



OPEN ACCESS

EDITED BY

Kuanquan Wang,
Harbin Institute of Technology, China

REVIEWED BY

Laura Burattini,
Marche Polytechnic University, Italy
Dongping Du,
Texas Tech University, United States
Muhammad Adib Uz Zaman,
University of Cincinnati, United States
Yang Liu,
Baidu, China

*CORRESPONDENCE

Bing Yao,
✉ byao3@utk.edu

RECEIVED 27 December 2023

ACCEPTED 25 March 2024

PUBLISHED 09 April 2024

CITATION

Xie J, Stavrakis S and Yao B (2024), Automated identification of atrial fibrillation from single-lead ECGs using multi-branching ResNet.

Front. Physiol. 15:1362185.

doi: 10.3389/fphys.2024.1362185

COPYRIGHT

© 2024 Xie, Stavrakis and Yao. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automated identification of atrial fibrillation from single-lead ECGs using multi-branching ResNet

Jianxin Xie¹, Stavros Stavrakis² and Bing Yao^{3*}

¹School of Data Science, University of Virginia, Charlottesville, VA, United States, ²Health Sciences Center, University of Oklahoma, Oklahoma City, OK, United States, ³Department of Industrial and Systems Engineering, University of Tennessee at Knoxville, Knoxville, TN, United States

Introduction: Atrial fibrillation (AF) is the most common cardiac arrhythmia, which is clinically identified with irregular and rapid heartbeat rhythm. AF puts a patient at risk of forming blood clots, which can eventually lead to heart failure, stroke, or even sudden death. Electrocardiography (ECG), which involves acquiring bioelectrical signals from the body surface to reflect heart activity, is a standard procedure for detecting AF. However, the occurrence of AF is often intermittent, costing a significant amount of time and effort from medical doctors to identify AF episodes. Moreover, human error is inevitable, as even experienced medical professionals can overlook or misinterpret subtle signs of AF. As such, it is of critical importance to develop an advanced analytical model that can automatically interpret ECG signals and provide decision support for AF diagnostics.

Methods: In this paper, we propose an innovative deep-learning method for automated AF identification using single-lead ECGs. We first extract time-frequency features from ECG signals using continuous wavelet transform (CWT). Second, the convolutional neural networks enhanced with residual learning (ResNet) are employed as the functional approximator to interpret the time-frequency features extracted by CWT. Third, we propose to incorporate a multi-branching structure into the ResNet to address the issue of class imbalance, where normal ECGs significantly outnumber instances of AF in ECG datasets.

Results and Discussion: We evaluate the proposed Multi-branching Resnet with CWT (CWT-MB-Resnet) with two ECG datasets, i.e., PhysioNet/CinC challenge 2017 and ECGs obtained from the University of Oklahoma Health Sciences Center (OUHSC). The proposed CWT-MB-Resnet demonstrates robust prediction performance, achieving an F1 score of 0.8865 for the PhysioNet dataset and 0.7369 for the OUHSC dataset. The experimental results signify the model's superior capability in balancing precision and recall, which is a desired attribute for ensuring reliable medical diagnoses.

KEYWORDS

convolutional neural network, residual network, wavelet transform, multi-branching outputs, ECG signal analysis, imbalanced data, atrial fibrillation

1 Introduction

Cardiovascular diseases have been the leading cause of mortality globally. The World Health Organization (WHO) states that about 17.9 million people perish due to cardiovascular disease each year (World Health Organization, 2024), contributing 32% to

the worldwide death toll (University of Washington, 2024). Atrial fibrillation (AF) is the most common cardiac arrhythmia caused by uncoordinated electrical activities in the atria (Nesheiwat et al., 2023). Although AF itself does not lead to a lethal condition, it will substantially increase the risk of catastrophic diseases such as heart failure, stroke, and sudden death (Lubitz et al., 2013; Bernstein et al., 2021). The prevalence of AF plagues over 2.7 million people in the United States, and this number is estimated to rise to 12.1 million in 2030, as the population ages (Colilla et al., 2013). In healthcare practice, the electrocardiogram (ECG) is a cost-effective and noninvasive medical approach to record the electrical signals on the body surface as a reflection of cardiac health conditions (Yao and Yang, 2016; Yao and Yang, 2020; Yao et al., 2021; Xie and Yao, 2023).

Historically, the utilization of ECG for cardiac monitoring has been substantially constrained by the need for expensive equipment and the involvement of specialized medical doctors to interpret complex ECG recordings. However, recent advancements in portable ECG sensors, such as the AliveCor (aliveCor, 2024), AD8232 (Analog Devices, 2024), and consumer-grade devices like the smartwatch (Isakadze and Martin, 2020), have revolutionized the way to detect heart abnormalities. These portable devices now enable the capture of high-fidelity ECG signals outside of traditional clinical settings. While multi-lead ECGs provide comprehensive cardiac activity information, single-lead ECGs make cardiac monitoring more accessible and less obtrusive for long-term rhythm surveillance or frequent measurements (Abdou and Krishnan, 2022). This is especially valuable in ambulatory settings, home monitoring, and situations where rapid and non-invasive monitoring is desired. Single-lead ECGs offer a simplified yet effective method for the early detection of AF and other cardiac anomalies (Boriani et al., 2021).

In conjunction with advanced sensing technologies, there has been a parallel development in machine learning methodologies. Given the prevalence of AF, a significant number of machine learning models have been developed specifically for the task of distinguishing AF from normal heart rhythms. Traditional machine learning models focus on extracting morphological features and heart rate variability from ECG signals to detect AF, which depends heavily on manual feature engineering (Ye et al., 2012; Da Silva-Filarder and Marzbanrad, 2017; Athif et al., 2018). Deep Neural Network (DNN), which does not require explicit feature engineering, is another powerful tool that has achieved promising results in data-driven disease detection. Various DNN-based models such as convolutional and recurrent neural networks (i.e., CNNs, RNNs) have been designed for AF detection and outperformed conventional machine learning methods (Andreotti et al., 2017; Schwab et al., 2017; Gao et al., 2021). Despite the performance improvement achieved by DNNs in detecting AF with single-lead ECG, there remains potential for further prediction enhancements. Four major challenges remain to be tackled: 1) ECG recordings collected from clinics are often in Portable Document Format (PDF). An effective preprocessing procedure is needed to retrieve digital ECG signals from the PDFs before being fed to the machine learning models. 2) ECG signals are generally composed of a wide spectrum of frequency components. DNN models built upon raw ECG time series may not fully exploit the time-frequency information inherent in the signals. 3) Note that the learning capacity for a DNN often increases when the network goes deeper.

However, the deeper structure can result in gradient dissipation problems, leading to unsatisfactory prediction performance. 4) Data-driven AF detection also suffers from the common issue of imbalanced data in machine learning (e.g., AF samples are much less compared to normal ECGs). The classifier directly built from the imbalanced data will generate biased and inaccurate predictions.

In this paper, we develop an automatic AF detector based on continuous wavelet transform (CWT) and 18-layer Residual Neural Network (ResNet18) with a multi-branching structure (CWT-MB-ResNet). We first develop a preprocessing procedure to extract ECG signals from ECG PDFs and leverage the CWT to transform the extracted signals into the time-frequency domain. Second, ResNet18 is engaged to alleviate the gradient dissipation problem in deep-structured networks, allowing it to learn deeper features from 2D time-frequency images and achieve better performance. Finally, we propose to incorporate a multi-branching output structure adapted from our prior work (Wang and Yao, 2021) into the ResNet to deal with the issue induced by the imbalanced dataset in AF identification. The multi-branching technique exempts artificial data augmentation and does not require any preassumptions in solving the imbalanced data issue. The performance of the proposed framework is evaluated by two real-world datasets: PhysioNet/CinC challenge 2017 (Goldberger et al., 2000; Clifford et al., 2017) and ECG data obtained from the University of Oklahoma Health Sciences Center (OUHSC). Experimental results show that our CWT-MB-ResNet significantly outperforms existing methods commonly used in current practice.

The rest of this paper is organized as follows: Section 2 presents the literature review of existing data-driven methods for AF detection. Section 3 introduces the data processing details and the proposed prediction method. Section 4 shows the experimental results in AF identification. Section 6 concludes the present investigation.

2 Research background

Traditional machine learning approaches focus on the extraction of ECG morphological features (De Chazal et al., 2004) and heart rate variability information (Park et al., 2009) to identify AF conditions. Those methods are mostly in light of two aspects of AF-altered ECG characteristics: 1) the absence of distinct P waves, which are replaced by irregular fibrillatory waves or F waves as oscillations in low amplitude around the baseline (Ladavich and Ghoraani, 2015); 2) irregular R-R intervals (Oster and Clifford, 2015). Multiple feature-based automation techniques have been proposed to classify AF-altered ECGs, such as linear discriminant analysis (De Chazal et al., 2004), support vector machine (Billeci et al., 2017; Islam et al., 2017), independent component analysis (Ye et al., 2012). When there exists a high level of noise or faulty detection, the performance of feature-extraction methods that solely study the P wave deteriorates significantly due to the chaotic signal baseline introduced by the noise (Larburu et al., 2011). Most R-R interval-based methods (Tateno and Glass, 2001; Lian et al., 2011) usually require long ECG segments to detect AF episodes, and become ineffective when it comes to short ECG signals (less than 60s) or in

the presence of significant sinus arrhythmia or frequent premature atrial contractions (Xia et al., 2018). Moreover, traditional methods require a separate feature extraction process before feeding the data into the classifier, as well as manually establishing the detection rules and threshold. This can be computationally expensive and may not generalize well when applied to a larger population.

In the past few decades, deep learning or deep neural network (DNN) has emerged as a powerful tool for pattern recognition that can learn the abstracted features from complex data and yield state-of-the-art predictions (Mousavi et al., 2019; Xie and Yao, 2022a; Xie and Yao, 2022b; Chen et al., 2022; Wang et al., 2022). As opposed to traditional machine learning, deep learning presents strong robustness and fault tolerance to uncertain factors, which makes it suitable for beat and rhythm classification from ECGs (Tutuko et al., 2021). Moreover, existing research has indicated that deep learning methods demonstrate more efficient and more potent predictive power than classical machine learning methods for AF identification (Cai et al., 2020; Murat et al., 2021). There has been a significant surge in leveraging deep learning for AF detection using single-lead ECGs, showing promising potential in enhancing diagnostic accuracy. We summarized four commonly used network structures in discerning AF samples using single-lead ECGs:

- 1) **Convolutional neural networks (CNNs):** CNNs, specifically 1-dimensional CNNs (1D-CNNs), have been widely applied to extracting hierarchical features from ECG data for distinguishing AF from normal heart rhythms (Andreotti et al., 2017; Fan et al., 2018; Lai et al., 2019; Phukan et al., 2023). For example, Andreotti et al. (2017) balanced the PhysioNet/CinC 2017 dataset by augmenting AF samples from various sources to address the class imbalance issue. They employed a ResNet model with 34 convolutional layers for AF detection, achieving a final F1 score of 0.79. Lai et al. (2019) developed a streamlined two-stream CNN with each stream containing only 8 layers. This model achieved a sensitivity of 89.5% and a specificity of 82.7% on the PhysioBank dataset (PhysioBank, 2000). The extracted cardiac rhythm features, specifically RR intervals and F-wave frequency spectra, served as dual inputs for the neural network. Similarly, Fan et al. (2018) developed a multi-scaled two-stream network with different filter sizes at each stream to capture features of different scales using single-lead ECGs from PhysioNet/Cinc 2017, achieving an F1 score of 0.8355. Phukan et al. (2023) did a systematic experiment on selections of filter size, number of layers, and activation function on multiple standard datasets. They concluded that the best 5-layer CNN with activation function of exponential linear unit and kernel size 4×1 provides the highest accuracy of 99.84% for 5s ECG segments.
- 2) **Recurrent Neural Networks (RNNs):** An RNN is a type of neural network designed to effectively process sequential data by maintaining a memory of previous inputs, making it suitable for classifying time-series signals, e.g., AF detection. For example, Schwab et al. (2017) built an ensemble of RNNs to jointly distinguish AF from normal ECGs, resulting in 0.79 of F1 score on the PhysioNet/Cinc 2017 dataset. Faust et al. (2018) utilized RNNs, specifically the long short-term memory (LSTM) architecture, to analyze ECGs from the MIT-BIH Atrial Fibrillation Database, achieving an accuracy rate of 99.77% for AF detection. Wang et al. (2023a) proposed a dual-path RNN which includes the intra- and inter-RNN modules to study the global and local aspects for end-to-end AF recognition. They used the PhysioNet/Cinc 2017 dataset to validate their model and achieved an F1 score of 0.842. More recently, bidirectional long short-term memory (Bi-LSTM), a type of RNN architecture capable of capturing both past and future context in sequential data, has been used to discern AF. Ramkumar et al. (2022) created an auto-encoder and Bi-LSTM-based network to detect AF among others. This method integrated a reconstruction error from the auto-encoder into the total loss function, leading to a sensitivity of 92% and specificity of 97% on the PhysioNet/Cinc 2017 dataset.
- 3) **CNN-RNNs:** CNN-RNN hybrids combine the morphological feature extraction capabilities of 1D-CNNs with the temporal pattern recognition strengths of RNNs to address complex tasks such as AF detection from ECG signals. For example, Limam et al. (2017) used dual CNNs to process the inputs consisting of both ECGs and heart rates independently, and then the processed features were merged into RNN to learn the temporal patterns, achieving a validated F1 score of 0.856 on the PhysioNet/CinC 2017 dataset. Wang et al. (2020) combined CNN with Bi-LSTM, exploring two concatenation strategies: a parallel concatenation of CNN and Bi-LSTM, and a sequential one where the CNN output feeds into the Bi-LSTM. They evaluated the methods on the MIT-BIH dataset, reporting a final F1 score of 0.82 for the sequential strategy. Zhang et al. developed a model that merges a multi-branch CNN (MCNN) with Bi-LSTM to improve AF detection from short ECG recordings (Zhang et al., 2022). Unlike our multi-branching approach for addressing the imbalanced data issue, their model extracted features from various segments of a single-lead ECG, which were then processed by the Bi-LSTM. They tested the model on the PhysioNet/CinC 2017 dataset, achieving an F1 score of 0.7894.
- 4) **Attention-based networks:** The attention mechanism (Bahdanau et al., 2014; Vaswani et al., 2017) in deep learning dynamically weighs the importance of different input features, allowing models to focus more on relevant data while processing a task. This special capability can facilitate pattern recognition in ECG signals, enhancing the accuracy and efficiency of AF detection. For example, Gao et al. (2021) designed a residual-based temporal attention CNN, generating temporal informative features related to AF, so as to consider the semantic information to achieve better performance. This model achieved an accuracy of 85.43% on the PhysioNet/CinC 2017 dataset. Nankani et al. (2022) investigated the transformer network for AF detection and underscored clinically relevant signal timestamps triggering the diagnosis, achieving an F1 score of 0.87 on the PhysioNet/Cinc 2017 dataset. Rohr et al.

Rohr et al. (2022) explored and assessed two advanced models for AF detection: a transformer-based DualNet architecture and a CNN-LSTM hybrid model, achieving F1 scores of 0.9127 and 0.9072, respectively, on the PhysioNet/CinC 2017 dataset.

As highlighted above (Andreotti et al., 2017; Fan et al., 2018; Lai et al., 2019; Phukan et al., 2023), 1D-CNNs have exhibited their effectiveness in identifying morphological features and comprehending temporal variations in time series data, demonstrating superior capability in AF detection using single-lead ECG signals. However, despite the promising utility of 1D-CNNs in time series analysis, comparative studies in the literature Ullah et al. (2021) and Wu et al. (2018) indicate that 1D-CNNs often yield lower prediction accuracies than their 2D counterparts under similar network configurations for ECG classification tasks.

This discrepancy can be attributed to the richer, more comprehensive information encapsulated in 2D input data, coupled with the inherently superior capacity of 2D CNNs for feature extraction and interpretation.

Owing to the outstanding performance and strong ability in pattern recognition, 2D CNN has been explored for ECG classification by virtue of its capacity to smartly suppress measurement noises and extract pertinent feature maps using convolutional and pooling layers (Huang et al., 2019). For example, Izci et al. Izci et al. (2019) engaged a 2D CNN model to investigate ECG signals for arrhythmia identification. They segmented the ECG signals by heartbeats and directly converted each heartbeat into grayscale images, which served as the input of the 2D CNN model. Similarly, Jun et al. Jun et al. (2018) proposed to combine 2D CNN and data augmentation with different image cropping techniques to classify 2D grayscale images of ECG beats. However, these end-to-end 2D CNNs are directly fed with original ECG beat segments without considering the possible noise contamination. Moreover, the 2D input data were created by directly plotting each ECG beat as a grayscale image with unavowed redundant information residing in the image background. This procedure requires extra storage space for training data and increases the computational burden without extracting relative features inherent in the ECG beats.

ECG signals generally consist of various frequency components, which can be used to identify disease-altered cardiac conditions. Wavelet transform (WT) (Daubechies, 1990; Yao et al., 2017; van Wyk et al., 2019) has been proven to be a useful technique for extracting critical time-frequency information pertinent to disease-altered ECG patterns (Kutlu and Kuntalp, 2012; He et al., 2018). As such, WT is favored as a feature-preprocessing procedure that converts 1D ECG signals into 2D images containing time-frequency features. The resulting 2D feature images then serve as the input of CNNs for ECG classification instead of the original 2D ECG plots. For instance, Xia et al. Xia et al. (2018) engaged the short-term Fourier transform (STFT) and stationary wavelet transform to convert ECG segments into 2D matrices which were then fed into a three-layer CNN for AF detection. Wang et al. Wang et al. (2021) combined the time-frequency features extracted by Continuous Wavelet Transform (CWT) and R-interval features to train a 2D CNN model for ECG signal classification. Wu et al. Wu et al. (2019)

built a 2D CNN based on time-frequency features of short-time single-lead ECGs extracted from three methods, i.e., STFT, CWT, and pseudo Wigner-Ville distribution, to detect arrhythmias. Huang et al. Huang et al. (2019) developed an ECG classification model by transforming ECG signals into time-frequency spectrograms using STFT and feeding them into a three-layer 2D CNN. Li et al. Li et al. (2019) included three different types of wavelet transform (i.e., Morlet wavelet, Paul wavelet, Gaussian Derivative) to create 2D time-frequency images as the input data to the 2D CNN-based ECG classifier. The above literature unequivocally demonstrates that incorporating frequency information through the WT can significantly enhance the efficacy of ECG classification, underscoring the vital role of frequency domain analysis in AF identification.

In addition to effective information extraction from ECG time series, the realization of the full data potential is heavily reliant on advanced analytical models. Although the abovementioned works have validated the superiority of 2D CNN-based approaches, the shallow network structures with a limited number of layers can potentially hinder the extraction of deeper features. Naturally, the capacity for a neural network to learn is enhanced by an increase in the number of layers. However, having a deeper network structure can result in a gradient dissipation problem, which impedes convergence during network training, leading to suboptimal prediction performance. To cope with this issue, the residual neural network (ResNet) has been developed with an important modification, i.e., identity mapping, induced by the skip connection technique (He et al., 2016), which has wide applications in classifying the ECG signals. For example, Jing et al. Jing et al. (2021) developed an improved ResNet with 18 layers for single heartbeat classification. Park et al. Park et al. (2022) used a squeeze-and-excitation ResNet with 152 layers and compared the model performance trained by ECGs from a 12-lead ECG system and single-lead ECG data. Guan et al. Guan et al. (2022) proposed a hidden attention ResNet to capture the deep spatiotemporal features using 2D images converted from ECG signals.

Automated ECG classification also suffers from the long-standing issue of imbalanced data in machine learning. Diverse sampling and synthetic strategies have been proposed to address the imbalanced data issue, which focuses on creating a balanced training dataset out from the original imbalanced data to mitigate the potential bias introduced by imbalanced data distribution during model training (He and Garcia, 2009). Frequently employed techniques consist of random over-sampling and under-sampling, informed adaptive undersampling, and synthetic minority over-sampling technique (SMOTE) (Gao et al., 2019; Wang and Yao, 2021; Qiu et al., 2022). For example, Luo et al. Luo et al. (2021) engaged SMOTE to synthesize minority samples and create a balanced training dataset for automated arrhythmia classification. Ramaraj et al. Ramaraj and Clement Virgeniya, (2021) incorporated an adaptive synthetic sampling process into the training of deep learning models built with gated recurrent units to address the class imbalance problem for ECG pattern recognition. Nurmaini et al. Nurmaini et al. (2020) compared sampling schemes of SMOTE and random oversampling with RNN and concluded that the balanced dataset created by SMOTE significantly improved the classification performance. In addition to fabricating balanced ECG datasets,

Gao et al. [Gao et al. \(2019\)](#) and Petmezas et al. [Petmezas et al. \(2021\)](#) proposed to engage dynamically-scaled focal loss function to suppress the weight of loss corresponding to the majority class, so that their contribution to the total loss is reduced to alleviate the class imbalance problem. However, this method requires the preassumption of a focusing parameter to modulate the effect of the majority class on the total loss. Existing methods mainly focus on using sampling and synthetic strategies or modifying the loss function, little has been done to create new network structures without making extra assumptions and feature engineering to cope with the imbalanced data issue in AF identification from ECG signals.

3 Materials and methods

3.1 Dataset

In this study, two AF databases from different sources, i.e., ECG recordings from PhysioNet/CinC challenge 2017 ([Goldberger et al., 2000](#); [Clifford et al., 2017](#)) and ECG PDFs from OUHSC, are used to evaluate the performance of data-driven detection methods. Both databases consist of short single-lead ECG recordings for AF and non-AF patients. PhysioNet/CinC Challenge 2017 is an open database including 8,528 single-lead ECG signals and their annotations. Among them, 5050 ECG recordings are labeled as normal sinus rhythm while 738 signals are annotated as AF. The sampling frequency of recordings is 300 Hz and the duration of ECG signals varies from 9s to 30s. The OUHSC database contains ECG signals in PDF format with 33 recordings from AF subjects and 227 normal samples, which are annotated by cardiologists from OUHSC. Each recording has a duration of around 30s with a sampling frequency of 60 Hz.

3.2 ECG signal preprocessing

Note that the original ECG recordings from OUHSC are in PDF format, as shown in [Figure 1A](#). It is necessary to accurately extract the numerical ECG readings from the PDF files for further data preprocessing and analysis, which is achieved by the following procedure:

- *Transforming PDF files into gray-scale images represented by 2D-pixel matrices:* We discretize the 2D image into a pixel matrix. Then, each pixel is converted to a fixed number of bits to represent the gray-scale intensity of the corresponding point in the image. As shown in [Figure 1A](#), the ECG signals are displayed in the darkest color on the plot with the color intensity of 1, i.e., $h(m,n) = 1$, while the grid lines appear in a lighter color, i.e., $0 < h(m,n) < 1$, where $h(m,n)$ denotes the color intensity of the pixel at column m and row n . Note that the background color intensity is 0.
- *Removing grid lines from the ECG plot:* We replace the pixel shade values of the grid lines with the background color value: i.e., $h(m,n | h(m,n) < 1) = 0$. This allows the ECG signals to distinguishably stand out, as illustrated in [Figure 1B](#). The quantized image is thus encoded into a binary digital format,

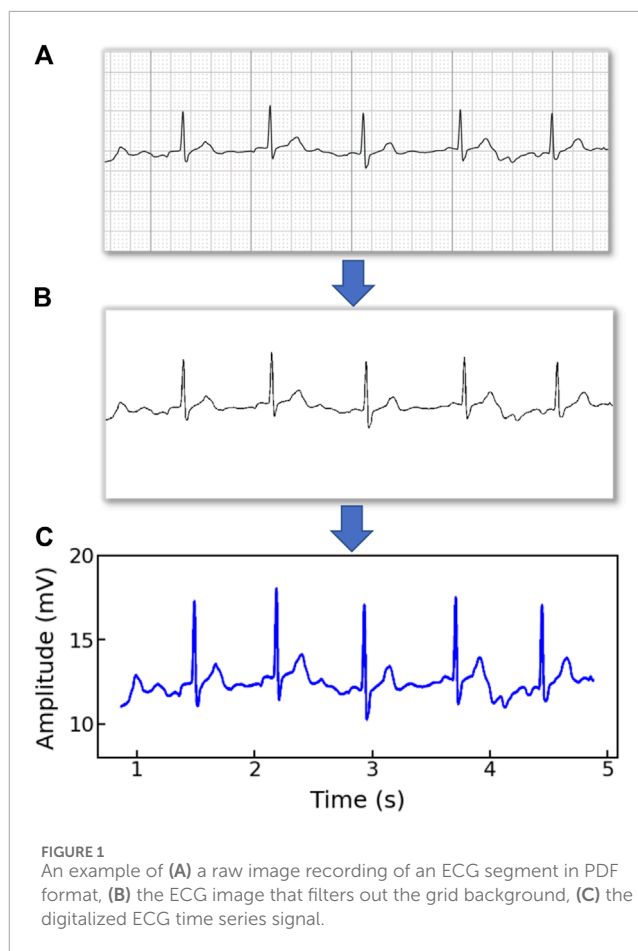


FIGURE 1 An example of (A) a raw image recording of an ECG segment in PDF format, (B) the ECG image that filters out the grid background, (C) the digitalized ECG time series signal.

i.e., black as “1” and white as “0”. As such, the entire ECG image is transformed into a binary digital matrix without the grid lines.

- *Extracting the digital ECG time series:* The positions of black pixels (i.e., ECG signal) in the binary matrix are further extracted, which are represented as a set of (m,n) pairs:

$$S = \{(m,n) | h(m,n) = 1\}$$

The resulting S is then used to reconstruct the digital ECG time series, where m stands for the time course, and n corresponds to the magnitude of the ECG signal. As such, we are able to extract the ECG recordings from the PDFs to digitalized ECG time series signals ([Figure 1C](#)), which will be used for further processing and model training.

Raw ECG recordings are often contaminated by noises, such as baseline wandering, electromyography disturbance, and power-line interference ([Mian Qaisar, 2020](#)), which will negatively impact the information extraction and model performance. In this work, we engage BioSPPy, a toolbox for biosignal processing written in Python, for ECG signal denoising. The BioSPPy library provides comprehensive functions for processing ECG signals including functions for importing ECGs, filtering out interfering components, and correcting baseline wandering ([PIA-Group, 2021](#)). Specifically, after loading the ECG data, we apply a high-pass filter to remove the low-frequency noise (e.g., baseline wandering), a notch filter to

remove power-line interference, and a low-pass filter to filter out the high-frequency noise.

3.3 Continuous wavelet transform

ECG signals encompass multiple feature components in both the time and frequency domains. In this study, we engage the continuous wavelet transform (CWT) to extract time-frequency features from ECGs due to its excellent performance in the analysis of transient and non-stationary time series signals (Keissar et al., 2009). CWT is the most popular tool for time-frequency analysis that reflects the frequency components of data changing with time. CWT is verified to outperform the traditional STFT due to its ability to provide multi-resolution decompositions of the signal, which allows for a trade-off between time and frequency resolution, i.e., higher frequency resolution for signals with sharp transients and higher time resolution for signals with slow-varying frequency content (Dokur and Ölmez, 2001). Additionally, compared to discrete wavelet transform (DWT), CWT remedies non-stationarity and coarse time-frequency resolution defects and supports the extraction of arbitrarily high-resolution features in the time-frequency domain (Addison, 2005).

The CWT of the ECG time-series signal denoted as $x(t)$ is achieved according to:

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (1)$$

where $T(a, b)$ stands for the intensity of transformed signals, $\psi(\cdot)$ is the wavelet basis (also known as the mother wavelet), a is the scale factor quantifying the compressed or stretched degree of a wavelet, and b is the time shift parameter defining the location of the wavelet. The scale can be used to derive the characteristic frequency of the wavelet as (Wu et al., 2019):

$$F = \frac{F_c \times f_s}{a} \quad (2)$$

where F_c is the center frequency of the mother wavelet and f_s is the sampling frequency of the signal. This relationship shows that smaller (larger) values of a correspond to higher (lower) frequency components. In CWT, the mother wavelet plays a critical role in time-frequency analysis, the choice of which depends on its similarity with the original signal (Ngui et al., 2013). Here, the Mexican hat wavelet (mexh) is chosen to serve as the mother wavelet because its shape is similar to the QRS waves and it is commonly used in ECG signal analysis (Wang et al., 2021). Specifically, the mexh is the second derivative of a Gaussian function (Addison, 2005), which is defined as:

$$\psi(t) = \frac{2}{\sqrt{3}\sqrt{\pi}} \exp\left(-\frac{t^2}{2}\right) (1-t^2) \quad (3)$$

Continuously changing the scale factor a and time shift parameter b generates the 2D wavelet coefficients $T(a, b)$, which can be viewed as a 2D scalogram of the ECG signal in both the time and frequency domain (Wang et al., 2021).

Figures 2A–D show the healthy and AF examples of the raw ECG signals obtained from PhysioNet and their 2D time-frequency patterns after CWT transformation with mexh wavelet,

respectively. The colors in the scalogram indicate the energy density of the signal component at the corresponding frequency and time (Addison, 2005; He et al., 2018). According to Figure 2A,C, two general differences can be observed: 1) The AF ECG signal lacks a distinct P wave, while it shows a fast and chaotic F wave due to the atrial fluttering (Figure 2C), in comparison to a normal ECG signal (Figure 2A); 2) Irregular RR intervals are observed in AF ECG (Figure 2C) caused by a non-synchronized ventricular response to the abnormal atrial excitation (He et al., 2018). The discriminative information in the time domain can also be captured by the CWT scalograms shown in Figures 2B,D. By using a 2D CNN to analyze the visual representation of 2D time-frequency scalograms, we can better understand the features that distinguish AF from normal heart rhythms and make more accurate predictions.

3.4 Convolutional neural network

We engage CNN to build a data-driven classifier for differentiating AF samples from normal ECG samples. CNN is a type of network architecture specifically designed to process data that has a grid-like structure such as images (Khan et al., 2020). As opposed to traditional multilayer perceptron networks (MLPs), where the input of each neuron consists of the outputs of all the neurons from the previous layer, the neuron in CNN only receives its input from a localized region of the previous layer, known as its receptive field. The main building blocks of a CNN are convolutional layers, pooling layers, and fully connected layers.

Convolutional layers are responsible for performing a convolution operation on the input data, using a set of filters to extract local features in the data, and producing a feature map that summarizes such local information. Let θ and X denote the filter (also known as the kernel) and the input. The convolution operation works as follows:

$$(X \otimes \theta)_{ij} = \sum_{m=0}^{s_1-1} \sum_{n=0}^{s_2-1} X(i+m, j+n) \theta(m, n) \quad (4)$$

where s_1 and s_2 denote the size of the 2D kernel, and (i, j) denotes the location on the 2D input (e.g., image). After being applied with the activation function, the feature map of the input is obtained as (LeCun and Bengio, 1995; Jing et al., 2021):

$$X_q^l = \sigma\left(\sum_p \theta_{pq}^l \otimes X_p^{l-1} + b_q^l\right) \quad (5)$$

where X_q^l is the q th feature at layer l , X_p^{l-1} is the p th input feature map of the previous $(l-1)$ -th layer, σ denotes the activation function to induce the non-linearity in the functional mapping, and b_q represents the bias. This procedure is repeated by applying multiple filters to generate multiple feature maps to capture different characteristics of the input. Note that kernels are shared across all the input positions, which is also called weight sharing, the key feature of CNN. The weight-sharing technique guarantees the extracted local patterns are translation invariant and increases computational efficiency by reducing the model parameters to learn compared with fully connected neural networks.

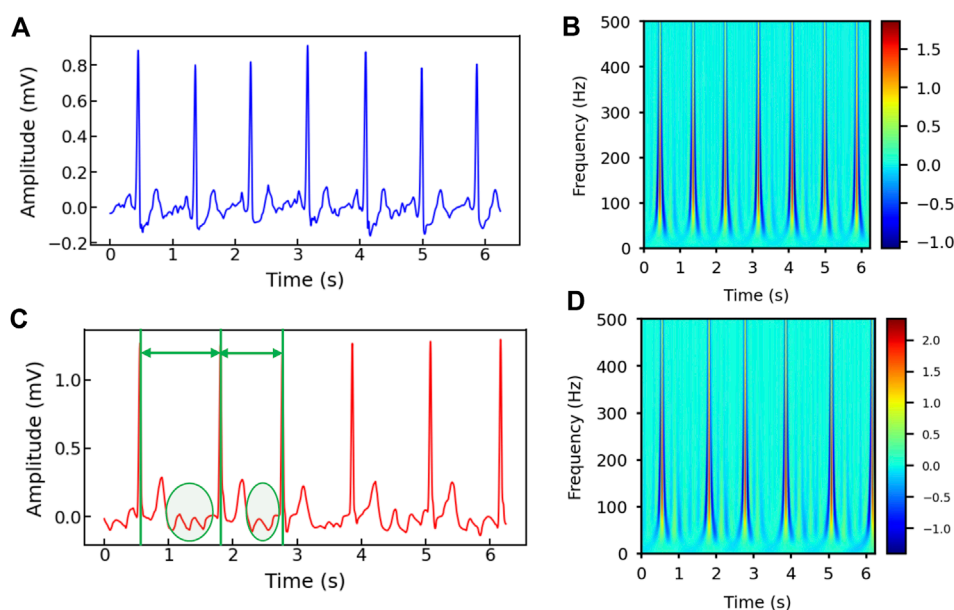


FIGURE 2

(A) The raw ECG signal from Physionet labeled as normal and (B) its corresponding 2D CWT scalogram. (C) The raw ECG signal from Physionet labeled as AF and (D) its corresponding 2D CWT scalogram. Note that the RR intervals are different in the AF sample and irregular F waves (circled) appear in (C).

The pooling layer mimics the human visual system by combining the outputs of multiple neurons (i.e., clusters) into a single neuron in the next layer, effectively creating a condensed representation of the input. The pooling significantly reduces the spatial resolution and only focuses on the prominent patterns of the feature maps, making the network more robust to small translations and distortion in the input data (Xia et al., 2018). Popular pooling techniques include maximum pooling, average pooling, stochastic pooling, and adaptive pooling. They are typically performed on the values in a sub-region of the feature map (Akhtar and Ragavendran, 2020).

The fully-connected layers form a dense network that can learn complex non-linear relationships between the inputs and outputs. It takes the output of the previous layer, which is typically a high-dimensional tensor containing discriminant features extracted by convolutional and pooling layers, and flattens it into a one-dimensional vector. This vector is then used as the input to a fully connected layer. The fully-connected layer is similar to an MLP in that every neuron in one layer is connected to every neuron in the next layer. By using a proper activation function, the neural network is able to produce classification decisions (Nurmaini et al., 2020). By stacking these building blocks (convolutional layers, pooling layers, and fully connected layers) in various combinations, CNN is able to learn complex features in the input data, allowing them to effectively solve a wide range of image and signal processing tasks (Andreotti et al., 2017).

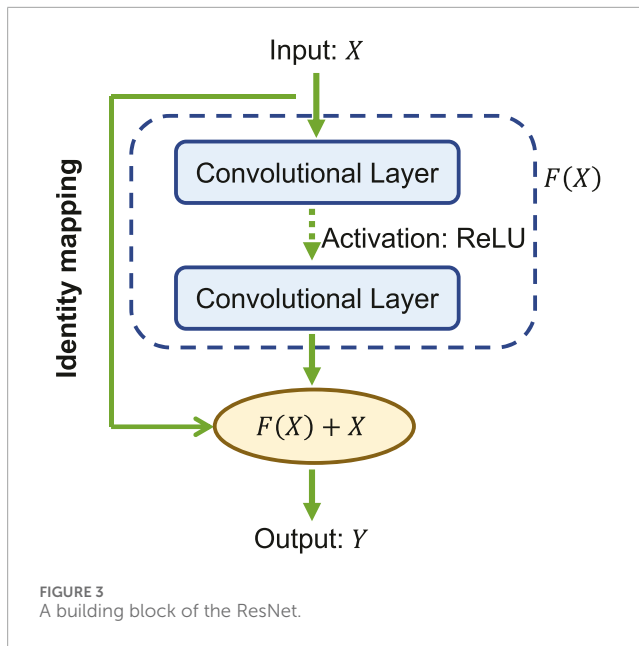
3.5 2D CNN with ResNet

We propose to engage 2D CNN to investigate the 2D time-frequency scalograms converted from denoised ECG signals by CWT for AF identification. It has been demonstrated that the

substantial depth of the convolutional network is beneficial to the network performance (Simonyan and Zisserman, 2014). However, as the number of convolutional layers increases, the training loss stops further decreasing and becomes saturated because of the gradient dissipation issue. As such, a CNN with a deeper architecture, counterintuitively, sometimes incurs a larger training error compared to its shallow counterpart upon convergence (He et al., 2016). To solve such network degradation and gradient vanishing problems, the residual network (ResNet) has been developed to improve the accuracy of CNNs with considerably increased depth.

The core of ResNet is the residual learning technique (He et al., 2016). Specifically, instead of using the stacked convolutional layers to directly fit the underlying mapping from the input to the output, ResNet focuses on fitting a residual mapping. Figure 3 shows a ResNet building block with input X and its corresponding output mapping Y . The residual block engages a shortcut connection that bypasses one or more convolutional layers and allows the information to flow directly from the input to the output. As such, the input X is added to the output of the block $F(X)$ (enclosed by the dashed circle in Figure 3, allowing the network to learn the residual mapping represented as $Y = F(X) + X$ instead of learning the direct mapping as $Y = F(X)$. This design mitigates the gradient vanishing problem and allows for deeper networks to be trained effectively.

In our study, we engage the ResNet with 18 layers (ResNet18) to build the AF classifier because ResNet18 has been proven to be able to generate a comparable result with a faster convergence compared to a deeper counterpart (He et al., 2016). Figure 4 shows the detailed structure of ResNet18. Note that the notation of $2DConv(n_{input}, n_{output}, n_{fdim1} \times n_{fdim2})$ denotes that, in the current 2D convolutional layer, there are n_{input} input



channels, n_{output} output channels (i.e., number of filters) with the 2D filter size of $n_{fdim1} \times n_{fdim2}$. For example, (64, 128, 3×3) indicates that this convolutional layer is composed of 128 filters with the filter size of 3×3 applied on the input data with 64 channels.

3.6 Multi-branching convolutional network

Data-driven identification of AF from ECG recordings generally suffers from imbalanced data issues. Figure 5A presents the distribution of AF and normal samples in Physionet/CinC 2017 and OUHSC datasets, illustrating a normal to AF sample ratio of approximately 7:1 for both. To address the data imbalance issue, we create N_b balanced datasets from the original data $D = \{D_-, D_+\}$, where D_- denotes the majority normal ECG samples and D_+ stands for the minority set, i.e., the entire AF training samples. D_- is partitioned into multiple subsets $D_- = \cup_{i=1}^{N_b} D_-^i$, where each subset D_i is roughly equivalent in size to D_+ . The normal subsets D_i for $i = 1, \dots, N_b$ are then paired with D_+ to formulate balanced sub-datasets. Each balanced subset, denoted as $D_i = \{D_-^i, D_+\}$ for $i = 1, \dots, N_b$, is processed through the ResNet core, with individual branches trained on their respective balanced sub-datasets. Figure 5B visualizes this method of partitioning the original dataset D into N_b balanced sub-datasets, i.e., D_i for $i = 1, \dots, N_b$, which serve as the balanced input in Figure 6. This strategic partitioning and training approach ensures a comprehensive model learning from a balanced representation of AF and normal ECG samples (Wang and Yao, 2021; Wang et al., 2022; Wang et al., 2023b).

In the current investigation, we aim to identify AF samples from normal ECG samples. The neural network is expected to produce high probabilities (close to 1) for AF samples and low probabilities (close to 0) for normal ECG samples. We choose the binary cross-entropy as the loss function for MB-ResNet, which is defined as:

$$\mathcal{L}(\omega; D) = - \sum_{j=1}^{N_d} \sum_{i=1}^{N_b} \mathcal{I}(j \in D_i) (y_j \log(\hat{P}_i(\omega; X^j)) + (1 - y_j) \log(1 - \hat{P}_i(\omega; X^j))) \quad (6)$$

where ω denotes the neural network parameter set, X^j and y_j stand for one input sample and its corresponding true label respectively, $\mathcal{I}(\cdot)$ denotes the indicator function, N_d is the total number of the training samples, and $\hat{P}_i(\omega; X^j)$ represents the predicted probability for AF at the i th branching output given the input signal X^j .

The adaptive momentum method (Adam) (Kingma and Ba, 2014) is adopted to minimize the loss function and update the network parameters. In the inference stage, the MB network generates N_b predictions for AF probability, which correspond to the N_b branching outputs. The final predicted probability for AF (\hat{P}) is determined by taking the average of the N_b outputs:

$$\hat{P} = \frac{1}{N_b} \sum_{i=1}^{N_b} \hat{P}_i$$

where \hat{P}_i is the predicted probability of i th branching output.

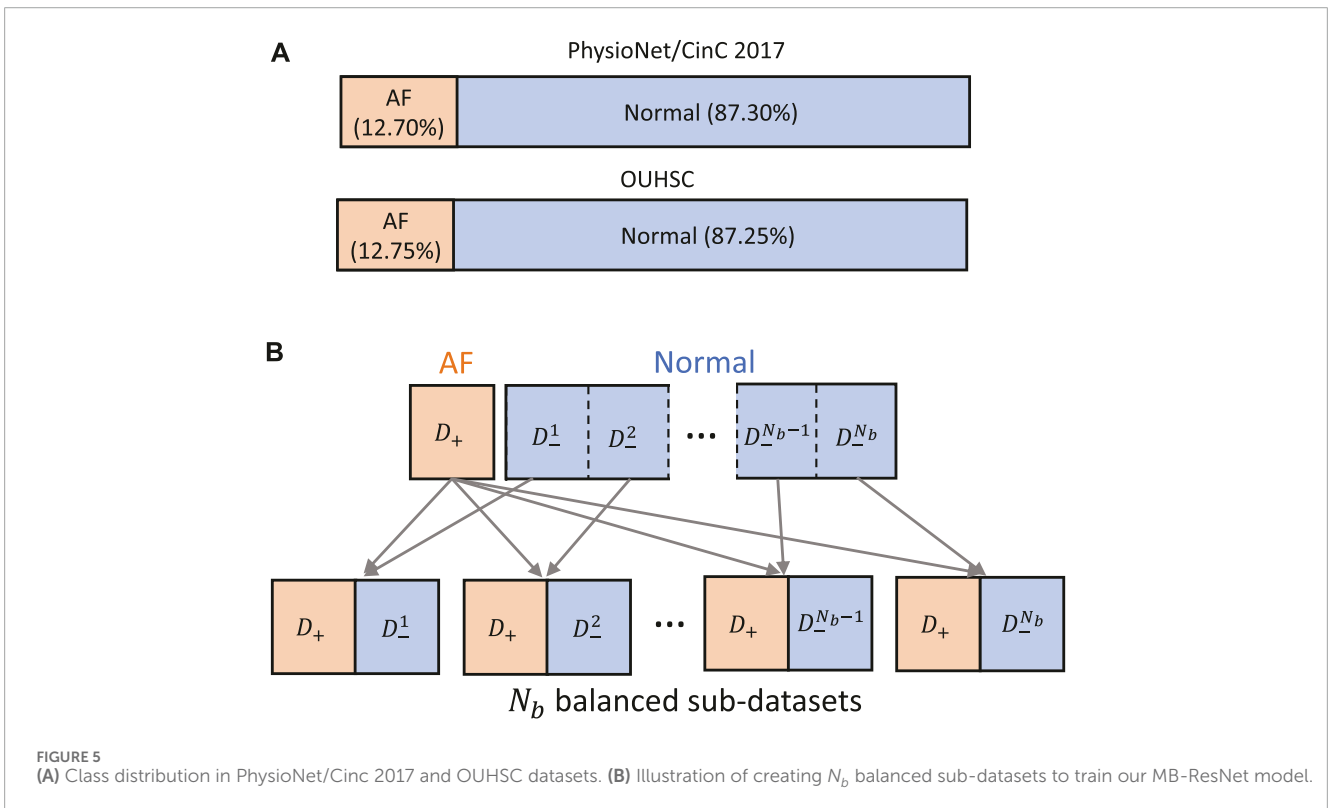
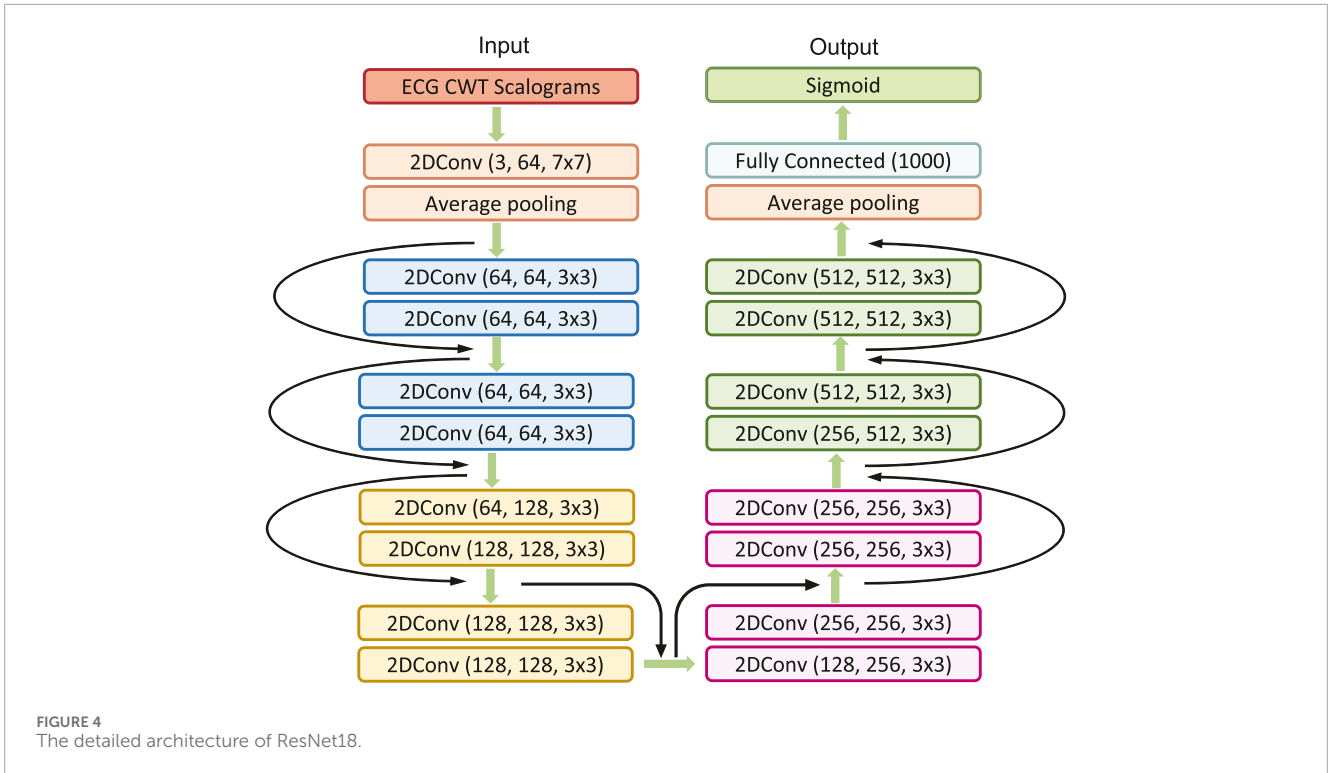
4 Experimental design and results

4.1 Experimental design

We validate and evaluate the performance of the proposed CWT-MB-ResNet framework using both OUHSC and Physionet Challenge datasets. In this study, the training and testing datasets are split interpatiently for both data sources. This ensures that no overlap exists between the patients in the training set and those in the testing set. We allocate 80% of the total samples for the training purpose and the remaining 20% for testing, applied on both datasets.

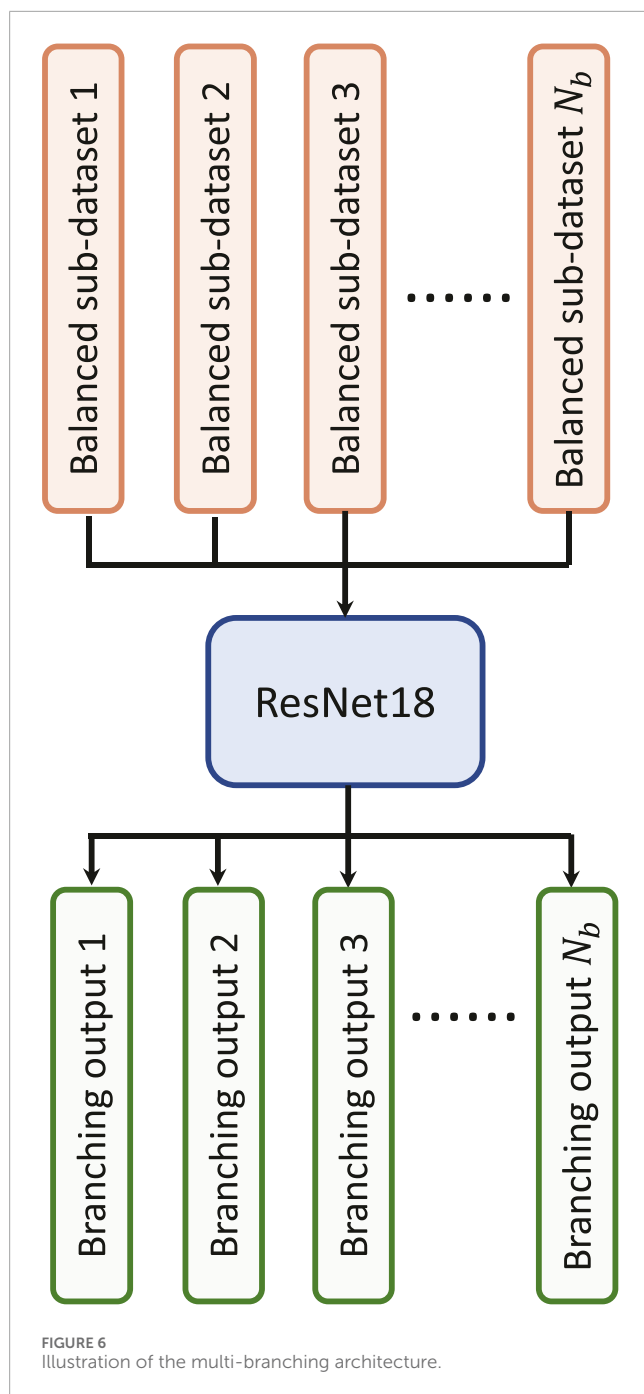
We first explore the impact of the learning rate on the training outcomes of the proposed CWT-MB-ResNet. We then conducted a comparison study to showcase the significance of ECG digitalization for the proposed multi-branching ResNet (MB-ResNet) model in identifying the AF samples. Next, we compare the performance of our CWT-MB-ResNet with 1D-CNN (Figure 7A), 1D-CNN with the multi-branching network (1D-MB-CNN) (Figure 7B), and ResNet with CWT features (CWT-ResNet). Note that the input of 1D-CNN and 1D-MB-CNN consists of the denoised ECG time series. The detailed 1D-CNN architecture is illustrated in Figure 8, including three convolutional layers followed by pooling layers to reduce the dimensionality of the data, a batch-normalization layer to stabilize the network training, and one fully connected layer to make the final prediction. Note that the notation of $1DConv(n_{input}, n_{output}, n_{fdim})$ indicates that, in the current 1D convolutional layer, there are n_{input} input channels and n_{output} output channels (i.e., number of filters) with a 1D filter size of n_{fdim} .

The classification performance will be evaluated with three metrics: Receiver-Operating-Characteristic (ROC) Curve, Precision-Recall (PR) Curve, and F1 score, which will be calculated using the test set. The ROC provides the graphic representation of the trade-off between the true positive rate (TPR) and the false



positive rate (FPR) for different threshold settings. The area under ROC (AUROC) is often used as a metric to compare different models, with a larger AUROC indicating a better-performing classifier. A good model typically has a ROC curve that is situated

toward the top-left corner of the graph. The PRC illustrates the interplay between a predictive model's precision and recall metrics across a range of probability thresholds. A good classifier has the PR curve towards the top-right corner. A higher area under PRC



(AUPRC) value suggests a more effective model. The F1 score quantifies the equilibrium between a model's precision and recall for a binary classifier system by computing their harmonic mean, which is defined as

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Note that the F1 score ranges from 0 to 1, where a score of 1 indicates a perfect balance between precision and recall and a good overall prediction performance.

4.2 The effect of the learning rate on CWT-MB-ResNet

In this study, we initiate the analysis by transforming ECG time series data into 2D scalograms utilizing CWT. These scalograms encapsulating both time and frequency information are input into our tailored MB-ResNet model. Specifically, we employ ResNet18 due to its proven efficacy in achieving results comparable to those of its deeper counterparts, while also ensuring faster convergence rates (He et al., 2016). The architecture of ResNet18, as adopted from He et al. (2016) and illustrated in Figure 4, comes with a predefined set of network architecture parameters, including number of layers, kernel size, and number of residual blocks.

In addition to selecting ResNet18 for its balance between efficiency and performance, the learning rate has a critical influence on the training outcomes. To further optimize our model, we conducted an experiment specifically focused on assessing the impact of various learning rates on the model's performance, particularly looking at the F1 score on the test set across both datasets used in our study. Table 1 summarizes the performance of the MB-ResNet given different learning rates. For both datasets, the highest F1 score achieved is 0.8865 for PhysioNet/CinC 2017 and 0.7396 for OUHSC datasets when the learning rate is set as 0.001. This indicates that a learning rate of 0.001 is the most effective in training our MB-ResNet model.

4.3 The effect of ECG digitalization from PDFs on CWT-MB-ResNet

We carry out a comparative analysis to demonstrate the importance of digitizing ECG records from their original PDF format. Specifically, we transform the original ECG PDFs into image files (i.e., Portable Network Graphic (.PNG) files) and apply segmentation to augment the sample sizes. Figure 9 illustrates examples of the resulting ECG images from normal and AF categories, which directly serve as inputs for our MB-ResNet without further preprocessing.

Figure 10 displays the ROC and PR curves generated by two variants of the MB-ResNet model: one trained on 2D scalograms derived from digitalized ECGs after undergoing denoising and CWT (referred to as CWT-MB-ResNet), and the other trained on pure ECG images converted directly from raw PDF files (denoted as PDF-MB-ResNet). Utilizing the same MB-ResNet model, we observed a substantial increase in the area under both ROC and PR curves when the model inputs were 2D scalograms processed from digitalized ECGs compared with using raw ECG images directly. Specifically, our CWT-MB-ResNet model demonstrates superior performance with an AUROC of 0.9351, AUPRC of 0.7930, and an F1 score of 0.7396. This performance significantly surpasses that of the PDF-MB-ResNet trained by raw ECG images with an AUROC of 0.8683, AUPRC of 0.6462, and an F1 score of 0.6257, highlighting the efficacy of our digitalization and preprocessing procedure. The enhanced performance of the MB-ResNet model trained with 2D scalograms from digitalized ECGs, as compared to training with raw ECG images, is attributed to several factors:

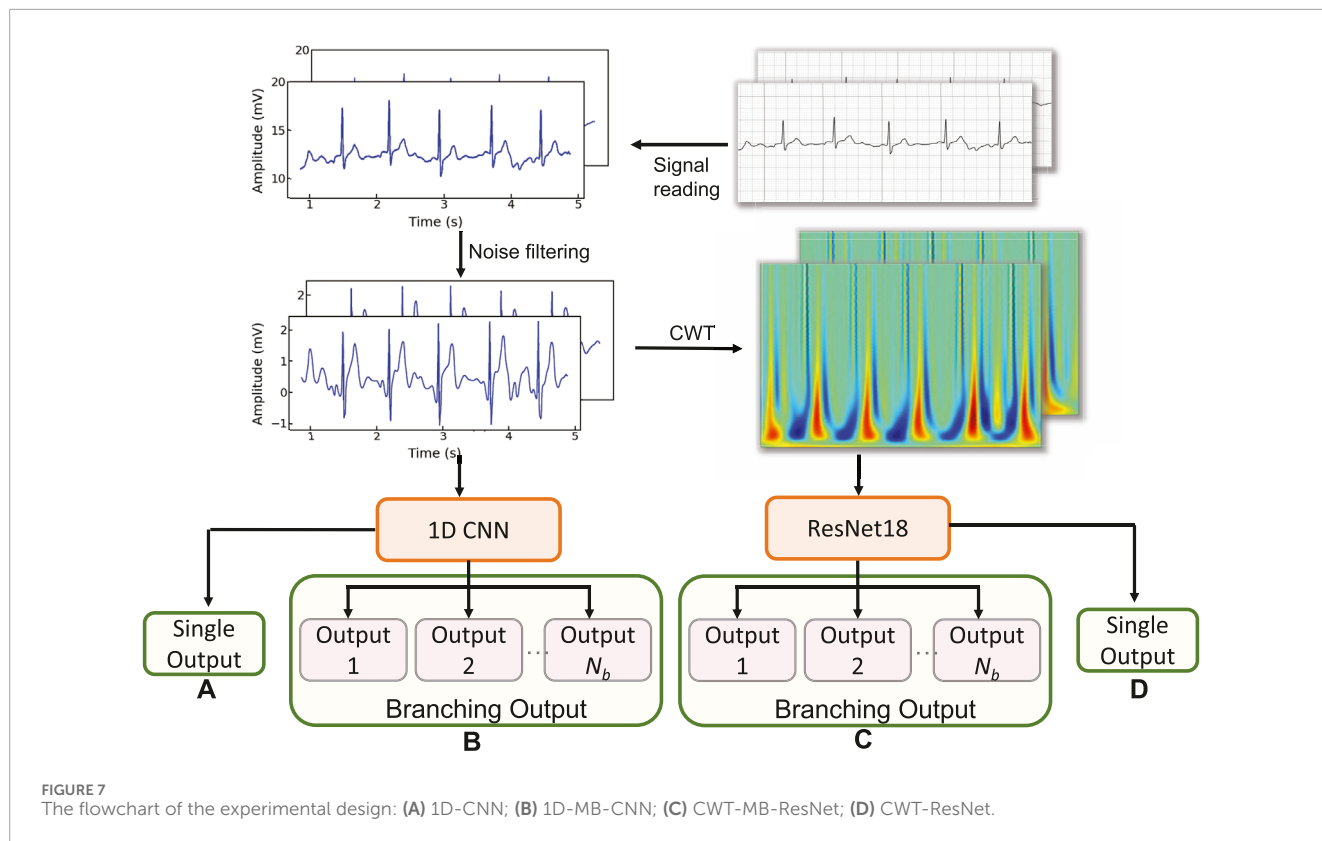


TABLE 1 F1 scores on the testing set given different learning rates for MB-ResNet training.

learning rate	0.0001	0.0005	0.001	0.005	0.01
F1 (PhysioNet)	0.8493	0.8759	0.8865	0.8652	0.8580
F1 (OUHSC)	0.6854	0.7273	0.7396	0.7385	0.7151

- The 2D scalograms provide a rich representation of temporal and frequency features, offering a more comprehensive dataset for the model to learn from.
- The raw ECG segmentation images contain large blank areas devoid of any ECG-related information, which do not contribute to learning discriminative features.
- The superimposed gridlines in the area could introduce noise into the data, potentially hindering the model's training efficiency.

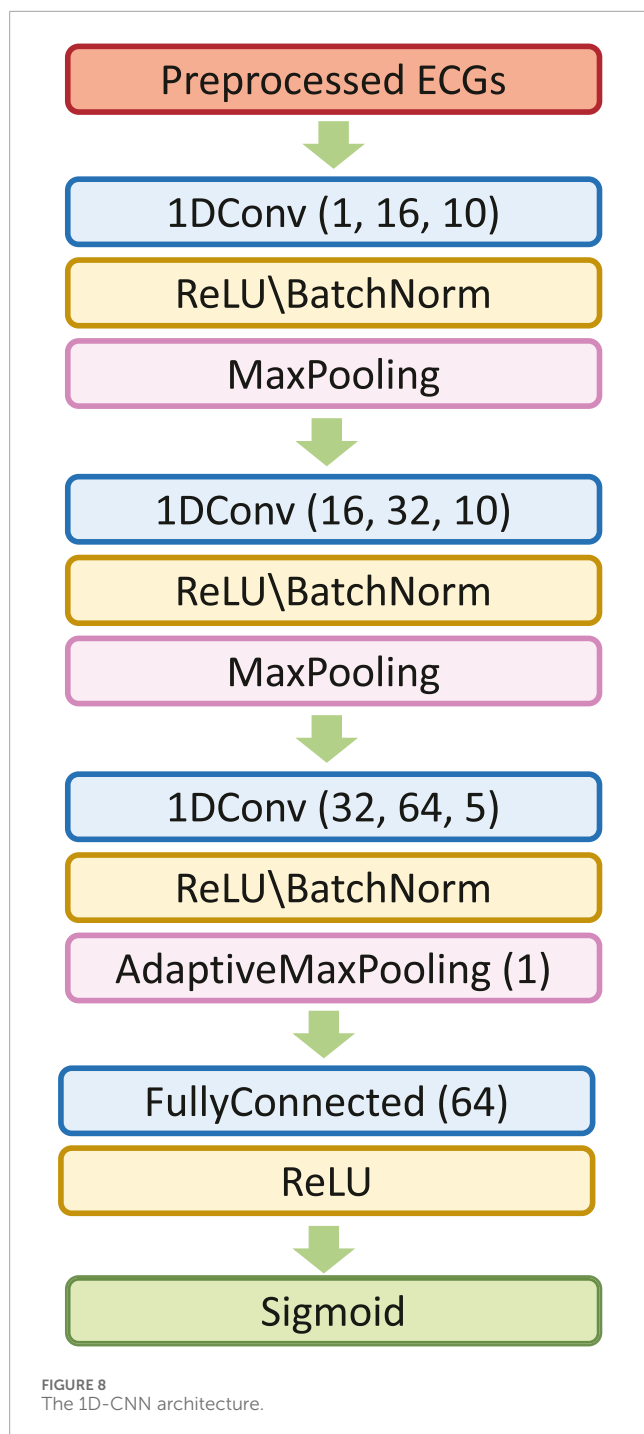
By training with 2D scalograms, the abovementioned issues are mitigated, allowing the MB-ResNet to focus on more relevant ECG features, leading to significant improvement in overall model performance.

4.4 Experimental results from the OUHSC dataset

Figure 11 displays the ROC and PR curves of all four models using the OUHSC dataset. The 2D ResNet models (i.e., CWT-ResNet

and CWT-MB-ResNet), which use 2D scalograms transformed from ECG signals as the input, produce a larger area under the curves (both ROC and PR) compared to their 1D counterparts (i.e., 1D-CNN and 1D-MB-CNN). This demonstrates the efficacy of using the CWT to extract time-frequency features in the ECG signal analysis. Additionally, the models with an MB architecture (i.e., 1D-MB-CNN and CWT-MB-ResNet) produce a larger AUROC and AUPRC compared to models without MB outputs (i.e., 1D-CNN and CWT-ResNet), which highlights the effectiveness of using the MB structure in addressing imbalanced data issues. The ROC and PR plots demonstrate the superiority and robustness of the proposed CWT-MB-ResNet framework for identifying the AF samples.

Table 2 shows AUROC, AUPRC, and F1 scores generated from the four methods using the OUHSC dataset. The proposed CWT-MB-ResNet method generates the best AUROC, AUPRC, and F1 scores with values of 93.51%, 79.30%, and 0.7396. Note that the MB technique demonstrates its effectiveness on both 1D-CNN and CWT-ResNet as the AUROC, AUPRC, and F1 scores provided by the MB-based neural network models are higher than their non-MB counterparts. Moreover, the AF classifier using 2D-CNN-based ResNet18 supported by the time-frequency transformation of ECG



time series presents a more potent predictive power than time sequence classification using 1D CNN. For example, CWT-MB-ResNet improves the AUROC, AUPRC, and F1 scores from 87.55%, 69.97%, and 0.6384 to 93.51%, 79.30%, and 0.7396 respectively compared with the 1D-MB-CNN.

4.5 Experimental results from the Physionet/CinC 2017 challenge dataset

Figure 12 further shows the ROC and PRC analysis for the Physionet/CinC 2017 challenge dataset. Similar to the results from

the OUHSC dataset, the 2D ResNet models (CWT-ResNet and CWT-MB-ResNet) outperform their 1D counterparts (1D-CNN and 1D-MB-CNN) in both the ROC and PR spaces. Furthermore, the MB-based models (1D-MB-CNN and CWT-MB-ResNet) effectively account for the imbalanced data issues, exhibiting better performance compared to the non-MB-based models (1D-CNN and CWT-ResNet). Table 3 demonstrates the comparison of AUROC, AUPRC, and F1 scores provided by 1D-CNN, 1D-MB-CNN, CWT-ResNet, and CWT-MB-ResNet. Our CWT-MB-ResNet yields the best classification performance among the four methods, generating the highest AUROC, AUPRC, and F1 scores of 97.41%, 93.53%, and 0.8865. Especially, our CWT-MB-ResNet model improves the F1 score by 46.2% percent compared to the pure 1D-CNN with no CWT transform or MB structure.

5 Discussion

5.1 Strengths of the proposed pipeline

This paper proposes a pipeline of CWT-MB-ResNet to identify the AF condition. The unique strengths of the proposed framework are:

- 1) **Digitalization of ECG readings in PDF:** This pipeline designed an ECG preprocessing method that can automatically convert ECG PDFs into digitalized, ready-to-use ECG time series data. This step is crucial for integrating machine learning models into clinical workflows, where ECGs are often archived in non-digitalized formats.
- 2) **Effectiveness of CWT representation:** The integration of CWT enhances feature extraction, enabling the model to better identify AF characteristics that might be missed by directly learning from raw time-series analysis alone. The resulted 2D ECG scalograms offer a rich representation of ECG data by encapsulating both time series and frequency components. The CWT-based feature reformulation can significantly enhance the model's performance by providing more comprehensive information for classifying ECG signals.
- 3) **Advantage of the network design:** The use of ResNet18 as the foundation allows our model to benefit from the strengths in deep residual learning, enabling it to learn from significantly deepened convolutional layers with improved accuracy. The ResNet18 has demonstrated comparable results to its deeper counterparts, meanwhile keeping its computational efficiency. This is further enhanced by our innovative multi-branching design, which addresses the class imbalance issue by training each branch on a balanced subset of the original dataset while the core network is exposed to the entire range of samples. This approach ensures that both AF and normal class is adequately represented and learned during the training process, significantly enhancing the network's ability to generalize across the imbalanced classes.

5.2 Discussion on the limitations

The proposed CWT-MB-ResNet framework, while effective, is not devoid of limitations. In our study, ECG segments were

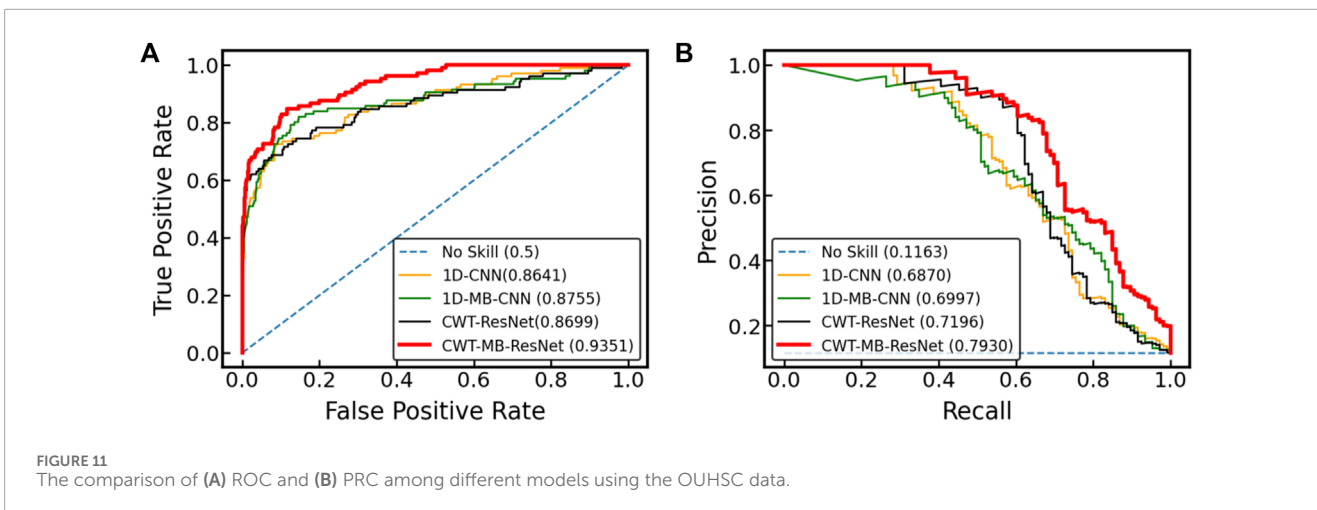
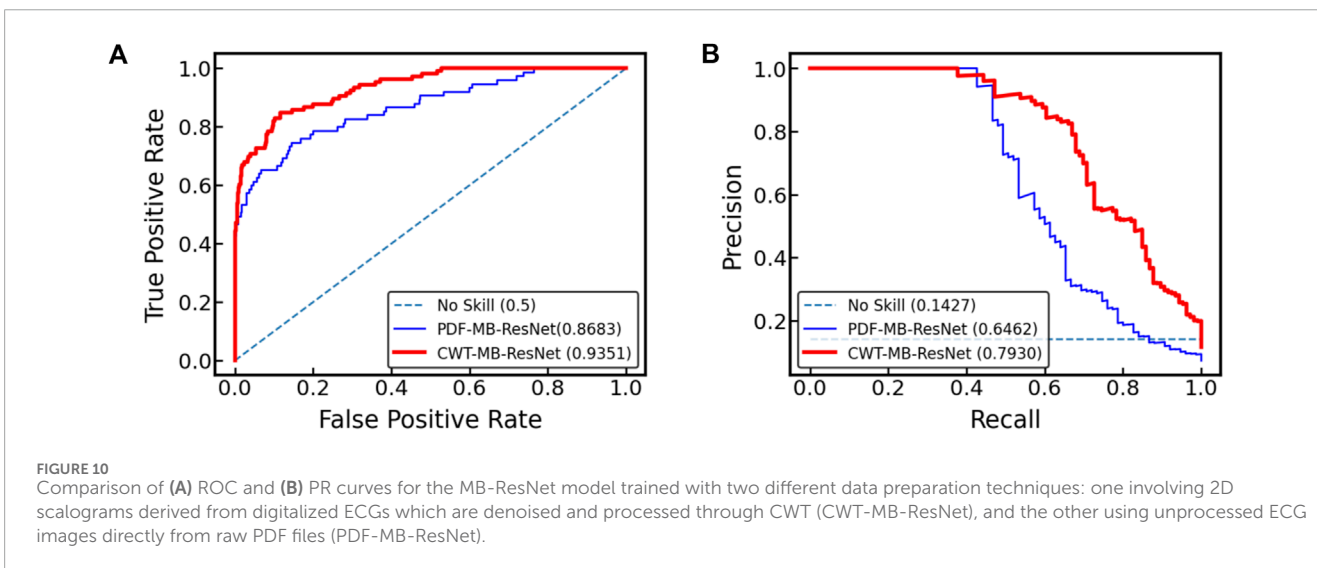
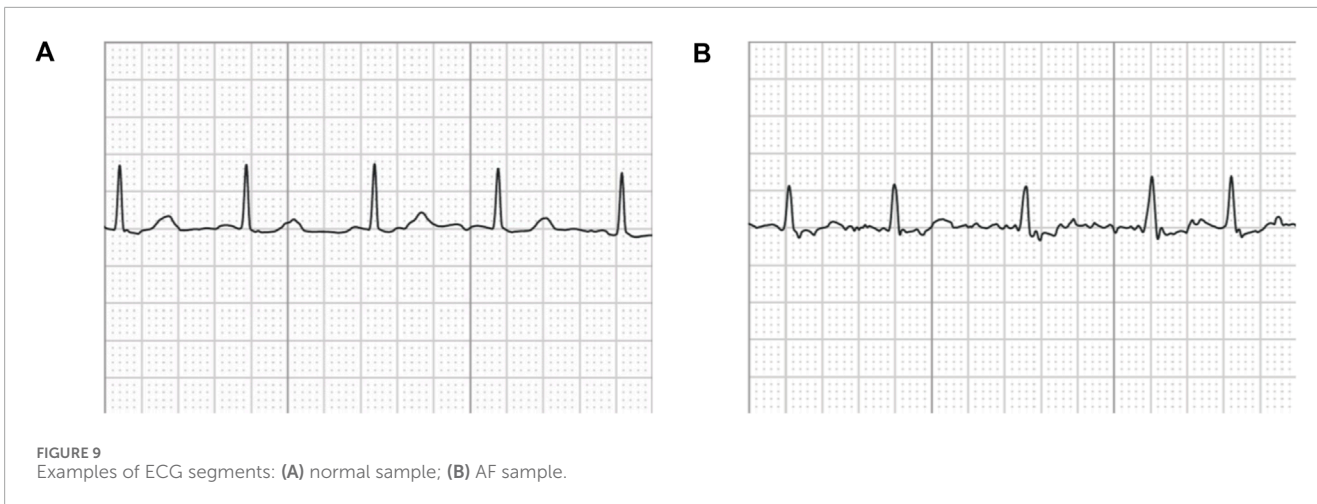


TABLE 2 The comparison of AUROC, AUPRC, and F1 scores generated from 1D-CNN, 1D-MB-CNN, CWT-ResNet, and the proposed CWT-MB-ResNet using OUHSC data.

	1D-CNN	1D-MB-CNN	CWT-ResNet	CWT-MB-ResNet
AUROC	86.41%	87.55%	86.99%	93.51%
AUPRC	68.70%	69.97%	71.96%	79.30%
F1	0.6370	0.6384	0.7150	0.7396

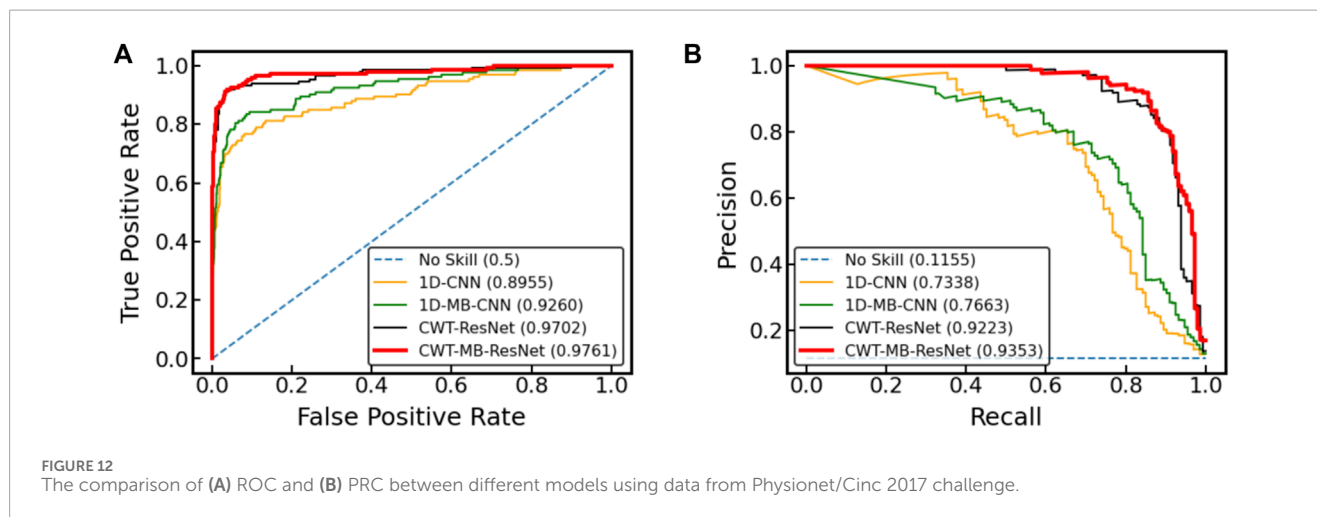


TABLE 3 The comparison of AUROC, AUPRC, and F1 scores generated from 1D-CNN, 1D-MB-CNN, CWT-ResNet, and the proposed CWT-MB-ResNet using data from Physionet/CinC 2017 challenge.

	1D-CNN	1D-MB-CNN	CWT-ResNet	CWT-MB-ResNet
AUROC	89.55%	92.60%	97.02%	97.61%
AUPRC	73.38%	76.63%	92.23%	93.53%
F1	0.7219	0.7380	0.8690	0.8865

around 5 s long. However, analyzing longer ECG recordings will significantly increase computational complexity. This is due to the CWT method of processing data across both time and frequency domains at various scales, demanding more computational resources. Additionally, while our method effectively addresses class imbalance, its performance remains influenced by the quality and diversity of the training data, which is a long-lasting limitation of most data-driven machine learning models. This is evident from the differing performances on the PhysioNet and OUHSC datasets. Specifically, PhysioNet, with its larger and more diverse pool of 5,788 subjects, provides a richer training environment compared to OUHSC, which is limited to ECG samples from only 260 subjects. Despite utilizing segmentation to expand the sample size of the OUHSC dataset to 5,809, notable differences in performance metrics remain, as detailed in Tables 2 and 3. This suggests that merely increasing the sample size by segmentation cannot fully address the limitations posed by data diversity and quality. Additionally, deep learning models, including the proposed CWT-MB-ResNet, are

often criticized for their “black box” nature. This means that while those models can make accurate predictions, the reasoning behind the predictions is not always clear or understandable to humans. This lack of interpretability can be a significant hurdle in clinical settings, making clinicians less confident in implementing machine learning models for automated diagnosis. One of our future research directions will focus on the development of interpretable models for AF detection.

5.3 Comparison with existing work

The direct comparison of our results with the values of performance metrics reported in other studies mentioned in Section 2 is neither fair nor feasible due to several factors: 1) variations in ECG duration used for training/testing data; 2) employment of non-unified metrics for evaluating model performance across studies; 3) variations in the proportions of

TABLE 4 The comparison of F1 scores between the proposed CWT-MB-ResNet method with existing literature using data from Physionet/CinC 2017 and OUHSC.

Authors	Methods	F1 (PhysioNet)	F1 (OUHSC)
Andreotti et al. Andreotti et al. (2017)	ResNet	0.8405	0.7054
Limam et al. Limam and Precioso, (2017)	CRNN	0.8310	0.7323
Wang et al. Wang and Li, (2020)	CNN-Bi-LSTM	0.7094	0.6996
Gao et al. Gao et al. (2021)	Residual-based temporal attention	0.8172	0.7368
This paper	CWT-MB-ResNet	0.8865	0.7396

training/testing data splits; 4) the model implementation on different databases. To enable a fairer and more meaningful comparison, we applied the ECG data from both the PhysioNet/CinC 2017 database and OUHSC to four deep learning models reviewed in [Section 2](#), ensuring that the comparison is based on consistent data and preprocessing steps.

[Table 4](#) summarizes the comparison results in terms of F1 score. Even though the proposed CWT-MB-ResNet model does not resort to complex neural network designs, it demonstrates the best F1 score compared with the other network structures developed in [Andreotti et al. \(2017\)](#); [Limam and Precioso, \(2017\)](#); [Wang and Li, \(2020\)](#); [Gao et al. \(2021\)](#). Specifically, the utilization of CWT distills both frequency and temporal insights from ECG signals, converting them into an image data format that significantly enriches the input information. We integrate the widely recognized image model, ResNet18 to achieve a robust interpretation of image data and meanwhile circumvent the gradient vanishing problem. Furthermore, the multi-branching structure is meticulously designed to address issues of data imbalance, ensuring that our model remains sensitive and accurate for both normal and AF classes.

6 Conclusion

In this paper, we develop a novel framework based on Continuous Wavelet Transform (CWT) and multi-branching ResNet for AF identification. We first transform the 1D ECG time series into 2D time-frequency scalograms to take into account various frequency components, which can serve as the input to the 2D CNN-based classifier. Second, we leverage the ResNet architecture to cope with the gradient dissipation problems in deep 2D CNN and increase the effectiveness of network training. Moreover, a multi-branching architecture is incorporated into the ResNet to mitigate the possible prediction bias caused by the imbalanced data issue. Finally, we implement the proposed CWT-MB-ResNet to predict AF using the ECG recordings from PhysioNet/CinC Challenge 2017 and the ECG PDFs from OUHSC. Experimental results show that the proposed CWT-MB-ResNet achieves the best prediction performance for both datasets in AF detection. The CWT-MB-ResNet framework has great potential to be applied in clinical practice to improve the accuracy in ECG-based diagnosis of heart disease.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: In this study, two AF databases from different sources, i.e., ECG recordings from PhysioNet/CinC challenge 2017 and ECG PDFs from OUHSC, are used to evaluate the performance of data-driven detection methods. PhysioNet/CinC challenge 2017 is open-source database. The ECG PDFs from OUHSC are provided by our cardiologist collaborators. Requests to access these datasets should be directed to Stavros-Stavrakis@ouhsc.edu.

Author contributions

JX: Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. SS: Conceptualization, Data curation, Resources, Supervision, Validation, Visualization, Writing—review and editing. BY: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Research reported in this publication was supported by the National Heart, Lung, And Blood Institute of the National Institutes of Health under Award Number R01HL172292. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abdou, A., and Krishnan, S. (2022). Horizons in single-lead ecg analysis from devices to data. *Front. Signal Process.* 2, 866047. doi:10.3389/frsip.2022.866047
- Addison, P. S. (2005). Wavelet transforms and the ecg: a review. *Physiol. Meas.* 26, R155–R199. doi:10.1088/0967-3334/26/5/R01
- Akhtar, N., and Ragavendran, U. (2020). Interpretation of intelligence in cnn-pooling processes: a methodological survey. *Neural Comput. Appl.* 32, 879–898. doi:10.1007/s00521-019-04296-5
- AliveCor (2024). AliveCor website. Available at: <https://www.alivecor.com/> (Accessed February 26, 2024).
- Analog Devices (2024). AD8232 single-lead, heart rate monitor front end. Available at: <https://www.analog.com/media/en/technical-documentation/data-sheets/ad8232.pdf> (Accessed February 26, 2024).
- Andreotti, F., Carr, O., Pimentel, M. A., Mahdi, A., and De Vos, M. (2017). "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ecg," in *2017 computing in cardiology (CinC)* (IEEE), 1–4.
- Athif, M., Yasawardene, P. C., and Daluwatte, C. (2018). Detecting atrial fibrillation from short single lead ecgs using statistical and morphological features. *Physiol. Meas.* 39, 064002. doi:10.1088/1361-6579/aac552
- Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- Bernstein, R. A., Kamel, H., Granger, C. B., Piccini, J. P., Sethi, P. P., Katz, J. M., et al. (2021). Effect of long-term continuous cardiac monitoring vs usual care on detection of atrial fibrillation in patients with stroke attributed to large- or small-vessel disease: the stroke-af randomized clinical trial. *Jama* 325, 2169–2177. doi:10.1001/jama.2021.6470
- Billeci, L., Chiarugi, F., Costi, M., Lombardi, D., and Varanini, M. (2017). "Detection of af and other rhythms using rr variability and ecg spectral measures," in *2017 computing in cardiology (CinC)* (IEEE), 1–4.
- Boriani, G., Palmisano, P., Malavasi, V. L., Fantecchi, E., Vitolo, M., Bonini, N., et al. (2021). Clinical factors associated with atrial fibrillation detection on single-time point screening using a hand-held single-lead ecg device. *J. Clin. Med.* 10, 729. doi:10.3390/jcm10040729
- Cai, W., Chen, Y., Guo, J., Han, B., Shi, Y., Ji, L., et al. (2020). Accurate detection of atrial fibrillation from 12-lead ecg using deep neural network. *Comput. Biol. Med.* 116, 103378. doi:10.1016/j.cmpbiomed.2019.103378
- Chen, S., Wang, Z., Yao, B., and Liu, T. (2022). "Prediction of diabetic retinopathy using longitudinal electronic health records," in *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)* (IEEE), 949–954.
- Clifford, G. D., Liu, C., Moody, B., Li-wei, H. L., Silva, I., Li, Q., et al. (2017). "Af classification from a short single lead ecg recording: the physionet/computing in cardiology challenge 2017," in *2017 computing in cardiology (CinC)* (IEEE), 1–4.
- Colilla, S., Crow, A., Petkun, W., Singer, D. E., Simon, T., and Liu, X. (2013). Estimates of current and future incidence and prevalence of atrial fibrillation in the us adult population. *Am. J. Cardiol.* 112, 1142–1147. doi:10.1016/j.amjcard.2013.05.063
- Da Silva-Filarder, M., and Marzbanrad, F. (2017). "Combining template-based and feature-based classification to detect atrial fibrillation from a short single lead ecg recording," in *2017 computing in cardiology (CinC)* (IEEE), 1–4.
- Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. theory* 36, 961–1005. doi:10.1109/18.57199
- De Chazal, P., O'Dwyer, M., and Reilly, R. B. (2004). Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE Trans. Biomed. Eng.* 51, 1196–1206. doi:10.1109/TBME.2004.827359
- Dokur, Z., and Ölmez, T. (2001). Ecg beat classification by a novel hybrid neural network. *Comput. methods programs Biomed.* 66, 167–181. doi:10.1016/s0169-2607(00)00133-4
- Fan, X., Yao, Q., Cai, Y., Miao, F., Sun, F., and Li, Y. (2018). Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ecg recordings. *IEEE J. Biomed. Health Inf.* 22, 1744–1753. doi:10.1109/JBHI.2018.2858789
- Faust, O., Shenfield, A., Kareem, M., San, T. R., Fujita, H., and Acharya, U. R. (2018). Automated detection of atrial fibrillation using long short-term memory network with rr interval signals. *Comput. Biol. Med.* 102, 327–335. doi:10.1016/j.cmpbiomed.2018.07.001
- Gao, J., Zhang, H., Lu, P., and Wang, Z. (2019). An effective lstm recurrent network to detect arrhythmia on imbalanced ecg dataset. *J. Healthc. Eng.* 2019, 6320651. doi:10.1155/2019/6320651
- Gao, Y., Wang, H., and Liu, Z. (2021). An end-to-end atrial fibrillation detection by a novel residual-based temporal attention convolutional neural network with exponential nonlinearity loss. *Knowledge-Based Syst.* 212, 106589. doi:10.1016/j.knsys.2020.106589
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation* 101, e215–e220. doi:10.1161/01.cir.101.23.e215
- Guan, Y., An, Y., Xu, J., Liu, N., and Wang, J. (2022). Ha-resnet: residual neural network with hidden attention for ecg arrhythmia detection using two-dimensional signal. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 20, 3389–3398. doi:10.1109/TCBB.2022.3198998
- He, H., and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Trans. Knowl. data Eng.* 21, 1263–1284. doi:10.1109/tkde.2008.239
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- He, R., Wang, K., Zhao, N., Liu, Y., Yuan, Y., Li, Q., et al. (2018). Automatic detection of atrial fibrillation based on continuous wavelet transform and 2d convolutional neural networks. *Front. physiology* 9, 1206. doi:10.3389/fphys.2018.01206
- Huang, J., Chen, B., Yao, B., and He, W. (2019). Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. *IEEE access* 7, 92871–92880. doi:10.1109/access.2019.2928017
- Isakadze, N., and Martin, S. S. (2020). How useful is the smartwatch ecg? *Trends Cardiovasc. Med.* 30, 442–448. doi:10.1016/j.tcm.2019.10.010
- Islam, S., Ammour, N., and Alajlan, N. (2017). "Atrial fibrillation detection with multiparametric rr interval feature and machine learning technique," in *2017 International Conference on Informatics, Health & Technology (ICIHT)* (IEEE), 1–5.
- Izci, E., Ozdemir, M. A., Degirmenci, M., and Akan, A. (2019). "Cardiac arrhythmia detection from 2d ecg images by using deep learning technique," in *2019 medical technologies congress (TIPTEKNO)* (IEEE), 1–4.
- Jing, E., Zhang, H., Li, Z., Liu, Y., Ji, Z., and Ganchev, I. (2021). Ecg heartbeat classification based on an improved resnet-18 model. *Comput. Math. Methods Med.* 2021, 6649970. doi:10.1155/2021/6649970
- Jun, T. J., Nguyen, H. M., Kang, D., Kim, D., Kim, D., and Kim, Y.-H. (2018). Ecg arrhythmia classification using a 2-d convolutional neural network. arXiv preprint arXiv:1804.06812
- Keissar, K., Davrath, L. R., and Akselrod, S. (2009). Coherence analysis between respiration and heart rate variability using continuous wavelet transform. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 367, 1393–1406. doi:10.1098/rsta.2008.0273
- Khan, A., Sohail, A., Zahoora, U., and Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* 53, 5455–5516. doi:10.1007/s10462-020-09825-6
- Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980
- Kutlu, Y., and Kuntalp, D. (2012). Feature extraction for ecg heartbeats using higher order statistics of wpd coefficients. *Comput. methods programs Biomed.* 105, 257–267. doi:10.1016/j.cmpb.2011.10.002
- Ladavich, S., and Ghoraani, B. (2015). Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity. *Biomed. Signal Process. Control* 18, 274–281. doi:10.1016/j.bspc.2015.01.007
- Lai, D., Zhang, X., Bu, Y., Su, Y., and Ma, C.-S. (2019). An automatic system for real-time identifying atrial fibrillation by using a lightweight convolutional neural network. *IEEE access* 7, 130074–130084. doi:10.1109/access.2019.2939822
- Larburu, N., Lopetegui, T., and Romero, I. (2011). "Comparative study of algorithms for atrial fibrillation detection," in *2011 computing in cardiology* (IEEE), 265–268.
- LeCun, Y., and Bengio, Y. (1995). "Convolutional networks for images, speech, and time series," in *The handbook of brain theory and neural networks*, 3361, 1995.

- Li, Q., Liu, C., Li, Q., Shashikumar, S. P., Nemati, S., Shen, Z., et al. (2019). Ventricular ectopic beat detection using a wavelet transform and a convolutional neural network. *Physiol. Meas.* 40, 055002. doi:10.1088/1361-6579/ab17f0
- Lian, J., Wang, L., and Muessig, D. (2011). A simple method to detect atrial fibrillation using rr intervals. *Am. J. Cardiol.* 107, 1494–1497. doi:10.1016/j.amjcard.2011.01.028
- Limam, M., and Precioso, F. (2017). "Atrial fibrillation detection and ecg classification based on convolutional recurrent neural network" in *2017 computing in cardiology (CinC) (IEEE)*, 1–4.
- Lubitz, S. A., Moser, C., Sullivan, L., Rienstra, M., Fontes, J. D., Villalon, M. L., et al. (2013). Atrial fibrillation patterns and risks of subsequent stroke, heart failure, or death in the community. *J. Am. Heart Assoc.* 2, e000126. doi:10.1161/JAHA.113.000126
- Luo, X., Yang, L., Cai, H., Tang, R., Chen, Y., and Li, W. (2021). Multi-classification of arrhythmias using a hcrnet on imbalanced ecg datasets. *Comput. Methods Programs Biomed.* 208, 106258. doi:10.1016/j.cmpb.2021.106258
- Mian Qaisar, S. (2020). Baseline wander and power-line interference elimination of ecg signals using efficient signal-piloted filtering. *Healthc. Technol. Lett.* 7, 114–118. doi:10.1049/htl.2019.0116
- Mousavi, S., Afghah, F., Razi, A., and Acharya, U. R. (2019). "Ecgnnet: learning where to attend for detection of atrial fibrillation with deep visual attention," in 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) (IEEE), 1–4.
- Murat, F., Sadak, F., Yildirim, O., Talo, M., Murat, E., Karabatak, M., et al. (2021). Review of deep learning-based atrial fibrillation detection studies. *Int. J. Environ. Res. Public Health* 18, 11302. doi:10.3390/ijerph182111302
- Nankani, D., and Baruah, R. D. (2022). "Atrial fibrillation classification and prediction explanation using transformer neural network," in 2022 International Joint Conference on Neural Networks (IJCNN) (IEEE), 01–08.
- Nesheiwat, Z., Goyal, A., and Jagtap, M. (2023). Atrial fibrillation. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK526072/> (Accessed February 22, 2024).
- Ngui, W. K., Leong, M. S., Hee, L. M., and Abdelrhman, A. M. (2013). Wavelet analysis: mother wavelet selection methods. *Appl. Mech. Mater.* 393, 953–958. doi:10.4028/www.scientific.net/amm.393.953
- Nurmaini, S., Tondas, A. E., Darmawahyuni, A., Rachmatullah, M. N., Partan, R. U., Firdaus, F., et al. (2020). Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. *Future Gener. Comput. Syst.* 113, 304–317. doi:10.1016/j.future.2020.07.021
- Oster, J., and Clifford, G. D. (2015). Impact of the presence of noise on rr interval-based atrial fibrillation detection. *J. Electrocardiol.* 48, 947–951. doi:10.1016/j.jelectrocard.2015.08.013
- Park, J., An, J., Kim, J., Jung, S., Gil, Y., Jang, Y., et al. (2022). Study on the use of standard 12-lead ecg data for rhythm-type ecg classification problems. *Comput. Methods Programs Biomed.* 214, 106521. doi:10.1016/j.cmpb.2021.106521
- Park, J., Lee, S., and Jeon, M. (2009). Atrial fibrillation detection by heart rate variability in poincare plot. *Biomed. Eng. online* 8, 38–42. doi:10.1186/1475-925X-8-38
- Petmezias, G., Haris, K., Stefanopoulos, L., Kilintzis, V., Tzavelis, A., Rogers, J. A., et al. (2021). Automated atrial fibrillation detection using a hybrid cnn-lstm network on imbalanced ecg datasets. *Biomed. Signal Process. Control* 63, 102194. doi:10.1016/j.bspc.2020.102194
- Phukan, N., Manikandan, M. S., and Pachori, R. B. (2023). Afibri-net: a lightweight convolution neural network based atrial fibrillation detector. *IEEE Trans. Circuits Syst. I Regul. Pap.* 70, 4962–4974. doi:10.1109/tcsi.2023.3303936
- PhysioBank, P., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220. doi:10.1161/01.cir.101.23.e215
- PIA-Group (2021). Biosppy: a toolbox for biosignal processing in python. Available at: <https://github.com/PIA-Group/BioSPpy>.
- Qiu, J., Zhu, J., Rosenberg, M., Liu, E., and Zhao, D. (2022). Optimal transport based data augmentation for heart disease diagnosis and prediction. arXiv preprint arXiv:2202.00567.
- Ramaraj, E., and Clement Virgeniya, S. (2021). A novel deep learning based gated recurrent unit with extreme learning machine for electrocardiogram (ecg) signal recognition. *Biomed. Signal Process. Control* 68, 102779. doi:10.1016/j.bspc.2021.102779
- Ramkumar, M., Kumar, R. S., Manjunathan, A., Mathankumar, M., and Pauliah, J. (2022). Auto-encoder and bidirectional long short-term memory based automated arrhythmia classification for ecg signal. *Biomed. Signal Process. Control* 77, 103826. doi:10.1016/j.bspc.2022.103826
- Rohr, M., Reich, C., Höhl, A., Lilienthal, T., Dege, T., Plesinger, F., et al. (2022). Exploring novel algorithms for atrial fibrillation detection by driving graduate level education in medical machine learning. *Physiol. Meas.* 43, 074001. doi:10.1088/1361-6579/ac7840
- Schwab, P., Sceba, G. C., Zhang, J., Delai, M., and Karlen, W. (2017). "Beat by beat: classifying cardiac arrhythmias with recurrent neural networks," in *2017 computing in cardiology (CinC) (IEEE)*, 1–4.
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
- Tateno, K., and Glass, L. (2001). Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of rr and δrr intervals. *Med. Biol. Eng. Comput.* 39, 664–671. doi:10.1007/BF02345439
- Tutuko, B., Nurmaini, S., Tondas, A. E., Rachmatullah, M. N., Darmawahyuni, A., Esafri, R., et al. (2021). Afibnet: an implementation of atrial fibrillation detection with convolutional neural network. *BMC Med. Inf. Decis. Mak.* 21, 216–217. doi:10.1186/s12911-021-01571-1
- Ullah, A., Rehman, S. u., Tu, S., Mehmood, R. M., and Ehatisham-Ul-Haq, M. (2021). A hybrid deep cnn model for abnormal arrhythmia detection based on cardiac ecg signal. *Sensors* 21, 951. doi:10.3390/s21030951
- University of Washington (2024). Learn more – global cardiovascular health program. Available at: <https://depts.washington.edu/globalcardio/about/learn-more/> (Accessed February 22, 2024).
- van Wyk, F., Khojandi, A., Williams, B., MacMillan, D., Davis, R. L., Jacobson, D. A., et al. (2019). A cost-benefit analysis of automated physiological data acquisition systems using data-driven modeling. *J. Healthc. Inf. Res.* 3, 245–263. doi:10.1007/s41666-018-0040-y
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* 30.
- Wang, J., and Li, W. (2020). Atrial fibrillation detection and ecg classification based on cnn-bilstm. arXiv preprint arXiv:2011.06187
- Wang, M., Rahardja, S., Fränti, P., and Rahardja, S. (2023a). Single-lead ecg recordings modeling for end-to-end recognition of atrial fibrillation with dual-path rnn. *Biomed. Signal Process. Control* 79, 104067. doi:10.1016/j.bspc.2022.104067
- Wang, T., Lu, C., Sun, Y., Yang, M., Liu, C., and Ou, C. (2021). Automatic ecg classification using continuous wavelet transform and convolutional neural network. *Entropy* 23, 119. doi:10.3390/e23010119
- Wang, Z., Liu, C., and Yao, B. (2022). "Multi-branching neural network for myocardial infarction prediction," in 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE) IEEE, 2118–2123.
- Wang, Z., Stavrakis, S., and Yao, B. (2023b). Hierarchical deep learning with generative adversarial network for automatic cardiac diagnosis from ecg signals. *Comput. Biol. Med.* 155, 106641. doi:10.1016/j.combiomed.2023.106641
- Wang, Z., and Yao, B. (2021). Multi-branching temporal convolutional network for sepsis prediction. *IEEE J. Biomed. Health Inf.* 26, 876–887.
- World Health Organization (2024). Cardiovascular diseases. Available at: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 (Accessed February 22, 2024).
- Wu, Y., Yang, F., Liu, Y., Zha, X., and Yuan, S. (2018). A comparison of 1-d and 2-d deep convolutional neural networks in ecg classification. arXiv preprint arXiv:1810.07088
- Wu, Z., Lan, T., Yang, C., and Nie, Z. (2019). A novel method to detect multiple arrhythmias based on time-frequency analysis and convolutional neural networks. *IEEE Access* 7, 170820–170830. doi:10.1109/access.2019.2956050
- Xia, Y., Wulan, N., Wang, K., and Zhang, H. (2018). Detecting atrial fibrillation by deep convolutional neural networks. *Comput. Biol. Med.* 93, 84–92. doi:10.1016/j.combiomed.2017.12.007
- Xie, J., and Yao, B. (2022a). Physics-constrained deep active learning for spatiotemporal modeling of cardiac electrodynamics. *Comput. Biol. Med.* 146, 105586. doi:10.1016/j.combiomed.2022.105586
- Xie, J., and Yao, B. (2022b). Physics-constrained deep learning for robust inverse ecg modeling. *IEEE Trans. Automation Sci. Eng.* 20, 151–166. doi:10.1109/tase.2022.3144347
- Xie, J., and Yao, B. (2023). Hierarchical active learning for defect localization in 3d systems. *IIEE Trans. Healthc. Syst. Eng.*, 1–15. doi:10.1080/24725579.2023.2233992
- Yao, B., Chen, Y., and Yang, H. (2021). Constrained markov decision process modeling for optimal sensing of cardiac events in mobile health. *IEEE Trans. Automation Sci. Eng.* 19, 1017–1029. doi:10.1109/tase.2021.3052483
- Yao, B., and Yang, H. (2016). Physics-driven spatiotemporal regularization for high-dimensional predictive modeling: a novel approach to solve the inverse ecg problem. *Sci. Rep.* 6, 39012–39013. doi:10.1038/srep39012
- Yao, B., and Yang, H. (2020). Spatiotemporal regularization for inverse ecg modeling. *IIEE Trans. Healthc. Syst. Eng.* 11, 11–23. doi:10.1080/24725579.2020.1823531
- Yao, B., Zhu, R., and Yang, H. (2017). Characterizing the location and extent of myocardial infarctions with inverse ecg modeling and spatiotemporal regularization. *IEEE J. Biomed. Health Inf.* 22, 1445–1455. doi:10.1109/JBHI.2017.2768534
- Ye, C., Kumar, B. V., and Coimbra, M. T. (2012). Heartbeat classification using morphological and dynamic features of ecg signals. *IEEE Trans. Biomed. Eng.* 59, 2930–2941. doi:10.1109/TBME.2012.2213253
- Zhang, H., Gu, H., Gao, J., Lu, P., Chen, G., and Wang, Z. (2022). An effective atrial fibrillation detection from short single-lead electrocardiogram recordings using mcn-bilstm network. *Algorithms* 15, 454. doi:10.3390/a15120454