, P. C. P. Khanh, and D. T. Trana, "An IoT-based Design Using Accelerometers in Animal Behavior Recognition Systems," *IEEE Sensors Journal*, pp. 1–1, 2021.

[21] E. S. Fogarty, D. L. Swain, G. M. Cronin, L. E. Moraes, and M. Trotter, "Behaviour classification of extensively grazed sheep using machine learning," *Computers and Electronics in Agriculture*, vol. 169, p. 105175, Feb. 2020.

[22] P. M. Graf, R. P. Wilson, L. Qasem, K. Hackländer, and F. Rosell, "The use of acceleration to code for animal behaviours; a case study in free-ranging Eurasian beavers Castor fiber," *PLoS ONE*, vol. 10, no. 8, Aug. 2015.

[23] K. Ren, J. Karlsson, M. Liuska, M. Hartikainen, I. Hansen, and G. H. Jørgensen, "A sensor-fusion-system for tracking sheep location and behaviour," *International Journal of Distributed Sensor Networks*, vol. 16, no. 5, p. 155014772092177, May 2020.

[24] Ö. Cangar *et al.*, "Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis," *Computers and Electronics in Agriculture*, vol. 64, no. 1, pp. 53–60, Jan. 2008.

[25] N. Krahnstoever, J. Rittscher, P. Tu, K. Chean, and T. Tomlinson, "Activity Recognition using Visual Tracking and RFID," in *Seventh IEEE Workshops on Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1*, 2005, vol. 1, pp. 494–500.

[26] X. F. Wang and D. S. Huang, "A novel multi-layer level set method for image segmentation," 2008.

[27] E. Schlecht, C. Hülsebusch, F. Mahler, and K. Becker, "The use of differentially corrected global positioning system to monitor activities of cattle at pasture," *Applied Animal Behaviour Science*, vol. 85, no. 3, pp. 185–202, Jan. 2004.

[28] E. D. Ungar, Z. Henkin, M. Gutman, A. Dolev, A. Genizi, and D. Ganskopp, "Inference of Animal Activity From GPS Collar Data on Free-Ranging Cattle," *Rangeland Ecology & Management*, vol. 58, no. 3, pp. 256–266, Jan. 2005.

[29] M. Schwager, D. M. Anderson, Z. Butler, and D. Rus, "Robust classification of animal tracking data," *Computers and Electronics in Agriculture*, vol. 56, no. 1, pp. 46–59, Dec. 2007.

[30] L. . Turner, M. . Udal, B. . Larson, and S. . Shearer, "Monitoring cattle behaviour and pasture use with GPS and GIS," *Can. J. Anim. Sci.*, vol. 80, pp. 405–413, 2000.

[31] H. Sheng *et al.*, "Construction of sheep forage intake estimation models based on sound analysis," *Biosystems Engineering*, vol. 192, pp. 144–158, Apr. 2020.

[32] A. Mason and J. Sneddon, "Automated monitoring of foraging behaviour in free ranging sheep grazing a biodiverse pasture," in *2013 Seventh International Conference on Sensing Technology (ICST)*, 2013, pp. 46–51.

[33] C. Umstatter, "The evolution of virtual fences: A review," *Computers and Electronics in Agriculture*, vol. 75, no. 1, pp. 10–22, Jan. 2011.

[34] D. M. Anderson, "Virtual fencing past, present and future," *Rangeland Journal*, vol. 29, no. 1, pp. 65–78, Jul. 2007.

[35] C. W. Simmons, "Fences, gates and cattle guards," *The Cattleman*, vol. 22, no. 32, 1935.

[36] L. D. Howery, A. F. Cibils, and D. M. Anderson, "Potential for using visual, auditory and olfactory cues to manage foraging behaviour and spatial distribution of rangeland livestock," *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, vol. 8, no. 049, Nov. 2013.

[37] D. M. Anderson, R. E. Estell, J. L. Holechek, S. Ivey, and G. B. Smith, "Virtual herding for flexible livestock management - A review," *Rangeland Journal*, vol. 36, no. 3, pp. 205–221, 2014.

[38] B. E. Norton, M. Barnes, and R. Teague, "Grazing Management Can Improve Livestock Distribution: Increasing accessible forage and effective grazing capacity," *Rangelands*, vol. 35, no. 5, pp. 45–51, Jan. 2013.

[39] S. M. Rutter, "Advanced livestock management solutions," in *Advances in Sheep Welfare*, D. M. Ferguson, C. Lee, and A. Fisher, Eds. Woodhead Publishing, 2017, pp.

245–261.

[40] J. E. Morris, A. D. Fisher, R. E. Doyle, and R. D. Bush, "Determination of sheep learning responses to a directional audio cue," *Journal of applied animal welfare science: JAAWS*, vol. 13, no. 4, pp. 347–360, May 2010.

[41] J. A. Vázquez Diosdado *et al.*, "Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system," *Animal Biotelemetry*, vol. 3, no. 1, p. 15, Dec. 2015.

[42] M. Moreau, S. Siebert, A. Buerkert, and E. Schlecht, "Use of a tri-axial accelerometer for automated recording and classification of goats' grazing behaviour," *Applied Animal Behaviour Science*, vol. 119, no. 3–4, pp. 158–170, Jul. 2009.

[43] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[44] S. Lomax, P. Colusso, and C. E. F. Clark, "Does Virtual Fencing Work for Grazing Dairy Cattle?," *Animals*, vol. 9, no. 7, p. 429, Jul. 2019.

[45] D. L. M. Campbell, J. M. Lea, S. J. Haynes, W. J. Farrer, C. J. Leigh-Lancaster, and C. Lee, "Virtual fencing of cattle using an automated collar in a feed attractant trial," *Applied Animal Behaviour Science*, vol. 200, no. October 2017, pp. 71–77, Mar. 2018.

[46] J. Barwick, D. W. Lamb, R. Dobos, M. Welch, D. Schneider, and M. Trotter, "Identifying sheep activity from tri-axial acceleration signals using a moving window classification model," *Remote Sensing*, vol. 12, no. 4, p. 646, Feb. 2020.

[47] D. D. Brown, R. Kays, M. Wikelski, R. Wilson, and A. Klimley, "Observing the unwatchable through acceleration logging of animal behavior," *Animal Biotelemetry*, vol. 1, no. 1, p. 20, 2013.

[48] K. M. Scheibe and C. Gromann, *Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis*, vol. 38, no. 3. Psychonomic Society Inc., 2006, pp. 427–433.

[49] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[50] J. Morales and D. Akopian, "Physical activity recognition by smartphones, a survey," *Biocybernetics and Biomedical Engineering*, vol. 37, no. 3. Elsevier, pp. 388–400, 01-Jan-2017.

[51] M. R. Borchers, Y. M. Chang, I. C. Tsai, B. A. Wadsworth, and J. M. Bewley, "A validation of technologies monitoring dairy cow feeding, ruminating, and lying behaviors," *Journal of Dairy Science*, vol. 99, no. 9, 2016.

[52] N. Kleanthous, A. Hussain, W. Khan, J. Sneddon, and A. Mason, "Feature Extraction and Random Forest to Identify Sheep Behavior from Accelerometer Data," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence*

and Lecture Notes in Bioinformatics), vol. 12465 LNAI, pp. 408–419, Oct. 2020.

[53] L. Guo, M. Welch, R. Dobos, P. Kwan, and W. Wang, "Comparison of grazing behaviour of sheep on pasture with different sward surface heights using an inertial measurement unit sensor," *Computers and Electronics in Agriculture*, vol. 150, pp. 394–401, Jul. 2018.

[54] A. Cardoso, J. Pereira, L. Nóbrega, P. Gonçalves, P. Pedreiras, and V. Silva, "SheepIT: Activity and Location Monitoring," *INForum 2018 - Simpósio de Informática*, no. October, pp. 1–12, 2018.

[55] E. Walton *et al.*, "Evaluation of sampling frequency, window size and sensor position for classification of sheep behaviour," *Royal Society Open Science*, vol. 5, no. 2, 2018.

[56] J. Barwick, D. W. Lamb, R. Dobos, M. Welch, and M. Trotter, "Categorising sheep activity using a tri-axial accelerometer," *Computers and Electronics in Agriculture*, vol. 145, pp. 289–297, Feb. 2018.

[57] N. Mansbridge *et al.*, "Feature selection and comparison of machine learning algorithms in classification of grazing and rumination behaviour in sheep," *Sensors (Switzerland)*, vol. 18, no. 10, pp. 1–16, Oct. 2018.

[58] J. A. Vázquez-Diosdado, V. Paul, K. A. Ellis, D. Coates, R. Loomba, and J. Kaler, "A combined offline and online algorithm for real-time and long-term classification of sheep behaviour: Novel approach for precision livestock farming," *Sensors (Switzerland)*, vol. 19, no. 14, 2019.

[59] F. A. P. Alvarenga, I. Borges, V. H. Oddy, and R. C. Dobos, "Discrimination of biting and chewing behaviour in sheep using a tri-axial accelerometer," *Computers and Electronics in Agriculture*, vol. 168, p. 105051, Jan. 2020.

[60] F. A. P. Alvarenga, I. Borges, L. Palkovič, J. Rodina, V. H. Oddy, and R. C. Dobos, "Using a three-axis accelerometer to identify and classify sheep behaviour at pasture," *Applied Animal Behaviour Science*, vol. 181, pp. 91–99, Aug. 2016.

[61] V. Giovanetti *et al.*, "Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer," *Livestock Science*, vol. 196, pp. 42–48, Feb. 2017.

[62] M. Decandia *et al.*, "The effect of different time epoch settings on the classification of sheep behaviour using tri-axial accelerometry," *Computers and Electronics in Agriculture*, vol. 154, pp. 112–119, 2018.

[63] M. Radeski and V. Ilieski, "Gait and posture discrimination in sheep using a tri-axial accelerometer," *animal*, vol. 11, no. 7, pp. 1249–1257, Jan. 2017.

[64] N. Kleanthous *et al.*, "Machine Learning Techniques for Classification of Livestock Behavior," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11304 LNCS, 2018, pp. 304–315.

[65] J. Marais *et al.*, *Automatic classification of sheep behaviour using 3-axis accelerometer data*, no. August. 2014.

[66] S. P. le Roux, J. Marias, R. Wolhuter, and T. Niesler, “Animal-borne behaviour classification for sheep (Dohne Merino) and Rhinoceros (Ceratotherium simum and Diceros bicornis),” *Animal Biotelemetry*, vol. 5, no. 1, p. 25, Dec. 2017.

[67] E. S. Nadimi, R. N. Jørgensen, V. Blanes-Vidal, and S. Christensen, “Monitoring and classifying animal behavior using ZigBee-based mobile ad hoc wireless sensor networks and artificial neural networks,” *Computers and Electronics in Agriculture*, vol. 82, no. Supplement C, pp. 44–54, Feb. 2012.

[68] J. W. Kamminga, H. C. Bisby, D. V. Le, N. Meratnia, and P. J. M. Havinga, “Generic Online Animal Activity Recognition on Collar Tags,” in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers on - UbiComp ’17*, 2017, no. October, pp. 597–606.

[69] S. P. le Roux, R. Wolhuter, and T. Niesler, “Energy-Aware Feature and Model Selection for Onboard Behavior Classification in Low-Power Animal Borne Sensor Applications,” *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2722–2734, Apr. 2019.

[70] L. Nóbrega *et al.*, “Assessing sheep behavior through low-power microcontrollers in smart agriculture scenarios,” *Computers and Electronics in Agriculture*, vol. 173, p. 105444, Jun. 2020.

[71] N. Kleanthous, A. Hussain, A. Mason, and J. Sneddon, “Data Science Approaches for the Analysis of Animal Behaviours,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11645 LNAI, 2019, pp. 411–422.

[72] R. M. Alexander, “Optimization and gaits in the locomotion of vertebrates,” *Physiological reviews*, vol. 69, no. 4, p. 1199—1227, Oct. 1989.

[73] A. M. de Passillé, M. B. Jensen, N. Chapinal, and J. Rushen, “Technical note: Use of accelerometers to describe gait patterns in dairy calves,” *Journal of Dairy Science*, vol. 93, no. 7, pp. 3287–3293, Jul. 2010.

[74] J. Hancock, “Studies in grazing behaviour of dairy cattle: II. Bloat in relation to grazing behaviour,” *The Journal of Agricultural Science*, vol. 45, no. 1, pp. 80–95, 1954.

[75] C. M. Dwyer and H. L. I. Bornett, “Chronic stress in Sheep: Assessment tools and their use in different management conditions,” *Animal Welfare*, vol. 13, no. 3. pp. 293–304, 2004.

[76] A. Bonn, T. Allott, K. Hubacek, and J. Stewart, *Drivers of environmental change in uplands*. 2008.

[77] R. Evans, “Reducing soil erosion and the loss of soil fertility for environmentally-sustainable agricultural cropping and livestock production systems,” *Annals of Applied Biology*, vol. 146, no. 2, pp. 137–146, 2005.

[78] K. Ito, M. A. G. von Keyserlingk, S. J. LeBlanc, and D. M. Weary, “Lying behavior as an indicator of lameness in dairy cows,” *Journal of Dairy Science*, vol. 93, no. 8, pp. 3553–3560, Aug. 2010.

[79] G. M. Borghart, L. E. O'Grady, and J. R. Somers, "Prediction of lameness using automatically recorded activity, behavior and production data in post-parturient Irish dairy cows," *Irish Veterinary Journal*, vol. 74, no. 1, p. 4, Dec. 2021.

[80] D. Jarchi, J. Kaler, and S. Sanei, "Lameness Detection in Cows Using Hierarchical Deep Learning and Synchrosqueezed Wavelet Transform," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9349–9358, Apr. 2021.

[81] N. C. Heglund and C. R. Taylor, "Speed, stride frequency and energy cost per stride: how do they change with body size and gait?," *Journal of Experimental Biology*, vol. 138, no. 1, 1988.

[82] "ELAN - The Language Archive." [Online]. Available: https://tla.mpi.nl/tools/tla-tools/elan/. [Accessed: 16-Dec-2019].

[83] A. Dehghani, O. Sarbishei, T. Glatard, and E. Shihab, "A quantitative comparison of overlapping and non-overlapping sliding windows for human activity recognition using inertial sensors," *Sensors (Switzerland)*, vol. 19, no. 22, pp. 10–12, Nov. 2019.

[84] Y. Wang, S. Cang, and H. Yu, "A survey on wearable sensor modality centred human activity recognition in health care," *Expert Systems with Applications*, vol. 137. Elsevier Ltd, pp. 167–190, 15-Dec-2019.

[85] Z. L. Sun, D. S. Huang, and Y. M. Cheun, "Extracting nonlinear features for multispectral images by FCMC and KPCA," in *Digital Signal Processing: A Review Journal*, 2005, vol. 15, no. 4, pp. 331–346.

[86] Z. L. Sun, D. S. Huang, Y. M. Cheung, J. Liu, and G. Bin Huang, "Using FCMC, FVS, and PCA techniques for feature extraction of multispectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, 2005.

[87] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *2014 Science and Information Conference*, 2014, pp. 372–378.

[88] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, vol. 181, pp. 108–115, Mar. 2016.

[89] S. J. Preece* *et al.*, "A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data - IEEE Xplore Document," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 871–879, Jun. 2009.

[90] P. Sarcevic, S. Pletl, and Z. Kincses, "Comparison of time- and frequency-domain features for movement classification using data from wrist-worn sensors," *SISY 2017 - IEEE 15th International Symposium on Intelligent Systems and Informatics, Proceedings*, pp. 261–265, Sep. 2017.

[91] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks*. 2010.

[92] W. Dargie, "Analysis of time and frequency domain features of accelerometer measurements," in *Proceedings - International Conference on Computer Communications and Networks, ICCCN*, 2009.

[93] D.-S. Huang, “Systematic theory of neural networks for pattern recognition,” *Publishing House of Electronic Industry of China, Beijing*, vol. 201, 1996.

[94] D. S. Huang, “Radial basis probabilistic neural networks: Model and application,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, no. 7, pp. 1083–1101, Nov. 1999.

[95] C. V Bouten, K. R. Westerterp, M. Verduin, and J. D. Janssen, “Assessment of energy expenditure for physical activity using a triaxial accelerometer,” *Medicine and Science in Sports and Exercise*, vol. 26, no. 12, pp. 1516–1523, 1994.

[96] T. Gneiting, H. Ševčíková, and D. B. Percival, “Estimators of fractal dimension: Assessing the roughness of time series and spatial data,” *Statistical Science*, pp. 247–277, 2012.

[97] L. Bao and S. S. Intille, “Activity recognition from user-annotated acceleration data,” *Pervasive Computing*, vol. 3001, pp. 1–17, Aug. 2004.

[98] M. Dash and H. Liu, “Feature selection for classification,” *Intelligent Data Analysis*, vol. 1, no. 3, pp. 131–156, Jan. 1997.

[99] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, Jan. 2014.

[100] C. Y. Lu and D. S. Huang, “Optimized projections for sparse representation based classification,” *Neurocomputing*, vol. 113, pp. 213–219, Aug. 2013.

[101] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, “A review of feature selection methods on synthetic data,” *Knowledge and Information Systems*, vol. 34, no. 3. Springer, pp. 483–519, 30-Mar-2013.

[102] A. Jović, K. Brkić, and N. Bogunović, “A review of feature selection methods with applications,” in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, 2015, pp. 1200–1205.

[103] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, “MIFS-ND: A mutual information-based feature selection method,” *Expert Systems with Applications*, vol. 41, no. 14, pp. 6371–6385, Oct. 2014.

[104] E. B. Hunt, P. J. Stone, and J. Marin, *Experiments in induction Earl B. Hunt, Janet Marin, Philip J. Stone*. 1966.

[105] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2016.

[106] L. Yu and H. Liu, “Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution,” in *Proceedings, Twentieth International Conference on Machine Learning*, 2003, vol. 2, pp. 856–863.

[107] R. O Duda, P. E Hart, and D. G.Stork, *Pattern Classification*, vol. xx. 2001.

[108] D. S. Modha and W. S. Spangler, “Feature weighting on k-means clustering,” *Machine Learning*, vol. 52, no. 3, pp. 217–237, 2003.

[109] K. Kira and L. A. Rendell, “Feature selection problem: traditional methods and a new algorithm,” in *Proceedings Tenth National Conference on Artificial Intelligence*, 1992, pp. 129–134.

[110] I. Kononenko, “Estimating attributes: Analysis and extensions of RELIEF,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1994, vol. 784 LNCS, pp. 171–182.

[111] M. Robnik-Šikonja and I. Kononenko, “Theoretical and Empirical Analysis of ReliefF and RReliefF,” *Machine Learning*, vol. 53, no. 1–2, pp. 23–69, 2003.

[112] S. Fan, Y. Jia, and C. Jia, “A feature selection and classification method for activity recognition based on an inertial sensing unit,” *Information (Switzerland)*, vol. 10, no. 10, p. 290, Sep. 2019.

[113] C. Freeman, D. Kulić, and O. Basir, “An evaluation of classifier-specific filter measure performance for feature selection,” *Pattern Recognition*, vol. 48, no. 5, pp. 1812–1826, May 2015.

[114] B. Li, C. Wang, and D. S. Huang, “Supervised feature extraction based on orthogonal discriminant projection,” *Neurocomputing*, vol. 73, no. 1–3, pp. 191–196, Dec. 2009.

[115] B. Li, D. S. Huang, C. Wang, and K. H. Liu, “Feature extraction using constrained maximum variance mapping,” *Pattern Recognition*, vol. 41, no. 11, pp. 3287–3294, Nov. 2008.

[116] J. Kittler, “Feature Set Search Alborithms,” in *Pattern Recognition and Signal Processing*, 1978, pp. 41–60.

[117] J. H. Holland, *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*. 1975.

[118] D. Asir, S. Appavu, and E. Jebamalar, “International Journal of Computer Mathematics,” *Technometrics*, vol. 7, no. 1, pp. 87–88, 1965.

[119] M. B. Kursa, A. Jankowski, and W. R. Rudnicki, “Boruta – A System for Feature Selection,” *Fundamenta Informaticae*, vol. 101, no. 4, pp. 271–285, Aug. 2010.

[120] M. B. Kursa and W. Rudnicki, *Feature Selection with Boruta Package*, vol. 36. 2010.

[121] B. H. Menze *et al.*, “A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data,” *BMC Bioinformatics*, vol. 10, no. 1, p. 213, Jul. 2009.

[122] J. Suto, S. Oniga, and P. P. Sitar, “Comparison of wrapper and filter feature selection algorithms on human activity recognition,” in *2016 6th International Conference on Computers Communications and Control, ICCCC 2016*, 2016, pp. 124–129.

[123] I. Inza, P. Larrañaga, R. Etxeberria, and B. Sierra, “Feature Subset Selection by Bayesian network-based optimization,” *Artificial Intelligence*, vol. 123, no. 1–2, pp. 157–184, Oct. 2000.

[124] M. Grimaldi, P. Cunningham, and A. Kokaram, “An Evaluation Of Alternative Feature Selection Strategies And Ensemble Techniques For Classifying Music,” *Workshop on*

Multimedia Discovery and Mining, vol. 2, 2003.

[125] J. G. Dy and C. E. Brodley, "Feature Subset Selection and Order Identification for Unsupervised Learning," *ICML 00 Proceedings of the Seventeenth International Conference on Machine Learning*, pp. 247–254, 2000.

[126] D. W. Aha and R. L. Bankert, "A Comparative Evaluation of Sequential Feature Selection Algorithms," Springer, New York, NY, 1996, pp. 199–206.

[127] F. Han and D. S. Huang, "A new constrained learning algorithm for function approximation by encoding a priori information into feedforward neural networks," *Neural Computing and Applications*, vol. 17, no. 5–6, pp. 433–439, Oct. 2008.

[128] S. P. le Roux, J. Marias, R. Wolhuter, and T. Niesler, "Animal-borne behaviour classification for sheep ({Dohne} {Merino}) and {Rhinoceros} ({Ceratotherium} simum and {Diceros} bicornis)," *Animal Biotelemetry*, vol. 5, p. 25, Nov. 2017.

[129] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Wadsworth International Group, 2017.

[130] J. R. Quinlan, *{C4}.5 - Programs for Machine Learning*. Kaufmann, 1993.

[131] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-Augu, pp. 785–794, Aug. 2016.

[132] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, no. 1–3, pp. 389–422, 2002.

[133] X.-F. Wang, J.-X. Du, H. Xu, and L. Heutte, "Classification of plant leaf images with complicated background," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 916–926, Nov. 2008.

[134] J. J. Valletta, C. Torney, M. Kings, A. Thornton, and J. Madden, "Applications of machine learning in animal behaviour studies," *Animal Behaviour*, vol. 124, pp. 203–220, Mar. 2017.

[135] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.

[136] A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, "k-Nearest Neighbor Classification," Springer, New York, NY, 2009, pp. 83–106.

[137] G. Xiao, K. Li, and K. Li, "Reporting l most influential objects in uncertain databases based on probabilistic reverse top-k queries," *Information Sciences*, vol. 405, pp. 207–226, Sep. 2017.

[138] J. Camargo and A. Young, "Feature Selection and Non-Linear Classifiers: Effects on Simultaneous Motion Recognition in Upper Limb.," *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 27, no. 4, pp. 743–750, 2019.

[139] G. Shakhnarovich, T. Darrell, and P. Indyk, *Nearest-Neighbor Methods in Learning*

and Vision. MIT Press, 2018.

[140] G. Xiao, K. Li, X. Zhou, and K. Li, “Efficient monochromatic and bichromatic probabilistic reverse top-k query processing for uncertain big data,” *Journal of Computer and System Sciences*, vol. 89, pp. 92–113, Nov. 2017.

[141] J. Hu, H. Peng, J. Wang, and W. Yu, “kNN-P: A kNN classifier optimized by P systems,” *Theoretical Computer Science*, vol. 817, pp. 55–65, May 2020.

[142] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, “EMG feature evaluation for improving myoelectric pattern recognition robustness,” *Expert Systems with Applications*, vol. 40, no. 12, pp. 4832–4840, Sep. 2013.

[143] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, “Support vector machines,” *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[144] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “Training algorithm for optimal margin classifiers,” in *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, 1992, pp. 144–152.

[145] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[146] A. Urtubia, R. León, and M. Vargas, “Identification of chemical markers to detect abnormal wine fermentation using Support Vector Machines,” *Computers & Chemical Engineering*, p. 107158, Nov. 2020.

[147] D. S. Huang and J. X. Du, “A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks,” *IEEE Transactions on Neural Networks*, vol. 19, no. 12, pp. 2099–2115, Dec. 2008.

[148] Y. Fan *et al.*, “Privacy preserving based logistic regression on big data,” *Journal of Network and Computer Applications*, vol. 171, p. 102769, Dec. 2020.

[149] S. B. Kotsiantis, “Decision trees: a recent overview,” *Artificial Intelligence Review*, vol. 39, no. 4, pp. 261–283, Sep. 2013.

[150] J. R. Quinlan, “Induction of decision trees,” *Machine Learning*, vol. 1, no. 1, pp. 81–106, Mar. 1986.

[151] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Aug. 2001.

[152] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[153] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

[154] P. Panov and S. Džeroski, “Combining bagging and random subspaces to create better ensembles,” in *Lecture Notes in Computer Science (including subseries Lecture Notes*

in Artificial Intelligence and Lecture Notes in Bioinformatics), 2007, vol. 4723 LNCS, pp. 118–129.

[155] X. Wu *et al.*, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.

[156] S. Chen, G. I. Webb, L. Liu, and X. Ma, "A novel selective naïve Bayes algorithm," *Knowledge-Based Systems*, vol. 192, p. 105361, Mar. 2020.

[157] D. S. Huang and S. De Ma, "Linear and nonlinear feedforward neural network classifiers: A comprehensive understanding," *Journal of Intelligent Systems*, vol. 9, no. 1, 1999.

[158] D. S. Huang, H. H. S. Ip, K. C. K. Law, and Z. Chi, "Zeroing polynomials using modified constrained neural network approach," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, 2005.

[159] L. Shang, D. S. Huang, J. X. Du, and C. H. Zheng, "Palmprint recognition using FastICA algorithm and radial basis probabilistic neural network," *Neurocomputing*, vol. 69, no. 13–15, 2006.

[160] S. Li, Z. H. You, H. Guo, X. Luo, and Z. Q. Zhao, "Inverse-free extreme learning machine with optimal information updating," *IEEE Transactions on Cybernetics*, vol. 46, no. 5, 2016.

[161] Z. Q. Zhao, D. S. Huang, and W. Jia, "Palmprint recognition with 2DPCA+PCA based on modular neural networks," *Neurocomputing*, vol. 71, no. 1–3, 2007.

[162] J. X. Du, D. S. Huang, G. J. Zhang, and Z. F. Wang, "A novel full structure optimization algorithm for radial basis probabilistic neural networks," *Neurocomputing*, vol. 70, no. 1–3, pp. 592–596, 2006.

[163] D. S. Huang and W. B. Zhao, "Determining the centers of radial basis probabilistic neural networks by recursive orthogonal least square algorithms," *Applied Mathematics and Computation*, vol. 162, no. 1, 2005.

[164] W. B. Zhao, D. S. Huang, J. Y. Du, and L. M. Wang, "Genetic optimization of radial basis probabilistic neural networks," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 8. 2004.

[165] D.-S. Huang, "A Constructive Approach for Finding Arbitrary Roots of Polynomials by Neural Networks," *IEEE Transactions on Neural Networks*, vol. 15, no. 2, pp. 477–491, Mar. 2004.

[166] B. Li, C. H. Zheng, and D. S. Huang, "Locally linear discriminant embedding: An efficient method for face recognition," *Pattern Recognition*, vol. 41, no. 12, 2008.

[167] Y. Zhao, D. S. Huang, and W. Jia, "Completed local binary count for rotation invariant texture classification," *IEEE Transactions on Image Processing*, vol. 21, no. 10, 2012.

[168] J. X. Du, D. S. Huang, X. F. Wang, and X. Gu, "Shape recognition based on neural networks trained by differential evolution algorithm," *Neurocomputing*, vol. 70, no. 4–6, pp. 896–903, 2007.

[169] X. F. Wang, D. S. Huang, and H. Xu, “An efficient local Chan-Vese model for image segmentation,” *Pattern Recognition*, vol. 43, no. 3, 2010.

[170] Z. Q. Zhao and D. S. Huang, “A mended hybrid learning algorithm for radial basis function neural networks to improve generalization capability,” *Applied Mathematical Modelling*, vol. 31, no. 7, 2007.

[171] F. Han and D. S. Huang, “Improved extreme learning machine for function approximation by encoding a priori information,” *Neurocomputing*, vol. 69, no. 16–18, pp. 2369–2373, Oct. 2006.

[172] B. M. Wilamowski, “Neural networks and fuzzy systems,” in *Mechatronic System Control, Logic, and Data Acquisition*, Kluwer Academic, 2018, pp. 12-1-12–26.

[173] J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61. pp. 85–117, 30-Apr-2015.

[174] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[175] R. A. Fisher, “The Use of Multiple Measurements in Taxonomic Problems,” *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, Sep. 1936.

[176] C. N. Li, M. Q. Shang, Y. H. Shao, Y. Xu, L. M. Liu, and Z. Wang, “Sparse L1-norm two dimensional linear discriminant analysis via the generalized elastic net regularization,” *Neurocomputing*, vol. 337, pp. 80–96, Apr. 2019.

[177] D. Tao, X. Li, X. Wu, and S. J. Maybank, “General Tensor Discriminant Analysis and Gabor Features for Gait Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.

[178] S. Le Roux *et al.*, *An Overview of Automatic Behaviour Classification for Animal-Borne Sensor Applications in South Africa*. 2017.

[179] F. Foerster, M. Smeja, and J. Fahrenberg, “Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring,” *Computers in Human Behavior*, vol. 15, no. 5, pp. 571–583, Sep. 1999.

[180] J. L. L. Hounslow *et al.*, “Assessing the effects of sampling frequency on behavioural classification of accelerometer data,” *Journal of Experimental Marine Biology and Ecology*, vol. 512, no. December 2018, pp. 22–30, Mar. 2019.

[181] G. J. Welk, J. A. Differding, R. W. Thompson, S. N. Blair, J. Dziura, and P. Hart, “The utility of the Digi-Walker step counter to assess daily physical activity patterns,” *Medicine and Science in Sports and Exercise*, vol. 32, no. 9 SUPPL., pp. S481–S488, Sep. 2000.

[182] C. H. Zheng, D. S. Huang, Z. L. Sun, M. R. Lyu, and T. M. Lok, “Nonnegative independent component analysis based on minimizing mutual information technique,” *Neurocomputing*, vol. 69, no. 7-9 SPEC. ISS., pp. 878–883, Mar. 2006.

[183] C. H. Zheng, D. S. Huang, and L. Shang, “Feature selection in independent component subspace for microarray data classification,” *Neurocomputing*, vol. 69, no. 16–18, pp.

2407–2410, Oct. 2006.

[184] C. G. Shaw *et al.*, “Ecological functionality of open spaces and physical planning in metropolitan areas: theoretical approaches and recent experiences in the Spanish context,” *Landscape and Urban Planning*, vol. 9, no. 1, 2018.

[185] N. A. D. Bourn and J. A. Thomas, “The challenge of conserving grassland insects at the margins of their range in Europe,” *Biological Conservation*, vol. 104, no. 3, 2002.

[186] JNCC, “Article 17 Habitats Directive Report 2019 (Species): JNCC - Adviser to Government on Nature Conservation,” *DEFRA*, 2019. [Online]. Available: https://jncc.gov.uk/our-work/article-17-habitats-directive-report-2019/. [Accessed: 22-Mar-2021].

[187] C. Umstatter, C. Tailleur, D. Ross, and M. J. Haskell, “Could virtual fences work without giving cows electric shocks?,” *Precision Livestock Farming 2009 - Papers Presented at the 4th European Conference on Precision Livestock Farming*, no. 2009, pp. 161–168, 2009.

[188] D. McSweeney *et al.*, “Virtual fencing without visual cues: Design, difficulties of implementation, and associated dairy cow behaviour,” *Computers and Electronics in Agriculture*, vol. 176, p. 105613, Sep. 2020.

[189] R. M. Peck, “Method and apparatus for controlling an animal,” 1973.

[190] P. K. Fay, V. T. McElligott, and K. M. Havstad, “Containment of free-ranging goats using pulsed-radio-wave-activated shock collars,” *Applied Animal Behaviour Science*, vol. 23, no. 1, pp. 165–171, May 1989.

[191] T. M. Quigley, H. R. Sanderson, A. R. Tiedemann, and M. L. McInnis, “Livestock control with electrical and audio stimulation.,” *Rangelands Archives*, vol. 12, no. 3, pp. 152–155, Nov. 1990.

[192] A. Rose, “An alternative to fences.,” *Rangelands Archives*, vol. 13, pp. 144–145, 1991.

[193] A. F. Rose, “Electronic nose-clip with solar cell,” 27-Jan-1996.

[194] “Managing stocking density in real time,” in *Proceedings of the VIIth International Rangeland Congress*, 2003, pp. 840–843.

[195] D. Anderson and B. Nolen, “Representing spatially explicit directional virtual fencing DVFTM data,” *The 24th Annual …*, no. January, pp. 1–14, 2004.

[196] C. Lee *et al.*, “Associative learning by cattle to enable effective and ethical virtual fences,” *Applied Animal Behaviour Science*, vol. 119, no. 1–2, pp. 15–22, Jun. 2009.

[197] C. Umstatter, S. Brocklehurst, D. W. Ross, and M. J. Haskell, “Can the location of cattle be managed using broadcast audio cues?,” *Applied Animal Behaviour Science*, vol. 147, no. 1–2, pp. 34–42, Nov. 2013.

[198] C. Umstatter, J. Morgan-Davies, and T. Waterhouse, “Cattle responses to a type of virtual fence,” *Rangeland Ecology and Management*, vol. 68, no. 1, pp. 100–107, Nov. 2015.

[199] D. L. M. Campbell, J. M. Lea, W. J. Farrer, S. J. Haynes, and C. Lee, “Tech-savvy

beef cattle? How heifers respond to moving virtual fence lines," *Animals*, vol. 7, no. 9, p. 72, Sep. 2017.

[200] E. I. Brunberg, K. E. Bøe, and K. M. Sørheim, "Testing a new virtual fencing system on sheep," *Acta Agriculturae Scandinavica A: Animal Sciences*, vol. 65, no. 3–4, pp. 168–175, Oct. 2015.

[201] "nofence.no – GPS basert lydgjerde for beitedyr." 29-Jan-2018.

[202] E. I. Brunberg, I. K. Bergslid, K. E. Bøe, and K. M. Sørheim, "The ability of ewes with lambs to learn a virtual fencing system," *Animal*, vol. 11, no. 11, pp. 1–6, Nov. 2017.

[203] D. L. M. Campbell, J. M. Lea, H. Keshavarzi, and C. Lee, "Virtual Fencing Is Comparable to Electric Tape Fencing for Cattle Behavior and Welfare," *Frontiers in Veterinary Science*, vol. 6, no. December, pp. 1–13, Dec. 2019.

[204] D. L. M. Campbell, S. J. Haynes, J. M. Lea, W. J. Farrer, and C. Lee, "Temporary exclusion of cattle from a Riparian zone using virtual fencing technology," *Animals*, vol. 9, no. 1, p. 5, Dec. 2019.

[205] D. L. M. Campbell, J. Ouzman, D. Mowat, J. M. Lea, C. Lee, and R. S. Llewellyn, "Virtual fencing technology excludes beef cattle from an environmentally sensitive area," *Animals*, vol. 10, no. 6, pp. 1–15, Jun. 2020.

[206] A. Muminov, D. Na, C. Lee, H. K. Kang, and H. S. Jeon, "Modern virtual fencing application: Monitoring and controlling behavior of goats using GPS collars and warning signals," *Sensors (Switzerland)*, vol. 19, no. 7, Apr. 2019.

[207] D. Marini, T. Kearton, J. Ouzman, R. Llewellyn, S. Belson, and C. Lee, "Social influence on the effectiveness of virtual fencing in sheep," *PeerJ*, vol. 8, pp. 1–16, 2020.

[208] D. Marini, F. Cowley, S. Belson, and C. Lee, "The importance of an audio cue warning in training sheep to a virtual fence and differences in learning when tested individually or in small groups," *Applied Animal Behaviour Science*, vol. 221, p. 104862, Dec. 2019.

[209] T. Kearton, D. Marini, F. Cowley, S. Belson, and C. Lee, "The effect of virtual fencing stimuli on stress responses and behavior in sheep," *Animals*, vol. 9, no. 1, 2019.

[210] P. I. Colusso, C. E. F. Clark, and S. Lomax, "Should dairy cattle be trained to a virtual fence system as individuals or in groups?," *Animals*, vol. 10, no. 10, pp. 1–19, 2020.

[211] M. Gordon, J. R. Kozloski, A. Kundu, and C. Pickover, "Specialized contextual drones for animal virtual fences and herding," US20180027772A1, 29-Jul-2018.

[212] C. Yinka-Banjo and O. Ajayi, "Sky-Farmers: Applications of Unmanned Aerial Vehicles (UAV) in Agriculture," in *Autonomous Vehicles*, 2020.

[213] C. Brose Peter W. (Harwinton, "Animal training and restraining system," no. 4898120. Feb-1990.

[214] T. M. Quigley, "Method and apparatus for controlling animals with electronic

fencing," 1995.

[215] R. E. Marsh, *Fenceless animal control system using GPS location information*. Google Patents, 1999.

[216] A. R. Tiedemann, T. M. Quigley, L. D. White, E. Al, and M. Cinnis, "Electronic (fenceless) control of livestock.," *Res. Pap. PNW-RP-510. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station*, vol. 510, no. January, Nov. 1999.

[217] R. D. Boyd, "Method and apparatus for controlling an animal," 1999.

[218] D. M. Anderson and C. S. Hale, *Animal control system using global positioning and instrumental animal conditioning*. Google Patents, 2001.

[219] D. M. Anderson, C. S. Hale, R. Libeau, and B. Nolen, "Managing Stocking Density in Real-Time," in *Proceedings of the VIIth International Rangelands Congress*, 2003, no. August, pp. 840–843.

[220] Z. Butler, P. Corke, R. Peterson, and D. Rus, "Virtual fences for controlling cows," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2004, vol. 2004, no. 5, pp. 4429–4436.

[221] Z. Butler, P. Corke, R. Peterson, and D. Rus, "Dynamic virtual fences for controlling cows," *Springer Tracts in Advanced Robotics*, vol. 21, no. April, pp. 513–522, Dec. 2006.

[222] C. Lee, K. Prayaga, M. Reed, and J. Henshall, "Methods of training cattle to avoid a location using electrical cues," *Applied Animal Behaviour Science*, vol. 108, no. 3–4, pp. 229–238, Dec. 2007.

[223] G. J. Bishop-Hurley, D. L. Swain, D. M. Anderson, P. Sikka, C. Crossman, and P. Corke, "Virtual fencing applications: Implementing and testing an automated cattle control system," *Computers and Electronics in Agriculture*, vol. 56, no. 1, pp. 14–22, Mar. 2007.

[224] D. M. Anderson, D. Rus, C. Detweiler, M. Schwager, and I. Vasilescu, "Husbandry of free-ranging cows using virtual fencing concepts," in *Proceedings of the Multifunctional Grasslands in a Changing World. XXI International Grassland Congress, VIII International Rangeland Congress*, 2008, p. 660.

[225] C. Lee, K. C. Prayaga, A. D. Fisher, and J. M. Henshall, "Behavioral aspects of electronic bull separation and mate allocation in multiple-sire mating paddocks," *Journal of Animal Science*, vol. 86, no. 7, pp. 1690–1696, Jul. 2008.

[226] M. Jouven, H. Leroy, A. Ickowicz, and P. Lapeyronie, "Can virtual fences be used to control grazing sheep?," *Rangeland Journal*, vol. 34, no. 1, pp. 111–123, Nov. 2012.

[227] D. Marini, R. Llewellyn, S. Belson, and C. Lee, "Controlling within-field sheep movement using virtual fencing," *Animals*, vol. 8, no. 3, p. 31, May 2018.

[228] D. Marini, M. D. Meuleman, S. Belson, T. B. Rodenburg, R. Llewellyn, and C. Lee, "Developing an ethically acceptable virtual fencing system for sheep," *Animals*, vol. 8, no. 3, pp. 1–9, 2018.

[229] "The Animal Welfare (Electronic Collars) (Wales) Regulations 2010."

[230] M. Festa-Bianchet, "The social system of bighorn sheep: grouping patterns, kinship and female dominance rank," *Animal Behaviour*, vol. 42, no. 1, pp. 71–82, Jul. 1991.

[231] R. F. F. Hunter and C. Milner, "The behaviour of individual, related and groups of South Country Cheviot hill sheep," *Animal Behaviour*, vol. 11, no. 4, pp. 507–513, Oct. 1963.

[232] A. B. Lawrence and D. G. M. Wood-Gush, "Home-Range Behaviour and Social Organization of Scottish Blackface Sheep," *The Journal of Applied Ecology*, vol. 25, no. 1, p. 25, Apr. 1988.

[233] J. C. Pollard and R. P. Littlejohn, "Shelter for lambing in Southern New Zealand. II. sheltering behaviour and effects on productivity," *New Zealand Journal of Agricultural Research*, vol. 42, no. 2, pp. 171–177, Jan. 1999.

[234] G. W. Arnold, S. R. Wallace, and W. A. Rea, "Associations between individuals and home-range behaviour in natural flocks of three breeds of domestic sheep," *Applied Animal Ethology*, vol. 7, no. 3, pp. 239–257, Jun. 1981.

[235] A. Stolba, G. N. Hinch, J. J. Lynch, D. B. Adams, R. K. Munro, and H. I. Davies, "Social organization of Merino sheep of different ages, sex and family structure," *Applied Animal Behaviour Science*, vol. 27, no. 4, pp. 337–349, Oct. 1990.

[236] V. R. Squires and G. T. Daws, "Leadership and dominance relationships in Merino and Border Leicester sheep," *Applied Animal Ethology*, vol. 1, no. 3, pp. 263–274, Jul. 1975.

[237] A. Boissy and B. Dumont, "Interactions between social and feeding motivations on the grazing behaviour of herbivores: Sheep more easily split into subgroups with familiar peers," *Applied Animal Behaviour Science*, vol. 79, no. 3, pp. 233–245, Nov. 2002.

[238] C. G. Winfield, G. J. Syme, and A. J. Pearson, "Effect of familiarity with each other and breed on the spatial behaviour of sheep in an open field," *Applied Animal Ethology*, vol. 7, no. 1, pp. 67–75, Jan. 1981.

[239] J. R. Franklin and G. D. Hutson, "Experiments on attracting sheep to move along a laneway. I. Olfactory stimuli," *Applied Animal Ethology*, vol. 8, no. 5, pp. 439–446, Nov. 1982.

[240] R. F. Parrott, "Physiological responses to isolation in sheep.," *Social stress in domestic animals*. Kluwer Academic Publishers, Dordrecht, pp. 212–226, 1990.

[241] L. Marino and D. Merskin, "Intelligence, complexity, and individuality in sheep," *Animal Sentience*, vol. 4, no. 25, p. 1, 2019.

[242] A. Boissy, A. Aubert, L. Désiré, L. Greiveldinger, E. Delval, and I. Veissier, "Cognitive sciences to relate ear postures to emotions in sheep," *Animal welfare*, vol. 20, Feb. 2011.

[243] K. M. Kendrick, "Social cognition in sheep: Welfare implications Commentary on Marino & Merskin on Sheep Complexity," 2019.

[244] F. Sebe, T. Aubin, A. Boue, and P. Poindron, “Mother-young vocal communication and acoustic recognition promote preferential nursing in sheep,” *Journal of Experimental Biology*, vol. 211, no. 22, pp. 3554–3562, 2008.

[245] E. Shillito Walser, P. Hague, and E. Walters, “Vocal Recognition of Recorded Lambs Voices By Ewes of Three Breeds of Sheep,” *Behaviour*, vol. 78, no. 3, pp. 260–271, Jan. 1981.

[246] A. J. Morton and L. Avanzo, “Executive decision-making in the domestic sheep,” *PLoS ONE*, vol. 6, no. 1, p. e15752, Jan. 2011.

[247] F. Knolle, S. D. McBride, J. E. Stewart, R. P. Goncalves, and A. J. Morton, “A stop-signal task for sheep: introduction and validation of a direct measure for the stop-signal reaction time,” *Animal Cognition*, vol. 20, no. 4, pp. 615–626, 2017.

[248] H. E. Heffner, “Auditory awareness,” *Applied Animal Behaviour Science*, vol. 57, no. 3, pp. 259–268, May 1998.

[249] L. H. Harbers, D. R. Ames, A. B. Davis, and M. B. Ahmed, “Digestive responses of sheep to auditory stimuli,” *Journal of Animal Science*, vol. 41, no. 2, pp. 654–658, 1975.

[250] J. . C. Talling, N. . K. Waran, C. . M. Wathes, and J. . A. Lines, “Sound avoidance by domestic pigs depends upon characteristics of the signal,” *Applied Animal Behaviour Science*, vol. 58, no. 3, pp. 255–266, Jul. 1998.

[251] J. Auty, “Carrots and sticks: Principles of animal training - By P McGreevy and R Boakes,” *Australian Veterinary Journal*, vol. 87, no. 5, p. 174, Jan. 2009.

[252] D. Sander, D. Grandjean, and K. R. Scherer, “A systems approach to appraisal mechanisms in emotion,” *Neural Networks*, vol. 18, no. 4, pp. 317–352, 2005.

[253] I. Veissier, A. Boissy, L. Désiré, and L. Greiveldinger, “Animals’ emotions: Studies in sheep using appraisal theories.,” *Animal Welfare*, vol. 18, no. 4, pp. 347–354, 2009.

[254] J. W. Kamminga, “Generic online animal activity recognition on collar tags.” Data Archiving and Networked Services (DANS), 2017.

[255] C. M. Salgado, C. Azevedo, H. Proença, and S. M. Vieira, “Missing Data,” in *Secondary Analysis of Electronic Health Records*, Cham: Springer International Publishing, 2016, pp. 143–162.

[256] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 2nd ed. McGraw-Hill School Education Group, 2001.

[257] L. R. Rabiner and B. Gold, *Theory and application of digital signal processing*. 1975.

[258] I. Guyon and A. Elisseeff, “An Introduction to Variable and Feature Selection,” *Journal of Machine Learning Research*, vol. 3, no. 3, pp. 1157–1182, Mar. 2003.

[259] C. M. Bishop, *Neural networks for pattern recognition*. Oxford university press, 1995.

[260] O. Kramer, “K-Nearest Neighbors,” in *Dimensionality Reduction with Unsupervised Nearest Neighbors*, Springer, Berlin, Heidelberg, 2013, pp. 13–23.

[261] R. Rifkin, A. Klautau, K. S. Publication, and R. E, "In Defense of One-Vs-All Classification," *J. Mach. Learn. Res.*, vol. 5, pp. 101–141, Sep. 2004.

[262] K. M. Scheibe *et al.*, "ETHOSYS (R)—new system for recording and analysis of behaviour of free-ranging domestic animals and wildlife," *Applied Animal Behaviour Science*, vol. 55, no. 3, pp. 195–211, Feb. 1998.

[263] C. Umstätter, A. Waterhouse, and J. P. Holland, "An automated sensor-based method of simple behavioural classification of sheep in extensive systems," *Computers and Electronics in Agriculture*, vol. 64, no. 1, pp. 19–26, Nov. 2008.

[264] "HyperIMU – ianovir.com." [Online]. Available: https://ianovir.com/works/mobile/hyperimu/. [Accessed: 06-Mar-2019].

[265] Xiuxin Yang, Anh Dinh, and Li Chen, "Implementation of a wearerable real-time system for physical activity recognition based on Naive Bayes classifier," in *2010 International Conference on Bioinformatics and Biomedical Technology*, 2010, pp. 101–105.

[266] J. W. Kamminga, D. V. Le, J. P. Meijers, H. Bisby, N. Meratnia, and P. J. M. Havinga, "Robust Sensor-Orientation-Independent Feature Selection for Animal Activity Recognition on Collar Tags," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–27, Mar. 2018.

[267] MBIENTLAB INC, "MetaMotionR – MbientLab," 2018. [Online]. Available: https://mbientlab.com/metamotionr/. [Accessed: 15-Dec-2019].

[268] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, 1987.

[269] L. Breiman, "out-of-bag estimation, Technical Report," pp. 1–13, 1996.

[270] L. A. González, G. J. Bishop-Hurley, R. N. Handcock, and C. Crossman, "Behavioral classification of data from collars containing motion sensors in grazing cattle," *Computers and Electronics in Agriculture*, vol. 110, pp. 91–102, Jan. 2015.

[271] B. Robert, B. J. J. White, D. G. G. Renter, and R. L. L. Larson, "Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle," *Computers and Electronics in Agriculture*, vol. 67, no. 1–2, pp. 80–84, Jun. 2009.

[272] A. Rahman, D. V. Smith, B. Little, A. B. Ingham, P. L. Greenwood, and G. J. Bishop-Hurley, "Cattle behaviour classification from collar, halter, and ear tag sensors," *Information Processing in Agriculture*, 2018.

[273] D. Smith *et al.*, "Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems," *Computers and Electronics in Agriculture*, vol. 131, pp. 40–50, Dec. 2016.

[274] R. Dutta *et al.*, "Dynamic cattle behavioural classification using supervised ensemble classifiers," *Computers and Electronics in Agriculture*, vol. 111, pp. 18–28, Feb. 2015.

[275] A. L. H. Andriamandroso *et al.*, "Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors," *Computers and Electronics in Agriculture*, vol. 139, pp. 126–

137, Jun. 2017.

[276] L. Riaboff *et al.*, "Evaluation of pre-processing methods for the prediction of cattle behaviour from accelerometer data," *Computers and Electronics in Agriculture*, vol. 165, p. 104961, Oct. 2019.

[277] J. Werner *et al.*, "Evaluation and application potential of an accelerometer-based collar device for measuring grazing behavior of dairy cows," *Animal*, vol. 13, no. 9, pp. 1–10, Feb. 2019.

[278] X. Gou, A. Tsunekawa, F. Peng, X. Zhao, Y. Li, and J. Lian, "MEthod for classifying behavior of livestock on fenced temperate Rangeland in Northern China," *Sensors (Switzerland)*, vol. 19, no. 23, pp. 1–15, 2019.

[279] D. Gutierrez-Galan *et al.*, "Embedded neural network for real-time animal behavior classification," *Neurocomputing*, vol. 272, pp. 17–26, Jan. 2018.

[280] M. A. Ladds *et al.*, "Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry," *Animal Biotelemetry*, vol. 5, no. 1, p. 8, Dec. 2017.

[281] S. Navon, A. Mizrach, A. Hetzroni, and E. D. Ungar, "Automatic recognition of jaw movements in free-ranging cattle, goats and sheep, using acoustic monitoring," *Biosystems Engineering*, vol. 114, no. 4, pp. 474–483, Apr. 2013.

[282] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1717–1724, 2014.

[283] "ImageNet." [Online]. Available: http://www.image-net.org/. [Accessed: 03-Oct-2020].

[284] S. Xia *et al.*, "Transferring Ensemble Representations Using Deep Convolutional Neural Networks for Small-Scale Image Classification," *IEEE Access*, vol. 7, pp. 168175–168186, 2019.

[285] G. Xiao, Q. Wu, H. Chen, D. Da, J. Guo, and Z. Gong, "A Deep Transfer Learning Solution for Food Material Recognition Using Electronic Scales," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2290–2300, 2020.

[286] A. Akbari and R. Jafari, "Transferring activity recognition models for new wearable sensors with deep generative domain adaptation," in *IPSN 2019 - Proceedings of the 2019 Information Processing in Sensor Networks*, 2019, pp. 85–96.

[287] S. González, J. Sedano, J. R. Villar, E. Corchado, Á. Herrero, and B. Baruque, "Features and models for human activity recognition," *Neurocomputing*, vol. 167, pp. 52–60, Nov. 2015.

[288] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, May 2018.

[289] J. Wu, "Introduction to Convolutional Neural Networks," 2017.

[290] J. Bo Yang, M. Nhut Nguyen, P. Phyo San, X. Li Li, and S. Krishnaswamy, "Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition."

[291] P. P. San, P. Kakar, X. L. Li, S. Krishnaswamy, J. B. Yang, and M. N. Nguyen, "Deep Learning for Human Activity Recognition," in *Big Data Analytics for Sensor-Network Collected Intelligence*, 2017.

[292] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *2017 International Conference on Communication and Signal Processing (ICCSP)*, 2017, pp. 0588–0592.

[293] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, p. 9, Dec. 2016.

[294] V. Lankin, "Factors of diversity of domestic behaviour in sheep," *Genetics Selection Evolution*, vol. 29, no. 1, pp. 73–92, 1997.

[295] P. A. R. Hawken, C. Fiol, and D. Blache, "Genetic differences in temperament determine whether lavender oil alleviates or exacerbates anxiety in sheep," *Physiology and Behavior*, vol. 105, no. 5, pp. 1117–1123, Mar. 2012.

[296] M. González, X. Averós, I. B. de Heredia, R. Ruiz, J. Arranz, and I. Estevez, "The effect of social buffering on fear responses in sheep (Ovis aries)," *Applied Animal Behaviour Science*, vol. 149, no. 1–4, pp. 13–20, Dec. 2013.

[297] S. D. McBride and A. J. Morton, "Visual attention and cognitive performance in sheep," *Applied Animal Behaviour Science*, vol. 206, pp. 52–58, Sep. 2018.

[298] M. Baxter, P. McCullagh, and J. A. Nelder, "Generalised Linear Models," *The Mathematical Gazette*, vol. 74, no. 469, p. 320, 1990.

[299] O. A. Quijano Xacur and J. Garrido, "Generalised linear models for aggregate claims: to Tweedie or not?," *European Actuarial Journal*, vol. 5, no. 1, pp. 181–202, 2015.

[300] G. Dunteman and M.-H. Ho., *An Introduction to Generalized Linear Models*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc., 2011.

[301] L. Fahrmeir, T. Kneib, S. Lang, and B. Marx, *Regression: Models, methods and applications*, vol. 9783642343. 2013.

[302] H. Keshavarzi, C. Lee, J. M. Lea, and D. L. M. Campbell, "Virtual Fence Responses Are Socially Facilitated in Beef Cattle," *Frontiers in Veterinary Science*, vol. 7, 2020.

[303] M. B. . Schilder and J. A. . van der Borg, "Training dogs with help of the shock collar: short and long term behavioural effects," *Applied Animal Behaviour Science*, vol. 85, no. 3–4, pp. 319–334, Mar. 2004.

[304] P. A. Dufour, "Effects of noise on wildlife and other animals, review of research since 1971.," 1980.

[305] R. Baumont, S. Prache, M. Meuret, and P. Morand-Fehr, "How forage characteristics influence behaviour and intake in small ruminants: a review," *Livestock Production*

Science, vol. 64, no. 1, pp. 15–28, May 2000.

[306] F. C. İskenderoğlu, M. K. Baltacioğlu, M. H. Demir, A. Baldinelli, L. Barelli, and G. Bidini, "Comparison of support vector regression and random forest algorithms for estimating the SOFC output voltage by considering hydrogen flow rates," *International Journal of Hydrogen Energy*, Aug. 2020.