Attention-enhanced Convolutional Neural Network for Inter-patient Classification of Premature Ventricular Contraction from ECG

Meng Chen[1], Yongjian Li[1], Wenzhuo Shi[2], Shoushui Wei[1]

[1]School of Control Science and Engineering, Shandong University, Jinan, China
[2]Institute of Marine Science and Technology, Shandong University, Qingdao, China

Abstract

Premature ventricular contraction (PVC) is a common type of cardiac arrhythmia characterized by an early and irregular heartbeat originating from the ventricles. This study presents a dual-path CNN network (WACNN) designed based on wavelet decomposition and attention mechanisms for detecting PVCs from single-lead electrocardiograms (ECGs). All ECG recordings are firstly resampled to 256 Hz. The R-peak positions detected using the Pan-Tompkins method are firstly used to extract a 256 points fixed-duration as training set A, and then used to extract two complete RR intervals as another training set (B). Subsequently, all segments in training set B are resampled to 512 points, resulting in a rescaled training set C. Set B and set C are fed into the model through two separate paths for training. The ratio of the original segment length in B to 512 is considered as a critical infection feature (heart rate feature) for PVC detection in the model. The proposed model is trained on 11 recordings from the MIT-BIH Arrhythmia Database (MITAD) and validated on another 11 recordings, excluding pacemaker data. Testing is conducted on the remaining 22 recordings. On the test dataset, the model achieved a precision of 95.41%, recall of 98.72%, F1 score of 97.04%, and overall accuracy of 99.63%, outperforming most existing methods.

1. Introduction

Premature ventricular contraction (PVC) is a relatively common type of cardiac arrhythmia characterized by the presence of ectopic foci in the ventricles. These ectopic centers alter the normal propagation pathway of the activation front, resulting in the generation of QRS complexes with wide and atypical waveforms [1]. The presence of PVCs has been linked to increased total mortality in certain patient subgroups, indicating that a high frequency of PVCs may reflect a more severe underlying condition rather than being merely a precursor to a critical electrical event [2]. Consequently, accurate and timely detection of PVCs holds substantial clinical importance. Computer-aided ECG analysis can significantly alleviate the clinical burden on physicians, enhance diagnostic efficiency, and improve the overall quality of medical care.

Early automated cardiac devices utilized rule-based algorithms and simple computational models based on ECG knowledge [3]. Recently, there has been a surge in studies applying artificial intelligence methods for PVC detection [4], [5], [6], [7]. These methods typically involve four key stages: preprocessing, heartbeat segmentation, feature extraction, and classification. Each of these stages plays a crucial role in accurately identifying PVCs. Signal quality assessment and denoising, are essential for minimizing misclassification as preprocessing steps. Heartbeat segmentation is critical for beat-by-beat PVC detection, though segment-based anomaly detection methods are less applicable in this context. Heartbeat segmentation involves detecting R-peaks and defining the heartbeat. Due to the small amplitudes of other features, such as the T-wave and P-wave, accurately identifying a complete heartbeat can be challenging. Currently, beat-by-beat PVC detection often relies on fixed-length segments centered around the R-peak, such as 250 ms before and 390 ms after the peak [8]. Feature extraction, a core component of automatic recognition algorithms, is divided into manual and deep learning-based methods. Manual feature extraction involves identifying time-domain, frequency-domain, and nonlinear features of ECG signals, which requires expert knowledge. With the advent of deep learning, many methods now leverage deep learning models to automatically extract features for arrhythmia detection and perform end-to-end recognition, eliminating the need to input extracted features into traditional machine learning models. Among these methods, convolutional neural networks (CNNs) are widely utilized by researchers [4], [7], [9]. Additionally, other studies have explored models based on recurrent neural networks (RNNs) [10], and Transformers [11], [6], to detect arrhythmias from ECG signals.

Despite the availability of numerous methods for PVC detection, many algorithms rely on R-peak positions

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.463

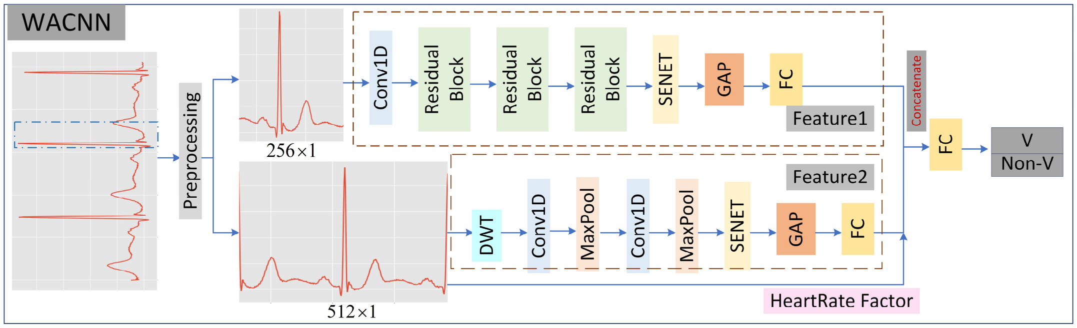

Figure 1. Flowchart of the proposed method.

annotated in databases during development, which can limit their practical applicability. Additionally, these methods typically focus on individual heartbeats without considering the role of temporal information, such as the compensatory pauses associated with PVCs, which could enhance the accuracy of the algorithms. In this study, a dual-path CNN network is designed based on wavelet decomposition and attention mechanisms for detecting PVCs from single-lead ECGs. All ECG recordings are firstly resampled to 256 Hz. The R-peak positions detected using the Pan-Tompkins method are firstly used to extract a 256 points fixed-duration as training Set A, and then used to extract two complete RR intervals as another training set (B). Subsequently, all segments in training set B are resampled to 512 points using a resampling method, resulting in the rescaled training set C. Set B and set C are fed into the model through two separate paths for training. The ratio of the original segment length in B to 512 (RR ratio) considered a critical feature for PVC detection in the detection model. We utilize the morphological characteristics of the heartbeat itself, along with the features of surrounding heartbeats, to effectively identify PVCs.

2. Methods

2.1. Problem Formulation

Most machine learning methods anchor around the R peak, extracting fixed-duration signals before and after as a heartbeat length for classification. However, this method fails to accommodate variations in heart rate and individual differences; some segments may not fully encompass the T-wave, while others might include two heartbeats. Moreover, segmenting continuous ECG signals into individual heartbeat segments results in loss of temporal information. To address these issues, this study proposes the use of heart rate recalibration to segment heartbeat segments, including a complete cycle of the heartbeat in addition to the features extracted by fixed-duration segments.

2.2. Preprocessing

All ECG recordings are first resampled to 256 Hz. The R-peak positions are detected using the Pan-Tompkins method [12]. We then extract segments of fixed lengths (100 samples before and 156 samples after) around each detected R-peak, excluding the first and last beats, to create dataset Set A. Based on the R-peak positions, we extract two complete RR intervals to define segment lengths, which we refer to as Set B. Subsequently, all segments in Set B are resampled to 512 points using piecewise cubic spline interpolation, resulting in Set C. A scaling factor (heart rate factor) is calculated between Set B and Set C, which is used as an additional feature for the proposed model. This process yields two sets of training data and an additional feature.

2.3. Wavelet-attention CNN

As illustrated in Figure 1, we design a dual-path architecture to extract signal features, with Set A and Set C as the inputs to each path, respectively. Each path employs a CNN as the core framework. Additionally, we incorporate the Squeeze-and-Excitation Network (SENET) [13] for channel attention, which assigns importance weights to features based on their significance.

Specifically, for the upper path (path1) for Set A, Initially, a one-dimensional convolutional layer (Conv1D) is employed to perform preliminary feature extraction from the input signals. This is followed by three residual blocks, each comprising multiple Conv1Ds with shortcut connections, as shown in Figure 2(a). These residual blocks facilitate the learning of complex mappings and enhance gradient flow during the training process. Subsequent to the residual blocks, the SENET module is integrated, the architecture of which is represented in Figure 2 (b). SENET dynamically recalibrates channel-wise feature responses, enabling the network to emphasize the most relevant features and thereby improve feature representation. The extracted features are then subjected to global average pooling (GAP), which aggregates the spatial dimensions of the feature maps into a single vector by averaging over all spatial locations. This reduces the feature dimensionality while preserving global information. Finally, a fully connected (FC) layer processes the pooled features, producing a 64-dimensional output vector (Feature1).

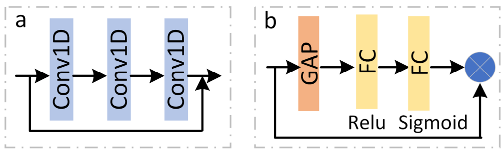

Figure 2. Architectures of Residual block (a) and SENET (b).

For the lower path (path2), the input in Set B initially undergoes a nine-level discrete wavelet transform (DWT) with a gap of 10Hz, decomposing the signal into various frequency components to provide a detailed time-frequency representation, as done in [8]. The transformed data is then processed through two Conv1Ds to extract local feature patterns, each followed by a max pooling layer that reduces dimensionality by selecting the maximum values within each pooling window. Subsequently, a SENET module is applied to dynamically recalibrate channel-wise feature responses and GAP aggregates the refined features into a single vector by averaging across all spatial dimensions. Finally, an FC layer processes this vector to produce a 32-dimensional feature output (Feature2).

The features extracted from the two paths, Feature1 and Feature2, combined with the heart rate factor obtained during the preprocessing phase, amounting to a total of 97 features, are concatenated. This concatenated feature vector is then fed into an FC layer to generate the final model output.

3. Experimental setup

3.1. Dataset and Evaluation Metrics

The MIT-BIH arrhythmia dataset (MITAD) [14], [15] serves as the gold-standard benchmark dataset for performance evaluation in this study. It contains 48 half-hour dual-lead ECG records from 47 subjects, with a sampling rate of 360 Hz.

In this study, we employ an inter-patient paradigm. Specifically, the ECG records are divided into DS1 and DS2, excluding those containing paced beats, as described in [16]. Additionally, the validation dataset comprises 11 separate records from DS1. The detailed data distribution is presented in Table 1. We categorize ventricular premature contractions and ventricular escape beats as V, while classifying all other types of heartbeats as non-V.

The performance indicators include the commonly used metrics such as accuracy (Acc), F1-score (F1), precision (Pre), and recall (Rec). For imbalanced datasets, F1 becomes a more pertinent metric for method comparison than Acc.

3.2. Hyperparameters

Due to the significant imbalance between V-type and non-V-type data, we used a focal loss function to address this classification issue. The batch size is set to 16, the initial learning rate is 0.001, and the Adam optimizer is utilized. Additionally, an early stopping mechanism is implemented to determine the number of training epochs: training is halted if the validation loss does not improve for 10 consecutive epochs.

4. Results and Discussion

4.1. Comparison with Baseline Methods

Table 1. PVC detection performance comparison with baseline methods.

Method	Year	F1	Pre	Rec	Acc
[17]	2014	88.97	92.75	85.48	98.63
[18]	2016	91.05	86.48	96.12	98.77
[1]	2017	/	/	97.59	**99.41**
[19]	2020	94.10	**95.58**	92.67	/
[20]	2021	92.60	92.00	95.80	/
[4]	2024	91.53	92.07	91.00	**98.90**
Ours	-	**97.04**	95.41	**98.72**	99.63

We compared our results with those from six other studies that utilized the same data for PVC detection. As indicated in Table 1, the proposed method in this study demonstrates superior performance with F1 and Rec metrics of 97.04% and 98.72%, respectively. The methods in [19] and [1], achieve the highest values in Pre and Acc, respectively. Furthermore, a high Rec value means the proposed model can correctly identify most of the PVC samples, resulting in a low miss rate. It is noteworthy that the innovative heartbeat segmentation method for path2 proposed in this study can effectively reflect the fully compensatory pause characteristic of most PVC beats, distinguishing them from other types of heartbeats. This approach, in combination with the WACNN model, yields robust results for PVC detection.

4.2. Determination of Scaling Length

When determining the length of each segment in Set C, we also examined the lengths of all two RR intervals in the database. As shown in Figure 3, the lengths of two RR intervals centered around PVC beats or non-PVC beats in the MITAD are generally within 512 samples. Most outliers, caused by noisy data, fall outside this range. Therefore, we have standardized the length of the rescaled segments to 512 samples.

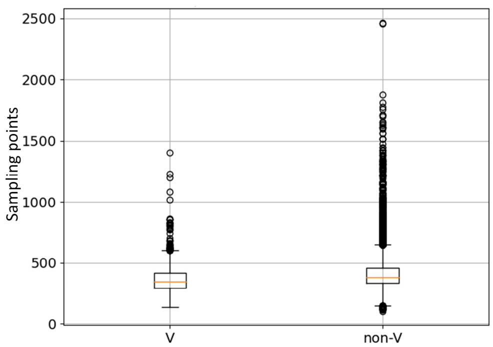

Figure 3. Distribution of the sample points length in consecutive RR intervals for all heart beats in MITAD.

5. Conclusion

In this study, a dual-path CNN network is designed based on wavelet decomposition and attention mechanisms for detecting PVCs from ECGs. We utilize the morphological characteristics of the heartbeat itself, along with the features of surrounding heartbeats, to effectively identify PVCs. The features extracted from the two paths, Feature1 and Feature2, combined with the heart rate factor obtained during the preprocessing phase are demonstrated to be effective in the propose WACNN model.

Acknowledgments

This work has been partially supported by National Natural Science Foundation of China under Grant 82072014 and the Shandong Province Natural Science Foundation under Grant ZR2020MF028.

References

[1] F.-y. Zhou, L.-p. Jin, and J. Dong, “Premature ventricular contraction detection combining deep neural networks and rules inference,” *Artif. Intell. Med.,* vol. 79, pp. 42-51, 2017/06/01/, 2017.

[2] Y. Shen, Z. Cai, L. Zhang *et al.*, “Efficient Premature Ventricular Contraction Detection based on Network Dynamics Features,” *IEEE Trans. Instrum. Meas.,* vol. 73, pp. 1-15, 2024.

[3] P. De Chazal, and R. B. Reilly, “A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features,” *IEEE Trans. Biomed. Eng.,* vol. 53, no. 12, pp. 2535-2543, 2006.

[4] Z. Wang, K. Wang, X. Chen *et al.*, “A deep learning approach for inter-patient classification of premature ventricular contraction from electrocardiogram,” *Biomed. Signal Process. Control.,* vol. 94, pp. 106265, 2024.

[5] H. Chen, G. Wang, G. Zhang *et al.*, “CLECG: A novel contrastive learning framework for electrocardiogram arrhythmia classification,” *IEEE Signal Process. Lett.,* vol. 28, pp. 1993-1997, 2021.

[6] R. Hu, J. Chen, and L. J. C. i. B. Zhou, “A transformer-based deep neural network for arrhythmia detection using continuous ECG signals,” *Comput. Biol. Med.,* vol. 144, pp. 105325, 2022.

[7] G. Wang, M. Chen, Z. Ding *et al.*, “Inter-patient ECG arrhythmia heartbeat classification based on unsupervised domain adaptation,” *Neurocomputing,* vol. 454, pp. 339-349, 2021.

[8] M. U. Zahid, S. Kiranyaz, and M. Gabbouj, “Global ECG classification by self-operational neural networks with feature injection,” *IEEE Trans. Biomed. Eng.,* vol. 70, no. 1, pp. 205-215, 2022.

[9] T. J. Jun, H. M. Nguyen, D. Kang *et al.*, “ECG arrhythmia classification using a 2-D convolutional neural network,” *arXiv preprint arXiv:.06812*, 2018.

[10] Ö. Yildirim, “A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification,” *Comput. Biol. Med.,* vol. 96, pp. 189-202, 2018.

[11] Y. Xia, Y. Xiong, and K. J. B. S. P. Wang, “A transformer model blended with CNN and denoising autoencoder for inter-patient ECG arrhythmia classification,” *Biomed. Signal Process. Control.,* vol. 86, pp. 105271, 2023.

[12] J. Pan, and W. J. Tompkins, “A Real-Time Qrs Detection Algorithm,” *IEEE Trans. Biomed. Eng.,* vol. 32, no. 3, pp. 230-236, 1985.

[13] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*, pp. 7132-7141, 2018.

[14] G. B. Moody, and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Eng. Med. Biol. Mag.,* vol. 20, no. 3, pp. 45-50, 2001.

[15] A. L. Goldberger, L. A. N. Amaral, L. Glass *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet - Components of a new research resource for complex physiologic signals,” *Circulation,* vol. 101, no. 23, pp. E215-E220, Jun 13, 2000.

[16] L. Sun, Y. Wang, Z. Qu *et al.*, “BeatClass: a sustainable ECG classification system in IoT-based eHealth,” *IEEE Internet Things J.,* vol. 9, no. 10, pp. 7178-7195, 2021.

[17] Z. Zhang, J. Dong, X. Luo *et al.*, “Heartbeat classification using disease-specific feature selection,” *Comput. Biol. Med.,* vol. 46, pp. 79-89, 2014.

[18] R. Zarei, J. He, G. Huang *et al.*, “Effective and efficient detection of premature ventricular contractions based on variation of principal directions,” *Digit. Signal Process.,* vol. 50, pp. 93-102, 2016.

[19] V. Kalidas, and L. S. Tamil, “Detection of premature ventricular complexes using semisupervised autoencoders and random forests,” *2020 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 337-340, 2020.

[20] J. Yang, W. Cai, and M. J. P. M. Wang, “Premature beats detection based on a novel convolutional neural network,” *Physiol. Meas.,* vol. 42, no. 7, pp. 075003, 2021.

Address for correspondence:

Shoushui Wei
Shandong University, #17923 Jingshi Rd, Lixia District, Jinan, Shandong, China
sswei@sdu.edu.cn