SPP-AENet: A New ECG Biometric Identification Approach Based on Spatial Pyramid Pooling and Autoencoder

Xin Liu [1], Di Wang [2,3,*], Ping Wang [1], Tianyue Sun [1], Qi Sun [1], Yihan Fu [2,3], and Yu Fu [4]

[1] Noncommissioned Officer Institute, Army Academy of Armored Forces, Changchun, China
[2] School of Electronic and Information Engineering, Tiangong University, Tianjin, China
[3] Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tianjin, China
[4] Computer Engineering Technical College (Artificial Intelligence College), Guangdong Polytechnic of Science and Technology, Zhuhai, China
Email: liuxin52419558@126.com (X.L.); wangdi@tiangong.edu.cn (D.W.); wang19040313@163.com (P.W.); suntianyue1@126.com (T.S.); 15662136669@163.com (Q.S.); 2330080943@tiangong.edu.cn (Y.F.); fuyu_fiona@163.com (Y.F.)
*Corresponding author

***Abstract*—In recent years, Electrocardiogram (ECG) signals have emerged as a promising modality for biometric recognition, and have attracted significant attention. In this work, a new ECG biometric identification approach based on Spatial Pyramid Pooling (SPP) and Autoencoder (AE), referred to as the SPP-AENet, is proposed. The SPP-AENet is a hybrid architecture consisting of two parts: the SPP layer and AE layer. The SPP layer allows non-fixed length input and by incorporating prior knowledge, can ensure that critical characteristics of ECG signals are retained while compressing redundant information. The AE layer, through its automatic feature learning mechanism, can discover latent patterns and relationships within the data. The proposed method is evaluated using three public databases acquired under different health conditions, and gives a signal identification accuracy of 98.88%, 100%, and 100% for the ECG-ID, Physikalisch-Technische Bundesanstalt (PTB), and ST-Change database, respectively. To facilitate a better understanding of the model learning process, features of the SPP layer and the AE layer are visualized using t-distributed Stochastic Neighbor Embedding (t-SNE).**

***Keywords*—Electrocardiogram (ECG) biometric, Spatial Pyramid Pooling (SPP), Autoencoder (AE), SPP-AENet, t-distributed Stochastic Neighbor Embedding (t-SNE)**

I. Introduction

Biometric identification refers to the process of using unique physiological or behavioral characteristics to identify a person. Over the past few decades, biometric systems have played an essential role in security and privacy applications. Common biometric modalities such as face, fingerprint, iris and speech have already been successfully deployed in real scenarios. Recently, data from Electrocardiographic (ECG) signals has also proved its effectiveness on person identification due to its properties such as universality, uniqueness, performance and acceptability, and has become a promising candidate for biometric system [1].

Besides that, when compared with traditional biometric methods such as fingerprint or facial recognition, ECG-based biometric identification exhibits some advantages: (1) Unlike external biometric features like fingerprints or faces, ECG signals, being internal physiological characteristics, are generally more difficult to steal or counterfeit. (2) The intrinsic liveness detection characteristic of ECG helps to make the ECG-based biometric systems more resistant to fraud attacks or falsification attempts than traditional methods. (3) ECG signals are typically one-dimensional and thus ECG-based biometric identification systems often requires fewer storage resources than those using two-dimensional fingerprint or face images. Due to the above advantages, ECG-based biometric systems have the potential to be applied in health monitoring scenarios for patient identification and managing clinical records, and high-security facilities or restricted areas where continuous identification is required.

To date, various methods for ECG biometric identification have been developed. Generally, a complete ECG biometric identification contains three steps: the preprocessing stage, the feature extraction stage and the classification stage. According to the feature extraction manner, traditional ECG identification methods can be typically divided into two categories [2]: the fiducial and the non-fiducial. In the fiducial methods, morphological [3] and temporal features [4] extracted based on characteristic points such as P, Q, R, S, and T are used. The non-fiducial methods, on the other hand, do not rely on these specific points and instead analyze the ECG signal using statistical features obtained by techniques such as local binary pattern Principal Component Analysis

Manuscript received August 23, 2024; revised September 23, 2024; accepted October 9, 2024; published January 15, 2025.

(PCA) [5], Latent Dirichlet Allocation (LDA) [6] and empirical mode decomposition [7]. Despite the achieved good performance, features of traditional methods are often manually designed and extracted through the feature engineering method, which has been identified as a heavy workload.

In recent years, deep learning with the merits of automatic feature engineering and nonlinear feature mapping provides a new way for ECG biometric identification. Unlike traditional feature-engineering based methods, deep learning approaches can automatically learn hierarchical representations from original ECG signals, eliminating the need for manual feature extraction [8–10]. These data-driven models can capture complex patterns and subtle characteristics that may be overlooked by conventional techniques in a layer-by-layer way, leading to improved accuracy. Labati *et al.* [11] presented a Convolutional Neural Network (CNN)-based scheme for ECG biometric identification, which employed multiple convolutional layers to learn abstract representations of ECG, followed by fully connected layers for classification. Boujnouni *et al.* [12] developed a new ECG biometric identification method, which is a combination of continuous wavelet transform, discrete wavelet transform and Capsule network.

While these deep-learning based approaches have demonstrated impressive performance in ECG biometric identification, certain limitations still exist. Firstly, in consideration of the real-time performance and good user experience, most existing research focused on increasing performance using single-heartbeat [13–15]. During the model training, researchers usually employ fixed-length ECG signal centered on the R peak as input data. Although ECG is a periodic signal, the duration of each beat can still vary due to factors such as heart rate change. Consequently, their fixed-length method may result in loss of critical information pertaining to certain waveforms, particularly the T and P waves, and the inconsistency in the physiological significance among heartbeats, thus leading to decreased performance when dealing with complex and variable ECG patterns. Secondly, the black-box nature of deep learning models often makes it challenging to understand the extracted features.

To address the above issues, a new ECG biometric identification method, referred to as the SPP-AENet, has been proposed and evaluated. The SPP-AENet mainly contains two parts: the prior knowledge-based SPP layer and the feature learning-based AE layer. Specifically, the SPP layer is designed based on the prior knowledge. It allows non-fixed length heartbeats, which contain the same physiological significance, as input, and can preserve key waveform features while compress redundant information. The AE layer employs automatic feature learning to extract discriminative representations. Furthermore, features of the two parts are respectively visualized to provide insights into the network's learning process, which may facilitate the potential future improvement in the network architecture or training process. The main contributions are as follows:

- We provide a new deep learning-based architecture for accurate ECG biometric identification. The SPP-AENet can take non-fixed length heartbeats with the same physiological significance as input, thus avoiding potential information loss of certain waveforms and enhancing the ECG biometric identification performance.
- Features of the SPP layer and the AE layer are visualized using interpretability techniques. This visualization provides insights into the network's learning process and feature representation, facilitating a deeper understanding of how the model processes ECG signals for biometric identification.
- Extensive experiments are conducted on multiple public ECG datasets acquired under different health and environment conditions to evaluate the performance of the proposed SPP-AENet. The results demonstrate its superior accuracy compared to state-of-the-art methods in ECG biometric identification tasks.

II. Methods

Fig. 1 depicts the overall flowchart of the proposed scheme for ECG biometric identification. As shown in Fig. 1, the proposed system mainly contains two stages: preprocessing and identification. Specifically, the preprocessing stage aims to separate the required biological information from the background noise and divide long signal into multiple cardiac cycles for subsequent identification. The identification stage is essentially the proposed SPP-AENet model, which intends to determine "who does this ECG belong to?"

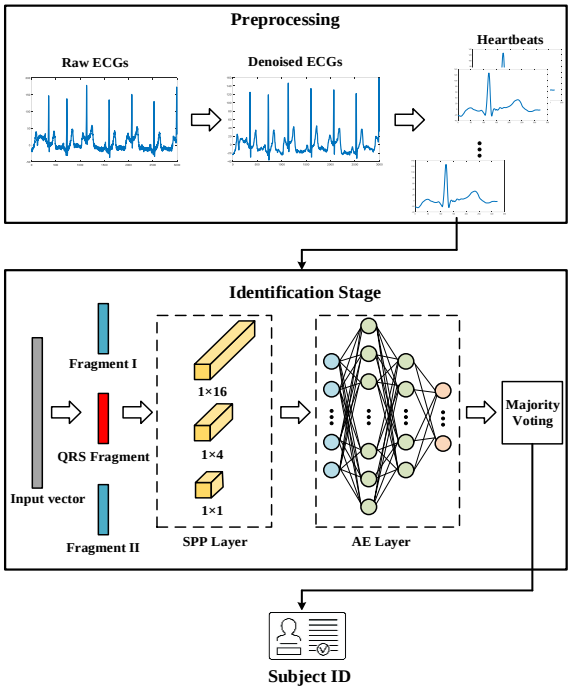


Fig. 1. The overall flowchart of the proposed scheme for ECG biometric identification.

A. Preprocessing

1) Denoising

Due to external factors, ECG signal contains various kinds of noises, e.g., baseline wandering, powerline interference, and electromyographic noise. Generally, signal quality is essential for neural network training. Therefore, denoising is necessary to construct a reliable ECG biometric identification model. In this study, a wavelet decomposition reconstruction algorithm along with wavelet threshold soft method [16] is applied for noise removal. Specifically, the wavelet decomposition reconstruction algorithm is firstly used to filter low-frequency noises (e.g., baseline wandering) occurring around 0.5–0.6 Hz. Then, through the wavelet thresholding, high-frequency noises such as powerline interference and electromyographic noise are removed.

2) Segmentation

After denoising, segmentation is performed to make the format of the data suitable for the subsequent identification stage. Fig. 2 shows the segmentation process in this study.

As shown in Fig. 2, one heartbeat is constructed based on the R points, which are detected using the Pan-Tompkins (PT) algorithm in this study. Specifically, taking the heartbeat centered on R2 as example, it covers 0.7×R1–R2 interval preceding the R2 peak and 0.7×R2–R3 interval length following the R2 peak. In this way, the segmentation approach allows for a comprehensive representation of the cardiac electrical activity including the P wave, QRS complex and T wave, within a normalized time frame relative to each individual heartbeat. It ensures the extracted waveform encompasses the primary components of the cardiac cycle while maintaining consistency across varying heart rates. After segmentation, all the heartbeats are normalized using the min-max normalization method.

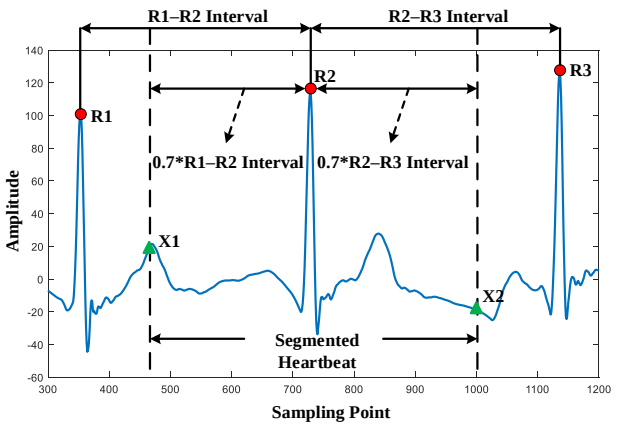


Fig. 2. ECG segmentation process.

B. Network Architecture

In this study, the proposed SPP-AENet mainly contains two parts: the SPP layer and the AE layer. The details of them are presented below.

1) SPP layer

After segmentation, heartbeats comprising the P wave, QRS complex and T wave etc. are obtained. It should be noted that due to the differences of R1–R2 and R2–R3 intervals among cardiac cycles, the length of the obtained heartbeats exhibits variability. To solve this problem, a new SPP layer, which is designed based on prior knowledge that not all waveform components within a heartbeat are equally crucial for ECG biometric identification, is proposed in this study. Specifically, based on the prior literature [17, 18], our main idea is that the QRS complex contributes most to ECG identification, and its temporal and morphological information is all important. For the rest parts in the heartbeat, since they are susceptible to heart rate change, only the most salient features are considered and preserved. Fig. 3 depicts the architecture of the proposed SPP layer. As shown in Fig. 3, a heartbeat is firstly segmented to three parts: Fragment I, QRS Fragment, and Fragment II. Specifically, the QRS Fragment is centered on the detected R peak, and it covers 0.2 s preceding the R peak and 0.2 s following the R peak. In the SPP layer, to preserve all the QRS information originally, the QRS Fragment does not need any process. For Fragment I and II, they are both processed by a three-level spatial pyramid pooling layer to extract features at multiple scales. Here, the pooling regions at each level are configured with dimensions of 1×1, 1×4, and 1×16, respectively. Finally, the QRS Fragment is concatenated with the output of spatial pyramid pooling layer to get the input of the AE layer. The output of the SPP layer can be described as Eq. (1):

$$Output = [\mathrm{SPP}(Fragment\ I),\ QRS\ Fragment,\ \mathrm{SPP}(Fragment\ II)] \quad (1)$$

where $SPP(\bullet)$ represents the spatial pyramid pooling operation. In this study, considering that the number of pyramid levels is 3 and the pooling regions at each level are configured with dimensions of 1×1, 1×4 and 1×16, the output dimensions of SPP(*Fragment* I) and SPP(*Fragment* II) are fixed at 21. Since the length of *QRS Fragment* is set to 0.4 s, the output dimension of SPP layer *Output* can be expressed as follows:

$$Output\ dimention = 0.4 \times Samplingrate + 21 \times 2 \quad (2)$$

where *Samplingrate* is the sampling rate of ECG. It can be seen that once the *Samplingrate* is determined, the SPP layer is able to generate a fixed-length output regardless of varying input sizes.

By preserving the QRS Fragment intact and applying SPP to the surrounding fragments, the proposed SPP layer aims to balance the preservation of critical QRS information with the extraction of multi-scale features from the rest of the heartbeat. This approach can potentially help improve the model's ability to detect subtle abnormalities in both the QRS complex and the surrounding P and T waves, leading to more accurate ECG classification and interpretation.

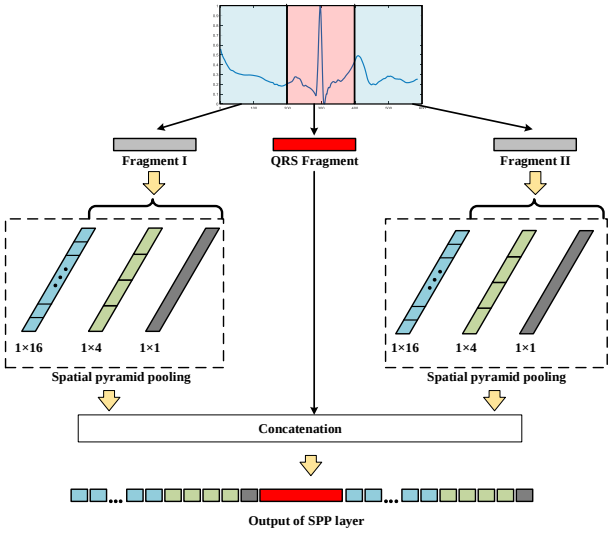


Fig. 3. The SPP layer of the proposed SPP-AENet.

2) AE layer

In recent years, autoencoder has become a good candidate for feature learning and been widely applied in various tasks such as image analysis and time series prediction. In this study, to further automatically learn discriminative features from the output of the SPP layer, an AE layer is designed. The proposed AE layer mainly contains two autoencoders and its training can be divided into two steps:

- Pretraining step: Once the parameters such as weights and biases are initialized, the AE layer is pretrained in a two-stage way as shown in Fig. 4.

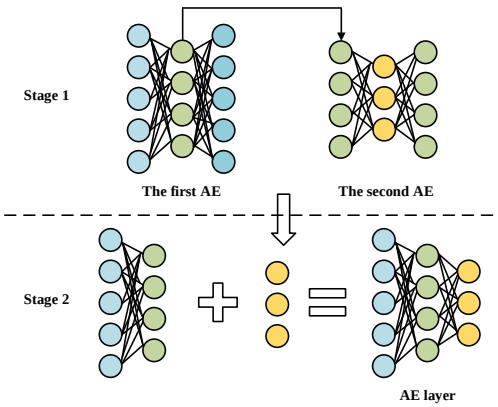


Fig. 4. Pretraining step.

During the first stage, the autoencoders are pretrained with the purpose of reconstruction error minimization between the input and output. In each autoencoder, it is noted that the input layer and the output layer have the same number neurons. Here, the input of the first autoencoder is the output of the SPP layer and that of the second autoencoder is the hidden layer output of the first autoencoder. Then in the second stage, output layer of the two autoencoders are both removed and the rest layers (the input layer and the hidden layer of the first AE, the hidden layer of the second AE) are stacked together to obtain the AE layer. During the pretraining, the cost function is defined as Eq. (3):

$$J(W,b)=\frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{2}\left\|h_{W,b}(x_i)-x_i\right\|^2\right)+\frac{\lambda}{2}\sum_{j=1}^{2}\sum_{k}\left\|W_{jk}\right\|^2+\beta\sum_{n=1}^{S_2}\mathrm{KL}(\rho\,\|\,\rho_n) \quad (3)$$

where W, b are respectively a weight matrix and a bias vector; $h_{W,b}(x_i)$ represents the result of an example x_i after passing through the whole transformation process of the autoencoder; m is the sample number; The second term $\frac{\lambda}{2}\sum_{j=2}^{2}\sum_{k}\left\|W_{jk}\right\|^2$ denotes the square sum of all weights with a hyperparameter λ ; The third term aims to add a sparsity constraint with a penalty weight β. S_2 is the neuron number of the hidden layer and $KL(\bullet)$denotes the Kullback-Leibler (KL) divergence. Specifically, the activation function used in the first and second hidden layers of the AE model is the sigmoid function, which is defined by Eq. (4):

$$f(x)=1/(1+\exp(-x)) \quad (4)$$

- Fine-tuning step: After pretraining, a softmax layer is stacked with the two previous layers, and the back-propagation algorithm is used to optimize the whole stacked network parameters. Details and parameter settings of the proposed AE layer can be seen in Table I.

TABLE I. DETAILS AND PARAMETER SETTINGS OF THE PROPOSED AE LAYER

Parameter Setting	Value
Input layer size	Heart rate × 0.4 + 42
First hidden layer size	200
Second hidden layer size	100
Output layer size	89(ECG-ID)/290(PTB)/28(ST-Change)
Hyperparameter λ	3×10^{-9}
Penalty weight β	0.9
Epoch number	800

C. Majority Voting

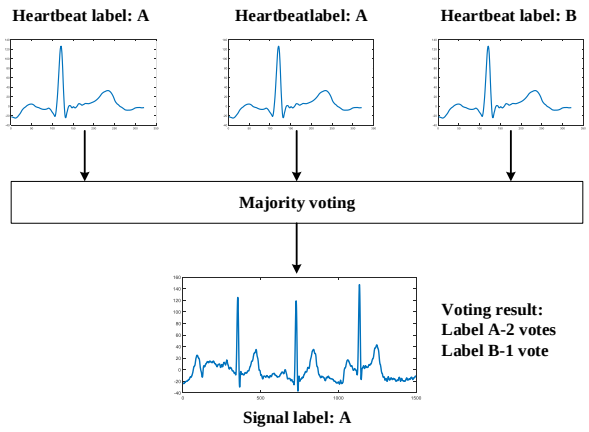


Fig. 5. The majority voting strategy.

The AE layer can output the identification results of the heartbeats. Considering a signal typically contains multiple heartbeats, a majority voting strategy is adopted to determine the overall signal label. As shown in Fig. 5,

taking a signal containing three heartbeats as example, the heartbeats are assigned label A, A, and B after the AE layer, respectively. As a result, label A gets 2 votes and label B gets 1 vote. With majority voting, we will take the label with the most votes as the signal label, so the overall signal label is A.

III. Result and Discussion

A. Database and Experimental Setup

In this study, the proposed scheme's performance was evaluated and analyzed using three public databases obtained from the Physionet: the ECG-ID Database, the MIT-BIH ST Change Database and the Physikalisch-Technische Bundesanstalt (PTB) Database. Descriptions of these three databases are summarized in Table II.

TABLE II. Descriptions of ECG Databases Used in This Paper

Database	Sample Rate (Hz)	Subject Quantity	Channel Adopted	Health Condition
ECG-ID	500	89	I-Lead	Healthy
PTB	1000	290	I-Lead	Mixed (healthy, myocardial infraction and other 7 heart diseases)
ST-Change	360	28	The first channel	ST depression

After preprocessing, a total of 4095, 10,486, and 16,800 ECG heartbeats were obtained from the ECG-ID, PTB and ST-Change databases, respectively. For the ECG-ID database, since only a small subset of subjects has more than two recordings, the first two recordings of each subject were adopted for the experiments. Specifically, a total of 178 records (two from each of the 89 subjects) were used, each with a length of 20 s. During the experiments, two recordings of each subject alternately served as the training set and the test set, and a two-fold cross-validation scheme was performed. For the PTB and ST-Change database, only the first recording of each subject was utilized, in accordance with previous studies [12, 14]. As a result, the number of recordings for the PTB and ST-Change was 290 (one from each of the 290 subjects) and 28 (one from each of the 28 subjects), respectively. For each recording, the entire signal duration was adopted. During the experiments, heartbeats of each signal were divided into ten folds according to the sequence of time, and a 10-fold cross validation was applied. Specifically, one fold comprising 10% of the heartbeats was used for testing, while the rest nine folds for training. The SPP-AENet model was implemented using MATLAB 2024a, and the training and testing were conducted on the computing platform with Intel core i7-4720HQ CPU, GeForce GTX 960 and 16G RAM.

B. Performance Metrics

In this section, five statistical performance metrics were introduced and used for the assessment of the model performance. These metrics are defined as below:

Heartbeat Identification (HI) accuracy is defined as Eq. (5):

$$\text{HIAccuracy} = \text{Heartbeat}_{\text{Correct}} / \text{Heartbeat}_{\text{all}} \quad (5)$$

The Sensitivity (Sen) is defined as Eq. (6):

$$\text{Sen} = \text{TP} / (\text{TP} + \text{FN}) \quad (6)$$

The precision (Pre) is defined as Eq. (7):

$$\text{Pre} = \text{TP} / (\text{TP} + \text{FP}) \quad (7)$$

The Specificity (Spe) is defined as Eq. (8):

$$\text{Spe} = \text{TN} / (\text{TN} + \text{FP}) \quad (8)$$

Signal identification accuracy is defined as Eq. (9):

$$\text{SIAccuracy} = \text{Signal}_{\text{Correct}} / \text{Signal}_{\text{all}} \quad (9)$$

where $\text{Hearbeat}_{\text{Correct}}$ denotes the number of correctly classified heartbeats, and $\text{Heartbeat}_{\text{all}}$ is the total number of heartbeats. TP, TN, FP, FN refer to the true positive, true negative, false positive and false negative heartbeat predictions, respectively. $\text{Signal}_{\text{Correct}}$ denotes the number of correctly classified signal, and $\text{Signal}_{\text{all}}$ is the total number of signals.

C. Performance Evalutaion

Table III depicts the heartbeat identification performance of the proposed SPP-AENet. Experimental results show that on ECG-ID (89 subjects), PTB (290 subjects) and ST-Change (28 subjects) database, SPP-AENet achieved a sensitivity of 91.43%, 91.93%, and 90.28%, a specificity of 99.89%, 99.97%, and 99.60%, a precision of 91.97%, 91.80%, and 91.16%, and a heartbeat accuracy of 91.95%, 92.04%, and 90.28%, respectively. The above results indicate that the relationship between single heartbeat and ID label can be modeled with good accuracy to some extent.

TABLE III. Heartbeat Identification Results Achieved by SPP-AENet on Three Public Databases

Database	Performance metrics			
	Sen.	Spe.	Pre.	HI Accuracy
ECG-ID	91.43%	99.89%	91.97%	91.95%
PTB	91.93%	99.97%	91.80%	92.04%
ST-Change	90.28%	99.60%	91.16%	90.28%

In addition, Table IV shows the signal identification accuracy under different ECG length conditions. It can be observed that with the increase of heartbeat number, the accuracy of signal identification correspondingly improves. That may be because that more heartbeats involved in the voting process can effectively reduce the impact of occasional distorted cardiac cycles on identification accuracy and allows for a more comprehensive analysis. Besides that, it is noticed that with only 5 heartbeats, the SPP-AENet can already achieve subject accuracy over 97% on ECG-ID and PTB databases, which indicates the potential of implementing the method on devices required for short-term ECG identification. Based on the above analysis, it can be noted that the proposed SPP-AENet can fit well with the ECG data and achieve an accurate classification of subject ID.

TABLE IV. SIGNAL IDENTIFICATION ACCURACY UNDER DIFFERENT ECG LENGTH CONDITIONS

ECG length	Database		
	ECG-ID	PTB	ST-Change
1 heartbeat	91.15%	92.04%	90.28%
3 heartbeats	94.86%	95.21%	91.53%
5 heartbeats	97.64%	97.31%	92.82%
7 heartbeats	97.67%	97.51%	93.18%
Whole Signal	98.88%	100%	100%

D. Ablation Study

To verify the effectiveness of each component in the SPP-AENet, ablation experiments are conducted. Table V shows the design of ablation experiments. It can be seen that the ablation experiments mainly focus on the effectiveness of different layers: the SPP layer and the AE layer. Here, the signal identification accuracies were evaluated using the whole signal length.

TABLE V. DESIGN OF ABLATION EXPERIMENTS

Method	Component	
	SPP Layer	AE Layer
SPP-Net	√	×
AE-Net	×	√
SPP-AENet	√	√

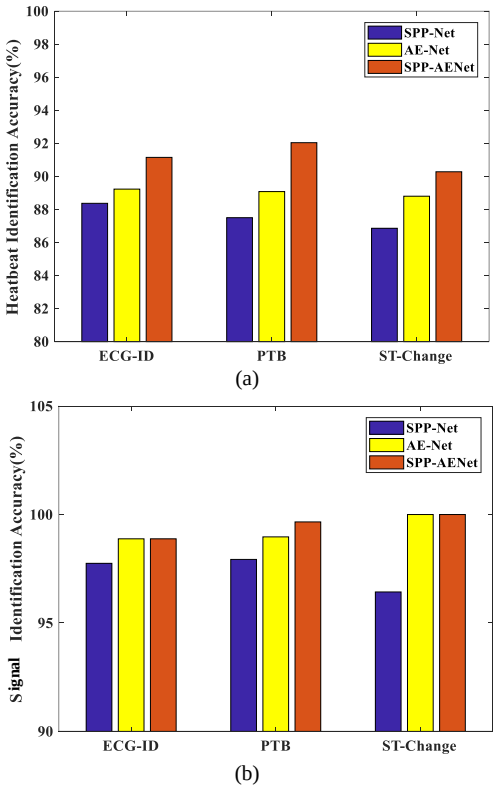
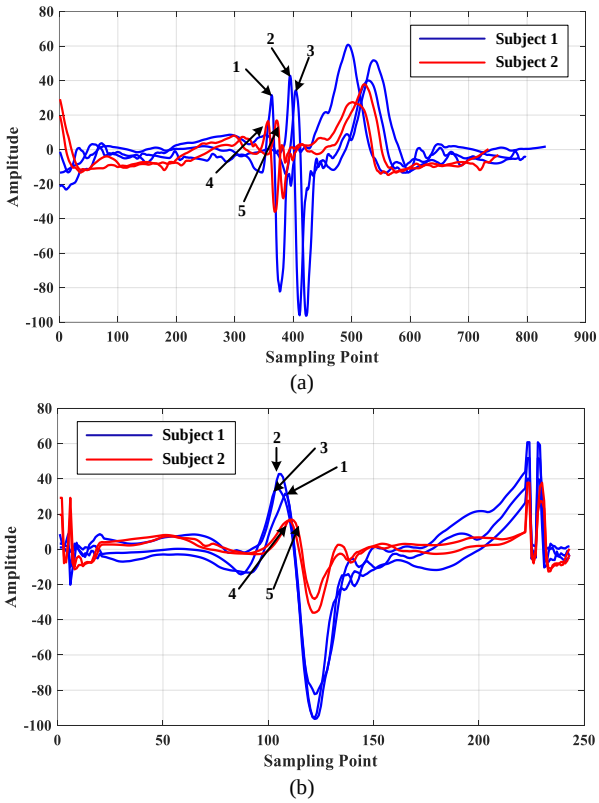


Fig. 6. Performance Comparison in ablation study.(a) Heartbeat identification accuracy comparison. (b) Signal identification accuracy comparison.

As shown in Fig. 6, it is noted that SPP-AENet achieves the highest heartbeat and signal identification accuracy on three databases. When either the SPP layer or the AE layer is removed, the identification accuracies all suffer a significant decrease, which demonstrates the necessity of both SPP layer and AE layer.

E. Feature Visualization

To further explain the functions of the SPP layer and AE layer, the visualization of features is shown in Fig. 6. Specifically, Fig. 7(a) displays the original temporal features of several heartbeats extracted from two subjects. It can be seen that for original temporal features, even heartbeats (e.g., 1, 2, 3) which come from the same subject do not exhibit the same ECG pattern, thus can hardly be directly used for ECG biometric. Fig. 7(b) shows the feature output of the SPP layer. By retaining the original QRS waveform and extracting multi-scale pooling information from the rest of time series, the SPP layer makes heartbeats from the same subject more similar to each other, while heartbeats from different subjects exhibit more significant discrepancy. The visualization of features of the SPP layer is also displayed using t-SNE in Fig. 7(c). It can be seen that although heartbeat samples from multiple subjects can be discriminated to some extent, there still exists overlap among different classes (e.g., Subject B, D, F). To solve this problem, SPP-AENet further adopts the AE layer to conduct automatic feature learning. Fig. 7(d) shows the t-SNE visualization for features of the SPP-AENet. According to Fig. 7(d), the AE layer optimizes the feature representation through the simultaneous minimization of intra-class distances and maximization of inter-class distances, thus resulting in more well-defined decision boundaries among different classes. In all, both the SPP layer and the AE layer are essential for ECG biometric identification.



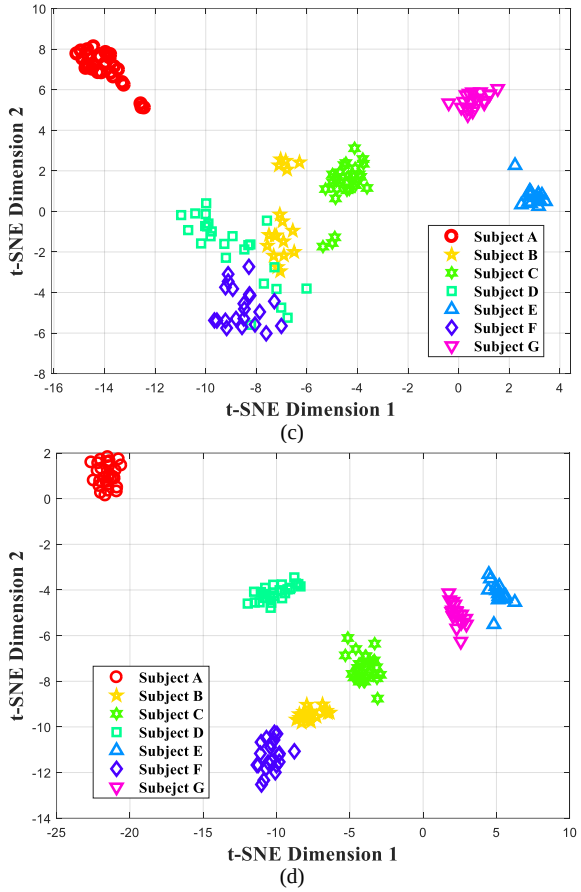


Fig. 7. Feature visualization. (a) Original temporal features; (b) Features of the SPP layer; (c) t-SNE visualization for features of the SPP layer; (d) t-SNE visualization for features of the SPP-AENet.

F. Comparison to Related Works

TABLE VI. PERFORMANCE COMPARISON WITH PREVIOUS WORKS

Method	Required Fiducial Point	Database	Subject	[1]SI Accuracy
LSTM [19]	R	ECG-ID	89	100%
TQRST-PCANet [20]	R, T	ECG-ID	89	97.75%
CapsNet [12]	R	ECG-ID	89	93.75%
TVPCAD [21]	R	ECG-ID	89	99.25%
SPP-AENet	R	ECG-ID	89	98.88%
CNN [22]	None	PTB	235	99.1%
Phase features-SVM [23]	T	PTB	54	95.4%
CapsNet [12]	R	PTB	235	94.59%
Spatial features-SVM [24]	None	PTB	200	95.15%
DCRN [25]	None	PTB	290	94.76%
TVPCAD [21]	R	PTB	290	99.8%
WT-CapsNet [14]	R	PTB	290	98.8%
SPP-AENet	**R**	**PTB**	**290**	**100%**
Spatial features-SVM [24]	None	ST-Change	28	98.93%
WT-CapsNet [14]	R	ST-Change	28	98.2%
SPP-AENet	**R**	**ST-Change**	**28**	**100%**

Note: [1]SI-Signal Identification

To further evaluate the proposed method, a comparison was made between the proposed SPP-AENet and related previous works, as presented in Table VI. Here it is noted that for techniques which were deemed difficult to reliably reproduce, their reported findings were directly adopted to compare with our results. As shown in Table VI, it can be seen that end-to-end approaches such as SPP-AENet, TQRST-PCANet tend to achieve higher signal identification accuracy than traditional methods adopting handcrafted features. That may be because end-to-end networks are able to conduct automatic feature representation on the ECG input, which avoids potential information loss during manual feature extraction. When further compared with TQRST-PCANet, another advantage of the SPP-AENet is that it achieves promising results using only the R points, which are easy to be detected in most cases, thus facilitating its potential application in real world scenarios. The results demonstrate the effectiveness of the SPP-AENet.

IV. CONCLUSION

In this work, we have proposed a new ECG biometric identification method based on spatial pyramid pooling and autoencoder. The proposed neural network is designed to learn the biometric representation by simultaneously taking advantage of the prior knowledge and automatic feature learning. Experimental results show that on the ECG-ID, PTB and ST-Change database, the proposed SPP-AENet achieves 98.88%, 100%, 100% signal identification accuracy, which outperforms related works in the literature. Furthermore, techniques including t-SNE are employed to help understand the learning process of the network better, which may facilitate the potential future improvement.

V. LIMITATION

Despite the achieved good performance, there are still limitations to be considered for ECG as a biometric identifier. Research has shown that having caffeinated beverage, ECG parameters including peak amplitude of waves and their corresponding intervals change significantly. To improve robustness, further studies are needed to assess how daily activities impact the accuracy of ECG biometric in the future. Besides that, multi-modal biometric fusion which combines ECG with other modalities such as face and fingerprint can also help to compensate the limitations of individual biometric methods.

Additionally, current studies mainly focus on ECG signals from standard 12-lead or Frank leads. Although these recordings have high signal-to-noise ratio, acquisitions of them are often inconvenient in practical applications. To improve the acceptability, off-the-person systems, such as wearable devices or non-contact sensors, have been proposed. However, these alternative methods often suffer from lower signal quality and increased motion artifacts, which may compromise the accuracy and reliability of ECG biometrics. Further research including developing more sophisticated signal processing algorithms is needed to enhance signal quality and extract relevant features from ECG recordings.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Xin Liu, Di Wang conducted the research; Xin Liu, Di Wang, Ping Wang analyzed the data; Xin Liu, Di Wang originally wrote the paper; Tianyue Sun, Qi Sun reviewed and revised the paper. Yihan Fu and Yu Fu edited the paper. All authors had approved the final version.

FUNDING

This research was funded by Opening foundation, Tianjin Key Laboratory of Optoelectronic Detection Technology and System under grant number 2024LODTS112.

REFERENCES

[1] Z. Zhao, Y. Zhang, Y. Deng, and X. Zhang, ''ECG authentication system design incorporating a convolutional neural network and generalized s-transformation," *Comput. Biol. Med*., vol. 102, pp. 168–179, Nov. 2018.

[2] Y. Li, Y. Pang, K. Wang, and X. Li, "Toward improving ECG biometric identification using cascaded convolutional neural networks," *Neurocomputing*, vol. 391, pp. 83–95, May 2020.

[3] D. Rezgui and Z. Lachiri, "ECG biometric recognition using SVM-based approach," *IEEJ Trans. Electr. Electron. Eng.*, vol. 11, pp. S94–S100, Jun. 2016.

[4] D. Jyotishi and S. Dandapat, "An ECG biometric system using hierarchical LSTM with attention mechanism," *IEEE Sens. J*., vol. 22, pp. 6052–6061, Mar. 2022.

[5] M. Hejazi, S. A. R. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. A. Aziz, "ECG biometric authentication based on non-fiducial approach using kernel methods," *Digit. Signal Prog*., vol. 52, pp. 72–86, May 2016.

[6] R. Srivastva and Y. N. Singh, "ECG analysis for human recognition using non-fiducial methods," *IET Biom*., vol. 8, pp. 295–305, Sep. 2019.

[7] H. Ferdinando, T. Seppanen, and E. Alasaarela, "Bivariate empirical mode decomposition for ECG-based biometric identification with emotional data," in *Proc. 39th Annu. IEEE Engn. Med. & Biol. Soc.*, South Korea, 2017, pp. 450–453.

[8] R. Srivastva, A. Singh, and Y. N. Singh, "PlexNet: A fast and robust ECG biometric system for human recognition," *Inf. Sci*., vol. 558, pp. 208–228, May 2021.

[9] D. Jyotishi and S. Dandapat, "An ECG biometric system using hierarchical LSTM with attention mechanism," *IEEE Sensors J*., vol. 22, no. 6, pp. 6052–6061, Mar. 2022.

[10] D. A. AlDuwaile and M. S. Islam, "Using convolutional neural network and a single heartbeat for ECG biometric recognition," *Entropy*, vol. 23, no. 6, 733, Jun. 2021.

[11] R. D. Labati, E. Munoz, V. Piuri, and R. Sassi. "Deep-ECG: Convolutional neural networks for ECG biometric recognition," *Pattern Recognit. Lett.*, vol. 126, pp. 78–85, Sep. 2019.

[12] E. Boujnouni and A. Tali, "Biometric identification of individuals based on scalograms of electrocardiogram and capsule network," in *Proc. 3rd Int. Conf. Adv. Intell. Syst. Sustain. Dev (AI2SD)*, Tangier, Morocco, 2022, pp. 29–42.

[13] L. Sun, Z. Zhong, and Z. Qu, "PerAE: An effective personalized autoEncoder for ECG-based biometric in augmented reality system," *IEEE J. Biomed. Health Inform.*, vol. 26, pp. 2435–2446 Jun. 2022.

[14] E. Boujnouni and H. Zili, "A wavelet-based capsule neural network for ECG biometric identification," *Biomed. Signal Process. Control*., vol. 76, Jul. 2022.

[15] E. Merdjanovska and A. Rashkovska. "Comprehensive survey of computational ECG analysis: Databases, methods and applications," *Expert Syst. Appl*., vol. 203, Oct. 2022.

[16] X. Wang, W. Cai, and M. Wang, "A novel approach for biometric recognition based on ECG feature vectors," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023.

[17] E. Maiorana, C. Romano, and E. Schena, "BIOWISH: Biometric recognition using wearable inertial sensors detecting heart activity," *IEEE Trans. Dependable Secur. Comput*., vol. 21, pp. 987–1000, Mar. 2024.

[18] D. R. Bhuva and S. Kumar, "A novel continuous authentication method using biometrics for IOT devices," *Internet Things*, vol. 24, Dec. 2023.

[19] R. Salloum and C. C. J. Kuo, "ECG-based biometrics using recurrent neural networks," in *Proc IEEE Int Conf on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, pp. 2062–2066.

[20] D. Wang, Y. Si, and W. Yang, "A novel heart rate robust method for short-term electrocardiogram biometric identification," *Applied Sciences*, vol. 9, Jan. 2019.

[21] H. Liu, H. Lin, and X. Wang, "Total variation PCA-based descriptors for electrocardiography identity recognition," *IEEE Access*, vol. 12, pp. 3815–3824, Jan. 2024.

[22] J.-K. Chiu, C.-S. Chang, and S.-C. Wu, "ECG-based biometric recognition without QRS segmentation: A deep learning-based approach," in *Proc. 43rd Annu. IEEE Engn. Med. & Biol. Soc.*, 2021, pp. 88–91.

[23] S. Hamza and Y. B. Ayed, "Recognition of person using ECG signals based on single heartbeat," in *Proc. Int. Conf. on Intelligent Systems Design and Applications (ISDA)*, World Wide Web, 2021, pp. 452–460.

[24] T. N. Alotaiby and S. A. Alshebeili, "ECG-based subject identification using common spatial pattern and SVM," *J. Sens.*, Apr. 2019.

[25] Z. Gong, Z. Tang, and Z. Qin, "Electrocardiogram identification based on data generative network and non-fiducial data processing," *Comput. Biol. Med*., vol. 173, pp. 108333–108333, May 2024.