



Integrating Manual ECG Feature Extraction with Ensemble Learning for Myocardial Infarction Diagnosis

Wen-Cheng Liu¹, Yu-Lan Liu¹, Da-Wei Chang¹, Chiao-Chin Lee¹, Chiao-Hsiang Chang¹, Dung-Jang Tsai^{2,3}

¹Division of Cardiology, Department of Internal Medicine, Tri-Service General Hospital, National Defense Medical University, ²Medical Technology Education Center, School of Medicine, College of Medicine, National Defense Medical University, ³Department of Artificial Intelligence and Internet of Things, Tri-Service General Hospital, National Defense Medical University, Taipei, Taiwan

Background: Myocardial infarction (MI) is a leading cause of mortality worldwide. Electrocardiograms (ECGs) are primary diagnostic tools but face limitations with high-dimensional, imbalanced data. Combining manual ECG waveform feature extraction with machine learning may improve diagnostic accuracy. **Aim:** The aim of this study is to improve the accuracy and efficiency of MI diagnosis by integrating manually extracted ECG waveform features with ensemble machine learning models and evaluating their performance on a comprehensive, large-scale ECG dataset. **Methods:** We utilized the PTB-XL dataset, which consists of 15,014 12-lead ECG recordings. Manual feature extraction was conducted on 94 waveform-specific variables, including the durations and areas of P-waves, QRS complexes, and T-waves. The preprocessing steps included signal filtering, detrending, and data augmentation to mitigate noise and address class imbalance issues. Three ensemble learning algorithms – XGBoost, Random Forest, and AdaBoost – were trained and validated using a 70%, 20%, and 10% split for training, validation, and testing, respectively. **Results:** Compared with the other models, the XGBoost model demonstrated superior performance, with an accuracy of 86.12%, a sensitivity of 84.81%, a specificity of 87.43%, and an area under the receiver operating characteristic curve (AUROC) of 93.59%. Subcategory analyses indicated strong diagnostic performance in detecting anterior MI (AUROC: 94.09%), inferior MI (AUROC: 90.14%), and lateral MI (AUROC: 89.76%). The ST interval emerged as the most influential feature for accurate diagnosis. **Conclusion:** Manual ECG feature extraction combined with ensemble learning improves MI diagnosis. Future work should explore automated deep learning and additional physiological features.

Key words: Electrocardiogram, ensemble learning, myocardial infarction

INTRODUCTION

Cardiovascular diseases have become the leading cause of death globally. According to the World Health Organization (WHO), approximately 17.9 million people die from cardiovascular diseases annually, accounting for 32% of all deaths worldwide. It is estimated that by 2030, the number of deaths due to cardiovascular diseases will increase to 23 million worldwide.¹ In Taiwan, according to the 2022 statistics from the Ministry of Health and Family Welfare, heart disease ranks second among the top 10 causes of death and causes 23,668 deaths, with an annual increase rate of 8.3%, second only to cancer.² Myocardial infarction (MI), one of the most common cardiovascular diseases, typically presents

with symptoms such as chest pain and difficulty breathing. However, some patients may be asymptomatic, especially elderly patients and those with diabetes,³ which makes early diagnosis even more crucial.

Electrocardiography (ECG) is the primary tool for diagnosing MI and is capable of recording cardiac electrical potential changes and analyzing waveform characteristics, such as ST-segment and T-wave changes, to detect diseases.⁴

Supplementary Tables are available on
<https://journals.lww.com/joms>

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Liu WC, Liu YL, Chang DW, Lee CC, Chang CH, Tsai DJ. Integrating manual ECG feature extraction with ensemble learning for myocardial infarction diagnosis. J Med Sci 2025;45:229-36.

Received: April 28, 2025; Revised: May 28, 2025;
Accepted: June 05, 2025; Published: August 22, 2025
Corresponding Author: Dr. Dung-Jang Tsai, Room 7201,
No. 161, Min-Chun East Road, Section 6, Neihu, Taipei
114, Taiwan. Tel: +886-2-87923100#774314; Fax: +886-2-
87924811. E-mail: oo800217@mail.ndmctsgh.edu.tw

However, traditional 12-lead ECGs are used mainly within medical institutions, which may be inconvenient for many people and thus reduce their willingness to undergo examination. Moreover, this detection method requires connecting multiple leads, which makes it less flexible for detecting sudden symptoms, and may fail to detect abnormalities when symptoms disappear. All of these factors contribute to delays in diagnosis and treatment.⁵

This study aims to explore improvements in the accuracy and efficiency of MI diagnosis through ECG feature extraction and ensemble learning models. ECG signal features, such as the wave peak amplitude, interval, duration, and area, are crucial for diagnosing heart diseases.⁶⁻⁹ Researchers¹⁰⁻¹³ have extensively used techniques such as wavelet transform, time-domain morphology, and gradient analysis to automatically extract key physiological indicators such as P-waves, QRS complexes, and T-waves that play important roles in early diagnosis and treatment.

However, traditional machine learning techniques face performance challenges when dealing with high-dimensional and imbalanced data. To address this, ensemble learning methods such as Random Forest, XGBoost, and AdaBoost have demonstrated excellent performance in handling imbalanced data and significantly improved prediction accuracy. Early or subtle manifestations of MI may be difficult to identify through standard ECGs, and the limitations of traditional machine learning models further highlight the necessity of using advanced techniques. Therefore, this study manually extracts waveform and area features from a publicly available 12-lead ECG dataset, combines ensemble learning models such as Random Forest, XGBoost, and AdaBoost to develop an assisted MI diagnosis model, and evaluates the impact of different data augmentation strategies on model prediction accuracy.

MATERIALS AND METHODS

Data source

This study uses the Physikalisch-Technische Bundesanstalt (PTB)-XL ECG dataset provided by PhysioNet. The dataset was compiled by the German Federal Institute for Physics and Technology (PTB) and included ECG data collected by Schiller AG equipment between 1989 and 1996. The PTB-XL dataset was made public in 2020 and included 21,837 12-lead ECG recordings from 18,885 patients. Each recording lasts 10 s and is sampled at 500 Hz and 100 Hz.¹⁴ The institutional review board (IRB) approval was exempted from our IRB as this study used publicly available and fully de-identified data from the PhysioNet PTB-XL dataset.

Data collection

The PTB-XL dataset records patients' ECG data, is sex balanced, covers a wide age range, and is annotated by cardiology experts. The recordings include five major diagnostic categories and are converted into up to 71 different diagnostic statements according to the Standard Communications Protocol (SCP)-ECG standard.¹⁴ In addition, the data are stored in WFDB format, with corresponding metadata provided for research analysis. This study focuses on MI data and selects 15,014 ECG recordings, including 9528 normal recordings and 5486 MI recordings.

Research methods

This study manually extracts 11 base waveform variables per lead (e.g., P-wave duration, PR interval, QRS width, QT interval, R-peak amplitude, etc.). For each of these 11 variables, we then compute seven summary statistics across the 12 leads – maximum, minimum, median, standard deviation (SD), 25th percentile, 75th percentile, and 99th percentile – yielding a total of 77 features. In addition, we include six clinical/demographic features (age, sex, heart rate, P-wave area, T-wave area, and QRS-wave area), bringing the grand total to 94 input features for our models. Multiple signal processing techniques are subsequently used to analyze the ECG data. First, the `convolve` function from the NumPy package is used to implement moving average filtering, with a Window size of 10, to smooth the data. Next, the `signal_filter` function from the NeuroKit2 toolkit is applied, using a second-order Butterworth bandpass filter with cutoff frequencies of 0.5 Hz to 50 Hz to enhance R-wave peak recognition. In addition, 10th-order polynomial detrending is performed using NeuroKit2's `signal_detrend` function to remove baseline drift. To detect R-wave peaks, this study adopts the `hamilton_segmenter` and `correct_rpeaks` functions from the BioSPPy toolkit package and further calculates RR intervals and heart rate. The QRS region is set to 0.14 s to precisely capture the QRS waveform by defining time Windows before and after the R-wave. We then apply a previously utilized methodology¹⁵ to locate the onset and offset points of the QRS complex and determine the onset of the S-wave and offset of the Q-wave based on slope changes. P-wave detection adopts a previously utilized strategy¹⁶ that identifies the P-wave peak, onset, and offset by analyzing the signal portion before QRS onset. T-wave detection is based on setting time Windows after each QRS offset to identify the T-wave peak, onset, and offset. Finally, this study uses the `trapz` function from the SciPy package to calculate the area of each waveform on the ECG.

Statistical analysis

This study uses Python for data analysis, first cleaning missing values and excluding data that do not meet the research criteria.

The descriptive statistical analysis includes variables such as sex, age, heart rate, waveform duration, and waveform area. For machine learning analysis, three classification algorithms are adopted: XGBoost, Random Forest, and AdaBoost. Data preprocessing includes missing value imputation, dataset splitting (70% training, 20% validation, and 10% testing), and the use of random undersampling techniques to handle class imbalance problems. According to Supplementary Table 1, the best model configuration is selected through a parameter grid search and multiple performance metric tests, and a final performance evaluation is conducted on an independent test set to verify the model's prediction accuracy and generalization ability, which ensures its reliability in practical applications.

RESULTS

After data augmentation and 1:1 data splitting, this study balanced the samples of MI and normal, with 5400 samples

for each category to ensure equal representation of both conditions in the analysis. The data were then split into 70% training, 20% validation, and 10% testing proportions.

As shown in Table 1, there were no significant differences in demographic variables such as age or heart rate between the training, validation, and testing groups, which indicated a consistent distribution ($P > 0.05$). In terms of waveform duration, none of the features reached statistical significance. With respect to the waveform area, there were no significant differences in the distributions of the P-wave, T-wave, and QRS wave areas. For the statistical values of various waveform durations in Supplementary Table 2, such as the maximum, minimum, median, standard deviation, and percentiles, all features did not reach statistical significance, which indicated a consistent distribution of these statistical values across different groups. The proportions of MI samples and NORM samples were similar ($P = 1.000$). These results show that the distributions of most features were consistent among the training, validation, and testing sets.

Table 1: Descriptive statistical analysis based on a 1:1 split strategy with augmented data

Variable	Training set ($n=7559$)	Validation set ($n=2160$)	Testing set ($n=1081$)	P
Demographic characteristics				
Age	59.68±16.79	58.85±16.97	59.28±16.13	0.116
Sex, n (%)				0.355
Male	4082 (54.0)	1199 (55.5)	600 (55.5)	
Female	3477 (46.0)	961 (44.5)	481 (44.5)	
Heart rate	73.51±15.87	73.82±16.32	74.03±16.46	0.498
Wave duration time				
P-wave duration	102.79±5.32	102.94±5.42	102.94±5.20	0.384
Q-wave duration	38.54±5.13	38.41±5.21	38.45±5.08	0.523
S-wave duration	37.80±5.03	37.83±4.78	37.83±4.93	0.959
T-wave duration	251.98±35.69	251.01±36.31	251.09±36.86	0.455
PR interval	151.64±14.32	152.07±14.49	152.10±14.41	0.348
QT interval	460.61±40.22	459.25±40.93	459.08±40.61	0.245
ST interval	328.96±39.36	327.73±39.96	327.45±39.72	0.275
RR interval	853.27±170.62	850.29±170.12	847.94±170.10	0.537
QRS interval	131.66±5.90	131.52±5.83	131.63±5.85	0.646
ST segment	76.97±24.21	76.72±24.62	76.36±25.79	0.708
PR segment	48.86±13.88	49.14±13.94	49.15±13.70	0.622
Wave area				
P-wave area	2.02±1.05	2.00±0.97	2.03±1.02	0.760
T-wave area	9.37±4.81	9.41±4.42	9.33±4.49	0.878
QRS-wave area	7.74±3.83	7.65±3.49	7.55±3.66	0.202
MI (%)				
NORM	3779 (50.0)	1080 (50.0)	541 (50.0)	0.999
MI	3780 (50.0)	1080 (50.0)	540 (50.0)	

* $P < 0.05$. NORM=Normal; MI=Myocardial infarction

Predictive ability of myocardial infarction and its subcategories in various machine learning models

To predict MI, this study compared three different machine learning models. As shown in Table 2, the XGBoost model achieved an accuracy of 86.12%, a sensitivity of 84.81%, a specificity of 87.43%, a precision of 87.07%, and an F1 score of 85.92%. The Random Forest model achieved an accuracy of 83.81%, a sensitivity of 80.92%, a specificity of 86.69%, a precision of 82.50%, and an F1 score of 83.31%. In addition, the AdaBoost model achieved an accuracy of 85.29%, a sensitivity of 83.14%, a specificity of 87.43%, a precision of 79.60%, and an F1 score of 80.32%. As shown in Figure 1, the area under the receiver operating characteristic curve (AUROC) for XGBoost, Random Forest, and AdaBoost reached 94%, 91%, and 90%, respectively. The XGBoost model had the best performance based on the AUROC.

Next, predictions for four subcategories were performed. For anterior MI (AMI), as shown in Table 2, the XGBoost model achieved an accuracy of 88.66%, a sensitivity of 85.33%, a specificity of 92.00%, a precision of 91.42%, and an F1 score of 88.27%. The Random Forest model achieved an accuracy of 82.00%, a sensitivity of 78.66%, a specificity of 85.33%, a precision of 84.28%, and an F1 score of 81.37%. In addition, the AdaBoost model achieved an accuracy of 86%, a sensitivity of 83.66%, a specificity of 88.33%, a precision of 87.76%, and an F1 score of 85.66%. According to Figure 2a, the AUROC for XGBoost, Random Forest, and AdaBoost reached 94%, 91%, and 92%, respectively. The XGBoost model had the best performance based on the AUROC.

Table 2: Evaluation metric results for myocardial infarction and subtypes across different machine learning models

Model	Category	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)
XGBoost	MI	86.12	84.81	87.43	87.07	85.92
	AMI	88.66	85.33	92.00	91.42	88.27
	IMI	80.00	80.00	80.00	80.00	80.00
	LMI	80.46	80.00	80.95	80.00	80.00
Random Forest	MI	83.81	80.92	86.69	82.50	83.31
	AMI	82.00	78.66	85.33	84.28	81.37
	IMI	77.81	76.87	78.75	78.34	77.60
	LMI	80.48	85	76.19	77.27	80.95
AdaBoost	MI	85.29	83.14	87.43	79.60	80.32
	AMI	86.00	83.66	88.33	87.76	85.66
	IMI	80.78	80.00	81.65	81.26	80.62
	LMI	78.04	75	80.95	78.94	76.92

MI=Myocardial infarction; AMI=Anterior MI; IMI=Inferior MI; LMI=Lateral MI

For the inferior MI (IMI), according to Table 2, the XGBoost model achieved an accuracy of 80%, a sensitivity of 80%, a specificity of 80%, a precision of 80%, and an F1 score of 80%. The Random Forest model achieved an accuracy of 77.81%, sensitivity of 76.87%, specificity of 78.75%, precision of 78.34%, and F1 score of 77.60%. In addition, the AdaBoost model achieved an accuracy of 80.78%, a sensitivity of 80%, a specificity of 81.65%, a precision of 81.26%, and an F1 score of 80.62%. As shown in Figure 2b, the AUROCs for XGBoost, Random Forest, and AdaBoost reached 90%, 88%, and 90%, respectively. The XGBoost and AdaBoost model had the best performance.

Finally, for the lateral MI (LMI), according to Table 2, the XGBoost model achieved an accuracy of 80.46%, a sensitivity of 80%, a specificity of 80.95%, a precision of 80%, and an F1 score of 80.00%. The Random Forest model achieved an accuracy of 80.48%, sensitivity of 85.0%, specificity of 76.19%, precision of 77.27%, and F1 score of 80.95%. In addition, the AdaBoost model achieved an accuracy of 78.04%, a sensitivity of 75.00%, a specificity of 80.95%, a precision of 75.94%, and an F1 score of 76.92%. As shown in Figure 2c, the AUROC for XGBoost, Random Forest, and AdaBoost reached 90%, 89%, and 85%, respectively. The XGBoost model had the best performance.

Next, feature importance ranking was performed. As shown in Figure 3a, for MI, the top 10 important features were the standard deviation of the S-wave duration, 75th percentile of the QT interval, the median QT interval, age, 75th percentile of the ST interval, 75th percentile of the S-wave duration, 75th percentile of the PR segment, sex, the median T-wave

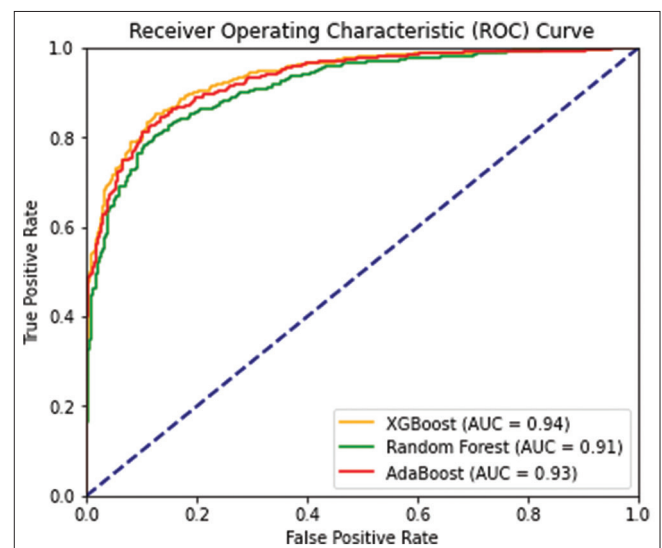


Figure 1: Area under the receiver operating characteristic curve of three different models under a 1:1 data split. ROC = Receiver operating characteristic; AUC = Area under the curve

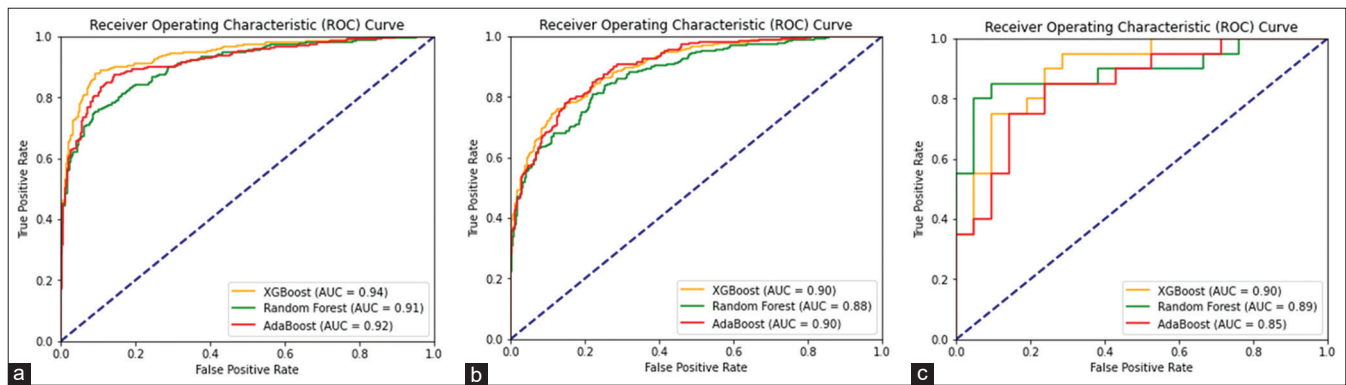


Figure 2: Area under the receiver operating characteristic curve (AUROC) analysis of each subclass under different machine learning models. (a) AUROC analysis for anterior myocardial infarction, (b) AUROC analysis for inferior myocardial infarction, (c) AUROC analysis for lateral myocardial infarction. ROC=Receiver operating characteristic; AUC= Area under the curve

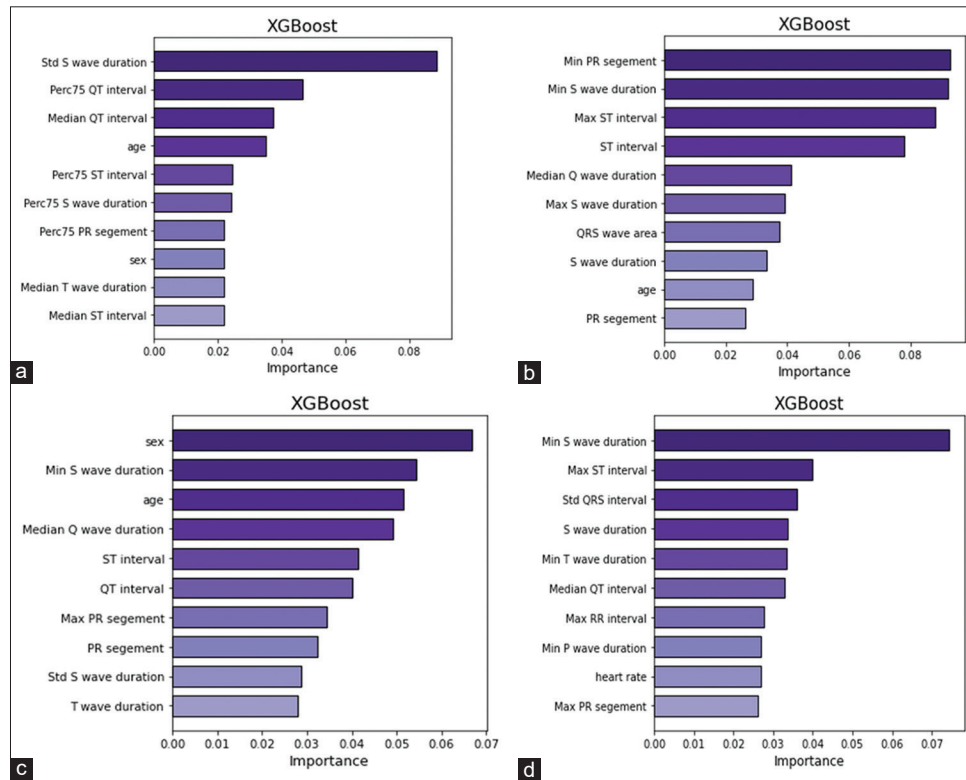


Figure 3: Top 10 feature importance rankings for myocardial infarction (MI) and its subcategories across three models: (a) MI, (b) anterior MI, (c) inferior MI, (d) lateral MI

duration, and the median ST interval. For AMI, the XGBoost model had the best performance. As shown in Figure 3b, the top 10 important features were the minimum PR segment, the minimum S-wave duration, the maximum ST interval, the ST interval, the median Q-wave duration, the maximum S-wave duration, the QRS-wave area, the S-wave duration, age, and the PR segment. For inferior MI, the XGBoost model had the best performance. As shown in Figure 3c, the top 10 important features were sex, the minimum S-wave duration, age, the

median Q-wave duration, the ST interval, the QT interval, the maximum PR segment, the PR segment, the standard deviation of the S-wave duration, and the T-wave duration. For LMI, the XGBoost model performed best. As shown in Figure 3d, the top 10 important features were the minimum S-wave duration, the maximum ST interval, the standard deviation of the QRS interval, the S-wave duration, the minimum T-wave duration, the median QT interval, the maximum RR interval, the minimum S-wave duration, heart rate, and the maximum PR segment.

DISCUSSION

This study employed manual feature extraction and waveform analysis techniques to locate and analyze P-waves, QRS complexes, and T-waves precisely in Electrocardiograms (ECGs). The duration and area of each waveform were calculated to predict MI. Through data augmentation, the overall AUROC for MI reached 93.59%, with an AUROC of 94.09% for AMI diagnosis and 90.14% and 89.76% for inferior and LMI, respectively. The study identified the ST interval as the most crucial feature, which is consistent with clinical diagnosis, because changes in the ST segment and T-wave are vital for MI diagnosis. The XGBoost model achieved an accuracy of 86.21% for MI prediction. In contrast to previous studies that used automatically generated ECG data, this study utilized the PTB-XL dataset provided by PhysioNet and employed manual feature extraction to increase model performance. The manual approach more accurately captured key ECG characteristics, reduced the impact of instrumental errors in automatic extraction, and improved model accuracy. However, manual feature extraction is tedious and time-consuming and heavily depends on expert knowledge and experience. The accuracy of feature extraction largely depends on the precise detection of reference points, and any deviation may affect diagnostic reliability.

Several previous studies have employed different methodologies and achieved varying results in MI detection. An earlier study by Ibrahim *et al.*¹⁷ used the ECG-VIEW II dataset, which was comprised 979,273 ECG records from 371,401 patients, and achieved an AUROC of 96.5% with XGBoost. The study used RobustScaler for standardization and synthetic minority oversampling technique (SMOTE) to address class imbalance, which are techniques not employed in the current study. Another study by Sharma and Sunkaria¹⁸ proposed a method based on stationary wavelet transform (SWT) and machine learning to detect inferior MI using ECG data from leads II, III, and aVF. Their support vector machine model achieved an impressive AUROC of 99.94% for inferior MI detection and significantly outperformed this study's XGBoost model (AUROC of 90.14%) for the same task. The superior performance of Sharma and Sunkaria's study may be attributed to their focus on specific leads most sensitive to changes in the inferior wall of the heart and their use of SWT for signal processing, which maintains signal translation invariance and allows multiscale analysis. In contrast, this study employed Butterworth and moving average filters, which may introduce phase delays. A study by Chumachenko *et al.*¹⁹ using heart rate variability (HRV) indicators for feature extraction achieved an accuracy of 99.629% with a Random Forest model, demonstrating

superior predictive capability compared with this study's 86.21% accuracy.

Recent research has increasingly utilized deep learning approaches for MI detection²⁰ and developed hybrid models that combine convolutional neural networks (CNNs) and recurrent neural networks (RNNs), specifically the CNN-long short-term memory (LSTM) and CNN- bidirectional LSTM (BiLSTM) models. When trained on the PTB-XL dataset, these models achieved accuracies ranging from 89% to 91%, which slightly outperformed this study's XGBoost model. The advantages of their approach lie in the automatic feature extraction capabilities of deep learning models, which can effectively learn important patterns in ECG data. The combination of CNNs and RNNs allows the learning of both spatial and temporal features, which is particularly important for ECG data processing. Another study by Zhang and Li²¹ employed a Bi-LSTM network based on a heartbeat-attention mechanism, which achieved an accuracy of 94.77% on the PTB diagnostic ECG database. Compared with the manual feature extraction approach in this study, the ability of the Bi-LSTM model to automatically extract features and weight important heartbeat signals significantly improved the predictive performance.²² and proposed a multichannel lightweight CNN (MCL-CNN) for detecting AMI using ECG data from leads v1, v2, and v3 and achieved an AUROC of 95.50%. While this study's XGBoost model demonstrated better performance for AMI detection, with an AUROC of 96.19%, it is likely due to the use of complete 12-lead ECG data, which provides more comprehensive information than the three-lead approach does.

This study identified the ST interval as the most important feature for MI detection, which aligns with clinical diagnosis practices. This finding contrasts with a previous study by Kalmady *et al.*,²³ which identified patient sex as a key feature, possibly because of significant sex distribution effects in their dataset. The current study relied primarily on ECG morphological features (such as P-waves, Q-waves, S-waves, T-waves, and ST segments), whereas other studies, such as Chumachenko *et al.*,¹⁹ utilized HRV indicators such as interbeat intervals and beats per minute. HRV indicators can consistently reflect short-term cardiac activity changes and have strong discriminative power for MI. However, they may require hospital visits for examination, which may delay diagnosis and treatment. The impact of applying different filters on feature extraction results is a subject for further investigation. This study used Butterworth and moving average filters, whereas others employed various types, such as notch filters, finite impulse response filters, and bandpass filters. Future research should explore the effects of different filters on machine learning model performance. In addition, focusing on specific ECG leads that are most sensitive to changes in

the heart (such as leads II, III, and aVF for inferior MI) could optimize model performance in future studies. Although manual feature extraction allows for precise control over the extracted features, it may not fully capture the complex patterns in the data compared with the automatic feature extraction methods used in deep learning approaches. Future studies may benefit from combining the strengths of manual and automatic feature extraction techniques, which may lead to more robust and accurate MI detection models.

CONCLUSION

This study employed manual feature extraction and waveform analysis techniques to accurately locate and analyze P-waves, QRS complexes, and T-waves in ECGs and calculate the duration and area of each waveform to successfully predict MI. Through data augmentation, the overall AUROC for MI reached 93.59%, with 94.09% and 90.14% for anterior and inferior MI diagnosis, respectively. However, the AUROC was 89.76% for lateral wall MI. This study suggests that future research could incorporate additional parameters such as HRV, explore more class imbalance handling techniques such as SMOTE or adaptive synthetic sampling, and compare the effects of different filtering techniques. Furthermore, the adoption of deep learning models to enhance the automation of feature extraction and improve diagnostic accuracy is recommended.

Data availability statement

The data that support the findings of this study are openly available in the PhysioNet repository at <https://physionet.org/content/ptb-xl/1.0.1/>.

Financial support and sponsorship

This study was supported by funding from the National Science and Technology Council, Taiwan (NSTC 112-2222-E-030-002-MY2 to D. J. Tsai).

Conflicts of interest

There are no conflicts of interest.

REFERENCES

- Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* 2006;3:e442.
- Wu YC, Lo WC, Lu TH, Chang SS, Lin HH, Chan CC. Mortality, morbidity, and risk factors in Taiwan, 1990-2017: Findings from the Global Burden of Disease Study 2017. *Journal of the Formosan Medical Association = Taiwan yi zhi* 2021;120:1340-9. Available from: <https://doi.org/10.1016/j.jfma.2020.11.014>.
- Yang Y, Li W, Zhu H, Pan XF, Hu Y, Arnott C, *et al.* Prognosis of unrecognised myocardial infarction determined by electrocardiography or cardiac magnetic resonance imaging: Systematic review and meta-analysis. *BMJ* 2020;369:m1184.
- Sgarbossa EB, Birnbaum Y, Parrillo JE. Electrocardiographic diagnosis of acute myocardial infarction: current concepts for the clinician. *American Heart Journal* 2001;141:507-17.
- Et-Tousy J, Et-Tousy S, Ait El Aouad S, Zyane A. Enhanced 12-Lead ECG Reconstruction from Single-Lead Data Using WaveNet. Vol. 69. In *ITM Web of Conferences*; 2024. p. 02008. EDP Sciences.
- Alexakis C, Nyongesa H, Saatchi R, Harris N, Davies C, Emery C, *et al.* Feature Extraction and Classification of Electrocardiogram (ECG) Signals Related to Hypoglycaemia. In *Computers in Cardiology*, 2003. IEEE; 2003.
- de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans Biomed Eng* 2004;51:1196-206.
- Khan TT, Sultana N, Reza RB, Mostafa R. ECG Feature Extraction in Temporal Domain and Detection of Various Heart Conditions. In *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*. IEEE; 2015.
- Zhang Z, Dong J, Luo X, Choi KS, Wu X. Heartbeat classification using disease-specific feature selection. *Comput Biol Med* 2014;46:79-89.
- Tekeste T, Bayasi N, Saleh H, Khandoker A, Mohammad B, Al-Qutayri M, *et al.* Adaptive ECG Interval Extraction. In *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE; 2015.
- Bayasi N, Tekeste T, Saleh H, Khandoker A, Mohammad B, Ismail M. Adaptive Technique for P and T Wave Delineation in Electrocardiogram Signals. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE; 2014.
- Bayasi N, Tekeste T, Saleh H, Mohammad B, Khandoker A, Ismail M. Low-power ECG-based processor for predicting ventricular arrhythmia. *IEEE Trans Very Large Scale Integr (VLSI) Syst* 2015;24:1962-74.
- Ramakrishnan S, Yogeswaran R. Design and Analysis of Feature Extraction Algorithm for ECG Signals Using Adaptive Threshold Method. In *2017 Trends in Industrial Measurement and Automation (TIMA)*. IEEE; 2017.

14. Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, *et al.* PTB-XL, a large publicly available electrocardiography dataset. *Sci Data* 2020;7:154.
15. Sadhukhan D, Mitra M. Detection of ECG Characteristic Features Using Slope Thresholding and Relative Magnitude Comparison. In 2012 Third International Conference on Emerging Applications of Information Technology. IEEE; 2012.
16. Mazomenos E, Chen T, Acharyya A, Bhattacharya A, Rosengarten J, Maharatna K. A Time-Domain Morphology and Gradient Based Algorithm for ECG feature extraction. In 2012 IEEE International Conference on Industrial Technology. IEEE; 2012.
17. Ibrahim L, Mesinovic M, Yang KW, Eid MA. Explainable prediction of acute myocardial infarction using machine learning and Shapley values. *IEEE Access* 2020;8:210410-7.
18. Sharma LD, Sunkaria RK. Inferior myocardial infarction detection using stationary wavelet transform and machine learning approach. *Signal Image Video Process* 2018;12:199-206.
19. Chumachenko D, Butkevych M, Lode D, Frohme M, Schmailzl KJ, Nechyporenko A. Machine learning methods in predicting patients with suspected myocardial infarction based on short-time HRV data. *Sensors (Basel)* 2022;22:7033.
20. Hasbullah S, Mohd Zahid MS, Mandala S. Detection of myocardial infarction using hybrid models of convolutional neural network and recurrent neural network. *BioMedInformatics* 2023;3:478-92.
21. Zhang Y, Li J. Application of heartbeat-attention mechanism for detection of myocardial infarction using 12-lead ECG records. *Appl Sci* 2019;9:3328.
22. Chen Y, Chen H, He Z, Yang C, Cao Y. Multi-Channel Lightweight Convolution Neural Network for Anterior Myocardial Infarction Detection. In 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). IEEE; 2018.
23. Kalmady SV, Salimi A, Sun W, Sepehrvand N, Nademi Y, Baine K, *et al.* Development and validation of machine learning algorithms based on electrocardiograms for cardiovascular diagnoses at the population level. *NPJ Digit Med* 2024;7:133.

Supplementary Table 1: Hyperparameter settings for various machine learning models

Model	Hyperparameters
XGBoost	n_estimators=200, learning_rate=0.1, reg_alpha=0.05, gamma=0.5, max_depth=4, subsample=0.8, colsample_bytrees=1.0
Random Forest	n_estimators=200, max_depth=8, min_samples_split=2, min_samples_leaf=1, bootstrap=True
AdaBoost	n_estimators=300, learning_rate=0.1

Supplementary Table 2: Expanded manually extracts waveform data

Variable	Training set	Validation set	Testing set	<i>P</i>
Wave duration time				
Maximum P-wave duration	116.08±5.12	116.21±5.12	116.28±5.13	0.334
Maximum Q-wave duration	65.17±10.36	65.16±10.40	65.13±10.09	0.994
Maximum S-wave duration	57.23±12.13	57.24±11.82	57.00±11.79	0.829
Maximum T-wave duration	341.95±24.82	342.29±24.45	341.53±24.84	0.700
Maximum PR interval	183.57±15.21	184.24±15.32	183.93±15.41	0.180
Maximum QT interval	549.90±46.29	549.78±46.96	549.21±45.79	0.901
Maximum ST interval	414.35±46.45	413.94±46.62	413.52±45.45	0.828
Maximum RR interval	875.62±183.15	870.68±181.84	870.59±187.44	0.434
Maximum QRS interval	152.59±9.30	152.50±9.51	152.47±9.15	0.866
Maximum ST segment	146.97±41.08	147.48±42.23	146.08±41.77	0.664
Maximum PR segment	77.11±15.56	77.62±15.77	77.56±15.46	0.312
Minimum P-wave duration	81.36±10.19	81.69±10.08	81.77±10.20	0.241
Minimum Q-wave duration	23.52±7.44	23.42±7.41	23.54±7.38	0.844
Minimum S-wave duration	27.04±6.69	26.98±6.60	27.09±6.78	0.900
Minimum T-wave duration	172.40±34.27	171.96±35.09	173.45±33.92	0.506
Minimum PR interval	119.82±16.00	120.09±16.06	119.75±16.12	0.757
Minimum QT interval	380.00±42.11	380.11±42.38	379.65±40.96	0.956
Minimum ST interval	250.30±38.72	250.83±38.39	249.71±37.60	0.721
Minimum RR interval	819.95±184.52	816.90±185.99	816.00±181.61	0.683
Minimum QRS interval	112.24±13.08	112.01±13.02	112.26±13.00	0.757
Minimum ST segment	12.71±17.11	12.26±16.68	12.64±18.13	0.553
Minimum PR segment	27.03±11.71	26.97±11.65	26.87±11.17	0.904
Median P-wave duration	104.72±6.02	104.81±6.13	104.83±5.92	0.731
Median Q-wave duration	35.08±5.30	34.93±5.28	34.98±5.14	0.480
Median S-wave duration	35.90±4.42	35.97±4.25	36.04±4.28	0.585
Median T-wave duration	247.23±47.59	245.19±48.47	246.25±49.19	0.209
Median PR interval	151.64±16.29	152.09±16.55	152.17±16.41	0.385
Median QT interval	457.96±49.19	455.38±50.13	456.03±49.69	0.070
Median ST interval	326.85±48.14	324.61±49.12	325.06±48.70	0.116
Median RR interval	853.80±171.80	850.97±170.88	848.13±170.44	0.522
Median QRS interval	131.30±6.24	131.22±6.18	131.28±6.21	0.888
Median ST segment	78.31±28.95	77.92±29.52	77.50±30.80	0.637
Median PR segment	47.03±15.97	47.41±15.99	47.45±15.75	0.506
Standard P-wave duration	10.02±3.02	9.95±2.97	9.99±3.06	0.644

Contd...

Supplementary Table 2: Contd...

Variable	Training set	Validation set	Testing set	<i>P</i>
Standard Q-wave duration	12.42±3.51	12.43±3.55	12.32±3.38	0.639
Standard S-wave duration	8.61±4.19	8.61±4.15	8.49±4.12	0.685
Standard T-wave duration	52.75±12.75	53.28±12.75	52.20±12.43	0.060
Standard PR interval	17.93±5.06	18.00±5.01	18.03±5.15	0.728
Standard QT interval	50.00±15.79	50.20±16.01	49.85±15.42	0.808
Standard ST interval	48.40±15.20	48.41±15.44	48.20±14.87	0.918
Standard RR interval	15.98±35.39	15.31±34.49	15.50±33.42	0.701
Standard QRS interval	11.40±4.34	11.42±4.48	11.33±4.19	0.872
Standard ST segment	39.65±12.33	40.23±12.71	39.37±12.39	0.094
Standard PR segment	14.65±4.41	14.79±4.36	14.87±4.42	0.161
Perc25 P-wave duration	97.95±7.02	98.17±7.16	98.11±6.86	0.384
Perc25 Q-wave duration	30.36±5.06	30.20±5.12	30.43±4.96	0.357
Perc25 S-wave duration	32.46±4.92	32.45±4.97	32.60±4.93	0.663
Perc25 T-wave duration	215.85±41.82	214.54±43.04	215.29±42.89	0.436
Perc25 PR interval	141.08±15.15	141.57±15.24	141.57±15.09	0.307
Perc25 QT interval	428.88±44.70	426.99±44.87	427.37±44.87	0.166
Perc25 ST interval	297.90±43.15	296.38±43.44	296.54±43.21	0.266
Perc25 RR interval	852.73±171.70	850.26±171.11	847.29±170.53	0.566
Perc25 QRS interval	125.14±7.55	125.03±7.53	125.25±7.13	0.717
Perc25 ST segment	51.62±27.05	50.76±27.02	51.02±28.21	0.384
Perc25 PR segment	39.11±14.37	39.38±14.43	39.27±14.31	0.725
Perc75 P-wave duration	109.63±5.23	109.66±5.34	109.74±5.15	0.79
Perc75 Q-wave duration	43.76±7.79	43.67±7.94	43.53±7.77	0.619
Perc75 S-wave duration	40.87±7.04	40.90±6.56	40.76±6.78	0.854
Perc75 T-wave duration	285.58±45.15	284.96±45.44	283.76±46.19	0.439
Perc75 PR interval	162.13±16.93	162.45±17.05	162.68±17.06	0.505
Perc75 QT interval	490.40±47.36	489.06±48.31	488.35±47.98	0.263
Perc75 ST interval	358.32±46.92	357.04±47.74	356.05±47.43	0.227
Perc75 RR interval	856.36±173.11	853.43±172.53	850.59±170.96	0.514
Perc75 QRS interval	137.62±6.34	137.48±6.38	137.46±6.42	0.518
Perc75 ST segment	100.41±31.85	100.70±32.87	99.88±33.31	0.791
Perc75 PR segment	56.99±16.44	57.28±16.49	57.59±16.37	0.465
Perc99 P-wave duration	115.76±5.05	115.89±5.06	115.95±5.05	0.342
Perc99 Q-wave duration	64.12±10.04	64.09±10.09	64.05±9.77	0.970
Perc99 S-wave duration	56.26±11.74	56.27±11.42	56.04±11.42	0.834
Perc99 T-wave duration	339.61±25.13	339.96±24.76	338.99±25.10	0.581
Perc99 PR interval	182.34±15.10	182.99±15.22	182.74±15.29	0.184
Perc99 QT interval	546.66±45.28	546.55±46.00	545.93±44.81	0.885
Perc99 ST interval	411.30±45.49	410.89±45.73	410.42±44.51	0.808
Perc99 RR interval	874.00±181.56	869.24±180.44	868.98±185.46	0.446
Perc99 QRS interval	151.79±8.93	151.68±9.11	151.67±8.81	0.845
Perc99 ST segment	144.14±39.70	144.65±40.92	143.20±40.45	0.622
Perc99 PR segment	76.04±15.38	76.55±15.61	76.50±15.30	0.307