



OPEN

Hybrid CNN-BLSTM architecture for classification and detection of arrhythmia in ECG signals

Yuguang Ye^{1,2,5}, Kavimbi Chipusu³✉, Muhammad Awais Ashraf³, Bijiao Ding⁴, Yifeng Huang⁴ & Jianlong Huang^{1,2,5}✉

This study introduces a robust and efficient hybrid deep learning framework that integrates Convolutional Neural Networks (CNN) with Bidirectional Long Short-Term Memory (BLSTM) networks for the automated detection and classification of cardiac arrhythmias from electrocardiogram (ECG) signals. The proposed architecture leverages the complementary strengths of both components: the CNN layers autonomously learn and extract salient morphological features from raw ECG waveforms, while the BLSTM layers effectively model the sequential and temporal dependencies inherent in ECG signals, thereby improving diagnostic accuracy. To further enhance training stability and non-linear representation capability, the Mish activation function is incorporated throughout the network. The model was trained and evaluated using a combination of the widely recognized MIT-BIH Arrhythmia Database and de-identified clinical ECG recordings sourced from collaborating healthcare institutions, ensuring both diversity and clinical relevance of the dataset. Notably, the framework operates with minimal preprocessing, underscoring its practical viability for real-time implementation. Experimental results demonstrate the model's exceptional performance, achieving an overall classification accuracy of 99.52%, sensitivity of 99.48%, and specificity of 99.85%. These outcomes highlight the model's robustness, generalizability, and strong potential for integration into clinical decision support systems, particularly in high-throughput or resource-constrained healthcare environments.

Keywords Arrhythmia detection, Long short-term memory, Electrocardiogram, Hybrid model application

According to the 2019 World Health Organization report, about 16% of the 55.4 million deaths worldwide are related to heart disease¹. ECG plays a critical role in monitoring cardiac activity and remains an essential diagnostic tool for identifying arrhythmias^{2,3}. Changes in the electrical activity of the heart are often precursors to serious cardiac conditions, making early detection through ECG vital. Timely and accurate classification of arrhythmias can improve clinical decision-making, prevent complications, and support real-time patient monitoring^{4,5}. However, traditional ECG analysis methods rely heavily on clinical expertise to identify and classify waveforms, which is not only labor-intensive but also susceptible to subjective variability. Manual feature extraction lacks scalability, and conventional algorithms often fail to generalize across patients with varying signal morphologies. These challenges create a pressing need for more automated, accurate, and generalizable methods for arrhythmia detection.

In recent years, the application of artificial intelligence (AI) in ECG classification has gained considerable traction^{6–8}. Numerous machine learning and deep learning models have been proposed to automate arrhythmia diagnosis, often demonstrating promising results in terms of classification performance. Many of these models suffer from three critical limitations: (1) the feature extraction and classification stages are often treated as separate processes, which reduces optimization efficiency; (2) a significant number of models still rely on manually crafted features, which may omit latent or complex signal characteristics; and (3) the performance of many existing models degrades in noisy or imbalanced datasets, limiting their real-world applicability. To address these limitations, this study aims to develop a robust, end-to-end hybrid deep learning model that can simultaneously learn morphological and temporal features from raw ECG signals. Specifically, we propose a

¹Faculty of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China. ²Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou 362000, China. ³Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada. ⁴Department of Diagnostic Radiology, Huaqiao University Affiliated Strait Hospital, Quanzhou 362000, Fujian, China. ⁵Key Laboratory of Intelligent Computing and Information Processing, Fujian Province University, Quanzhou 362000, China. ✉email: kavimbi.chipusu@usask.ca; robotics@qztc.edu.cn

novel framework that integrates CNN for spatial feature extraction and BLSTM networks for temporal sequence modeling. CNNs are capable of capturing morphological patterns in ECG signals by extracting hierarchical features^{9–11}, while BLSTMs are effective in modeling sequential dependencies and contextual information.

The central focus of this study is the design and validation of a CNN-BLSTM hybrid model for multi-class arrhythmia classification using minimally preprocessed ECG signals. By combining these two architectures, the model benefits from the strengths of both spatial and temporal feature learning, enhancing its classification capability. In addition, we introduce the Mish activation function into the hybrid model to improve learning stability and maintain information continuity, which is especially important in physiological signal processing. This activation function has shown better performance than traditional functions like ReLU in non-linear representation tasks. Several prior studies have established the individual success of CNNs and LSTMs in arrhythmia detection. For example, Saenz-Cogollo et al.¹² used a CNN-based system with signal transformation techniques to detect atrial fibrillation, achieving accuracies over 98%. Garcia et al.¹³ employed deep CNN models alongside focal loss to address class imbalance in heartbeat detection. Hassan et al.¹⁴ proposed a multi-view CNN for classifying heartbeat disorders using symbolic representations, reaching notable F1-scores for different arrhythmia types. On the temporal modeling side, Kiranyaz et al.¹⁵ applied LSTMs with Fourier-transformed ECG data, further underscoring the effectiveness of temporal deep learning techniques in ECG analysis.

Recent literature further demonstrates a growing interest in hybrid deep learning models and novel optimization techniques for ECG-based arrhythmia classification. A study by Kiyimac and Kaya (2023)¹⁶ introduced a novel approach for optimizing CNN-based arrhythmia classification using a memory-enhanced artificial hummingbird algorithm. Recognizing that cardiac arrhythmia is a major indicator of cardiovascular disease and that deep learning performance relies heavily on hyperparameter optimization (HPO), a known NP-hard problem, they proposed a metaheuristic solution tailored for efficient HPO. Their method incorporates a memory unit to store previous solution evaluations, significantly reducing computation time. Moreover, they designed a custom fitness function that balances accuracy and model complexity. Using the MIT-BIH arrhythmia database, their model achieved a classification accuracy of 98.87%, outperforming or matching other MH-based techniques. While their framework focuses on optimizing CNN hyperparameters to improve performance, our work takes a different approach by integrating CNN and BLSTM architectures in an end-to-end hybrid model. This enables both spatial and temporal feature learning and offers a simpler and more generalizable solution for arrhythmia detection. In parallel, cardiac-based modelling techniques^{17,18} are often employed to establish guidelines and standards for the accurate evaluation of heart diseases, reinforcing the importance of robust and interpretable diagnostic frameworks.

These studies reflect the rapid evolution of deep learning in ECG analysis, and underscore the need for integrated, end-to-end systems like ours that unify spatial and temporal learning while maintaining generalizability and clinical relevance. However, few studies have successfully fused CNN and BLSTM into a unified pipeline capable of end-to-end training for both morphological and sequential ECG feature learning^{19–22}. The novelty of our work lies in three key areas. First, we propose a fully integrated CNN-BLSTM architecture that enables robust arrhythmia classification directly from raw ECG signals, thereby eliminating the reliance on hand-crafted features. Second, we incorporate the Mish activation function into the model^{23,24}, which enhances training stability and overall classification performance by promoting smoother gradient flow and preserving important information. Third, the proposed framework is rigorously validated using both publicly available (MIT-BIH) and real-world clinical ECG datasets, ensuring strong generalizability and practical relevance across diverse patient populations. Further, recent studies have proposed advanced models such as AttBiLNet²⁵, which leverages attention mechanisms and bidirectional LSTM-FCNs for improved temporal feature learning; ADLNet²⁶, a lightweight dual-branch network optimized for efficient real-time ECG classification; and ECG-NET²⁷, a deep residual architecture designed to capture multi-scale signal characteristics. These works demonstrate the growing trend toward end-to-end architectures with enhanced interpretability and efficiency. Building upon this momentum, our study introduces a CNN-BLSTM hybrid model that uniquely incorporates bidirectional sequence modeling and the Mish activation function, operating effectively on minimally preprocessed raw ECG signals for robust arrhythmia detection.

While CNN-LSTM architectures have been widely applied in arrhythmia detection, our work introduces several distinct innovations that set it apart from existing models. First, we integrate a BLSTM layer within the hybrid framework, allowing the model to capture temporal dependencies in both forward and backward directions, a feature that enhances the contextual understanding of ECG sequences. Second, we incorporate the Mish activation function, a smoother and non-monotonic alternative to traditional functions such as ReLU and tanh, which contributes to improved convergence behavior and greater training stability. Third, the proposed model is designed to work effectively on minimally preprocessed raw ECG signals, eliminating the need for manual feature engineering and thereby reducing preprocessing overhead, a critical advantage for real-time clinical deployment. These aspects, in combination, position our framework as a robust and efficient end-to-end solution that advances beyond prior CNN-LSTM implementations.

The main contributions of this study are as follows. We introduce a novel deep learning framework that combines a one-dimensional convolutional neural network (1D-CNN) with a BLSTM network, forming a hybrid architecture tailored for accurate arrhythmia detection. Leveraging the strengths of both spatial and temporal feature learning, our model achieves remarkable classification performance, with accuracy, sensitivity, and specificity values of 99.52%, 99.48%, and 99.85%, respectively. By unifying the feature extraction and classification processes within a single deep learning model, we eliminate the need for manual feature engineering, thus streamlining and enhancing the overall efficiency of the classification pipeline. Furthermore, the integration of the Mish activation function contributes to better training convergence and more stable network behavior, particularly when processing noisy or variable ECG data. Lastly, the proposed model's performance is validated

not only on benchmark datasets but also on additional clinical ECG recordings, confirming its robustness and applicability in practical diagnostic scenarios.

Methods

The article proposes an automatic arrhythmia detection approach using a hybrid model that combines a CNN with a BLSTM network, as illustrated in Fig. 1. The CNN first learns the morphological characteristics of the ECG signal, after which the BLSTM captures the contextual dependencies from these extracted features. The softmax function is used to perform the classification task. To enhance model stability during training, the Mish activation function is employed. The model processes preprocessed ECG signal data directly, leveraging the CNN's feature extraction capabilities and the BLSTM's ability to capture temporal dependencies.

The approach involves selecting the dataset, applying preprocessing techniques such as wavelet transform, IIR zero-phase shift digital filtering, and R-peak detection^{28,29}, followed by a detailed description of the CNN and BLSTM architectures. The experimental setup includes the configuration of CNN and BLSTM parameters, as well as the evaluation of performance metrics such as accuracy, sensitivity, specificity, and positive predictive value. Results indicate that the proposed CNN-BLSTM hybrid model outperforms the individual CNN and BLSTM models, achieving a high level of classification accuracy for arrhythmias. Ablation studies further demonstrate the superiority of the Mish activation function over ReLU. The study concludes by emphasizing the significance of the proposed hybrid model in enabling fast and accurate diagnosis of various arrhythmia conditions, offering potential for further research and practical applications in arrhythmia detection.

Mathematical formulation of the CNN-BLSTM architecture

Let $X = [x_1, x_2, \dots, x_T] \in \mathbb{R}^T$ represent an input ECG segment of length T . The one-dimensional convolutional layer extracts local spatial features as follows:

$$f_i^{(1)} = \sigma \left(W_1^{(1)*} X + b_1^{(1)} \right), \quad i = 1, \dots, F \quad (1)$$

where $W_1^{(1)}$ and $b_1^{(1)}$ are the filter weights and bias for the i^{th} kernel, F is the number of filters, $*$ denotes the convolution operator, and σ is the activation function (Mish in our case). The output from the convolutional block is passed to a BLSTM, which processes the sequence in both forward and backward directions:

$$\vec{h}_t = LSTM \left(x_t, \vec{h}_{t-1} \right), \quad \overleftarrow{h}_t = LSTM \left(x_t, \overleftarrow{h}_{t-1} \right) \quad (2)$$

The combined hidden representation at each time step is given by:

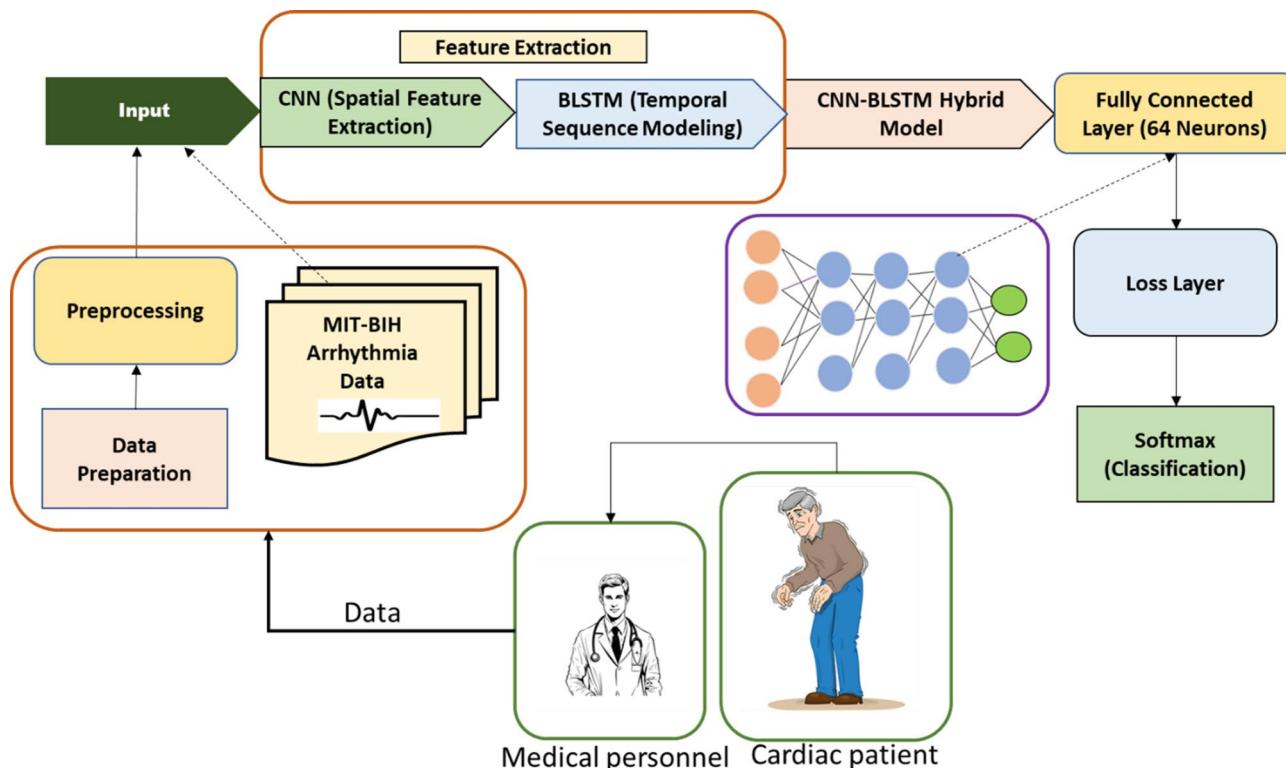


Fig. 1. Hybrid model for automatic arrhythmia detection: Integration of CNN for feature extraction and BLSTM for contextual dependency modeling.

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (3)$$

The Mish activation function is adopted to enhance learning dynamics and maintain gradient flow continuity. It is defined as:

$$\text{Mish}(x) = x \bullet \tanh(\ln(1 + e^x)) \quad (4)$$

layer produces class probabilities using the softmax function:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}, \quad i = 1, \dots, C \quad (5)$$

where \hat{y}_i is the predicted probability for class i , z_i is the input to the softmax layer, and C is the total number of arrhythmia classes. This mathematical formulation ensures a complete and transparent description of the end-to-end learning process performed by our model on raw ECG signals.

Datasets

In this study, ECG data were collected from our affiliated hospital and supplemented with records from the MIT-BIH Arrhythmia Database³⁰, of which a portion was used for model training. The hospital-acquired data were reserved exclusively for model testing to assess generalization performance. Among the clinical data, 60% originated from inpatients and 40% from outpatients³¹. The complete dataset encompasses a broad spectrum of arrhythmia types, with each ECG sample digitally annotated and labeled by experienced medical professionals. The recordings were obtained using a single-lead channel at a sampling rate of 360 Hz. Annotations include key ECG waveform components namely, the P wave, QRS complex, and T wave. Due to its distinct morphology and diagnostic importance, the QRS complex was used as the primary reference point for heartbeat detection. For R-wave identification, the Pan-Tompkins algorithm³², a well-established method for real-time and robust QRS detection, was employed. ECG recordings were categorized into five distinct classes based on characteristic features: normal heartbeat (N), left bundle branch block (L), right bundle branch block (R), atrial premature beat (A), and ventricular premature beat (V). The dataset comprises a total of 25,000 samples, with 5,000 samples allocated to each class to ensure balanced representation across the categories during both training and testing. Data segmentation was conducted at two temporal scales: large-scale beats (L), spanning 4 s, and small-scale beats (S), spanning 0.75 s, both defined relative to the QRS complex. Each heartbeat instance consists of the large-scale segment, the small-scale segment, and an associated label denoting the heartbeat class. The label is assigned independently of the large-scale segment. The 4-second segment was selected to preserve contextual information surrounding each heartbeat, enabling the model to capture inter-beat dependencies. Meanwhile, the 0.75-second segment, derived based on a median heart rate of 80 beats per minute, within the physiological range of 60–100 bpm, ensures inclusion of a complete individual heartbeat cycle.

To address the inherent class imbalance present in ECG datasets, a weighted categorical cross-entropy loss function was employed during model training. The class weights were computed as the inverse of class frequencies in the training set, thereby assigning greater emphasis to minority classes during backpropagation. This weighting strategy effectively mitigated bias toward majority classes and facilitated more balanced learning across all arrhythmia types.

Pre-processing

The ECG signal is inherently non-stationary and susceptible to noise, which can obscure its features and affect the accuracy of ECG interval measurements^{33,34}. The primary sources of interference in ECG signals include electromyographic (EMG) noise, power line interference, and baseline drift³⁵. To enhance classification efficiency and accuracy, it is essential to preprocess the ECG signal prior to classification. Several methods exist for removing such interference. First, wavelet transform is employed to suppress high-frequency components such as power line and EMG noise³⁶. Second, a Butterworth filter can be used to attenuate residual interference signals³⁷. In this study, wavelet transform is applied to denoise the signal. Subsequently, an infinite impulse response (IIR) zero-phase digital filter is used to correct baseline drift. Finally, QRS complex detection is performed to identify the R-peaks, which are then used to segment the ECG signal. The segmented signals are normalized to address the variability in signal amplitude across different phases. The normalization formula is as follows:

$$y = \frac{(y_{\max} - y_{\min}) \times (x - x_{\min})}{x_{\max} - x_{\min}} + y_{\min} \quad (6)$$

In the formula, y_{\min}, y_{\max} are the minimum and maximum values of the range of the normalized data, and the range is between 0 and 1. x_{\min}, x_{\max} are the minimum and maximum values of the data to be processed, respectively. In the signal denoising stage, we employed the Daubechies-4 (db4) wavelet due to its proven effectiveness in biomedical applications and its morphological similarity to the QRS complex. A 4-level discrete wavelet decomposition was applied, enabling the model to retain both high-frequency noise and low-frequency baseline components, thereby enhancing the quality of the ECG signal for downstream classification.

Convolutional neural network

CNNs are widely recognized deep learning architectures that have demonstrated excellent performance in various domains, including natural language processing, image segmentation, and biomedical signal analysis. A

typical CNN architecture comprises an input layer, multiple hidden layers, and an output layer. The hidden layers usually consist of convolutional and pooling layers. The convolutional layers play a crucial role in automatically extracting features from the input data by sliding convolutional filters, also known as kernels, across the input to perform local convolutions. These operations generate feature maps that capture high-level representations of the input. Each convolutional layer contains multiple filters, allowing the model to learn diverse features simultaneously. To introduce non-linearity and enable the network to learn complex mappings, activation functions are applied to the output of convolutional operations. Commonly used activation functions include the sigmoid function, hyperbolic tangent (tanh), and Rectified Linear Unit (ReLU). Following the convolutional layers, pooling layers, also referred to as downsampling layers, are employed to reduce the spatial dimensions of feature maps. Pooling helps decrease computational complexity, accelerates training, and promotes feature abstraction by summarizing regions of the feature maps. In ECG classification tasks, CNNs can directly take raw or minimally processed heartbeat signals as input, thereby significantly simplifying the preprocessing stage. Figure 2 illustrates the architecture of a CNN, including its input layer, convolutional and pooling layers, and output layer, while also depicting the feature extraction process and the role of activation and pooling functions in enhancing the network's ability to analyze ECG data effectively.

BLSTM

LSTM is a specialized form of RNN designed to overcome the limitations of traditional RNNs in handling long-term dependencies. Unlike standard RNNs, which are highly sensitive to short-term inputs but struggle with long-term information retention, LSTM introduces a gate mechanism and a cell state that significantly enhance its ability to model sequential data over extended timeframes. Specifically, LSTM networks incorporate three types of gates: the input gate, forget gate, and output gate. These gates regulate the flow of information into, out of, and within the memory cell, allowing the network to selectively retain or discard information over time. The input of LSTM not only includes the current input X_t , but also includes the output H_{t-1} at the previous moment and the cell state C_{t-1} of the hidden layer. These three parts together constitute the real input at this time, and after calculation, we get the output H_t and the cell state C_t is used as the input of the next moment. The forget gate is responsible for deciding whether to save or discard unnecessary information, the input gate is used to update the cell state, and the output gate is to obtain the output H_t through the cell state C_t and calculation at this moment. Figure 3 illustrates the LSTM architecture, highlighting the roles of the input, forget, and output gates in processing both short-term and long-term information. It demonstrates how the cell state and previous output are integrated into the input for improved sequence learning.

The calculation formula of each gate of LSTM is as follows:

- (1) The forget gate will discard historical information judged as useless or irrelevant, as shown in Eq. (7).

$$f_t = \sigma (W_f \times (h_{t-1}, X_t) + b_f) \quad (7)$$

- (2) The information retained at the last moment and the input information at this moment is taken together as the update state of the input gate, as shown in Eqs. (8) and (9).

$$i_t = \sigma (W_i \times (h_{t-1}, X_t) + b_i) \quad (8)$$

$$C_t = i_t \times \partial_t + f_t \times C_{t-1} \quad (9)$$

- (3) The state information at the current moment is output by the output gate, as shown in Eqs. (10) and (11).

$$\partial_t = \tanh (W_c \times (h_{t-1}, X_t) + b_c) \quad (10)$$

$$h_t = \sigma (W_0 \times (h_{t-1}, X_t) + b_0) \times \tanh (C_t) \quad (11)$$

In the formula, W_i , W_c , W_f and W_0 are the corresponding connection weights, b_i , b_c , b_f and b_0 are the corresponding biases, f_t corresponds to the activation value of the forget gate at time t, σ is a function of

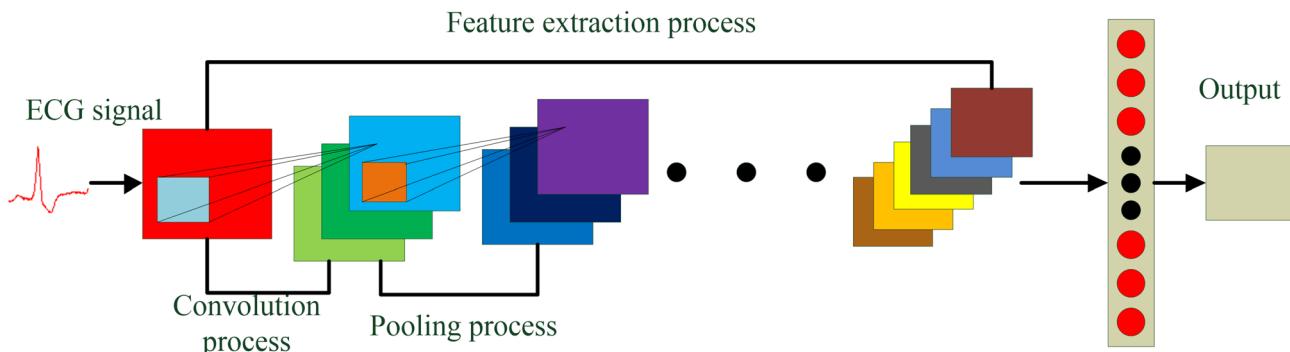


Fig. 2. Structure of the CNN for ECG-based arrhythmia detection.

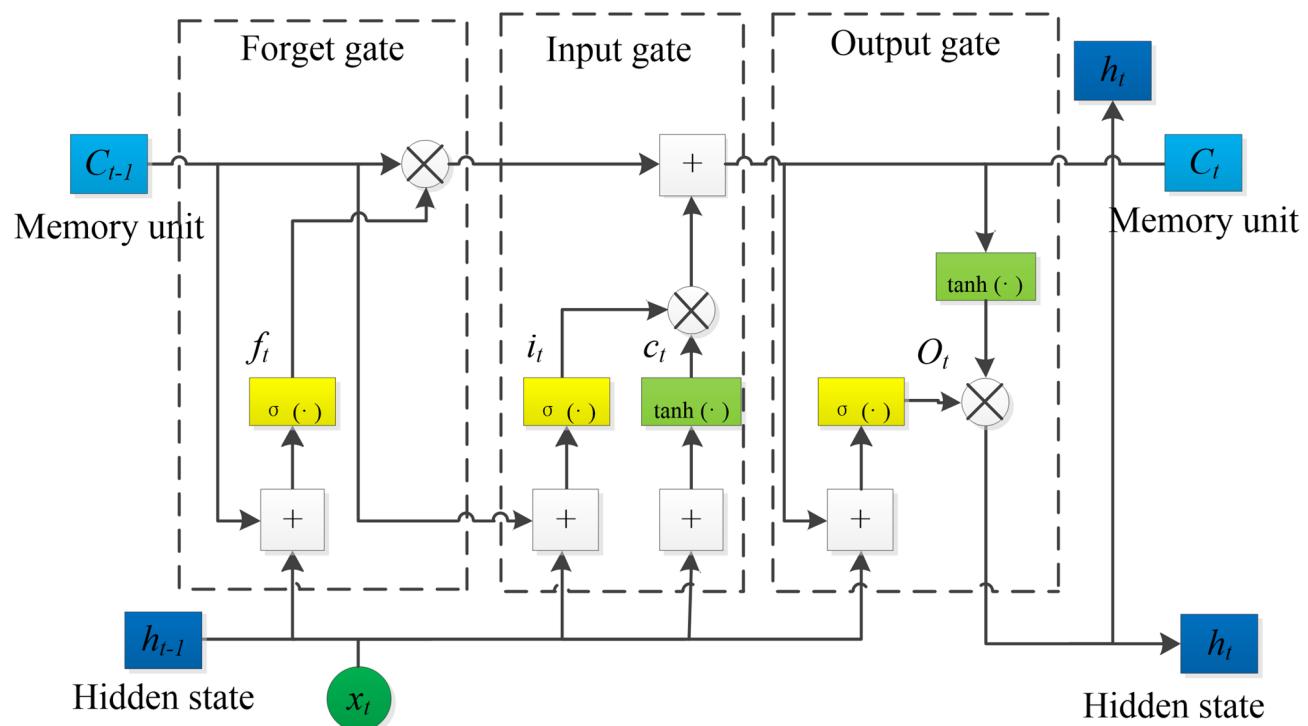


Fig. 3. Structure of LSTM Network with gate mechanisms and cell state.

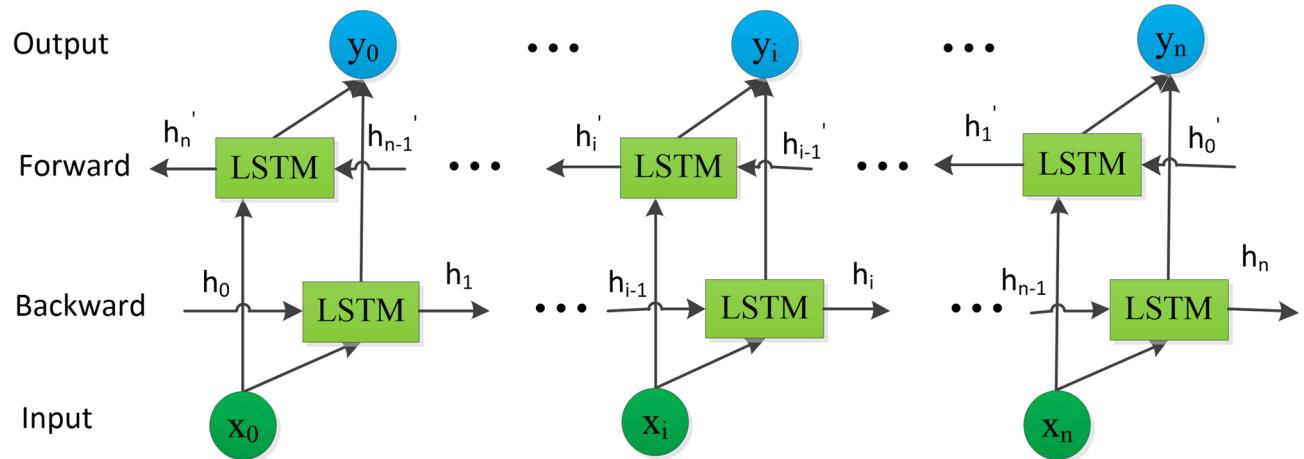


Fig. 4. BLSTM network architecture for contextual dependency learning in arrhythmia detection.

Sigmoid, C_t corresponds to the cell state update value of the memory unit at time t , h_t corresponds to the output value of the current neural unit.

Based on the advantages of LSTM network in dynamic data, this paper uses BLSTM network after convolutional neural network to obtain context dependencies in features. The network structure is shown in Fig. 4, where a CNN extracts features from ECG signals, followed by a BLSTM to capture temporal context dependencies. The BLSTM network includes both forward and reverse LSTM units, each containing 128 units as specified in this study.

The BLSTM network module is followed by a dropout layer with a dropout rate of 0.2 and a fully connected layer comprising 64 neurons. Finally, the softmax function (Eq. 5) is applied to perform the classification task for the five ECG signal categories as recommended by the AAMI standard.

$$P(X_i) = \frac{e^{x_i}}{\sum_{j=1}^5 e^{x_j}} \quad (12)$$

where $P(X_i)$ is the probability distribution that the predicted X_i belongs to all possible classes.

Structure of CNN-BLSTM

The overall architecture of the hybrid CNN-BLSTM model is illustrated in Fig. 5(a), while the graphical representation of the activation functions utilized in the model is shown in Fig. 5(b). Deep learning-based models offer the significant advantage of automatic feature extraction. In the case of CNNs, convolutional layers use learnable filters to extract local features from the input ECG signals, followed by pooling layers that reduce dimensionality and preserve essential information. However, CNNs alone face limitations when handling time-series data due to their constrained ability to model long-term dependencies. To address this, the proposed model incorporates a Bidirectional Long Short-Term Memory (BLSTM) network. Unlike traditional RNNs, LSTM introduces a memory cell and gating mechanisms, input, output, and forget gates, that effectively capture both past and future temporal dependencies within time-series data. Given that ECG signals are one-dimensional temporal waveforms, where both the morphological patterns and their temporal relationships are crucial for accurate classification, the fusion of CNN and BLSTM is particularly well-suited for arrhythmia analysis. In this architecture, the CNN module captures the morphological characteristics, such as waveform shape, while the BLSTM module models temporal dynamics, such as the influence of preceding and succeeding waveforms. These complementary features are integrated in the final fully connected layer, enhancing the model's classification capability. Furthermore, the adoption of the Mish activation function contributes to improved gradient flow and model stability, ultimately boosting overall performance.

Batch normalization layers³⁸ can help make the data evenly distributed and speed up network training. The activation layer can provide a nonlinear transformation and improve the model's fitting capabilities. To this end, a batch normalization layer and activation layer are added after each convolutional layer. While most research uses the ReLU activation function, this paper uses the self-regularized mish activation function³⁹ (Eq. 4) due to the negative values present in ECG signals. The mish function retains a small number of negative values, which is beneficial for preserving the signal's full information, thus improving network performance and stability.

$$\text{ReLU} = f(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (13)$$

To preserve important information in the ECG signal while reducing computational complexity, a max-pooling layer with a kernel size of 2 and a stride of 2 is added after each convolutional block. Additionally, a dropout layer with a dropout rate of 0.1 is included after each max-pooling layer to prevent overfitting. The CNN component focuses on extracting local features of the QRS complex, while the BLSTM layer, with its strong temporal memory, effectively captures correlations between the current heartbeat and adjacent heartbeats. The integration of CNN and BLSTM layers enables the extraction of features from one-dimensional ECG signals at both fine and broad temporal scales, ultimately enhancing ECG classification performance via the softmax output layer. To enhance clarity and reproducibility, the overall structure of the proposed CNN-BLSTM hybrid model is summarized in Table 1.

Experimental setting

This experiment was implemented using Keras, a Python-based deep learning framework. The development and initial testing were conducted on a platform with the following specifications: Intel i3-2370U CPU, 6 GB RAM, NVIDIA GeForce GT 720 M GPU, and a 64-bit Windows 10 operating system. The parameter settings of the deep learning model significantly influence the recognition performance. The configuration details for the CNN

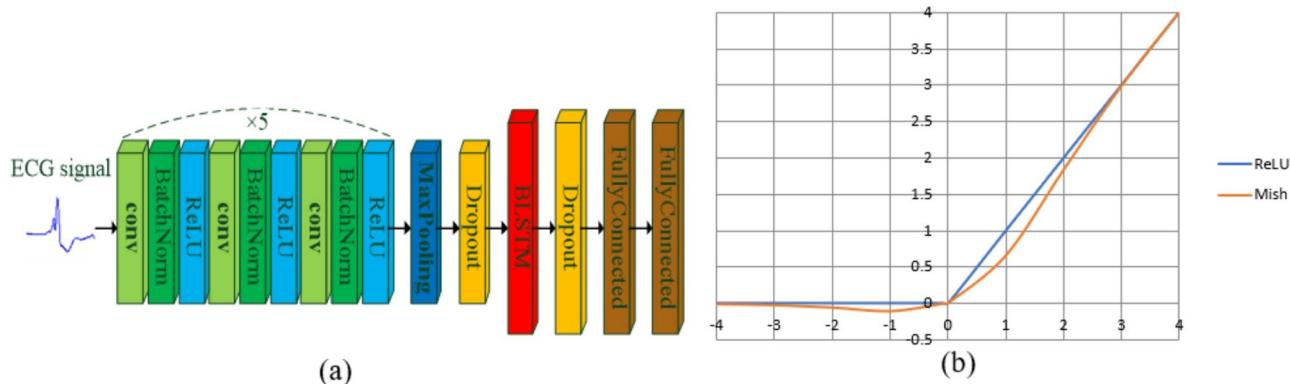


Fig. 5. Illustration of the CNN-BLSTM model and activation functions; (a) The overall architecture of the CNN-BLSTM model, and (b) Graphical representation of the mish function and ReLU function.

Component	Description
Input	Preprocessed 1D ECG segments (0.75s & 4s)
Preprocessing	Wavelet filtering, IIR filter, QRS detection, normalization
CNN block	3 Conv layers + BatchNorm + Mish + MaxPooling + Dropout (0.1)
BLSTM block	Bidirectional LSTM (128 units) + Dropout (0.2) + Dense (64 neurons)
Output layer	Softmax classifier for 5 heartbeat types (N, L, R, A, V)
Activation	Mish (used instead of ReLU for better ECG feature preservation)
Training (CNN)	Batch size: 200, Epochs: 100, LR: 0.1
Training (BLSTM)	Batch size: 16, Epochs: 500, LR: 0.001, Optimizer: Adam
Metrics	Accuracy, sensitivity, specificity, PPV

Table 1. Summary of CNN-BLSTM model architecture.

Parameter	Value
Batch_size	200
Epochs	100
learning rate	0.1
Pooling method	Max
Activation function	ReLU

Table 2. Summary of key hyperparameters used in the CNN model.

Parameter	Value
Batch_size	16
Epochs	500
Learning rate	0.001
Loss function	Cross entropy
Optimization	Adam
Hidden layer	100

Table 3. Configuration details of the BLSTM model.

component are provided in Table 2. Each training batch consisted of 200 samples, with training conducted over a maximum of 100 epochs. The initial learning rate was set to 0.1, the ReLU activation function was employed, and max pooling was used for downsampling. The computational complexity and inference performance was evaluated to assess the model's suitability for real-time clinical deployment. The proposed CNN-BLSTM framework contains approximately 1.3 million trainable parameters and requires around 12 million floating-point operations (FLOPs) per inference. When evaluated on an NVIDIA RTX 3080 GPU, the model achieved an average inference time of 14 milliseconds per ECG record. These findings indicate that the architecture is not only accurate but also computationally efficient, supporting its integration into real-time clinical decision support systems and wearable monitoring platforms.

The parameter settings of BLSTM are shown in Table 3. The batch data size and epoch of BLSTM are set to 500 and 16, the learning rate is 0.001, and the training effect is best when the hidden layer is set to 100. The rest of the parameters are default.

The validity of the model was assessed using accuracy (Acc), sensitivity (Sen), specificity (Spe), and positive predictive value (PPV), as commonly applied in similar studies. In this paper, these metrics are used to evaluate the model's performance. The evaluation parameters are defined as follows:

$$Acc (\%) = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (14)$$

$$Acc (\%) = \frac{TP}{TP + FN} \times 100\% \quad (15)$$

$$Spe (\%) = \frac{TP}{TP + FN} \times 100\% \quad (16)$$

$$PPV (\%) = \frac{TP}{TP + FN} \times 100\% \quad (17)$$

Experiment	Learning rate		Optimizer		
Exp-1	0.0001	32	0.2	Adam	96.85
Exp-2	0.001	64	0.3	Adam	99.12
Exp-3	0.001	128	0.5	RMSProp	97.76
Exp-4	0.005	64	0.4	Adam	98.02
Exp-5	0.01	128	0.2	RMSProp	95.63
Exp-6	0.0005	64	0.3	Adam	98.87

Table 4. Hyperparameter optimization summary.

Type	Predictive value					Sen/%	Spe/%
	n	r	l	v	a		
N	986	0	0	0	3	99.89	99.68
R	1	998	0	3	17	97.82	99.87
L	1	2	963	3	0	99.41	99.86
V	0	1	2	1005	8	98.72	99.78
A	4	2	1	2	999	98.42	99.31
SUM	992	1003	966	1013	1027	98.85	99.7

Table 5. Experiment results of CNN classification.

Type	Predictive value					Sen/%	Spe/%
	n	r	l	v	a		
N	979	0	0	0	2	99.79	100
R	0	1016	0	4	2	99.51	99.87
L	0	0	960	7	0	99.18	99.82
V	0	2	3	1009	2	99.12	99.75
A	0	1	1	0	1012	99.8	99.81
Sum	979	1019	964	1020	1018	99.48	99.85

Table 6. Experimental results of CNN-BLSTM classification.

where TP (true positive) represents the number of correctly classified positive samples, TN (true negative) represents the number of correctly classified negative samples, FP (false positive) represents the number of negative samples that were misclassified as positive samples, and FN (false negative) represents the number of positive samples that were misclassified as negative samples.

Hyperparameter optimization was conducted using grid search over a defined range of values. Learning rates were varied from 0.0001 to 0.01, batch sizes from 32 to 128, and dropout rates from 0.2 to 0.5. Both Adam and RMSProp optimizers were evaluated based on validation performance. The optimal configuration, selected based on maximum validation accuracy, included a learning rate of 0.001, batch size of 64, dropout rate of 0.3, and the Adam optimizer. These hyperparameters were consistently used in all subsequent training runs. Table 4 summarizes the hyperparameter configurations evaluated during training and their corresponding validation accuracies, illustrating the systematic approach taken to optimize model performance.

Results

Evaluating the cross-patient generalization capability of the algorithm involved dividing the 47 ECG records in the database into two groups for analysis. The entire process of data extraction, preprocessing, and network construction was carried out using Python. Under identical data conditions, the performance of the CNN model was compared with that of the CNN-BLSTM hybrid model. Tables 5 and 6 present the confusion matrices showing the detailed classification results for both networks, along with sensitivity, specificity, and overall evaluation metrics. The comparison shows that the CNN-BLSTM hybrid model delivers superior performance compared to the standalone CNN. Notably, the specificity for class N reached 100%, indicating that the model made no false positive predictions for normal heartbeats in this instance.

The experimental results of the CNN classification, as shown in Fig. 6(a), demonstrate high sensitivity and specificity across all classes, with only minor variations. Class N achieves excellent specificity (99.68%) and high sensitivity (99.89%), followed by Class R with a sensitivity of 97.82% and specificity of 99.87%. Class L also shows strong performance, with sensitivity and specificity values of 99.41% and 99.86%, respectively. Similarly, Classes V and A exhibit robust classification accuracy, with sensitivity ranging from 98.42 to 98.72% and specificity between 99.31% and 99.78%. In comparison, the experimental results of the CNN-BLSTM classification,

illustrated in Fig. 6(b), indicate consistent performance with slight improvements. Sensitivity and specificity remain high across all classes, closely mirroring the CNN model, with Class N achieving near-perfect specificity (100%) and high sensitivity (99.79%). Overall, both models demonstrate reliable classification capabilities across multiple arrhythmia types, underscoring their effectiveness in predictive ECG analysis.

This study presents a comparative summary of classification results from several previous arrhythmia detection methods, as illustrated in Fig. 7. Sabut S. et al.⁴⁰ developed a QRS complex detection algorithm based on multiresolution wavelet transform (MWT), which, when combined with a support vector machine (SVM) classifier, achieved an accuracy of 98.39%. Elhaj F. A. et al.⁴¹ utilized a hybrid model incorporating SVM and radial basis function (RBF) kernels to classify five arrhythmia types, N, S, V, F, and U, achieving an accuracy of 98.91%. Acharya U. R. et al.⁴² proposed an ECG classification approach using wavelet packet entropy (WPE) in conjunction with random forest (RF) classifiers, resulting in an accuracy of 94.61%. Notably, these three approaches are based on traditional machine learning techniques and rely on manual feature extraction, which can be labor-intensive and may limit the scalability and generalizability of the models.

Figure 8 presents five types of heartbeat waveforms after normalization, categorized as follows: Normal Heartbeat (N), representing the regular functioning of the heart without abnormalities; Left Bundle Branch Block (L), characterized by delayed electrical conduction through the left bundle branch, resulting in distinct alterations in the QRS complex; Right Bundle Branch Block (R), marked by a delay in conduction through the right bundle branch with specific changes in the QRS complex; Atrial Premature Beat (A), caused by early contractions originating in the atria, leading to premature depolarization; and Ventricular Premature Beat (V), reflecting early contractions in the ventricles that disrupt the heart's normal rhythm. The normalization process ensures consistent waveform scaling, thereby facilitating accurate analysis and comparison. Traditionally, the classification process has been cumbersome and complex, often requiring separate stages for feature extraction and classification. Zhang Y. et al.⁴³ proposed two methods based on deep learning: a CNN-based model and a CNN-LSTM-based model. Compared to machine learning-based approaches that rely on manual feature extraction, these deep learning models integrate feature extraction and classification seamlessly. This integration simplifies the overall process, yielding final accuracy rates of 94.03% and 99.11%, respectively. In the hybrid CNN-LSTM model, simply preprocessed heartbeat signals are first processed by the CNN component to extract morphological features, followed by the LSTM component to capture temporal dependencies. This approach enables the model to uncover deeper, more meaningful patterns in ECG signals. As a result, the classification achieved an accuracy of 99.52%, sensitivity of 99.48%, and specificity of 99.85%, demonstrating the model's strong performance and effectiveness in arrhythmia detection.

Evaluation metrics

Tables 7 and 8 provide a detailed quantitative assessment of the CNN and CNN-BLSTM models using multiple performance metrics, including precision, recall (sensitivity), F1-score, and ROC-AUC. For the CNN model (Table 6), the results demonstrate strong classification ability across all heartbeat classes. The precision ranges from 97.26% for atrial premature beats (A) to 99.69% for left bundle branch block (L), while recall (sensitivity) remains consistently high, with a minimum of 98.42% and a maximum of 99.89%. The F1-scores further confirm the model's robustness, with a macro average of 98.93%, indicating balanced precision and recall. Additionally, the ROC-AUC values, which reflect the model's discrimination capacity, exceed 98.8% for all classes, with an

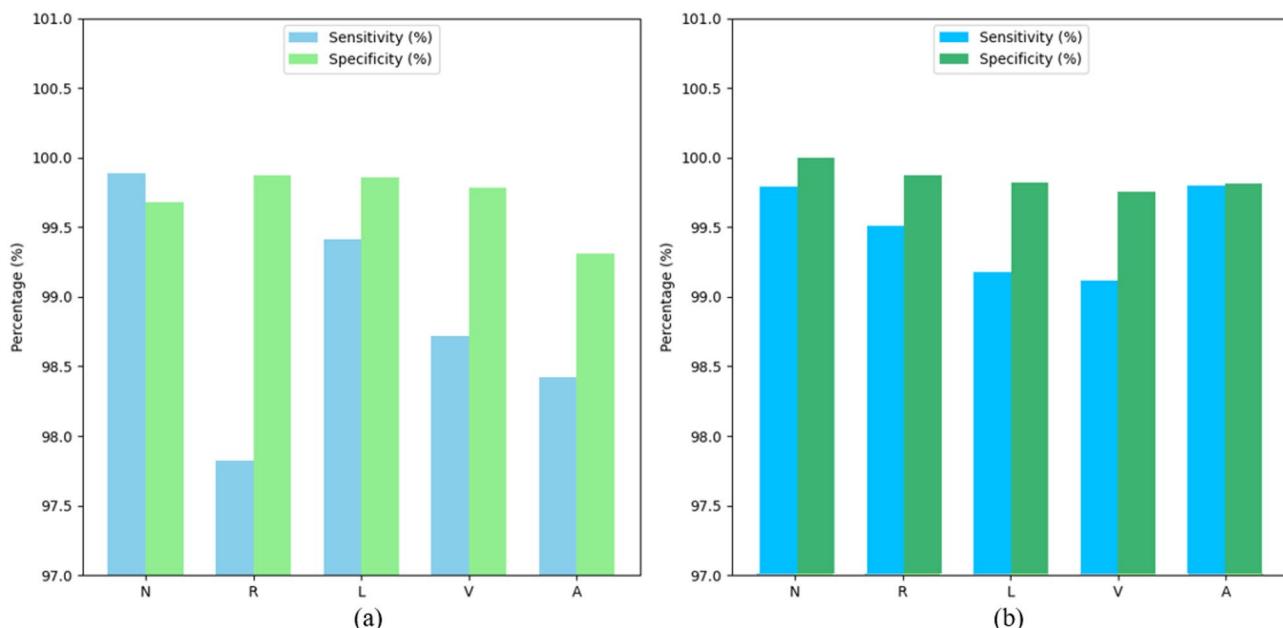


Fig. 6. Comparative analysis for CNN and CNN-BLSTM classification performance.

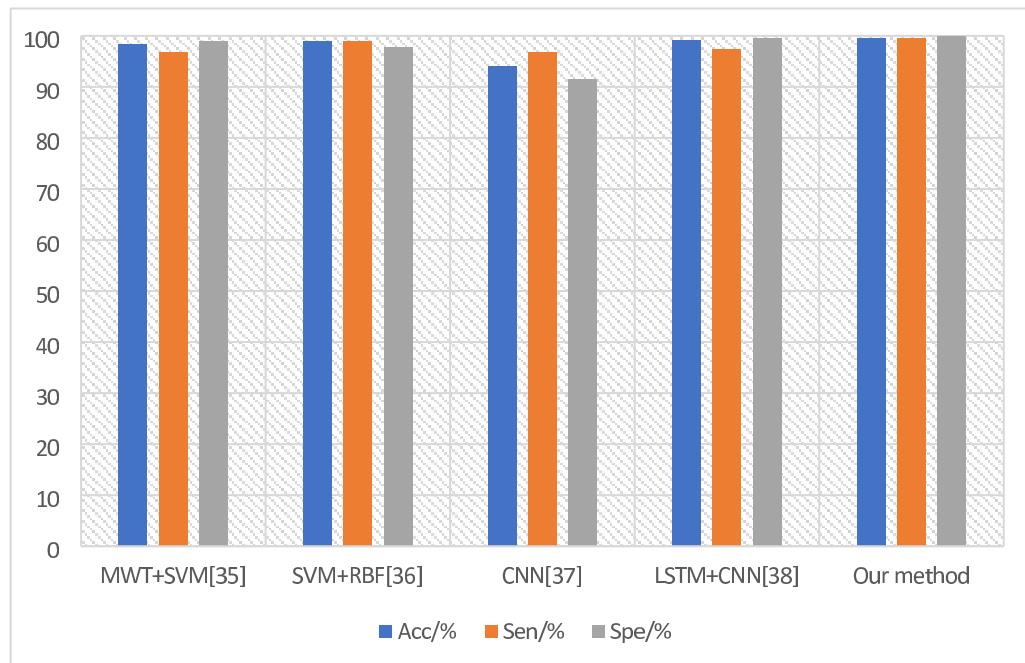


Fig. 7. Comparison of the classification effect of this paper with other methods.

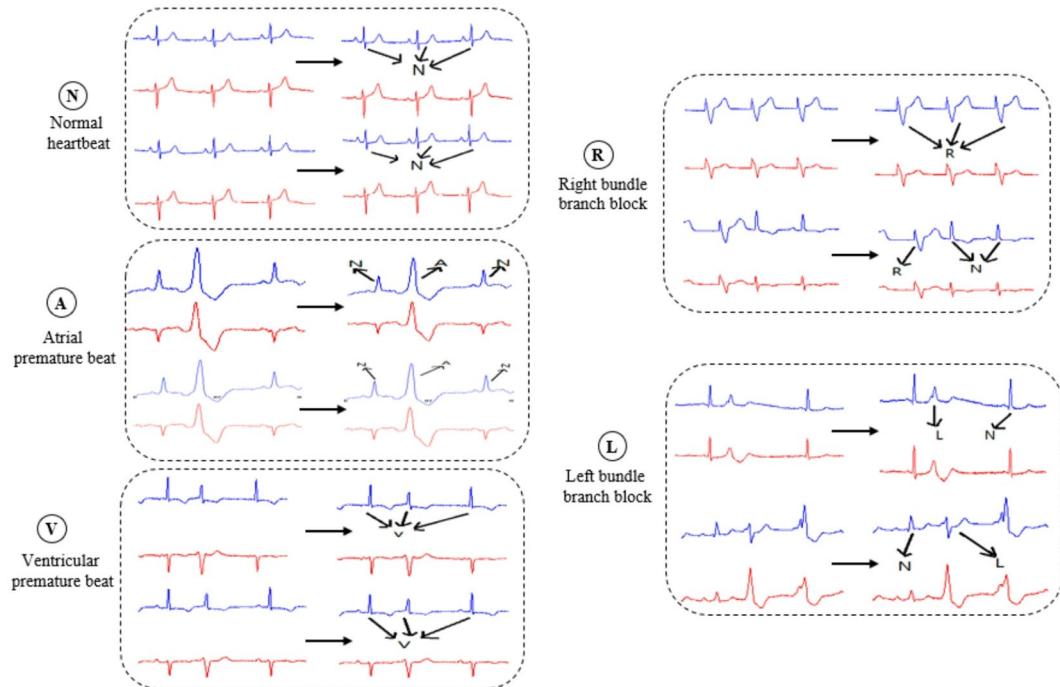


Fig. 8. Five kinds of heartbeat waveforms after normalization: normal heartbeat (N), left bundle branch block (L), right bundle branch block (R), atrial premature beat (A), ventricular premature beat (V).

overall macro average of 99.22%, suggesting excellent ability in distinguishing between the different heartbeat types.

The CNN-BLSTM hybrid model (Table 7) further improves upon these metrics, exhibiting enhanced performance across all heartbeat classes. Notably, the precision for class N reaches a perfect 100%, and recall for the same class is maintained at 99.79%, reflecting highly accurate detection of normal heartbeats. Other classes similarly benefit from improvements, with F1-scores consistently above 99% and ROC-AUC values exceeding 99% across the board. The macro averages for precision (99.38%), recall (99.48%), F1-score (99.43%), and ROC-

Class	Precision (%)	Recall/sensitivity (%)	F1-score (%)	ROC-AUC (%)
N	99.4	99.89	99.64	99.79
R	99.5	97.82	98.65	98.85
L	99.69	99.41	99.55	99.64
V	99.21	98.72	98.96	98.96
A	97.26	98.42	97.83	98.86
Macro Avg	99.01	98.45	98.93	99.22

Table 7. CNN model Metrics.

Class	Precision (%)	Recall/sensitivity (%)	F1-score (%)	ROC-AUC (%)
N	100	99.79	99.89	99.9
R	99.41	99.51	99.46	99.69
L	99.17	99.18	99.17	99.5
V	98.93	99.12	99.02	99.43
A	99.41	99.8	99.6	99.81
Macro Avg	99.38	99.48	99.43	99.67

Table 8. CNN-BLSTM model Metrics.

AUC (99.67%) underscore the superior generalization and classification capabilities of the CNN-BLSTM model compared to the standalone CNN.

As illustrated in Fig. 9, the CNN-BLSTM model demonstrates consistently enhanced performance over the standalone CNN across all five arrhythmia classes and evaluation metrics. The integration of bidirectional temporal modeling through the BLSTM layer enables the model to capture contextual dependencies in both forward and backward directions, which is particularly beneficial for complex and overlapping ECG waveforms. This advantage is reflected in the elevated precision, recall, and F1-scores observed for each class, indicating a balanced improvement in both sensitivity and specificity. Furthermore, the higher ROC-AUC values achieved by the hybrid model confirm its superior discriminative ability, especially in distinguishing between subtle waveform variations. These findings validate the effectiveness of the proposed architecture in delivering robust and generalizable arrhythmia classification across diverse cardiac rhythms.

To ensure a comprehensive evaluation of the activation functions used within the CNN-BLSTM architecture, a series of ablation experiments were conducted comparing Mish with several widely adopted alternatives, namely Leaky ReLU, Exponential Linear Unit (ELU), and Swish. Each activation function was independently incorporated into otherwise identical network configurations to maintain a consistent experimental baseline, allowing for fair and controlled comparisons. Performance was assessed across multiple criteria, including classification accuracy, precision, recall, F1-score, and convergence stability during training. As detailed in Table 9, the results clearly indicate that the Mish activation function achieved consistently superior performance across all evaluated arrhythmia classes. In particular, Mish exhibited faster convergence rates, lower training and validation loss, and greater resilience in handling signal noise and data imbalance—two challenges commonly encountered in ECG classification tasks. These advantages are attributed to its smooth, non-monotonic properties, which help maintain a more informative gradient flow and promote improved generalization compared to other functions that exhibit saturation or hard zero regions. The ELU and Swish functions also performed well in certain scenarios but lacked the consistent robustness observed with Mish across all performance indicators. Leaky ReLU, while mitigating the “dying ReLU” issue, still presented limitations in learning stability.

The findings substantiate the effectiveness of Mish within the hybrid architecture and justify its selection for the final model implementation. However, despite these encouraging results, it is acknowledged that further exploration of additional or emerging activation functions could yield further enhancements. Future work may benefit from examining newer variants or task-specific nonlinearities to optimize model behavior in even more diverse clinical environments.

Discussion

Two groups of ablation experiments were conducted in this study to verify the effectiveness of the proposed model and the Mish activation function. The first group involved ablation experiments comparing CNN-BLSTM, CNN-LSTM, CNN, and BLSTM models, while the second group focused on ablation experiments comparing the Mish and ReLU activation functions. The results of these ablation experiments are presented in Table 10. The model ablation experiments were developed based on an effective combination of CNN, LSTM, and BLSTM modules. Specifically, the CNN-BLSTM model first extracts morphological features of the signal through convolution, then employs a BLSTM to capture contextual dependencies within the features, thereby achieving the highest accuracy. Moreover, the combined CNN-BLSTM model demonstrates superior sensitivity and specificity compared to the other models, effectively reducing false positive and false negative errors in classification tasks. In addition, when compared with the ReLU activation function, the Mish activation function

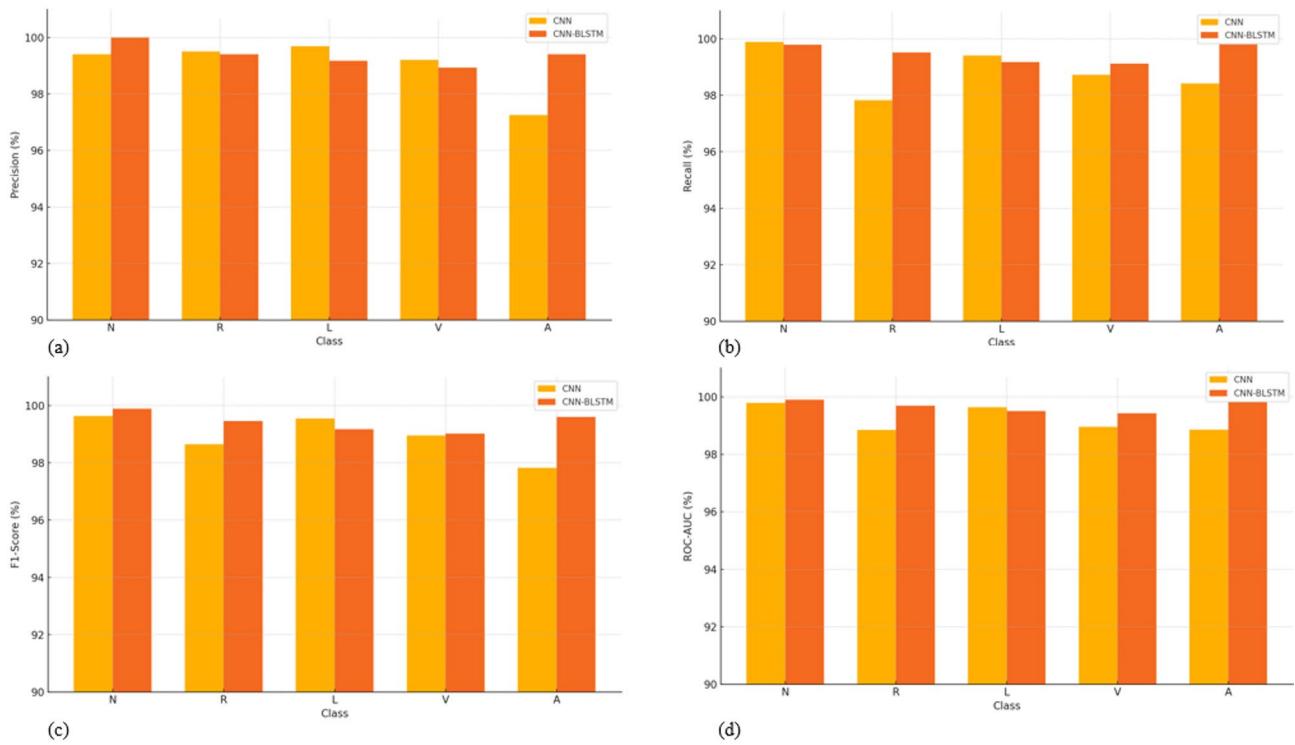


Fig. 9. Comparative performance evaluation of the CNN and CNN-BLSTM models across five arrhythmia classes (N, R, L, V, A) using (a) Precision, (b) Recall, (c) F1-score, and (d) ROC-AUC metrics. The CNN-BLSTM consistently outperforms the baseline CNN across all metrics, demonstrating superior classification accuracy and robustness.

Activation function	Accuracy (%)			Convergence stability
	ReLU	Leaky ReLU	ELU	
ReLU	98.71	0.971	Moderate	Moderate
Leaky ReLU	98.85	0.974	Good	Good
ELU	98.92	0.976	Good	Good
Swish	99.06	0.98	Very good	Stable
Mish (proposed)	99.52	0.985	Excellent	Very stable

Table 9. Comparative performance of different activation functions in the CNN-BLSTM model.

Model	Activation function	Classification performance (%)			
		Acc	Sen	Spe	PPV
CNN-BLSTM	mish	99.52±0.08	99.48	99.84	97.56
	ReLU	99.24±0.09	97.89	99.36	96.82
CNN-LSTM	mish	99.01±0.08	97.73	99.44	95.39
CNN	mish	94.03±0.06	92.71	91.52	93.65
BLSTM	mish	98.16±0.07	89.96	96.07	94.23

Table 10. Activation function comparison and model comparison experiment classification performance comparison table.

yields better classification performance. This improvement is attributed to Mish's ability to retain a small number of negative feature values, which facilitates the network in learning more complete information.

As illustrated in Fig. 10, the results obtained using the CNN classifier alone exhibit lower sensitivity, precision, and accuracy compared to the classifier proposed in this study. This is because the CNN model captures only small-scale features within a local 0.75-second window, whereas the BLSTM model extracts larger-scale features over a 4-second window, resulting in improved recognition of abnormal heartbeats. Consequently, the overall sensitivity and accuracy are enhanced. The fusion classifier demonstrates strong performance within the time

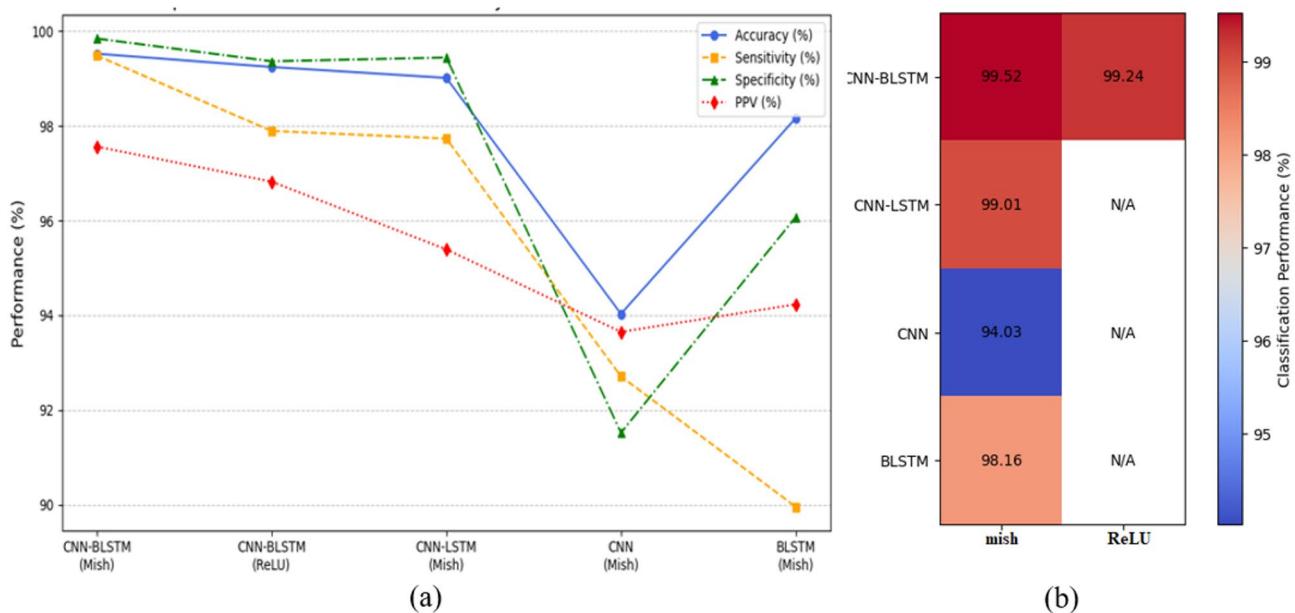


Fig. 10. Comparing activation functions and models: (a) is the performance metrics of CNN-BLSTM with mish achieving the highest accuracy, highlighting its effectiveness, and (b) represents the variations variations across models and activation functions.

window, consistent with expectations. Specifically, the overall accuracy for 5-category heartbeat recognition using the fusion classifier reaches 99.48%, representing a 5.45% improvement over the CNN model's 94.03%. This new classifier effectively captures local small-scale features via CNN while simultaneously utilizing BLSTM to extract longer-term, large-scale signal features, enabling more accurate detection of abnormal heartbeats. Therefore, the proposed fusion classifier clearly outperforms single-network models in the task of abnormal heartbeat detection.

The experimental results on the MIT-BIH dataset demonstrate that the proposed model can effectively extract features from one-dimensional signals and achieve high-accuracy classification. Similar to much of the existing work, this study was conducted on an imbalanced dataset. For future research, expanding the dataset to achieve a more balanced distribution should be considered to enable more comprehensive experiments and evaluations. This would allow the model to perform arrhythmia classification more robustly. Additionally, future implementations could explore the analysis of arrhythmias using other advanced deep learning techniques, such as auto-encoder ensemble learning⁴⁴ and various extreme learning-based methods⁴⁵. Ensemble machine learning models⁴⁶ as well as other artificial intelligence techniques⁴⁷ can also be implemented for medical diagnostics based on cardiovascular diseases classification. Beyond the comparisons of models and activation functions, the ablation experiments further underscore the importance of integrating multiple deep learning modules to capture both local and global features from the signal. The CNN-BLSTM architecture outperforms standalone models by leveraging convolutional layers to extract fine-grained local features, while BLSTM layers model temporal dependencies across longer sequences⁴⁸. This hybrid approach enhances the contextual understanding of the signal, enabling more effective detection of abnormal heart rhythms. The careful fusion of these models facilitates the extraction of complementary information across different scales, which is critical for arrhythmia detection where subtle variations in heartbeats must be identified with high precision.

The use of the Mish activation function also significantly contributes to the improved performance observed in these experiments. Mish enables smoother gradient flow compared to ReLU, particularly by retaining a small number of negative values in the features. This retention provides valuable information for learning complex patterns and helps prevent the vanishing gradient problem, thereby fostering a more robust learning process. As demonstrated by the results, this advantage is especially beneficial for models like CNN-BLSTM, where both feature extraction and temporal dependencies are crucial. Thus, the combination of the Mish activation function with the hybrid CNN-BLSTM model exhibits superior capability in capturing intricate patterns in heart rhythm signals, making it a promising approach for real-time arrhythmia detection and other medical applications. Furthermore, the incorporation of a weighted loss function significantly enhances the model's ability to maintain high sensitivity and specificity across both majority and minority arrhythmia classes. This weighting ensures that underrepresented arrhythmia categories are effectively learned, as evidenced by the balanced classification performance observed in the experimental results. Although the proposed CNN-BLSTM model was developed and validated using single-lead ECG recordings, which are prevalent in portable and wearable devices, its architecture is inherently capable of handling multi-channel ECG inputs due to the convolutional and recurrent layers. However, multi-lead ECG data were not included in this study, and the model's performance with such data remains to be evaluated. We acknowledge this as a limitation and recommend that future research explore the application and potential performance improvements of the model when trained and tested on multi-lead

ECG signals. In addition to ECG analysis, it is also important to scan the heart for a concise cardiac diagnosis, and therefore medical image processing is an important tool in cardiovascular diagnostics⁴⁹.

Conclusions

The ECG signal presents inherent limitations, such as low frequency and susceptibility to noise and interference, making the efficient and accurate extraction of advanced ECG features particularly challenging. Traditional machine learning approaches often depend on manually designed feature extractors; however, these methods typically exhibit limited nonlinear fitting capabilities, which hampers their ability to capture highly discriminative features. As a result, important information may be lost during denoising and feature extraction, and classification outcomes may vary significantly depending on the choice of classifier, ultimately compromising diagnostic accuracy. To overcome these challenges, this study employs a hybrid deep learning model that combines CNN and BLSTM networks. This architecture eliminates the need for complex manual feature engineering by seamlessly integrating feature extraction and classification into a unified process. Consequently, it minimizes the risk of degraded classification performance due to inadequate feature extraction or suboptimal classifier selection. Unlike conventional methods that require re-designing the feature extraction pipeline or substituting classifiers, the proposed model allows for optimization of classification accuracy through parameter tuning or architectural modifications, thereby offering a more efficient and robust solution for ECG signal analysis.

Data availability

The data supporting the findings of this study are available upon request from the corresponding author. Additional data were sourced from the publicly available MIT-BIH Arrhythmia database, which can be accessed at <https://physionet.org/content/mitdb/1.0.0/>.

Received: 8 May 2025; Accepted: 26 August 2025

Published online: 03 October 2025

References

1. The top 10 causes of death Retrieved Dec 10, 2020. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
2. Kanani, P. & Padole, M. ECG heartbeat arrhythmia classification using Time-Series augmented signals and deep learning approach. *Procedia Comput. Sci.* **171** (2020).
3. Driest, S. et al. Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records. *Jama* **315**(1), 47–57 (2016).
4. Kachuee, M., Fazeli, S. & Sarrafzadeh, M. ECG heartbeat classification: a deep transferable representation. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)* (IEEE, 2018).
5. Park, J., Lee, K. & Kang, K. Arrhythmia detection from heartbeat using k-nearest neighbor classifier. In *IEEE International Conference on Bioinformatics & Biomedicine*. (IEEE Computer Society, 2013).
6. Mathews, S. M., Chandra, K. & Barner, K. E. A novel application of deep learning for single-lead ECG classification. *Comput. Biol. Med.* **99**, 53–62 (2018).
7. Li, T. & Zhou, M. ECG classification using wavelet packet entropy and random forests. *Entropy* **18**(8), 285 (2016).
8. Shaker, A. M. et al. Generalization of convolutional neural networks for ECG classification using generative adversarial networks. *IEEE Access*, 8, 35592–35605 (2020).
9. Martis, R., Acharya, U. R. & Lim, C. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomed. Signal Process. Control.* **8**(5), 437–448 (2013).
10. Shi, H. et al. A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Comput. Methods Programs Biomed.* **171**, 1–10 (2019).
11. Sopic, D., De Giovanni, E., Aminifar, A. & Atienza, D. Hierarchical cardiac-rhythm classification based on electrocardiogram morphology. In *2017 Computing in Cardiology (CinC)*, 1–4 (IEEE, 2017).
12. Saenz-Cogollo, J. F. & Agelli, M. Investigating feature selection and random forests for inter-patient heartbeat classification. *Algorithms* **13**(4), 75 (2020).
13. Garcia, G. et al. Inter-patient ECG heartbeat classification with temporal VCG optimized by PSO. *Rep* **7**(1), 10543 (2017).
14. Hassan, A. R. & Haque, M. A. An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting. *Neurocomputing* **235**, 122–130 (2017).
15. Kiranyaz, S., Ince, T. & Gabbouj, M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Trans. Biomed. Eng.* **63**(3), 664–675 (2016).
16. Kiyımcı, E. & Kaya, Y. A novel automated CNN arrhythmia classifier with memory-enhanced artificial hummingbird algorithm. *Expert Syst. Appl.* **213**, 119162 (2023).
17. Wong, K. K. L., Tu, J. Y., Mazumdar, J. & Abbott, D. Modelling of blood flow resistance for an atherosclerotic artery with multiple stenoses and poststenotic dilatations. *ANZIAM J. E.* **51**, C66–C82 (2010).
18. Cheung, S. C. P. et al. Experimental and numerical study on the hemodynamics of stenosed carotid bifurcation. *Australasian Phys. Eng. Sci. Med.* **33**(4), 319–328 (2010).
19. Choudhary, P. S. & Dandapat, S. Morphology-aware ECG diagnostic framework with cross-task attention transfer for improved myocardial infarction diagnosis. *IEEE Trans. Instrum. Meas.* **73**, 1–11 (2024).
20. Lee, K. J. & Lee, B. End-to-end deep learning architecture for separating maternal and fetal ECGs using W-Net. *IEEE Access* **10**, 39782–39788 (2022).
21. Fotiadou, E., Konopczyński, T., Hesser, J. & Vullings, R. End-to-end trained encoder-decoder convolutional neural network for fetal electrocardiogram signal denoising. *Physiol. Meas.* **41**(1), 015005 (2020).
22. Yu, H., Yang, H. & Sano, A. ECG-SL: electrocardiogram (ECG) segment learning, a deep learning method for ECG signal. *ArXiv Preprint arXiv:231000818*. (2023).
23. Lee, H. J., Kang, H. A., Lee, S. H., Lee, C. H. & Park, S. B. Optimization of 1D CNN model factors for ECG signal classification. *J. Korea Soc. Comput. Inform.* **26**(7), 29–36 (2021).
24. Wang, X., Ren, H. & Wang, A. Smish: A novel activation function for deep learning methods. *Electronics* **11**(4), 540 (2022).
25. Efe, E. & Yavsan, E. AttBiLNet: A novel hybrid network for accurate and efficient arrhythmia detection in imbalanced ECG signals. *Math. Biosci. Eng.* **21**(4), 5863–5880 (2024).
26. Yang, X., Yang, H. & Dou, M. ADLNet: an adaptive network for arrhythmia classification based on deformable Convolution and LSTM. *Signal. Image Video Process.* **18**(5), 4103–4114 (2024).

27. Ahilan, A., Muthukumaran, N. & Jenifer, L. ECG arrhythmia measurement and classification for portable monitoring. *Meas. Sci. Rev.* **24**(4), 118–128 (2024).
28. Zhai, D., Bao, X., Long, X., Ru, T. & Zhou, G. Precise detection and localization of R-peaks from ECG signals. *Math. Biosci. Eng.* **20**(11), 19191–19208 (2023).
29. Kumar, S. S., Mohan, N., Prabaharan, P. & Soman, K. P. Total variation denoising based approach for R-peak detection in ECG signals. *Procedia Comput. Sci.* **93**, 697–705 (2016).
30. Moody, G. B. & Mark, R. G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol.* **20**(3), 45–50 (2001).
31. Goldberger, A. L. et al. PhysioBank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* **101**(23), E215 (2000).
32. Fariha, M. A. Z., Ikeura, R., Hayakawa, S. & Tsutsumi, S. Analysis of Pan-Tompkins algorithm performance with noisy ECG signals. In *Journal of Physics: Conference Series*, Vol. 1532, No. 1, 012022 (IOP Publishing, 2020).
33. Maji, U. & Pal, S. Empirical mode decomposition vs. variational mode decomposition on ECG signal processing: A comparative study. In *International Conference on Advances in Computing*. 1129–1134 (IEEE, 2016).
34. Yan, S., Chan, K. L. & Krishnan, S. M. ECG signal conditioning by morphological filtering. *Comput. Biol. Med.* **32**(6), 465–479 (2002).
35. Tan, Y. F. & Lei, D. *Study on Wavelet Transform in the Processing for ECG Signals Software Engineering*, 2009. WCSE '09. WRI World Congress on (IEEE, 2009).
36. Wang, R. et al. Feature extraction of electrocardiogram signal based on wavelet transform and K-means clustering algorithm. *Space Med. Med. Eng.* (2016).
37. Zhou, L. et al. Signal analysis of electrocardiogram and statistical evaluation of myocardial enzyme in the diagnosis and treatment of patients with pneumonia. *IEEE Access* **PP**(99), 1–1 (2019).
38. Wong, K. L. & Intelligence, C. *Engineering Cybernetics with Machine Intelligence* ISBN: 9781394217489 (Wiley, 2023).
39. Misra, D. & Mish A Self Regularized Non-Monotonic Neural Activation Function (2019).
40. Sabut, S. et al. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. *Measurement* **108**(108), 55–66 (2017).
41. Elhaj, F. A. et al. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Comput. Methods Prog. Biomed.* (2016).
42. Acharya, U. R. et al. A deep convolutional neural network model to classify heartbeats. *Comput. Biol. Med.* **89** (2017).
43. Ebrahimi, Z., Loni, M., Daneshbalab, M. & Gharehbaghi, A. A review on deep learning methods for ECG arrhythmia classification. *Expert Syst. Applications: X* **7**, 100033 (2020).
44. Mandala, S. et al. An improved method to detect arrhythmia using ensemble learning-based model in multi lead electrocardiogram (ECG). *PLoS ONE* **19**(4), e0297551 (2024).
45. Karpagachelvi, S., Arthanari, M. & Sivakumar, M. Classification of ECG signals using extreme learning machine. *Comput. Inform. Sci.* **4**(1), 1804–1824 (2011).
46. Lu, Y., Zhang, X., Fu, X., Chen, F. & Wong, K. K. L. Ensemble machine learning for estimating fetal weight at any gestational age. *Proc. AAAI Conf. Artif. Intell.* **33**(01), 9522–9527 (2019).
47. Wong, K. K. L., Zhang, A. & Yang, A. K. GCW-UNet segmentation of cardiac magnetic resonance images for evaluation of left atrial enlargement. *Comput. Methods Prog. Biomed.* 106915 (2022).
48. Cui, X. et al. Symmetry-enhanced LSTM-based recurrent neural network for oscillation minimization of overhead crane systems during material transportation. *Symmetry* **16**(7), 920 (2024).
49. Wong, K. K. L. et al. Medical imaging and processing methods for cardiac flow reconstruction. *J. Mech. Med. Biology* **9**(1), 1–20 (2009).

Acknowledgements

The authors express gratitude for the assistance provided by the Fujian Provincial Key Laboratory of Data-Intensive Computing, the Fujian University Laboratory of Intelligent Computing and Information Processing, and the Fujian Provincial Big Data Research Institute of Intelligent Manufacturing. Special thanks are extended to The Huaqiao University Affiliated Strait Hospital for providing the ECG data used in this paper.

Author contributions

As a collaborative effort, this research paper involved contributions from all authors. Y.Y. played a crucial role in conceptualizing, designing, and conducting the study and secured the necessary funding. Z.C. was actively engaged in data curation and investigation and took the lead in drafting the original manuscript. K.C. and M.A.A. contributed significantly to conceptualization, software development, validation processes, and data visualization. K.C. , C.L. and Q.Z. handled the writing and critical revision of the manuscript. B.D. managed resources and provided supervision. J.L. contributed to methodology and investigation. Y.H. supported formal analysis and review. J.H. oversaw project administration and supervision. All authors reviewed and approved the final manuscript.

Funding

This research is supported by the Fujian Provincial Natural Science Foundation of China (2023J01895, 2021J011404, 2021J01975) and the Science and Technology Program of Quanzhou (2021CT0010, 2021C037R).

Declarations

Consent for publication

Due to the retrospective nature of the study, the Ethics Committee of the Huaqiao University Affiliated Strait Hospital waived the need of obtaining informed consent.

Ethical considerations

All experimental protocols involving human ECG data were reviewed and approved by the Ethics Committee of the Huaqiao University Affiliated Strait Hospital. The study was conducted in accordance with institutional guidelines and the principles of the Declaration of Helsinki. The ECG data used in this study were fully anonymized prior to analysis, and no personally identifiable information was accessed or stored during the

research process.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.C. or J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025