Bellfield, RAA, Ortega-Martorell, S, Lip, GYH, Oxborough, D and Olier, I

Impact of ECG data format on the performance of machine learning models for the prediction of myocardial infarction

Journal of Electrocardiology 84 (2024) 17–26

Contents lists available at ScienceDirect

Journal of Electrocardiology

journal homepage: www.jecgonline.com

Impact of ECG data format on the performance of machine learning models for the prediction of myocardial infarction

Ryan A.A. Bellfield, Msc [a,b], Sandra Ortega-Martorell, PhD [a,b], Gregory Y.H. Lip, MD [b,c], David Oxborough, PhD [b,d], Ivan Olier, PhD [a,b,*]

[a] Data Science Research Centre, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[b] Liverpool Centre for Cardiovascular Science at University of Liverpool, Liverpool John Moores University and Liverpool Heart & Chest Hospital, Liverpool, UK
[c] Department of Clinical Medicine, Aalborg University, Denmark
[d] School of Sport and Exercise Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

ARTICLE INFO

Keywords:
Machine learning
Artificial intelligence
Myocardial infarction
Electrocardiogram

ABSTRACT

Background We aim to determine which electrocardiogram (ECG) data format is optimal for ML modelling, in the context of myocardial infarction prediction. We will also address the auxiliary objective of evaluating the viability of using digitised ECG signals for ML modelling. Methods Two ECG arrangements displaying 10s and 2.5 s of data for each lead were used. For each arrangement, conservative and speculative data cohorts were generated from the PTB-XL dataset. All ECGs were represented in three different data formats: Signal ECGs, Image ECGs, and Extracted Signal ECGs, with 8358 and 11,621 ECGs in the conservative and speculative cohorts, respectively. ML models were trained using the three data formats in both data cohorts. Results For ECGs that contained 10s of data, Signal and Extracted Signal ECGs were optimal and statistically similar, with AUCs [95% CI] of 0.971 [0.961, 0.981] and 0.974 [0.965, 0.984], respectively, for the conservative cohort; and 0.931 [0.918, 0.945] and 0.919 [0.903, 0.934], respectively, for the speculative cohort. For ECGs that contained 2.5 s of data, the Image ECG format was optimal, with AUCs of 0.960 [0.948, 0.973] and 0.903 [0.886, 0.920], for the conservative and speculative cohorts, respectively. Conclusion When available, the Signal ECG data should be preferred for ML modelling. If not, the optimal format depends on the data arrangement within the ECG: If the Image ECG contains 10s of data for each lead, the Extracted Signal ECG is optimal, however, if it only uses 2.5 s, then using the Image ECG data is optimal for ML performance.

Introduction

The electrocardiogram (ECG) is a simple, non-invasive test used globally to detect numerous cardiovascular issues. ECGs measure the electrical activity of the heart using electrodes, known as leads, attached to different parts of the body. Older ECG machines directly record the electrical signals from each lead onto graph paper, which are then stored as physical copies and then manually scanned so they can be viewed electronically. With newer machines, signals can be directly recorded and stored electronically as portable document format (PDF) files [1]. In some instances, the electrical signals from each lead are recorded and stored digitally (as a signal, not as a PDF) [2], yet this is rare as the machines that provide the raw digital signals are more expensive and usually research-based.

There is an ever-increasing amount of machine learning (ML) research being completed whereby ECG data is used to develop models to address a variety of cardiovascular conditions [3–6]. Models have been developed using both digital ECG signals [7] and ECGs in an image format [8]. Image ECGs are records that were either physically recorded and scanned or were recorded electronically and stored in a PDF format. While it has been reported that analysing ECGs in a digital signal format is preferable [9], it is often the case that the choice of format is dictated by the data available. There are several studies [9–13] that focus on solving this problem, providing methods of digitising image ECGs by extracting the signals from the image and storing them as a multivariate time series. These extracted signals show promise for ML model development [9] but have not yet seen widespread adoption. Nevertheless, and to the best of our knowledge, there has been no research thus far that confirms a tangible benefit to developing ML models using one ECG data format over another.

* Corresponding author at: Data Science Research Centre, Liverpool John Moores University, James Parsons Building, 3 Byrom Street, Liverpool L3 3AF, UK.
E-mail address: I.A.OlierCaparroso@ljmu.ac.uk (I. Olier).

https://doi.org/10.1016/j.jelectrocard.2024.03.005

Available online 7 March 2024

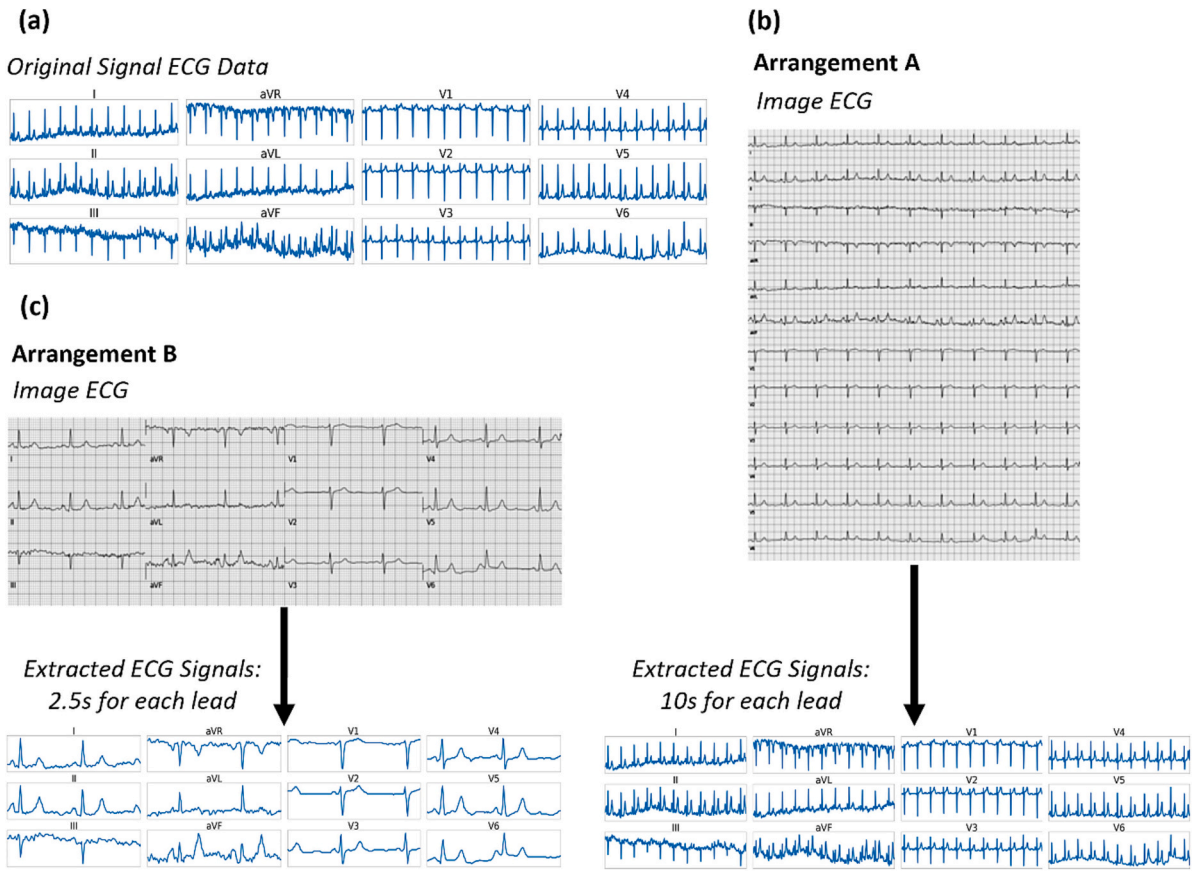

Fig. 1. Displays the same ECG in each of the three different data formats being evaluated, for both Image ECG arrangements. (a) Signal ECG data format; (b) Image ECG data and Extracted ECG Signal data format for arrangement A; (c) Image ECG data and Extracted ECG Signal data format for arrangement B.

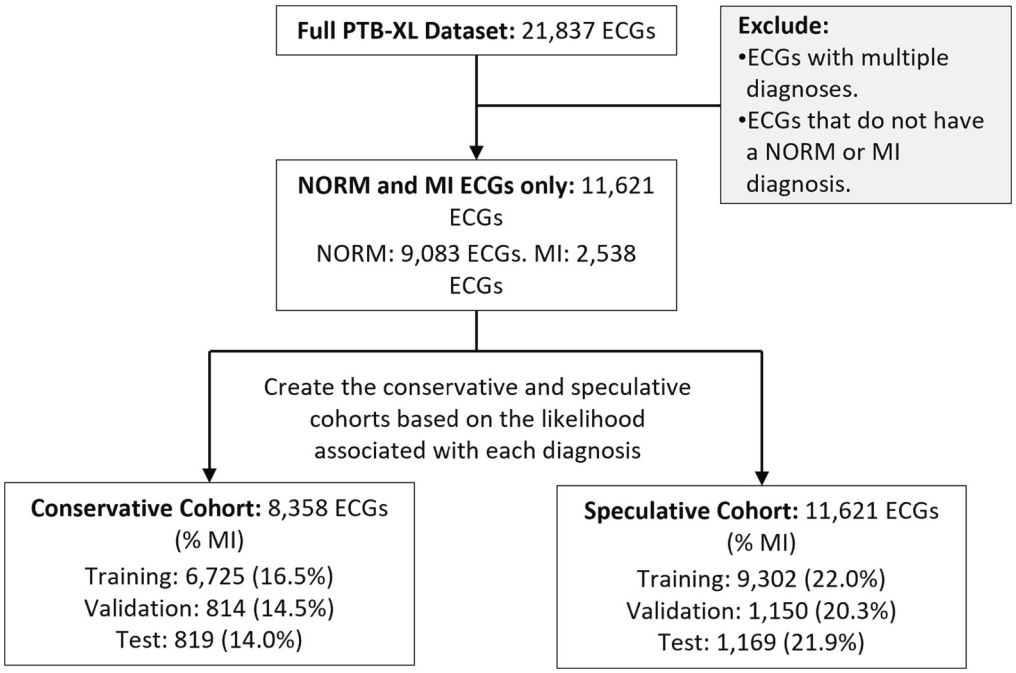

Fig. 2. Flowchart showing how the criteria was applied to the full PTB-XL dataset to generate both data cohorts. Values in brackets indicate the prevalence of MI within each of the training, validation, and test datasets respectively.

Table 1
Displays the modelling results using the arrangement A data. The AUCs of the best models trained using each ECG data format for both the conservative and speculative cohort are presented.

Arrangement A ECGs						
	Conservative Cohort (AUC [95% CI])			Speculative Cohort (AUC [95% CI])		
Data Format	Signal ECG Data	Image ECG Data	Extracted Signal ECG Data	Signal ECG Data	Image ECG Data	Extracted Signal ECG Data
Training	0.999 [0.998, 0.999]	0.998 [0.998, 0.999]	0.995 [0.993, 0.996]	0.949 [0.945, 0.953]	0.918 [0.913, 0.924]	0.954 [0.95, 0.958]
Validation	0.962 [0.951, 0.974]	0.944 [0.929, 0.959]	0.97 [0.96, 0.981]	0.921 [0.906, 0.937]	0.893 [0.874, 0.911]	0.911 [0.895, 0.928]
Testing	0.971 [0.961, 0.981]	0.952 [0.938, 0.966]	0.974 [0.965, 0.984]	0.931 [0.918, 0.945]	0.89 [0.871, 0.908]	0.919 [0.903, 0.934]

Table 2
Displays the modelling results using the arrangement B data. The AUCs of the best models trained using each ECG data format for both the conservative and speculative cohort are presented. ** The Signal ECG data used matches the same 2.5 s of signal used for the Extracted Signal ECG data to ensure a relevant comparison.

Arrangement B ECGs						
	Conservative Cohort (AUC [95% CI])			Speculative Cohort (AUC [95% CI])		
Data Format	Signal ECG Data**	Image ECG Data	Extracted Signal ECG Data	Signal ECG Data	Image ECG Data	Extracted Signal ECG Data
Training	0.985 [0.983, 0.988]	0.978 [0.975, 0.981]	0.979 [0.976, 0.982]	0.966 [0.963, 0.969]	0.963 [0.96, 0.967]	0.951 [0.947, 0.955]
Validation	0.946 [0.931, 0.961]	0.933 [0.916, 0.951]	0.949 [0.933, 0.963]	0.907 [0.89, 0.924]	0.900 [0.882, 0.918]	0.903 [0.886, 0.910]
Testing	0.938 [0.921, 0.954]	0.960 [0.948, 0.973]	0.937 [0.921, 0.953]	0.886 [0.867, 0.905]	0.903 [0.886, 0.92]	0.864 0.843, 0.884]

To that end, we collated a large dataset of ECGs represented in three different data formats: original digital ECG signal recordings (Signal ECGs); the ECGs in an image format (Image ECGs); and ECG signals extracted by digitising the Image ECGs (Extracted Signal). The main objective of this study therefore is to quantify the affect ECG data format choice has on ML model performance in the context of myocardial infarction (MI) prediction, thereby identifying the optimal format. In addressing the main objective, we also address an auxiliary objective by validating the feasibility of using Extracted Signal ECGs for ML cardiac outcome modelling.

Materials and methods

Data source

We selected the PTB-XL database [14] for use in this analysis for several reasons. First, it is, to date, the largest open-source ECG dataset available, hosted by PhysioNet [15]. The dataset consists of the digital signals for 21,837 ECG records from 18,885 patients, with most records being assigned at least one of five main diagnoses (or "superclasses"): Normal, Myocardial Infarction, ST/T wave Change, Conduction Disturbance, and Hypertension. Each diagnostic superclass was assigned based on the written notes in the original ECG report. Each class received a likelihood score between 0 and 100, which represented the cardiologist's certainty of the diagnosis. The signal data provided within the PTB-XL database represent a 10 s ECG recording sampled at two frequencies, 100 Hz and 500 Hz. For this analysis, the data sampled at 100 Hz was used as it falls within the common frequency range used by modern ECG machines [16].

Data extraction

Since 5 diagnostic superclasses are represented in these ECGs, this dataset has led to diverse study designs [17,18]. For this study, we designed a two-class classification task using only the two largest classes within the dataset: normal ECG (NORM) and MI. In this way, we could limit any source of variability that would interfere with evaluating the impact the data format has on model performance.

Another consideration was given to the diagnosis likelihoods. To ensure a thorough evaluation two subsets of the data were created: The 1st subset, referred to henceforth as the "conservative cohort", only included ECGs where the likelihood score equalled 100; the 2nd subset, referred to henceforth as the "speculative cohort", included all ECGs regardless of the likelihood score. Developing models on both data subsets allowed us to add a controlled amount of variability into our testing to provide a richer understanding of the optimal data format.

To provide transparency within our analysis, allow for straightforward external validation and comparable inter-model results, we followed the suggested data splitting as defined in the original PTB-XL study [14], which recommends using ECGs assigned to folds 1–8 for model training, fold 9 for validation and fold 10 for testing.

Signal ECG data preparation

As previously mentioned, each Signal ECG recording within the PTB-XL database contains 12 signals, ten seconds in length, with each signal representing one of the 12 standard sets of leads used in ECG recordings (I, II, III, aVL, aVF, aVR, V1, V2, V3, V4, V5, V6). With the data being sampled at 100 Hz, each ten-second recording consists of 1000 samples, giving a data dimension for each Signal ECG sample of 12×1000.

Image ECG data preparation

The Image ECG data was generated manually using the Signal ECG recordings by leveraging the "wfdb" and "ecg-plot" python packages. We formed the Image ECGs so they would resemble genuine, commonly found ECG recordings hence providing a realistic understanding of the performance ML models can achieve if deployed in a real-world application. To that end, we generated two sets of Image ECGs, with each set having different lead arrangements and displaying a different amount of the original signal. The first set of Image ECGs (arrangement A) arranged the 12 leads in a single column, with the full 10 s of data used for each lead (as shown in Fig. 1b). One Image ECG was created each set of the 12 lead ECG signals with dimension 1200×1000. The second set (arrangement B) had the 12 leads arranged in a 3×4 grid, with 2.5 s of the full 10 s available used for each lead (see Fig. 1c). The 2.5 s used for each lead was also staggered based on the column in which the lead was present such that:

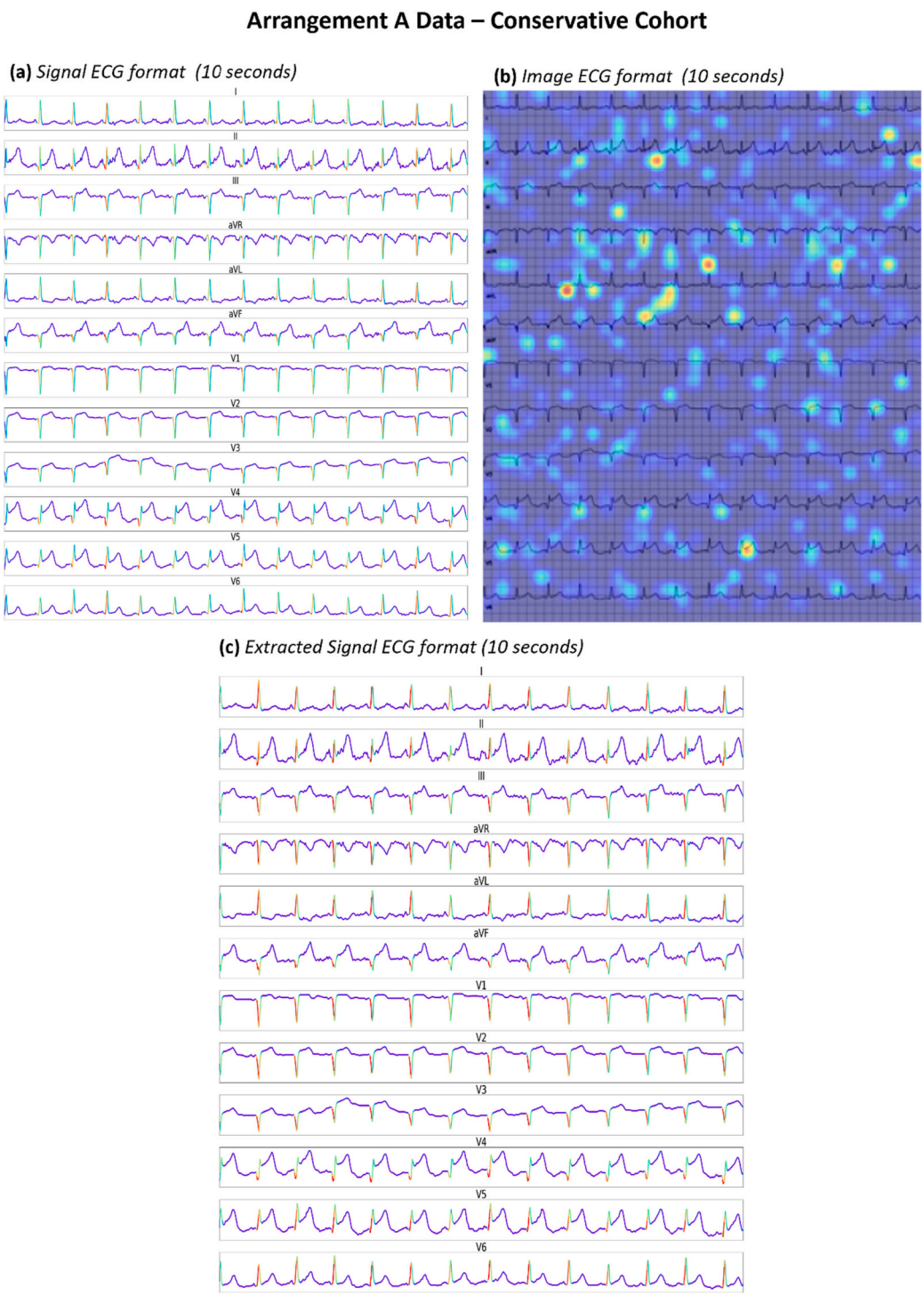

Fig. 3. Displays the HiResCAM activation maps generated using the three best models for each of the data formats in the conservative cohort for arrangement A data. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

- 0 s - 2.5 s used for leads I, II and III
- 2.5 s - 5 s used for leads aVR, aVL and aVF
- 5 s - 7.5 s used for leads V1, V2 and V3
- 7.5 s - 10s used for leads V4, V5 and V6

One image ECG was created from each set of the 12 lead ECG signals with dimensions 300 × 1000. However, to ensure computational tractability for the proposed experiments, the images for both formats were reduced prior to model development. The Image ECGs were therefore analysed using a dimension of 165 × 500 and 330 × 275 for arrangement A and arrangement B respectively.

Extracted signal ECG data preparation

We followed Fortune et al. [11] ECG digitisation algorithm to extract the signals from the Image ECGs. They created an open-source application that allows a user to import an Image ECG, manually draw borders around each lead, then extract the ECG signal contained within each border and export the signals to a CSV file. The manual nature of this application meant it was not feasible for use in our study due to the

Fig. 4. Displays the HiResCAM activation maps generated using the three best models for each of the data formats in the conservative cohort for arrangement B data. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

volume of ECGs, as it would take too long and be prone to potential human errors. To overcome this, we extended their approach and implemented a semi-automatic signal extraction algorithm (S1 within the supplementary information). Although our approach still requires the border positions to be manually set, this is performed only once as the Image ECGs are identical in layout and dimensionality. In addition, we added functionality that removed lead labels from the images, as they interfered with the extraction algorithm. The 10 s signal contained within the arrangement A Image ECGs was then extracted into a 12 × 1000 array, with the 2.5 s signal contained within the arrangement B Image ECGs being extracted into a 12 × 250 array, with both then being exported as a CSV file.

Machine learning modelling

Convolutional neural networks (CNN) were used for this analysis due to their ability to be applied to contextual datasets of varying forms. Specifically, we utilised different structures to allow for both the 2-D image and 1-D signal inputs. Other ML methodologies (or even more complex versions of the selected methodology) could have been better suited for analysing the different formats being compared. However, the use of different methodologies would introduce a source of variation to the experiments, which would have detracted away from the direct comparison of the data formats, which was the aim of this study. Hyperparameter tuning was used to develop the models that will be applied to the different data formats. For details regarding the software and hardware used for the model development and tuning, please refer to S2 (i) within the supplementary information. For completeness, each data format passed through three rounds of hyperparameter tuning, with each round using a different hyperparameter search space method: random search [19]; hyperband [20] and Bayesian optimisation [21]. For further details, please refer to S2 (ii) within the supplementary information.

Two hyperparameter search spaces were defined: one to develop 2-D CNN models to be applied to the Image ECGs; One to develop 1-D CNN models to be applied to both the Signal ECG and Extracted Signal ECG data. For full details of the search spaces used to develop both the 2-D and 1-D CNNs, please refer to S2 (iii) within the supplementary information.

After we identified the best model for each data format, to help interpret the models generated we utilised the high-resolution class activation mapping (HiResCAM). HiResCAM is a technique used for visualising what areas of an input are considered most important by a CNN model when making a prediction [22]. It also addresses issues with other commonly used techniques such as Grad-CAM [23] whereby they could lead to spurious correlations. For further details on this method, please refer to S2 (iv) within the supplementary information. All models were evaluated using the results from the testing data split, with the performance metric used throughout being AUC.

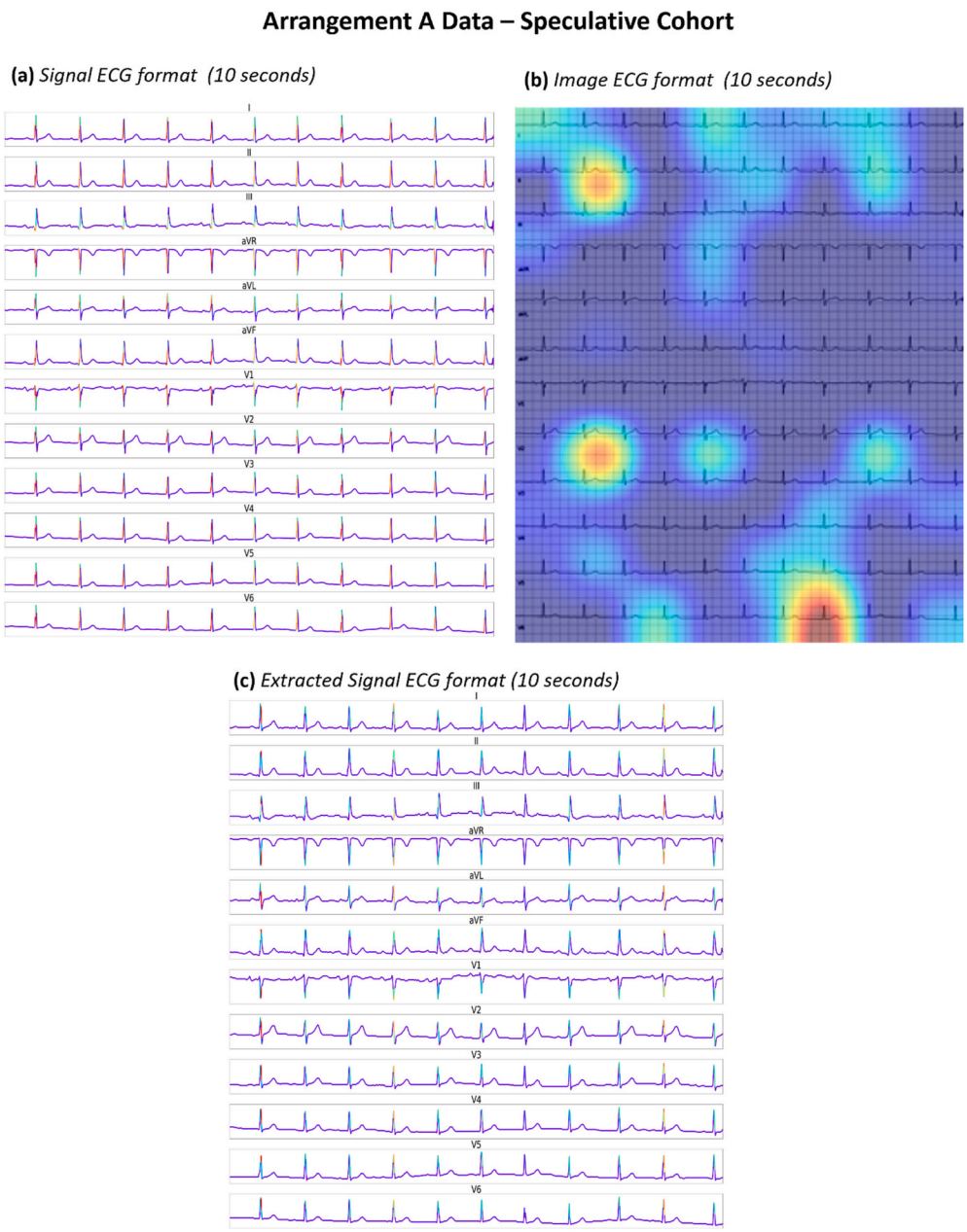

Fig. 5. Displays the HiResCAM activation maps generated using the three best models for each of the data formats in the speculative cohort for arrangement A data. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

Results

Dataset generation

Applying the criterion set out in the Data Extraction section to the 21,837 ECGs, we were left with a total of 11,621 eligible ECGs: 9083 of which were NORM and 2538 were MI (21.7% prevalence). The conservative and speculative cohort subsets were then created from the eligible ECGs. The conservative cohort contained a total of 8358 ECGs: 7017 of which were NORM and 1341 are MI (16% prevalence), whilst the speculative cohort contained the full 11,621 ECGs. The ECGs were then grouped according to the fold they were assigned in the original PTB-XL dataset to generate the training, validation, and testing data splits. The data pre-processing steps outlined for the Signal ECGs, Image ECG and Extracted Signal ECGs data formats were then applied to generate the final datasets that would be used for the ML model developments. The full process is outlined in Fig. 2.

Model comparisons

Following the aforementioned framework for model development, we trained and tested models for both arrangement A and B data. The results of the best-performing models for each data format within each cohort are listed in Tables 1 and 2 for arrangement A and B data respectively. Starting with arrangement A, the Signal ECG and Extracted Signal ECG formats performed the best, with both also significantly outperformed the Image ECG format for both the conservative and speculative cohort tests. Additionally, the Signal ECG and Extracted ECG signal formats did not perform significantly different from each other. Moving to the arrangement B data, here the Image ECG format performed the best, significantly outperforming the Signal and Extracted Signal ECG formats, in both the conservative and speculative cohort tests. Like with the arrangement A data however, the Signal ECG and Extracted ECG signals did not perform significantly differently from one another. Across both tests with the arrangement A and B data, we see a

Fig. 6. Displays the HiResCAM activation maps generated using the three best models for each of the data formats in the speculative cohort for arrangement B data. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

drop in performance across all the data formats.

Class activation maps

We applied HiResCAM to the outputs of the best performing models for each data format. This provided a visual heat map that we overlaid onto the inputted data to analyse the areas important to the decision making of the model. Figs. 3 and 4 display the same ECG of a patient with MI, represented in the three different formats for both data arrangements, with their respective activation maps superimposed on top, for the conservative cohort data. Figs. 5 and 6 display the same ECG of a patient deemed normal, represented in the three different for both data arrangements, with the activation maps superimposed on top, this time for the speculative cohort data. For Figs. 3–6, the red sections of the activation map represent regions of the input the model deemed most important, with the blue sections representing areas deemed less relevant to determining the outcome.

Discussion

The results highlight the very real presence of ML performance differences between the three different data formats, as well as between how the data is represented within each format. As expected, Signal ECG should be the preferred choice for ML modelling, provided that such format is available. If this is not the case, the decision would depend on the particular needs. Starting with arrangement A (10-s ECGs), the Extracted Signal ECG format seems to offer better performance results when using either the conservative or speculative cohort tests. Remarkably, the performance results of the Extracted Signal ECG were comparable to those of the Signal ECG format. This provides key quantifiable evidence that Extracted Signal ECGs are not only feasible for ML modelling, but in some situations (such as with arrangement A data), that it could be the preferred choice.

However, a drop in performance was observed when arrangement B (2.5-s ECGs) was used either with conservative or speculative cohort tests. Performance drop was particularly significant if the Extracted Signal ECG format was used, which was outperformed by Image ECG models when this arrangement was used. Interestingly, a similar drop in performance was observed when the original Signal ECG data was modelled. This suggests that the drop in performance seen by the Extracted Signal ECG format results from the shorter ECG duration, and not because in inherent issue with the ECG digitisation.

Overall, models developed using the conservative cohort subset performed better than the models developed using the speculative cohort subset. This an expected result; the added uncertainty brought about by using noisier data in the speculative cohort was naturally harder to model than in the conservative cohort. Therefore, based on model performance alone, then the Extracted Signal ECG format would be the preferred choice, should the Image ECG data contain 10 s of data per lead. Should the Image ECG contain less data per lead, then this becomes the preferred format. However, there may arise conditions brought on by external factors whereby choosing a format with slightly

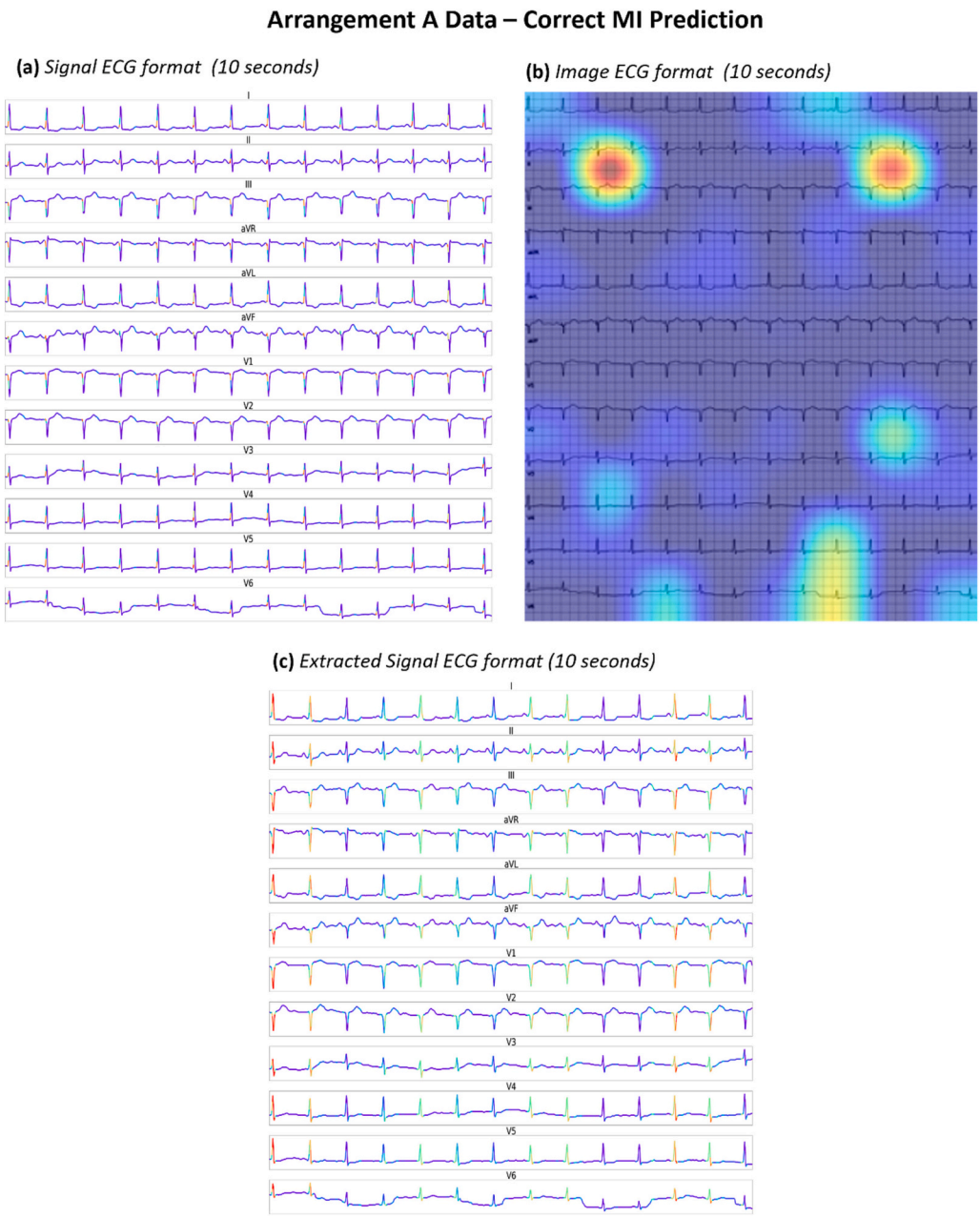

Fig. 7. Displays the HiResCAM activation maps generated for an ECG that represents MI, whereby the best models for each format all correctly predicted MI. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

lower performance could yield more meaningful results.

One such example would be the interpretability of model output using techniques such as HiResCAM activation maps. Using the activation maps described in Figs. 3 to 6 initially, the Image ECG data the maps highlight general regions of the signal that the model found important, making it difficult to precisely ascertain the key information. The maps for both the Signal ECG and Extracted Signal ECG data are much clearer, providing specific time points of interest on each digital signal that their respective models deemed important. For example, in Fig. 3 the regions that have been deemed important relate primarily to the onset / upslope of the QRS signal alongside the known impact on the ST segment. The early part of the QRS signal is not routinely evaluated using conventional interpretation algorithms for MI. This also demonstrates the unique ability of the technique to identify new and novel patterns. Using a further example, Figs. 7 and 8 contains the activation maps for a participant where the correct prediction was made for every data format to demonstrate the difference in interpretability. All three maps show that similar areas of the signal are considered by both the Signal ECG, Extracted Signal ECG, and Image ECG data formats. However, the maps associated with the Image ECGs show that the model has considered areas in between two signals as very important, implying that the model learning is far less intuitive. In contrast, the maps for the Signal ECGs and Extracted Signal ECGs show the models focus primarily on the peaks, and more specifically, what sections of the peaks were more important than others. This allows the user to understand clearly what led to the prediction and if the correct point of the signal is being considered.

Further examples of situations that would favour the Extracted ECG Signals would be if key features from the ECG, such as QRS duration and P wave duration, need to be extracted. Extracting these features from the images is difficult, however previous studies have seen success by first digitising the image and then extracting the features [24]. There are also

Fig. 8. Displays the HiResCAM activation maps generated for an ECG that represents MI, whereby the best models for each format all correctly predicted MI. (a) Signal ECG data format; (b) Image ECG data format; (c) Extracted ECG data format.

open-source Python packages such as "neurokit2" [25] that can automatically detect key points on a digital ECG signal. Also, as briefly mentioned in the results section, the dimension of the Image ECGs had to be manually reduced to allow for ML modelling to be carried out. This highlights an inherent benefit of the Extracted Signal ECGs again over the Image ECGs in that they are computationally more efficient to process and model when compared to the images when working with large datasets, as was the case in this study.

One instance which would favour the Image ECG format over the Extracted Signal ECG format would be in the event whereby the recorded ECG was noisy. The ECG digitisation algorithm used in this study works by removing background noise within the image, isolating the ECG recording within a desired window which will be converted to a digital signal. For this analysis, as the Image ECGs were generated manually using the Signal ECG data, we intentionally cultivated a perfect scenario whereby we had fully clean Image ECGs. An Image ECG could be considered noisy if there is a significant overlap between the recordings of different leads, or artifacts on the image such as a coffee stain (should the Image ECG be a scanned version of a physical copy). The presence of these could lead to the digitisation algorithm failing to extract the signal and therefore excluding that ECG from any further analysis. This favours the Image ECG format as it allows for the ECG to be analysed regardless of the state of the original image, reducing the chance data is removed.

The application of AI and ML in the utilisation of the 12-lead ECG has evolved in tandem with technological developments. Recent studies have demonstrated the role of AI on the ECG being able to predict disease that is not achievable through routine individual scrutiny [26] and hence there is a significant immediate and long-term clinical impact. The 12-lead ECG is the fundamental and primary cardiac investigation for patients presenting with symptoms and hence the ability of AI / ML technology to provide insight into structural and functional cardiac adaptation will improve patient diagnosis, management and reduce downstream costs secondary to a reduction in unnecessary investigations. It is apparent from our study that to build up large datasets with sufficient accuracy the signal format is important and should be considered when developing ML studies going forward. That aside, as hospital environments continue a transition to a full digital set-up the likelihood of securing widespread Signal ECGs is unlikely. Our data highlights the importance of digitally storing pdfs and refining methodology to better handle these image files and subsequently allowing more robust predictive models.

Conclusion

The analysis conducted in this study provides an evaluation of three different data formats that can feasibly be used to analyse ECGs. Signal ECGs, Image ECGs and Extracted Signal ECGs were all compared using two different ECG arrangements and two data subsets: the first contained best-case scenario data with a clear separation between the classes; the second had more noise and less confident diagnoses. The results of the analysis showed that should the Signal ECG data be available, then this should always be used for any ML modelling. In the absence of data in this format, we showed that the optimal data regarding model performance is dependent on the way the data is arranged within the ECG: If the Image ECG contains 10 s of data for each lead, the digitising the signal and using the Extracted Signal ECGs is optimal; If the Image ECG contains 2.5 s of data per lead, then using the Image ECG data is optimal for ML performance. As highlighted in the discussion, the decision may become situational with certain criteria, such as noisy Image ECGs, meaning one is more effective than the other. What these results also speak to is the viability of extracting digital ECG signals from image ECGs and using those for ML model development. However, further analyses will be needed to investigate how factors such as changes in image resolution and in extraction algorithms influence model performance.

Author statement

On behalf of the co-authors, I declare that this manuscript is original, has not been published.

before and is not currently considered for publication elsewhere.

Funding

This work has been funded by the Liverpool John Moores University Vice Chancellors PhD Scholarship Fund.

Disclosures

The authors declare that there is no conflict of interest.

CRediT authorship contribution statement

Ryan A.A. Bellfield: Data curation, Formal analysis, Writing – original draft, Methodology. **Sandra Ortega-Martorell:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Gregory Y.H. Lip:** Writing – review & editing. **David Oxborough:** Writing – review & editing. **Ivan Olier:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Acknowledgements

RAAB, IO and SOM conceptualised and led the development of the analysis presented within the manuscript. RAAB drafted the article with all authors critically reviewing and refining the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jelectrocard.2024.03.005.

References

[1] Randazzo V, Puleo E, Paviglianiti A, et al. Development and validation of an algorithm for the digitization of ECG paper images. Sensors 2022;22:7138. https://doi.org/10.3390/s22197138.
[2] Bond RR, Finlay DD, Nugent CD, et al. A review of ECG storage formats. Int J Med Inform 2011;80:681–97. https://doi.org/10.1016/j.ijmedinf.2011.06.008.
[3] Somani S, Russak AJ, Richter F, et al. Deep learning and the electrocardiogram: review of the current state-of-the-art. Europace 2021;23:1179–91. https://doi.org/10.1093/europace/euaa377.
[4] Ebrahimi Z, Loni M, Daneshtalab M, et al. A review on deep learning methods for ECG arrhythmia classification. Expert Syst with Appl X 2020;7:100033. https://doi.org/10.1016/j.eswax.2020.100033.
[5] Bellfield RAA, Ortega-Martorell S, Lip GYH, et al. The Athlete's heart and machine learning: a review of current implementations and gaps for future research. J Cardiovasc Dev Dis 2022;9:382. https://doi.org/10.3390/jcdd9110382.
[6] Olier I, Ortega-Martorell S, Pieroni M, et al. How machine learning is impacting research in atrial fibrillation: implications for risk prediction and future management. Cardiovasc Res 2021;117:1700–17. https://doi.org/10.1093/cvr/cvab169.
[7] Kwon J, Kim KH, Medina-Inojosa J, et al. Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography. J Hear Lung Transplant 2020;39:805–14. https://doi.org/10.1016/j.healun.2020.04.009.
[8] Makimoto H, Höckmann M, Lin T, et al. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction. Sci Rep 2020;10:8445. https://doi.org/10.1038/s41598-020-65105-x.
[9] Mishra S, Khatwani G, Patil R, et al. ECG paper record digitization and diagnosis using deep learning. J Med Biol Eng 2021;41:422–32. https://doi.org/10.1007/s40846-021-00632-0.
[10] Swamy P, Jayaraman S, Girish Chandra M. An improved method for digital time series signal generation from scanned ECG records. In: ICBBT 2010–2010 Int Conf Bioinforma Biomed Technol; 2010. p. 400–3. https://doi.org/10.1109/ICBBT.2010.5478930.
[11] Fortune JD, Coppa NE, Haq KT, et al. Digitizing ECG image: a new method and open-source software code. Comput Methods Programs Biomed 2022:221. https://doi.org/10.1016/j.cmpb.2022.106890.
[12] Li Y, Qu Q, Wang M, et al. Deep learning for digitizing highly noisy paper-based ECG records. Comput Biol Med 2020;127:104077. https://doi.org/10.1016/j.compbiomed.2020.104077.
[13] Baydoun M, Safatly L, Hassan OKA, et al. High precision digitization of paper-based ECG records: a step toward machine learning. IEEE J Transl Eng Heal Med 2019;7:1–9. https://doi.org/10.1109/JTEHM.2019.2949784.
[14] Wagner P, Strodthoff N, Bousseljot R-D, et al. PTB-XL, a large publicly available electrocardiography dataset. Sci Data 2020;7:154. https://doi.org/10.1038/s41597-020-0495-6.
[15] Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000:101. https://doi.org/10.1161/01.cir.101.23.e215.
[16] Tereshchenko LG, Josephson ME. Frequency content and characteristics of ventricular conduction. J Electrocardiol 2015;48:933–7. https://doi.org/10.1016/j.jelectrocard.2015.08.034.
[17] Strodthoff N, Wagner P, Schaeffter T, et al. Deep learning for ECG analysis: benchmarks and insights from PTB-XL. IEEE J Biomed Heal Informatics 2021;25:1519–28. https://doi.org/10.1109/JBHI.2020.3022989.
[18] Śmigiel S, Pałczyński K, Ledziński D. ECG signal classification using deep learning techniques based on the PTB-XL dataset. Entropy 2021;23:1121. https://doi.org/10.3390/e23091121.
[19] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res 2012;13:281–305.
[20] Li L, Jamieson K, DeSalvo G, et al. Hyperband: a novel bandit-based approach to Hyperparameter optimization. J Mach Learn Res 2016;18:1–52. https://doi.org/10.48550/ariXiv.1603.06560.
[21] Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. Adv Neural Inf Process Syst 2012;4:2951–9. https://doi.org/10.48550/arXiv.1206.2944.
[22] Draelos RL, Carin L. Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks. arXiv Prepr 2020:1–20. http://arxiv.org/abs/2011.08891.
[23] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vis 2020;128:336–59. https://doi.org/10.1007/s11263-019-01228-7.
[24] Wang S, Zhang S, Li Z, et al. Automatic digital ECG signal extraction and normal QRS recognition from real scene ECG images. Comput Methods Programs Biomed 2020:187. https://doi.org/10.1016/j.cmpb.2019.105254.
[25] Makowski D, Pham T, Lau ZJ, et al. NeuroKit2: a Python toolbox for neurophysiological signal processing. Behav Res Methods 2021;53:1689–96. https://doi.org/10.3758/s13428-020-01516-y.
[26] Ito S, Cohen-Shelly M, Attia ZI, et al. Correlation between artificial intelligence-enabled electrocardiogram and echocardiographic features in aortic stenosis. Eur Hear J - Digit Heal 2023:1–6. https://doi.org/10.1093/ehjdh/ztad009.