ECG Authentication Method Based on Parallel Multi-scale One-dimensional Residual Network with Center and Margin Loss

3 authors, including:

Kejie Huang
Zhejiang University
30 PUBLICATIONS   **79** CITATIONS

SEE PROFILE

ECG Authentication Method Based on Parallel Multi-scale One-dimensional Residual Network with Center and Margin Loss

YIFAN CHU[1], HAIBIN SHEN[2], AND KEJIE HUANG[3]
[1, 2, 3]Zhejiang University, Institute of VLSI Design, Hangzhou, China, 310058 (e-mails: {yfchu, shen_hb, huangkejie}@zju.edu.cn)

Corresponding author: Kejie Huang (e-mail: huangkejie@zju.edu.cn).

ABSTRACT To enhance the security level of digital information, the biometric authentication method based on Electrocardiographic (ECG) is gaining increasing attention in a wide range of applications. Compared with other biometric features, e.g., fingerprint and face, the ECG signals have several advantages, such as higher security, simpler acquisition, liveness detection, and health information. Therefore, various methods for ECG-based authentication have been proposed. However, the generalization ability of these methods is limited because the feature extraction for the ECG signals in conventional methods is data dependent. To improve the generalization ability and achieve more stable results on different datasets, a parallel multi-scale one-dimensional residual network is proposed in this paper. This network utilizes three convolutional kernels with different kernel sizes, achieving better classification accuracy than the conventional schemes. Moreover, two loss functions named center loss and margin loss are used during the training of the network. Compared with the conventional softmax loss, these two loss functions can further improve the generalization ability of the extracted embedding features. Furthermore, we evaluate the effectiveness of our proposed method thoroughly on the ECG-ID database, the PTB Diagnostic ECG database, and the MIT-BIH Arrhythmia database, achieving 2.00%, 0.59%, 4.74% Equal Error Rate (EER), respectively. Compared with other works, our proposed method improves 1.61% and 4.89% classification accuracy on the ECG-ID database and the MIT-BIH Arrhythmia database, respectively.

INDEX TERMS ECG, authentication, feature extraction, residual network, multi-scale, center loss, margin loss

I. INTRODUCTION

WITH the rapid development of information technology, information security is facing increasing threats. From the personal phone number to the payment password, the leaking of digital information can cause huge losses. Different identification and authentication methods are used to enhance the security level, including token-based, knowledge-based and biometric-based methods. The token may be lost and the knowledge may be forgotten, but the physiological signals are unforgettable and unique. Thus, the biometric authentication method is attracting great attention in security applications.

Among all physiological signals, the Electrocardiographic (ECG) signal has four unique advantages [1]: 1) The ECG signal is an intrinsic physiological signal, which means it has higher security than some external features, such as fingerprint and face. 2) With the maturity of acquisition technology, the ECG signals can be acquired easily by mobile equipment. 3) The liveness detection prevents the ECG signals from counterfeiting. 4) The ECG signals contain some health information, which can be used in some medical fields.

Many methods for ECG-based authentication have been proposed. For example, fiducial methods, non-fiducial methods, and partially fiducial methods [1]. Recently, deep neural networks are widely used for feature extraction and classification. These methods can achieve effective results in the authentication. However, the biggest issue in these methods is that the generalization ability is limited, especially in the matching task. For the individuals who have not been seen before, these methods usually predict unconvinced results.

In this paper, we present an ECG biometric authentication method based on parallel multi-scale one-dimensional residual network, which can improve the generalization ability of the method on different ECG signals sampled in the different environment for the matching task. Specifically, we perform a signal preprocessing based on R point detection on the raw ECG signals and enlarge the dataset to generate the training and test sets. After that, we propose a parallel multi-scale one-dimensional residual network including three convolutional kernels with different sizes to extract more detailed features.

The widely used softmax loss function does not require the intra-class compactness and inter-class dispersion, and thus limits the generalization ability in the matching task. To address this issue, we utilize the center loss and the margin loss to train our residual network. After training, we use the network to generate an embedding vector, which represents the features extracted from the input ECG signals. For new users, we can also obtain the corresponding embedding vectors by the trained network, which means that the re-training is unnecessary. The result of the matching task in authentication is the similarity between the test signal and the enrolled signal, which is calculated by the Euclidean distance between two embedding vectors.

Compared with other works, our proposed method is more suitable for matching tasks in identity recognition, and also has good performance in classification tasks. The effectiveness of the proposed method is thoroughly evaluated on the ECG-ID database, the PTB Diagnostic ECG database, and the MIT-BIH Arrhythmia database, achieving 0.59-4.74% Equal Error Rate (EER). We also fine-tuned the network and test the closed-set classification performance, achieving 95.99-100% accuracy on these three databases.

The remainder of the paper is organized as follows: section II introduces the related studies and achievements on ECG biometrics; section III illustrates the detailed design of the proposed ECG authentication method; section IV presents the experiments to demonstrate the advantage of our method and section V provides the result of experiments compared with the state-of-the-art works. Finally, section VI draws the conclusion.

II. RELATED WORKS

An ECG-based biometric authentication method usually contains four procedures: data collection, preprocessing, feature extraction, and identification [1], where feature extraction and identification are two important procedures. The ECG feature extraction approaches can be classified into three strategies: fiducial, non-fiducial, and partially fiducial methods. While the identification approaches can be divided into classification task and matching task.

A. FEATURE EXTRACTION
1) Fiducial methods
Fiducial methods extract features based on the morphological characteristics of ECG signals, such as amplitude, width,

angle, or slope of the sub-waves [2]–[4]. The morphological characteristics are determined by the fiducial points (P, Q, R, S, T, and U) [5]–[7], and thus these methods heavily rely on the quality of heartbeat segment and the precision of the fiducial points detection techniques, which means these methods are not robust enough for noisy and variable heartbeat waveforms [8]. Moreover, the fiducial features need to be selected manually. For example, J.S. Arteaga-Falconi [9] selects only one feature point and reports a relatively poor accuracy, while Liu [10] selects four feature points and obtains a higher accuracy rate.

2) Non-fiducial methods
To reduce the complexity of fiducial points detection, some researchers propose non-fiducial methods. Non-fiducial methods usually convert the ECG signal from time domain to other domains, e.g., frequency domain, to extract features. For example, Ting [11] uses the coefficients of the Kalman filter as the features of ECG signals. Convolutional Neural Networks (CNNs) are widely used for the deeper feature extraction in non-fiducial methods recently. For example, Zhang [12] uses autocorrelation operation to remove the phase difference in the wavelet coefficients caused by the blind segmentation. The coefficients of the autocorrelation operation are then given to CNN for further feature extraction. Zhao [13] proposes a CNN-based non-fiducial method. He uses the generalized S-transformation to generate a 2D ECG trajectory first and then uses CNN to automatically extract deeper features from the trajectory.

3) Partially fiducial methods
Partially fiducial methods combine fiducial and non-fiducial methods. The search of R point is generally easier than the search of other points because it is the highest and sharpest peak in a heartbeat. Therefore, partially fiducial methods usually locate the R point to segment the heartbeat and then extract the time domain or frequency domain information as the features [14]–[17]. Similar to non-fiducial methods, partially fiducial methods can also use neural networks for further extraction. For instance, Labati [18] detects QRS complex and forms each four QRS complex to a QRS vector. Then the QRS vector is sent to CNN to extract deeper features.

B. IDENTIFICATION
The identification procedure can be classified into two categories: classification task and matching task. The conventional classifiers include Support Vector Machine (SVM) [19], autoencoders [20], [21], deep belief network [22], and softmax classifier [23]–[25], which mainly focus on the classification of healthy and non-healthy heartbeats [26]. However, it is obviously not appropriate to train a static classifier for large ECG-based authentication systems with matching task, where the test samples are usually unknown individuals. One solution for matching task is using the similarity of two test signals [18]. Another solution is
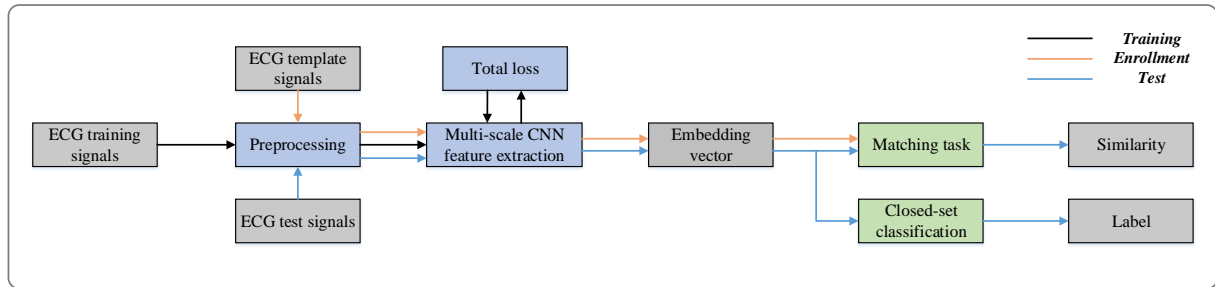
FIGURE 1. The system diagram of the proposed ECG authentication method.

treating the matching task as a one-class classification task. For instance, Hejazi [27] uses one-class SVM classifier for matching. However, it is not easy to train a one-class SVM model for a large training set.

III. PROPOSED METHOD

As shown in Fig. 1, our proposed method, can be divided into three steps: signal preprocessing, CNN feature extraction, and test. The CNN feature extraction step includes the network training and the template enrollment, while the test step includes the matching task and the closed-set classification.

Firstly, we locate the R points and use the whole heartbeats from the ECG signals to extract features. Compared with the QRS complex used in other works, the whole heartbeat contains more intact information, such as the information contained in P and T waveforms. Considering the powerful ability of CNN to extract the features automatically, more detailed features can be obtained.

Secondly, to improve the generalization ability for the preprocessed ECG heartbeats, we proposed the parallel multi-scale one-dimensional residual network. The size of the convolutional kernel used by the original residual network in each residual block is 3×3, which has been proved to be efficient in 2D image classification. In addition, the ECG waveform also contains a lot of timing information because it is a segment of signal essentially. To find the structure that can extract most detailed features from the ECG signals, we choose several convolutional kernels with different sizes. The result shows that the combination of three convolutional kernels with size 1×3, 1×5, and 1×7 can achieve the best performance.

Finally, we propose a new loss function to train the parallel multi-scale one-dimensional residual network. Conventional CNNs are usually trained by the softmax loss, which divides the whole hyperspace or hypersphere by the number of classes to ensure that the categories are separable. However, it does not require the intra-class compactness and inter-class dispersion, resulting in a larger distance between an inter-class samples pair than the distance between an intra-class samples pair, which is not allowed in the matching task. Based on the softmax loss, we propose center loss and margin

loss to improve the generalization ability of the network for the samples whose categories have not been seen in the training set. The center loss can constrain the compactness of the intra-class samples and the margin loss can keep the local neighbor relationships to constrain the dispersion of the inter-class samples. The parallel multi-scale one-dimensional residual network trained by proposed loss functions can work more accurately for the matching task.

A. SIGNAL PREPROCESSING

As shown in Fig. 2, the signal preprocessing includes four steps: resampling, filtering, R point detection, and heartbeat concatenation.
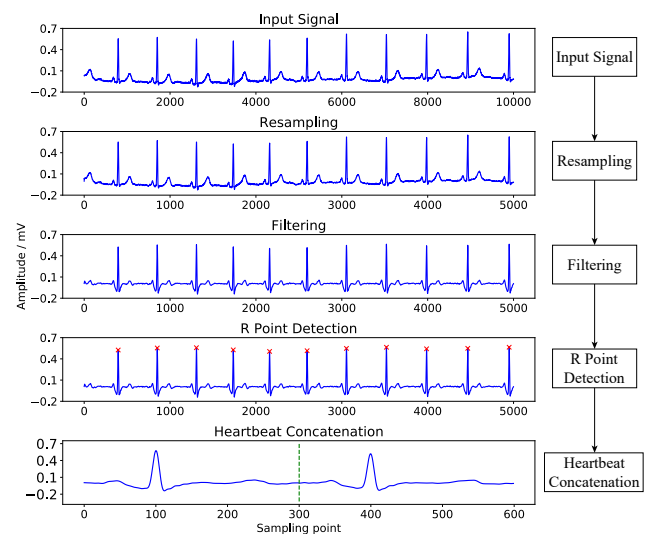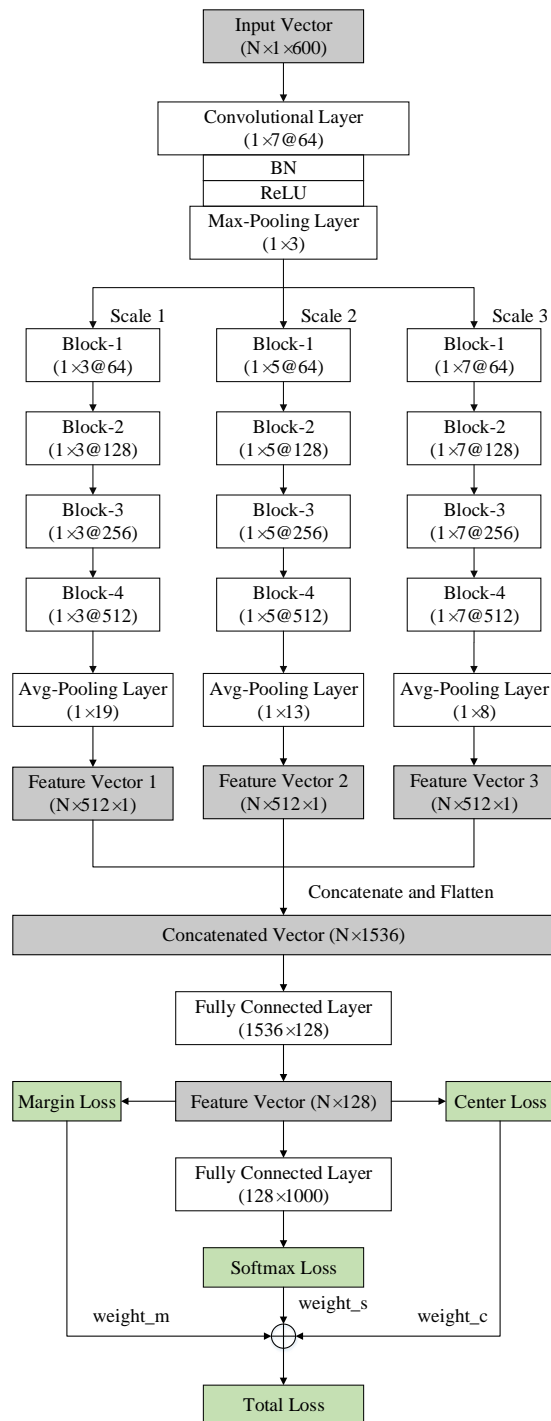

FIGURE 2. The signal preprocessing includes four steps: resampling, filtering, R point detection, and heartbeat concatenation.

The ECG signals in different databases usually have different sampling rates because these databases are acquired by different types of equipment. Therefore, to improve the generalization ability of the parallel multi-scale one-dimensional residual network, all ECG signals are resampled to 500 Hz.

After resampling, we use the filtering method proposed by Qiu [28] to eliminate the powerline interference, myoelectric interference, and the baseline wander. Finally, we locate the

FIGURE 3. The architecture of the parallel multi-scale one-dimensional residual network.

set by forming two random heartbeat signals to a heartbeat vector. The size of the heartbeat vector is 600 sampling points and 200 heartbeat vectors are used for one subject.

B. CNN FEATURE EXTRACTION

1) Overview

CNN is a powerful deep neural network inspired by visual neuroscience. The convolutional kernels in multiple hidden layers enable CNN to spontaneously extract higher-level features. ResNet [29] is one of the most effective CNN frameworks. Original ResNet uses 3×3 convolutional kernels in each residual block, which has been proven to be effective in image processing. However, the ECG waveform also contains a lot of timing information because it is a segment of signal essentially. Convolutional kernels with different sizes should be used to extract more timing features. Therefore, we propose a new structure: parallel multi-scale one-dimensional residual network.

2) Parallel multi-scale one-dimensional residual network architecture

As shown in Fig. 3, our parallel multi-scale one-dimensional residual network includes an input layer, a preprocessing convolutional block, three parallel residual blocks, an embedding fully connected layer, and an extra fully connected layer.

The input of parallel multi-scale one-dimensional residual network is the one-dimensional heartbeat vector, whose size is formatted as 1×600.

The first preprocessing convolutional block contains a convolutional layer, a batch normalization layer, a ReLU layer, and a max-pooling layer. The kernel sizes of the convolutional layer and max-pooling layer are 1×7 and 1×3, respectively. This convolutional block can extract the low-level pattern of the input vector.
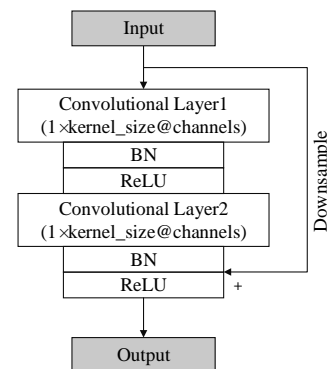
FIGURE 4. The block structure used in the residual network.

Three residual network blocks with different kernel size are used after the first convolutional block. The detail of the residual block is shown in Fig. 4. The residual block utilizes identity shortcut connection to skip one or more layers. Because of these skip connections in the block, our

R-peak points through thresholding and extract 99 sampling points before each R point and 200 sampling points after each R point to form a one-dimensional heartbeat signal with a fixed length of 300 sampling points. In order to prevent from overfitting of the neural network, we enlarge the training
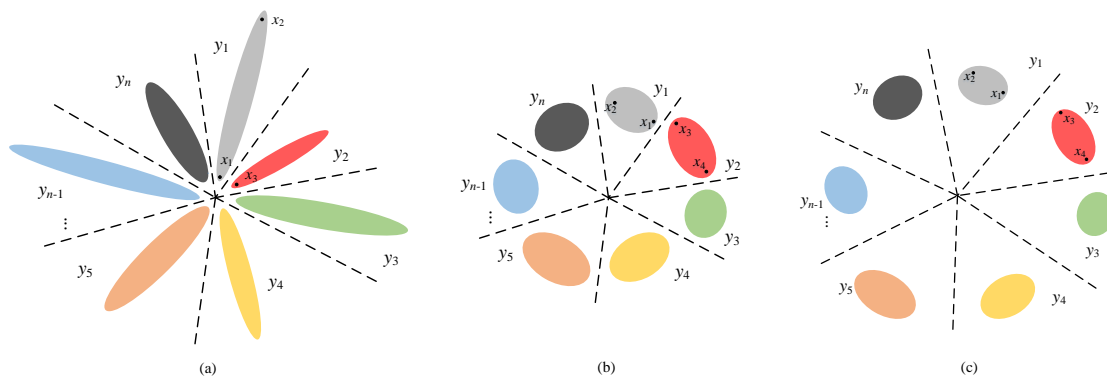
IEEE *Access*

FIGURE 5. (a) Illustration of the features extracted by softmax loss function; (b) Adding center loss function; (c) Total loss function.

proposed network can propagate larger gradients to initial layers, and learn these initial layers as fast as the final layers. The network only needs to learn the residual of output and input, rather than only the input, which means we can train the network easier and faster. The kernel sizes applied in three scales are 1×3, 1×5 and 1×7. Four blocks with same kernel size are connected in series to extract the higher-level abstract features. The input feature map is downsampled in the last three blocks and the number of channels increases after each block. Finally, an average pooling layer is utilized for each scale to downsample all feature maps to the same size 512×1.

Three feature vectors from different scales are concatenated and flatten to form one vector as the input of the embedding fully connected layer. The size of the formed vector is 1536. The embedding fully connected layer with size 1536×128 then generates an embedding vector with size 128. This embedding vector is the final feature of the initial heartbeat vector and is used in the matching task to compute the similarity. In the training, an extra fully connected layer with size 128×1000 is used to compute the softmax loss. Moreover, the embedding vector is also used to compute the center loss and margin loss. All these three losses are weighted and added, and the result is the final loss of our proposed network.

3) Loss function

Based on the softmax loss, we propose center loss and margin loss to train our parallel multi-scale one-dimensional residual network together.

As mentioned above, the simple softmax loss cannot satisfy the requirement of the matching task. However, softmax loss has a good performance on classification. When the subjects in the training set are sufficient, a CNN train by softmax loss has relatively good generalization ability. Therefore, the softmax loss is used as the basic loss function to ensure the basic classification ability of the network. The softmax loss

is computed as:

$$\mathcal{L}_S = -\sum_{i=1}^{m} \log \frac{e^{W_{y_i}^T f(x_i)+b_{y_i}}}{\sum_{j=1}^{n} e^{W_j^T f(x_i)+b_j}}. \quad (1)$$

where m is the number of samples, n is the dimension of the extra fully connected output layer, W and b are the weight and bias of the extra fully connected layer, respectively. x_i is the ith sample and $f(x_i)$ is the embedding vector of x_i. y_i is the label of the ith sample.

As shown in Fig. 5a, the softmax divides the whole hyperspace or hypersphere by the number of classes to ensure the extracted features are separable. There are obvious boundaries between each of the two types of extracted features. However, the softmax loss does not require the intra-class compactness and inter-class dispersion. For example, x_1 and x_2 are features with same class y_1, and x_3 is a feature with class y_2. The distance between x_1 and x_2 is obviously larger than the distance between x_1 and x_3, which means if x_1 and x_2 are authorized, x_1 and x_3 will be authorized, too.

To solve the disadvantage of softmax loss which can not be used for the intra-class compactness and inter-class dispersion, a loss function called center loss is used in training. Center loss constrains the distance between each sample and the average of the samples with the same class, which means that the features extracted from the intra-class sample are more compact in the embedding space. The center loss is defined as:

$$\mathcal{L}_C = \frac{1}{2}\sum_{i=1}^{m} ||f(x_i) - c_{y_i}||_2^2. \quad (2)$$

where c_{y_i} is the average embedding vector with label y_i.

As shown in Fig. 5b, after adding the center loss function, the extracted features with the same class are more compact than using softmax loss only. The distance between the features x_1 and x_2 is greatly reduced. Although the center loss can improve the intra-class compactness, it does not consider the relationships between the inter-class samples. For example, the distance between the homogeneous class

features x_3 and x_4 are still larger than the distance between the heterogeneous class features x_1 and x_3.

We propose a margin loss to describe the local neighbor relationships between the features extracted from each sample by the network. For each sample, the neighbor relationships between the farthest K intra-class samples and the nearest K inter-class samples in the original Euclidean space are maintained in the embedding space, which means that the distance between the specified sample and the farthest K intra-class samples is shorter than the distance between the specified sample and the nearest K inter-class samples. A constant distance α is used as the margin between these two distances like the support margin used in SVM. The margin loss is given by:

$$
\begin{aligned}
\mathcal{L}_M &= \sum_{i=1}^{m} \sum_{a}^{K} \sum_{b}^{K} [D_{ia}^P - D_{ib}^N + \alpha]; \\
D_{ia}^P &= ||f(x_i) - f(x_{ia}^P)||_2^2; \\
D_{ib}^N &= ||f(x_i) - f(x_{ib}^N)||_2^2.
\end{aligned}
\tag{3}
$$

where x_{ia}^P is the ath sample in the farthest K intra-class samples of x_i, and x_{ib}^N is the bth sample in the nearest K inter-class samples of x_i in the original Euclidean space, and α is the margin constant.

The total loss of the parallel multi-scale one-dimensional residual network is the weighted sum of the three losses above.

$$
\mathcal{L} = \mathcal{L}_S + \lambda_C \cdot \mathcal{L}_C + \lambda_M \cdot \mathcal{L}_M.
\tag{4}
$$

where λ_C is the weight of the center loss, and λ_M is the weight of the margin loss.

Fig. 5c shows an example of the features extracted by the total loss function.

4) Training strategy

An extra fully connected layer is added after the embedding vector to compute the softmax loss. The embedding vector is used to compute the center loss and margin loss. To train the network faster, an Adam optimizer is used and the weights of CNN filters are initialized using random sampling from a Gaussian distribution with zero mean and 0.01 standard deviation. The weight of center loss is set to 1.0 and the margin loss is set to 0.1. The learning rate of the model is 1e-4 and the batch size is 1024. The constant K used in the margin loss is set to 5 and α is set to 1.0. Finally, the parallel multi-scale one-dimensional residual network is trained for 300 epochs.

C. TEST

The test procedure is divided into two tasks: matching task and closed-set classification task. The performance metrics of the former is the Equal Error Rate (EER), and later is the accuracy of classification. Moreover, we use the confusion matrix to show the results of closed-set classification more intuitively.

1) Matching task

The matching task is divided into two steps: enrollment and matching. Normally, the test sample consists of one input signal and one corresponding label. At first, a signal with the given test label has been enrolled to generate a template vector. Then the test input signal is sent to the network to generate the corresponding embedding vector. Finally, the similarity between the embedding vector and the template vector is calculated. This similarity is compared to a fixed threshold defined by the user to determine whether the matching is successful or not. Since the thresholds used in different situations are usually different, we evaluate the performance of the matching task by EER.

2) Closed-set classification

In the closed-set classification task, an extra fully connected layer is used to fine-tune the network for different test sets. The size of the output layer in the extra fully connected layer is replaced by the number of categories in the test set. We use the softmax loss to train the extra fully connected layer to classify samples. To maximize contrast and create a balanced database, a 10-fold cross-validation method is used for 10 times to reduce the generalization error in the training set. The whole cross-validation test generates 10 different results and the average is the final accuracy of the closed-set classification task.

IV. EXPERIMENTS
A. DATASETS

To improve the generalization ability of the parallel multi-scale one-dimensional residual network, the training dataset should have a relatively large scale. The number of subjects should be maximized, and the sampling signals should be acquired during relatively long periods of time continuously for more heartbeats. For these purposes, we choose the public dataset, the PhysioNet/Cinc Challenge 2017 training dataset as our training set. This database contains 8,528 ECG recordings that were provided as a public training set for use in the 2017 PhysioNet/Computing in Cardiology Challenge. Compared with other datasets in the literature, this dataset includes more users and relatively long sessions, which is suitable for the proposed training strategy.

For the test, we consider three public databases, namely ECG-ID database, PTB Diagnostic ECG database (PhysioNet) and MIT-BIH Arrhythmia database. The sampling rates, leads, resolutions, and lengths of these three databases are all different. Moreover, these three databases contain both healthy and unhealthy signals (e.g. arrhythmia signals), which is useful for testing the generalization performance of the network.

The ECG-ID database contains 310 ECG recordings obtained from 90 healthy individuals (44 men and 46 women aged from 13 to 75 years). Each recording is ECG lead I, recorded for 20 seconds, digitized at 500 Hz with 12-bit resolution over a nominal ±10 mV range. Some records are
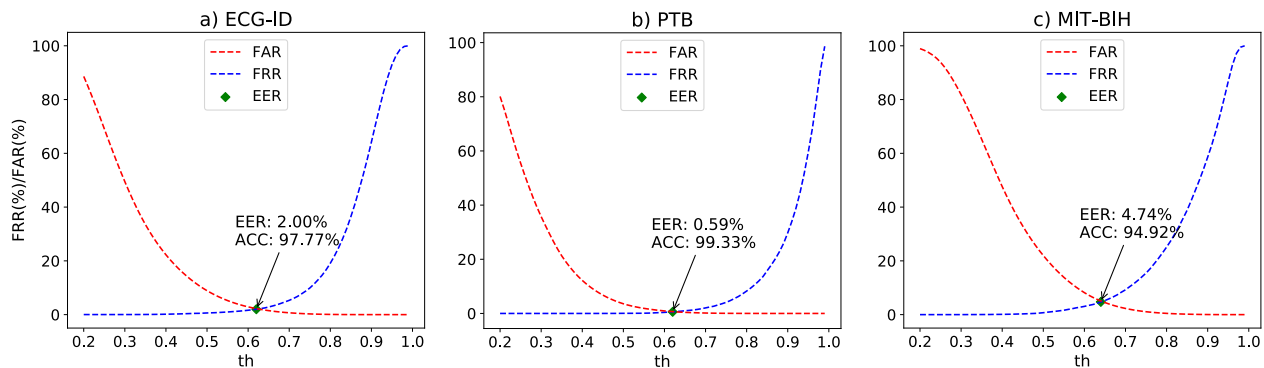
FIGURE 6. The FAR-FRR curve of ECG-ID, PTB, and MIT-BIH Arrhythmia ECG database.

collected during one day and others are collected periodically over 6 months.

The PTB Diagnostic ECG database contains 549 records from 290 subjects (aged 17 to 87, 209 men and 81 women), with different time durations of around 2 minutes. Only 52 of 290 individuals are healthy and the rest of them suffer from heart diseases. Each recording contains 15 leads (12 conventional lead and 3 Frank orthogonal leads X, Y, and Z), digitized at 1000 Hz with 16-bit resolution over a nominal ± 16.384 mV range. To perform test simulating realistic application scenarios and compare the results with other methods proposed in the literature, we only use the signals acquired from 52 healthy individuals.

The MIT-BIH Arrhythmia database is the most common ECG database. It contains 48 half-hour excerpts of two-channel ambulatory ECG recordings obtained from 47 subjects. The recordings are digitized at 360 Hz with 11-bit resolution over a nominal ± 10 mV range. All these individuals were suffered from heart disease like arrhythmia.

B. TRAINING PROCEDURE

We use the PhysioNet/Cinc Challenge 2017 training dataset to train our parallel multi-scale one-dimensional residual network. In order to save training time, we randomly select 1000 ECG records from total 8528 records and locate all R points to segment the heartbeats in each record. The number of selected training records is large enough for our test databases. For each record, we randomly selected two heartbeats to form a heartbeat vector, getting a total of 200 heartbeat vectors. Any two heartbeat vectors have to be different. After selection, we get a training set with the size of $200,000 \times 600$. In each iteration, we randomly select 1024 training samples to compute the loss and train the Adam optimizer. Finally, we train the network for 300 epochs.

C. TEST PROCEDURE

We use all three datasets, ECG-ID database, PTB Diagnostic ECG database, and MIT-BIH Arrhythmia database to evaluate both matching task and closed-set classification task. For each dataset, we first locate the heartbeats and form the heartbeat vectors. In order to do a thorough test, we generate 1000 heartbeat vectors for each individual. After the location and combination, we get the test set with the size of 90,000 $\times$ 600 for the ECG-ID database, 52,000 $\times$ 600 for the PTB database and 47000 $\times$ 600 for the MIT-BIH Arrhythmia database.

To test the performance of the matching task, we enroll 50% of samples to generate the templates and test the rest 50% samples. For each test sample, we calculate the similarity between the test sample and the template of each individual. Therefore, we have 45,000, 26,000, and 23,500 genuine comparisons, and 4,050,000, 1,326,000, and 1,081,000 impostors for the ECG-ID database, the PTB Diagnostic database, and the MIT-BIH Arrhythmia, respectively. Finally, we use the False Acceptance Rate (FAR) - False Rejection Rate (FRR) curve and the EER to evaluate the performance.

To test the performance of the closed-set classification, we fine-tune the network and use the 10-fold cross-validation method for 10 times. The average of the classification accuracy is the final evaluation criterion.

V. RESULTS
A. MATCHING TASK

The most common performance metrics in the matching task are the FAR and the FRR. The FAR is the probability that the system incorrectly authorizes a non-authorized person, due to the mismatch between the biometric input and the template. The FRR is the probability that the system incorrectly rejects access to an authorized person because the biometric input is failed to match the template. The EER is defined as the crossover point of the FAR and FRR curves on a graph. The test performances of the FAR and FRR of the above three datasets are shown in Fig. 6, where the value of the point that the FAR and FRR curves intersected is the EER value.

As shown in Fig. 6, when applied in these three test datasets, the EER and accuracy for the ECG-ID database are 2.00% and 97.7%, respectively. These two values are 0.59% and 99.33% in the PTB Diagnostic ECG database, and 4.74% and 94.92% in the MIT-BIH Arrhythmia database. It is worth mentioning that the thresholds corresponding to the
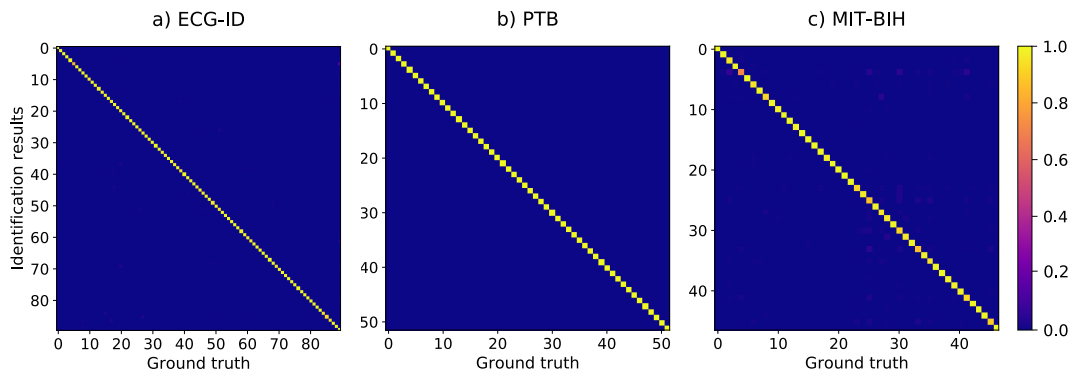
FIGURE 7. The confusion matrix of ECG-ID, PTB, and MIT-BIH Arrhythmia ECG database.

EER point in three datasets are all 0.63, which proves that our parallel multi-scale one-dimensional residual network has superior generalization performance.

B. CLOSED-SET CLASSIFICATION

To compare with the conventional methods, we use the softmax classifier to predict the label for each test sample and compute the classification accuracy. Moreover, we test the accuracy of NSRDB and STDB used in Zhang [12]. The results of the closed-set classification are shown in Table 1.

TABLE 1. Closed-set classification results of the proposed method and other reference papers

Datasets	Identification Rate (%)	Reference Paper (%)
ECG-ID	**98.24**	96.63 [13]
PTB Diagnostic	**100**	100 [18]
MIT-BIH Arrhythmia	**95.99**	91.1 [12]
NSRDB	**97.17**	95.1 [12]
STDB	**95.04**	90.3 [12]

In addition, the testing performance in terms of the confusion matrix of all the above three databases is shown in Fig. 7. The performance visualization presented in Fig. 7 clearly shows that the fine-tuned network works efficiently for the close-set classification. The trained network can identify the correct individuals (diagonal entries with a bright yellow color) with very little false positives or false negatives (non-zero off-diagonal entries).

To verify the superiority of our parallel multi-scale structure, we also test the performance for the networks with different structures. We build several one-dimensional residual networks with variant combinations of kernel sizes and test the close-set classification for these networks.

As shown in Table 2, our proposed parallel multi-scale structure with kernel size 1×3, 1×5 and 1×7 has the best performance. We infer that the results may be caused by the coupling relationship between features of different length. Different kernels may extract different features. Therefore, the multi-scale structure can extract features more flexible than the structure with only one kernel size.

TABLE 2. Closed-set classification results of proposed multi-scale structure and other different structures

Structure	ECG-ID (%)	PTB (%)	MIT-BIH Arrhythmia (%)
Proposed	**98.24**	**100**	**95.99**
1×1 only	93.24	98.03	89.12
1×3 only	96.07	99.96	93.15
1×5 only	96.02	99.92	93.41
1×7 only	95.36	99.89	92.90
1×9 only	95.11	99.82	91.88
$1 \times 1 + 1 \times 3$	97.32	100	94.24
$1 \times 3 + 1 \times 5$	97.91	100	95.08
$1 \times 5 + 1 \times 7$	97.89	99.96	95.12

As shown in Table 3, compared with other works, our proposed method achieves the best classification accuracy on the PTB database and a stable EER performance on different databases. In another word, our proposed method has better data-independence and generalization ability.

VI. CONCLUSION

This paper has introduced an ECG biometric authentication system based on the parallel multi-scale one-dimensional residual network. This method is trained for two common biometric tasks: matching task and closed-set classification. The system locates the R points from ECG signals to form heartbeat vectors and uses the parallel multi-scale one-dimensional residual network to extract the features from the heartbeat vector. The features extracted by our network have the following advantages: 1) the combination of multi-scale kernel size can extract more detailed features from the signal segment with different length; 2) the center loss and margin loss can make up the shortcomings of softmax loss, and improve the intra-class compactness and inter-class dispersion of features after extraction; 3) all features are embedding to a vector of fixed size and a fixed threshold can be used to authorize the features from different users in different environment, thus improving the generalization ability of the authentication system.

To evaluate the performance of our method, we have tested it using different datasets of samples including healthy

TABLE 3. Comparative tabulation of experimental results for different ECG biometrics methods

Methods	Database	Classifier	Performance (%)
Proposed	ECG-ID	CNN	Acc = 98.24, EER = 2.00
	PTB		Acc = 100, EER = 0.59
	MIT-BIH		Acc = 95.99, EER = 4.74
Labati [18]	IDEAL	CNN	EER = 1.36-5.95
	PTB		Acc = 100
Zhao [13]	Physionet ECG	CNN	Acc = 96.18-96.63
	ECG-ID		
Zhang [12]	8 normal/abnormal databases	CNN	Acc = 93.50
Islam [7]	multisession in-house database	Unimodal fusions of information	EER = 9.58
Mesin [6]	self-built database	SVM	Acc = 95.60
Acharya [5]	PTB	KNN	Acc = 98.90
Sharma [19]	PTB	SVM	Acc = 96.00
Liu [10]	PTB	Decision tree	Acc = 94.40

and unhealthy individuals for matching task and closed-set classification. In all cases, our method has obtained the best accuracy with the smallest EER. Since the training and test sets we used are completely different, the results prove that our system is impressively robust and generalized.

Furthermore, we have compared the performance of our method with other studies in the literature. Our accuracy of closed-set classification is better than the state-of-the-art works.

REFERENCES

[1] I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, and J. W. Rohrbaugh, "ECG biometric recognition: A comparative analysis," IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1812–1824, 2012.

[2] S. I. Safie, J. J. Soraghan, and L. Petropoulakis, "ECG biometric authentication using pulse active width (PAW)," in Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2011 IEEE Workshop on. IEEE, 2011, pp. 1–6.

[3] N. Laplante, P. Laplante, and J. Voas, "Caring: An undiscovered "super ility" of smart healthcare," IEEE software, vol. 33, no. 6, p. 16, 2016.

[4] X. Tang and L. Shu, "Classification of electrocardiogram signals with RS and quantum neural networks," International Journal of Multimedia and Ubiquitous Engineering, vol. 9, no. 2, pp. 363–372, 2014.

[5] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. L. Oh, M. Adam, J. E. Koh, J. H. Tan, D. N. Ghista, R. J. Martis, C. K. Chua et al., "Automated detection and localization of myocardial infarction using electrocardiogram: a comparative study of different leads," Knowledge-Based Systems, vol. 99, pp. 146–156, 2016.

[6] L. Mesin, A. Munera, and E. Pasero, "A low cost ECG biometry system based on an ensemble of support vector machine classifiers," in Advances in Neural Networks. Springer, 2016, pp. 425–433.

[7] S. Islam, N. Ammour, N. Alajlan, and M. Abdullah-Al-Wadud, "Selection of heart-biometric templates for fusion," IEEE Access, vol. 5, pp. 1753–1761, 2017.

[8] J. R. Pinto, J. S. Cardoso, and A. Lourenço, "Evolution, current challenges, and future possibilities in ECG biometrics," IEEE Access, vol. 6, pp. 34 746–34 776, 2018.

[9] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, "ECG authentication for mobile devices," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 3, pp. 591–600, 2016.

[10] B. Liu, J. Liu, G. Wang, K. Huang, F. Li, Y. Zheng, Y. Luo, and F. Zhou, "A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection," Computers in biology and medicine, vol. 61, pp. 178–184, 2015.

[11] C.-M. Ting and S.-H. Salleh, "ECG based personal identification using extended kalman filter," in Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on. IEEE, 2010, pp. 774–777.

[12] Q. Zhang, D. Zhou, and X. Zeng, "HeartID: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications," IEEE Access, vol. 5, pp. 11 805–11 816, 2017.

[13] Z. Zhao, Y. Zhang, Y. Deng, and X. Zhang, "ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation," Computers in biology and medicine, vol. 102, pp. 168–179, 2018.

[14] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," EURASIP journal on Advances in Signal Processing, vol. 2008, no. 1, p. 148658, 2007.

[15] A. Bonissi, R. D. Labati, L. Perico, R. Sassi, F. Scotti, and L. Sparagino, "A preliminary study on continuous authentication methods for photoplethysmographic biometrics," in Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2013 IEEE Workshop on. IEEE, 2013, pp. 28–33.

[16] R. D. Labati, V. Piuri, R. Sassi, and F. Scotti, "HeartCode: A novel binary ECG-based template," in Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings, 2014 IEEE Workshop on. IEEE, 2014, pp. 86–91.

[17] M. Tantawi, K. Revett, A.-B. Salem, and M. F. Tolba, "ECG based biometric recognition using wavelets and RBF neural network," in Proc. 7th Eur. Comput. Conf.(ECC), 2013, pp. 100–105.

[18] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional neural networks for ECG biometric recognition," Pattern Recognition Letters, 2018.

[19] L. Sharma, R. Tripathy, and S. Dandapat, "Multiscale energy and eigenspace approach to detection and localization of myocardial infarction," IEEE transactions on biomedical engineering, vol. 62, no. 7, pp. 1827–1837, 2015.

[20] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," Information Sciences, vol. 345, pp. 340–354, 2016.

[21] A. Eduardo, H. Aidos, and A. L. N. Fred, "ECG-based biometrics using a deep autoencoder for feature learning - an empirical study on transferability," in ICPRAM, 2017.

[22] M. Huanhuan and Z. Yue, "Classification of electrocardiogram signals with deep belief networks," 2014 IEEE 17th International Conference on Computational Science and Engineering, pp. 7–12, 2014.

[23] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," IEEE Transactions on Biomedical Engineering, vol. 63, no. 3, pp. 664–675, 2016.

[24] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, "Convolutional neural networks for patient-specific ECG classification," in Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. IEEE, 2015, pp. 2608–2611.

[25] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Feature leaning with deep convolutional neural networks for screening patients with paroxysmal atrial fibrillation," in Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE, 2016, pp. 5057–5064.

[26] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," Briefings in bioinformatics, vol. 18, no. 5, pp. 851–869, 2017.

[27] M. Hejazi, S. A. R. Al-Haddad, S. J. Hashim, A. F. A. Aziz, and Y. P. Singh, "Non-fiducial based ECG biometric authentication using one-

IEEE *Access*

class support vector machine," in 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2017, pp. 190–194.

[28] Y. Qiu, F. Xiao, and H. Shen, "Elimination of power line interference from ECG signals using recurrent neural networks," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 2296–2299.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

YIFAN CHU is a MA.Eng. Candidate in Zhejiang University, China. He received his Bachelor Degree in Electrical & Information Engineering from Zhejiang University, in 2015. His research interests include computer vision and signal processing.

HAIBIN SHEN is a professor of Zhejiang University, deputy director of HPEC (High Performance Embedded Computing) Key Lab of Ministry of Education, member of the second level of 151 talents project of Zhejiang Province. His research interests are focused on learning algorithm, processor architecture, and modeling. His research achievement has been used by many authority organizations, and he was awarded the first prize of Electronic Information Science and Technology Award from the Chinese Institute of Electronics.
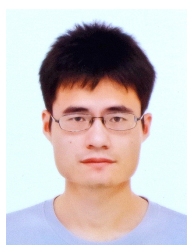
KEJIE HUANG(M'12-SM'18) received his Ph.D. degree from the Department of Electrical Engineering, National University of Singapore (NUS), Singapore, in 2014. Now he is a principal investigator at College of Information Science & Electronic Engineering, Zhejiang University. He has authored/co-authored more 30 papers in international peer-reviewed journals and conference proceedings. He holds four granted international patents, and another three pending ones. His research interests include low power circuits and systems design using emerging non-volatile memories, architecture and circuit optimization for reconfigurable computing systems and deep learning accelerator, computer vision and machine learning.

• • •