IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 40, No. 2, April, 2023, pp. 683-691
Journal homepage: http://iieta.org/journals/ts

Multi-Grained Deep Cascade Learning for ECG Biometric Recognition

Sujuan Wang, Ruili Zhang*

School of Computer, Heze University, Heze 274015, China

Corresponding Author Email: zhangruili@hezeu.edu.cn

https://doi.org/10.18280/ts.400227

Received: 19 January 2023
Accepted: 26 March 2023

Keywords:
electrocardiogram, biometric recognition, deep neural networks, deep cascade learning, multiple granularity scanning

ABSTRACT

Recently, electrocardiogram (ECG) biometric recognition has become an emerging area of interest in the identity identification technology. However, robust and accurate of ECG biometric recognition are challenging. Sparse representation-based classification (SRC) and deep neural networks (DNNs) have achieved significant success in biometric recognition, but there are some problems, and SRC-based learning methods are one-step models where the latent discriminative information cannot be fully exploited. DNNs are complex models and require large amounts of training data. The proposed method describes the design of a multi-grained deep cascade learning for ECG biometric recognition that addresses above problems. First, the global and local features are generated by principal components analysis (PCA) and the use of one-dimensional multi-resolution local binary pattern (1DMRLBP) method. Second, we obtain new features of class coding based on the sparse representation by multi-granularity scanning. Third, we propose an end-to-end deep cascade learning model without back-propagation to seek more discriminative information. This approach not only effectively solves the one-step model problem of sparse representation, but also effectively reduces the ECG signal noise. Extensive experiments are conducted on four ECG datasets. The results demonstrate that the proposed method can outperform other state-of-the-art methods in terms of both accuracy and efficiency.

1. INTRODUCTION

Biometric recognition is one of the most popular and challenging technologies for identification [1]. At present, many kinds of biological characteristics can be used for personal identification and they can be mainly divided into two categories: physiological characteristics and behavioural characteristics. Common physiological characteristics include human faces [2], fingerprints [3], hands [4, 5], irises [6, 7], and ears [8, 9]; behavioural characteristics include voices [10], gaits [11], signatures [12, 13], and keystrokes [14, 15]. However, these human biological characteristics are easy to falsify, which makes this type of approach to personal identification vulnerable. For example, presenting a photograph of a subject may deceive a system involved in face recognition, and a recording of a voice of a person may similarly deceive a speaker recognition technology. In a few cases, access to biometric systems has also been achieved using synthetic fingerprints made out of wax.

Forsen et al. [16] first proposed the use of electrocardiograms (ECG) for human feature recognition. Biel et al. [17] reported their highly influential work on ECG biometrics in 2001. ECG biometrics has since attracted increasing research interest and is regarded as one of the most promising biometric techniques. ECG biometrics has several unique advantages compared with other biometric traits and they can be summarized as follows: (1) the ECG signal is common to all organisms and it is not forgotten or lost; (2) the ECG signal is hard to replicate or spoof, which leads to the high security of ECG biometrics; (3) the ECG signal is a one-dimensional signal, which makes it easy to process and has less calculation data; (4) the ECG is used in the diagnosis of heart disease worldwide. As a result, the ECG signal can be gathered using many inexpensive sensing devices.

In recent years, sparse representation has been widely used in the ECG biometric recognition. The main idea of the sparse representation is to seek the optimal approximation of the signal in a certain space. In the transform domain, the original signal is represented by as few atoms as possible, which allows the basic information of the original signal to be grasped as a whole. Sparse representation learning can efficiently handle noise and many related methods have been proposed for ECG biometric recognition. Wang et al. [18] explored a novel framework to extract compact and discriminative features from ECG signals for human identification based on sparse representation of local segments. Tan et al. [19] presented a kernel sparse representation classifier to enhance system performance in a high-dimensional feature space for ECG biometric recognition. Li et al. [20] described a robust ECG biometrics method based on the graph regularized non-negative matrix factorization and sparse representation. Goshvarpour and Goshvarpour [21] developed an identification system using a non-fiducial one-lead ECG feature set based on the sparse representation. Huang et al. [22] proposed a unified sparse representation framework that collaboratively exploits joint and specific patterns for ECG biometric recognition. However, the sparse representation-based coding method is a one-step model, and hence the latent discriminative information of the coding error vectors cannot be fully exploited.

Many researchers are inclined to use neural networks to handle ECG signals because of the considerable progress that

has been made in the development and application of deep neural network (DNN) technology. The recognition rate is significantly higher using this type of approach compared with other recognition methods [23-25]. The advantage of neural networks lies in their unique ability to learn new knowledge and perform parallel data processing at high speed. Neural networks can still be corrected automatically through a large amount of training to generate accurate results even if the eigenvalues are somewhat inaccurate. Although neural networks have significant advantages, they also have some disadvantages. First, training requires a huge amount of data, which makes it impossible for deep neural networks to be directly applied to tasks that have small-scale data. Second, deep neural networks are complex models and powerful computational facilities are usually required for the training process. More importantly, deep neural networks have too many hyper-parameters, and the learning performance depends on careful tuning of these factors. The multi-grained cascade forest (gcForest) proposed by Zhou and Feng [26] in 2017 provided new ideas for pattern recognition and it overcomes some of the shortcomings of a DNN. The gcForest has fewer hyper-parameters than the DNN and can be easily trained without too many parameter adjustment skills. In many areas, the gcForest can achieve recognition results comparable to or better than the DNN.

We designed a multi-grained deep cascade learning model with global and local features for ECG biometric recognition, which can solve the one-step model problem of sparse representation, and also reduce the ECG signal noise. The main contributions of this paper are summarized as follows: (1) Global and local features are generated by principal components analysis (PCA) and one-dimensional multi-resolution local binary patterns (1DMRLBPs). (2) After the multi-granularity scanning, we obtain some new features through the getting new class vector (GNCV) function based on sparse representation. (3) We propose an end-to-end deep cascade learning model that can effectively reduce ECG signal noise and help find more discriminative information.

The paper is organized as follows: Section 2 reviews the literature on the relevant technologies used in ECG biometric recognition. Section 3 describes the proposed method. Section 4 reports the experimental results and provides a comprehensive analysis. Finally, we report the conclusion of our study and describe areas for future work.

2. RELATED WORK

2.1 ECG biometric recognition

Research on ECG biometrics can be divided into two categories: fiducial-based and non-fiducial–based approaches [27, 28]. Healthy ECG heartbeat has six fiducial points, namely P, Q, R, S, T, and U. These fiducial points determine the start and end positions of each ECG signal waveform, the peak value, the amplitude of the waveform, and the slope of the waveform. In the fiducial-points–based approaches, the distance, angle, area, and amplitude between these fiducial points are usually extracted as features. Biel et al. [17] proposed a new approach in human identification that uses a standard 12-lead to record the ECG signals of 20 people aged 20–55 and obtains a 100% recognition rate. Israel et al. [29] noted the lack of automatic feature extraction in the method presented by Biel and introduced an ECG biometric recognition method that does not require any waveform detections based on classification of coefficients from the discrete cosine transform (DCT) of the autocorrelation (AC) sequence of ECG data segments. Kyoso and Uchiyama [30] proposed that the feature parameters were sampled from the intervals and durations of the electrocardiographic waves. The interval is extracted using the characteristic points appearing on the waveform of the second-order derivative and discriminant analysis is used for identification. The challenge in the fiducial-based approaches is that the recognition rate depends on the positioning accuracy of the fiducial points. At present, there is no unified standard for positioning the ECG signal boundary, which somewhat affects the accuracy of feature extraction.

The detection of fiducial points increases the complexity of ECG identification to a certain extent. However, the non-fiducial point feature extraction approach does not need to locate the peak points and boundary points of the waveform. A non-fiducial approach is a holistic approach that extracts the characteristic parameters from the ECG segments of multiple single-cycle heartbeat signals as a whole. Plataniotis et al. [31] proposed to use autocorrelation characteristics of the ECG signal to perform identity recognition in 2006. The ECG signal was divided into non-overlapping windows and the standard autocorrelation was calculated. Jung and Lee [32] presented an ECG identification method based on non-fiducial feature extraction using window removal, nearest neighbor (NN), support vector machine (SVM), and linear discriminant analysis (LDA).

Currently, DNNs have been used in ECG biometrics to improve performance. Since DNN has the ability to effectively learn features from data, it has become increasingly popular for many tasks that require ECG signals for analysis. Page et al. [33] devised a streamlined approach that utilized neural networks to both identify QRS complex segments of the ECG signal and to subsequently perform user authentication on these segments. da Silva Luz et al. [34] described an ECG-based biometric system using a deep autoencoder to learn a lower-dimensional representation of heartbeat templates.

2.2 PCA and 1DMRLBP features

PCA is a classic feature extraction technique that can eliminate the correlation of the original signal to help find the main influencing factors of the internal connection of the signal. It reveals the essence of things and simplifies the problem. PCA is mainly suitable for data with a relatively large linear correlation of parameters and the ECG signal is a time-continuous signal. PCA can therefore use all ECG signals as feature vectors to obtain the global feature of the ECG. Chen et al. introduced a PCA method in exacting the heartbeat features [35]. Zhang and Zhang [36] determined the principal characteristics of the signal by means of the PCA technique and obtained satisfactory classification results.

In the field of computer vision and image processing, the local binary pattern (LBP) features were first proposed in an earlier study [37] and classically applied on two-dimensional (2D) images. Houam et al. [38] proposed a new approach that adapts the two-dimensional classical LBP to one-dimensional signals to classify textures from osteoporotic and control cases. One-dimensional multi-resolution local binary patterns (1DMRLBP) are inspired by image-based LBP [39], but this method has been modified and enhanced to be suitable for one-dimensional signals that can tolerate noise, preserve ECG

heartbeat morphology, and resolve the temporal variations of ECG signals through an extraction mechanism. In addition, the 1DMRLBP feature type is an online feature extraction, which can be applied in real-time and only depends on past and current observations. Wang et al. [40] proposed to extract the multi-scale differential feature (MSDF) from the one-dimensional ECG signal and subsequently fuse MSDF with 1DMRLBP to generate the MSDF-1DMRLBP method. Su et al. [41] used 1DMRLBP as an ECG feature descriptor to capture the local features of the ECG signal.

2.3 Cascade forest

The multi-grained cascade forest is a decision tree ensemble approach. Here, the cascade structure of deep nets is retained but the opaque edges and node neurons are replaced by groups of random forests paired with completely random tree forests. In this case, there are usually two classes to predict in each cascade layer and four in total. There are two important stages in the gcForest: the multi-grained scanning stage and the deep cascade learning stage. The gcForest performs well in various domains of feature extraction compared with DNN even with fewer training samples. Adjusting the parameters is a tedious process. The gcForest can achieve excellent performance with almost the same set of hyperparameters and is easy to train. In terms of processing speed, the parallel implementation of gcForest is much faster than DNNs. The gcForest adaptively decides its model complexity by terminating training when there is no significant performance improvement unlike most DNNs where model complexity is fixed.

At present, many researchers have adopted the gcForest model to obtain ideal recognition results. Ding et al. presented a multi-grained scanning-based weighted cascade forest that has been applied to fault diagnosis in chemical processes [42]. Lev et al. proposed a modified version of the confidence screening mechanism based on an adaptive weighting of every training instance at each cascade level of the deep cascade forest [43]. Liu and Yang [44] explored a novel and effective visual object tracking based on gcForest.

3. PROPOSED METHODS

3.1 Base feature extraction

3.1.1 PCA feature

Suppose that $X = [x_1, x_2, \ldots \ldots x_n] \in R^m$ denotes n heartbeats. The goal of PCA is to reduce the dimension of the data matrix by finding p new variables smaller than m. The principal component projects the high-dimensional data into the subspace spanned by the eigenvector with the largest eigenvalue p, while remaining uncorrelated and orthogonal. Each principal component is a linear combination of the original variables. We now summarize the PCA method as the following steps:

Step 1: Calculate the mean m of the original feature vectors:

$$m = \frac{1}{n}\sum_{i=1}^{N} x_i \tag{1}$$

where, N represents the number of samples and x_i represents the original feature vectors.

Step 2: Compute the covariance matrix S:

$$S = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)(x_i - m)^T \tag{2}$$

Step 3: Process the following decomposition:

$$\lambda_i e_i = S e_i \tag{3}$$

Step 4: Compose transformation matrix W via choosing the first p eigenvectors that are sorted according to their corresponding eigenvalue λ_i in descending order:

$$W = \{\mathrm{e}_i\}_{i=1}^{p} \tag{4}$$

Step 5: Finally, the original feature vectors x_i can be transformed to the reduced feature vectors y_i by

$$y_i = W^T x_i \tag{5}$$

3.1.2 1DMRLBP feature

1DMRLBP is a hybrid approach that sums local ECG vectors as feature vectors to extract local features. 1DMRLBP does not simply capture texture based on a fixed number of points. It considers different distances d, and points p, where d represents how far from the desired time sample $x(t)$, the features start being extracted, and p is how many time samples are considered for 1DMRLBP feature extraction from each side. Previous work [45] extracts binary patterns (BPs) of time sample $x(t)$ as follows:

$$\begin{aligned} BP(x(t)) = &\sum_{i=0}^{p-1} sign(x(t+i-p-d+1) \\ &-x(t))2^i + sign(x(t+i+d)-x(t))2^{i+p} \end{aligned} \tag{6}$$

where, t represents the time index of the heartbeat, p represents the number of points (time samples) to be considered on each side of $x(t)$, and $x(t)$ represents the desired time sample for extraction using 1DMRLBP. $sign(.)$ is defined as in Eq. (7). $BP(x(t))$ is assigned a value of zero when its parameters require information that is out of bounds. $BP(x(t))$ becomes out of bounds when $t+i+d>k$, where k represents the heartbeat length, and when $t+i<p+d$.

The level of quantization for the acquisition instrument is known. However, the highest peak in the ECG signal remains unknown. Consequently, a modified equation is designed to capture the morphology of the ECG signal by adding a margin in the binary pattern extraction to accommodate this issue. A parameter ε is included in Eq. (7) to account for quantization error. In addition, the inclusion of this margin reduces the influence of ripples (noise) on the ECG signals. Thus, $sign(.)$ is defined as follows:

$$sign(x) = \begin{cases} 1 \;\; if \; x + \varepsilon \geq 0 \\ 0 \;\; otherwise. \end{cases} \tag{7}$$

3.2 Getting new class vector (GNCV)

In the multi-grained deep cascade model, we formulate a getting new class vector (GNCV) function, which in the ECG signal coding part, that transforms the whole ECG signal into softmax vectors activated by the softmax function on the representation error. In this study, two different representation (coding) models, namely, a sparse representation and nuclear

norm matrix regression are used in the GNCV function to obtain softmax vectors.

It is worth pointing out that more suitable coding methods that are beneficial to recognition can be freely selected and integrated into GNCV in the proposed multi-grained deep cascade model.

GNCV based on sparse representation is described as follows:

Suppose that we have C classes of subjects, d represents a query sample and $D=[D_1,D_2,\ldots,D_n]$ represents the dictionary (a basis group). The representation model can be transformed into the following minimization problem in terms of the sparse representation-based classifier (SRC) and dictionary learning:

$$\min_{\alpha} = \|d - D_{\alpha}\|_2^2 + \lambda\|\alpha\|_1 \tag{8}$$

where, the parameter λ is a scalar constant and $\|\bullet\|_2$ and $\|\bullet\|_1$ represent the Frobenius norm of two matrices: l_2-norm and l_1-norm, respectively. After solving the coding coefficients α, the representation error of each class can be computed as follows:

$$r_c = \|d - D_C\alpha_C\|_2^2 \tag{9}$$

where, D_C denotes the sample set with respect to class C, and α_C is the coefficient vector associated with class C. Next, the proposed softmax vector $S_V \in R^C$ can be computed as follows using the softmax function:

$$S_V = \left[\frac{e^{-r_1}}{\sum_{c=1}^{C} e^{-r_c}}, \frac{e^{-r_2}}{\sum_{c=1}^{C} e^{-r_c}}, \ldots, \frac{e^{-rc}}{\sum_{c=1}^{C} e^{-r_c}}\right]^T \tag{10}$$

where, C represents the number of classes. S^i_D should be bigger than other atoms in the softmax vector S_V, which shows class discrimination, if the testing sample d belongs to class $i(\leq C)$. For clarity, this process of obtaining the softmax vector S_V is defined as GNCV conditioned on dictionary D (i.e., $GNCV_D$). For convenience, we define the whole procedure of computing the softmax vector S_V for a given query sample y as

$$S_V = GNCV_D(\mathrm{y}, \alpha, \mathrm{C}) \tag{11}$$

3.3 Multi-grained deep cascade model

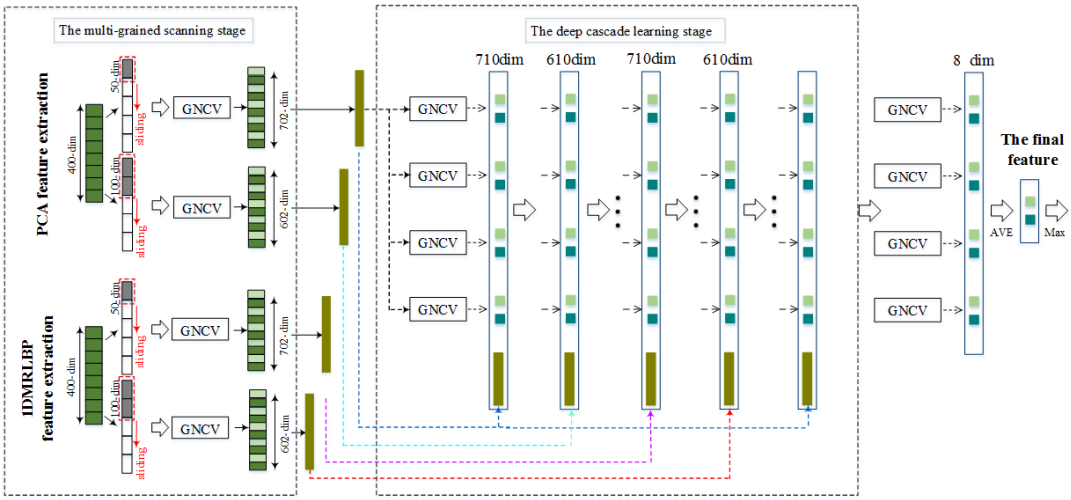

Figure 1. The overall framework of the multi-grained deep cascade learning model. The model supposes that there are two classes for prediction and the size of the sliding window is 50 dimensional and 100 dimensional. Multiple cascade layers are used for the final prediction

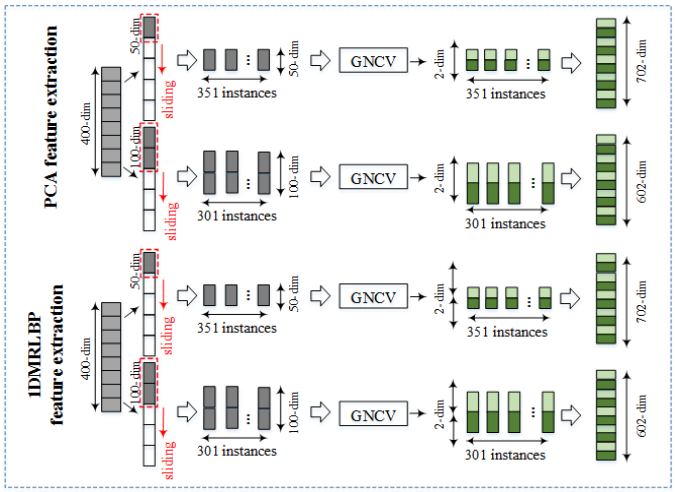

Figure 2. The feature re-representation example using sliding window scanning. Each group of signals has a window size of 50 dimensional and 100 dimensional

The main structure of the multi-grained deep cascade model includes two main components shown in Figure 1. The first component is the multi-grained scanning stage, which transforms the original feature vector to a high dimensional feature vector. The second component is the deep cascade learning stage, which performs the final feature classification using the high-dimensional feature vectors and the class vectors obtained from previous levels.

3.3.1 Multi-grained scanning stage

In the multi-grained scanning stage, we use two types of feature vectors as an original feature for input: global ECG features extracted by PCA and local ECG features extracted by 1DMRLBP. We have used sliding windows of multiple sizes to generate differently grained feature vectors. In addition, we introduce the GNCV function based on sparse representation into the deep cascade model. This step is beneficial not only because sparse representation can effectively remove the noise of ECG signals, but also because it can obtain a better classification result.

An example of the general procedure for the multi-grained scanning stage is shown in Figure 2. We suppose that the raw features are m-dimensional, the size of the sliding window is n-dimensional, the training set includes k categories for prediction, and $(m\text{-}n\text{+}1)*k$-dimensional feature vectors are generated by scanning each raw instance sequentially. The raw features are 400-dimensional, the size of the sliding windows was 50-dimensional, and there are two classes to predict as shown in Figure 2. Then, 351 two-dimensional class vectors are produced by each classifier, which generates a 702-dimensional transformed feature vector that finally corresponds to the original 400-dimensional feature vector. Similarly, using a sliding window with a size of 100-dimensional will produce a 602-dimensional transformed feature vector of the original 400-dimensional feature vectors. The transformed feature vector has many more dimensions and an enhanced feature representation compared with the original vector. After multi-grained scanning, the original features are converted into high-dimensional feature vectors.

3.3.2 Deep cascade learning stage

We propose an end-to-end deep cascade learning model without back-propagation to seek more discriminative information. The most significant feature of the deep cascade learning stage is the layer-by-layer processing. In this stage, we use the sparse representation-based GNCV function to transform the ECG features into softmax vectors. The transformed softmax vectors of 702-dimension are fed into the classifier and converted into class vectors, which are connected with the raw features as the first level input of the cascade layer, as shown in Figure 1. This procedure will be repeated until convergence of the validation performance, which means that the deep cascade model can adaptively determine the complexity of the model. The cascade layers stop growing if the addition of a new layer does not improve performance. In this way, we obtain the final prediction feature vector. Users can try more grains if computational resources are sufficient for difficult tasks.

3.4 Recognition

In the deep cascade learning stage, we obtain the final prediction feature vector $V \in R^C$, where C is the number of the class. In recognition of a query heartbeat y, we can obtain its class label by the final prediction feature vector V as follows:

$$Label(y) = \arg\max_i V \tag{12}$$

We can also determine whether the probe is an impostor or a genuine factor by comparing the distance $dis(V,T)$ to the threshold, where $dis(V,T)$ is the distance of the final prediction feature vector and the corresponding template T.

4. EXPERIMENTS

4.1 Datasets

We wanted to validate comprehensively the effectiveness of the proposed multi-grained deep cascade model with global and local features for ECG biometric recognition. We therefore conducted extensive experiments on the following four widely used benchmark datasets: MIT-BIH Arrhythmia (MITDB), CYBHiDB, PTB Diagnostic ECG Datasets, and the University of Toronto ECG Datasets (UofTDB).

The MIT-BIH Arrhythmia dataset has become the experimental data used as a test standard in most of the current electrocardiogram analysis. The dataset comprises ECG data recorded from 47 subjects at Beth Israel Hospital in Boston, USA, from 1975 to 1979. The dataset has 48 records, taken every 30 minutes, and the sampling frequency is 360 Hz. These ECGs are first marked by two or more ECG research experts and finally reach a unified marking result. Each marked result includes the location of the wave peak of R and the type of ECG heartbeat. In this study, we chose 60% of the data for training and 40% for testing the MITDB dataset in this study.

ECG records in CYBHiDB were collected from hand palms and fingertips. These hand palms and fingertips are acquired at a 1,000-Hz sample frequency with 12-bit resolution on a BioPLUX device. CYBHiDB contains two types of short-term and long-term datasets. The short-term dataset was collected from 65 healthy participants consisting of 49 males and 16 females, with ages between 22 and 41. The sessions were taken once every other day. The long-term dataset was collected from 63 healthy participants, consisting of 14 males and 49 females, with ages between 18 and 23. The group includes two distinct sessions obtained within a period of three months. In the long-term dataset, there are more inter-class variations in the heartbeat waveform over time. We therefore chose the long-term dataset as our experimental data. We chose the first and second sessions from the long-term dataset T1 and T2, respectively. We chose T1 for training and T2 for testing in this study as we did for the CYBHiDB dataset.

The PTB Diagnostic ECG dataset includes 549 recordings from 290 subjects and individuals with various cardiac conditions (including myocardial infarction, dysrhythmia, hypertrophy, or heart failure). There were 1–5 records for each topic, which ranged from 38.4 to 104.2 seconds, taken from all 12 criteria and 3 Frank clues. In this study, we chose 248 subjects whose range was longer than 100 seconds, and every subject had one record. We also chose 60% of the data for training and 40% for testing as for the PTB dataset.

The ECG signals in UofTDB were captured from the thumbs of both hands at a sample rate of 200 Hz and 12-bit resolution. UofTDB is the largest off-the-person public dataset. It contains ECG data captured from 1,020 subjects, and consists of 398 males and 622 females with ages 18 to 52 years.

ECG sessions were acquired from five postures. We use all the instances of UofTDB S1 data for training and evaluate on the S2, S3, S4, and S6 in this study for the UofTDB dataset.

4.2 Performance metrics

In the proposed model, each sample is represented by three heartbeats that are segmented in the preprocessing stage. The subject recognition rate is used as the evaluation criterion for the identification problem. This recognition rate is the percentage of correctly recognized probe samples and it is defined as follows:

$$Subject\ Recognition\ Rate = \frac{N\_correct\_samples}{N\_test\_samples} \tag{13}$$

where, $N\_test\_samples$ is the total number of probe samples and $N\_correct\_samples$ is the number of probe samples that are correctly identified. The equal error rate (EER) is the measurement for the verification problem. EER is acquired from the FRR and FAR, which are defined as follows:

$$FRR = \frac{FP}{TP + FP} \tag{14}$$

$$FAR = \frac{FN}{TN + FN} \tag{15}$$

where, *FRR* is the false reject rate, *FAR* is the false accept rate, *TP* is the true positive, *FP* is the false positive, *TN* is the true negative, and *FN* is the false negative. EER is defined by the value where *FRR* is equal to *FAR*.

4.3 Parameter evaluation

First, we set the regularization coefficient in the sparse representation. The expression of λ is as follows:

$$\lambda = \frac{1.2}{\sqrt{m}} \tag{16}$$

where, m is the characteristic dimension.

Next, we evaluate the number of cascade levels as follows:

It is apparent from Figure 3 that the subject recognition rate increases as the number of layers increases. However, when the number of layers reaches 5, the result is stable. Therefore the number of layers finally obtained is five.

Finally, we evaluate the number of iterations.

The subject recognition rate of the algorithm is more or less stable as the number of iteration times increases (see Figure 4). We therefore set the number of iterations to 70 after which the subject recognition rate gradually stabilizes.

4.4 Influence of the deep cascade learning model

We conducted three groups of feature experiments on four data sets to more intuitively illustrate the influence of the cascade model proposed in this study on the feature extraction process. The first group used PCA and 1DMRLBP approaches. The second group combined the two methods with the deep cascade learning model. The last group used both PCA and 1DMRLBP approaches combined with the deep cascade learning model. The recognition rates corresponding to the various methods are shown in Table 1.

The 1DMRLBP method has distinct advantages over PCA in recognition performance due to its multi-resolution concept. The two methods are combined with the deep cascade learning model and the performance is clearly improved. The best performance is obtained by combining the two methods with the deep cascade learning model, which is also the feature extraction method adopted in this study.

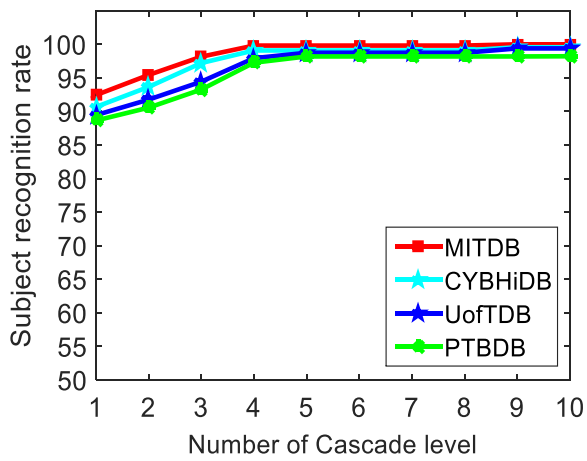

Figure 3. Results with different numbers of cascade levels

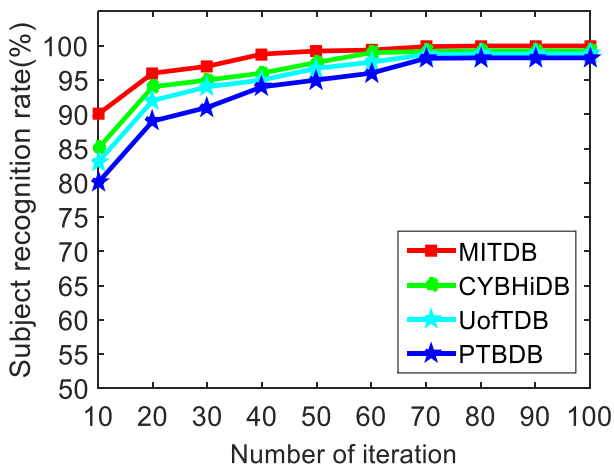

Figure 4. Influence of iteration number on performance

Table 1. Recognition rates for the different approaches

	PCA (%)	1DMRLBP (%)	PCA+Cascade (%)	1DMRLBP+Cascade (%)	Ours (%)
MIT-BIH	63.41	93.6	98.04	98.18	99.87
CYBHiDB	61.81	94.78	98.02	98.43	99.06
UofTDB	62.17	90.2	97.83	97.65	98.74
PTBDB	62.45	98.17	97.24	97.61	98.17

Table 2. Comparison with the state-of-the-art methods

Dataset	Method	Subject recognition rate (%)	EER (%)
MIT-BIH	Wang et al. [40]	94.68	2.73
	Louis et al. [45]	93.6	5.03
	Ergin et al. + PCA [28]	95.0	-
	Ours	99.87	0.68
PTBDB	Pal and Singh [46]	97.1	-
	Paiva et al. [47]	97.5	-
	Ours	98.17	1.92
CYBHiDB	da Silva Luz et al. [34]	-	1.3
	da Silva et al. [48]	94.4	-
	Lourenço et al. [49]	95.2	-
	Ours	99.06	1.34
UofTDB	Louis et al. [45]	90.2	2.55
	da Silva Luz et al. [34]	-	14.27
	Rehman et al. [50]	-	3.7
	Ours	98.74	2.56

4.5 Comparison with the state-of-the-art methods

Comparisons between our method and other state-of-the-art methods on the ECG dataset are summarized in Table 2.

The equal error rate (EER) of our method is significantly lower compared with the state-of-the-art methods. We specifically compared with the methods using PCA [28, 47] and 1DMRLBP [40, 45] for the MIT-BIH data set and the results indicate that our better performance is more significant. Our method also had satisfactory performance for the other datasets compared with the method of using the deep neural network for feature extraction [48, 49]. The results indicate that our method achieves satisfactory recognition performance on the four datasets. Recognition rates of 99.87, 98.18, 99.06 and 98.74% were achieved for MIT-BIH, PTBDB, CYBHiDB, and UofTDB, respectively. The main reason why our method is superior to the state-of-the-art methods is that both PCA and 1DMRLBP are used to extract the ECG features and convert the ECG signals into softmax vectors using the GNCV function based on spare representation. More importantly, the combination of our method with the deep cascade learning model can significantly improve the recognition performance.

5. CONCLUSIONS

In this study, we propose a multi-grained deep cascade learning model for ECG biometric recognition using global and local features. First, we combined PCA and 1DMRLBP methods to extract the global and local features of the ECG signals, respectively, thereby improving the accuracy. Second, we performed a multi-granularity scanning on the extracted ECG features and converted them to softmax vectors. Class differentiation was achieved using the GNCV method based on sparse representation. Finally, we proposed a multi-grained deep cascade learning model without back propagation and combined it with the GNCV function to greatly improve the classification accuracy. Extensive experimental results on the four datasets demonstrate that our approach is superior to several state-of-the-art techniques. In addition, our method is general and not only applicable to ECG biometrics but also to other traits biometric recognition. The good performance for ECG biometrics has been achieved with our approach. However, there is still room for improvement of the proposed ECG biometric recognition model, especially for large-scale ECG applications. In future work, we will further study the attribute information of the ECG signals to improve the accuracy of classification.

DATA AVAILABILITY

The simulated data used to support the simulation part of this study are available from the corresponding author upon request, and the real-world ECG data can be obtained from http://ecg.mit.edu/, https://www.physionet.org/content/ptbdb/1.0.0/, and https://zenodo.org/record/2381823#.X-7UZDMorRR.

CONFLICTS OF INTEREST

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

FUNDING

This work was supported in part by the Natural Science Foundation of Shandong Province (Grant No.: ZR2020MF052).

ACKNOWLEDGMENT

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

REFERENCES

[1] Jain, A.K., Ross, A., Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, 14(1): 4-20. https://doi.org/10.1109/TCSVT.2003.818349

[2] Schroff, F., Kalenichenko, D., Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, pp. 815-823. https://doi.org/10.1109/CVPR.2015.7298682

[3] Bansal, R., Sehgal, P., Bedi, P. (2011). Minutiae extraction from fingerprint images-a review. arXiv

preprint arXiv:1201.1422. https://doi.org/10.48550/arXiv.1201.1422
[4] Sarkar, I., Alisherov, F., Bhattacharyya, D., Kim, T.H. (2010). Palm vein authentication system: A review. International Journal of Control and Automation, 3(1): 27-34.
[5] Krishneswari, K., Arumugam, S. (2010). A review on palm print verification system. International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), 2: 113-120.
[6] Bowyer, K.W., Hollingsworth, K., Flynn, P.J. (2008). Image understanding for iris biometrics: A survey. Computer Vision and Image Understanding, 110(2): 281-307. https://doi.org/10.1016/j.cviu.2007.08.005
[7] Bourlai, T., Whitelam, C., Kakadiaris, I. (2011). Pupil detection under lighting and pose variations in the visible and active infrared bands. In 2011 IEEE International Workshop on Information Forensics and Security, Iguacu Falls, Brazil, pp. 1-6. https://doi.org/10.1109/WIFS.2011.6123139
[8] Abaza, A., Bourlai, T. (2013). On ear-based human identification in the mid-wave infrared spectrum. Image and Vision Computing, 31(9): 640-648. https://doi.org/10.1016/j.imavis.2013.06.001
[9] Pflug, A., Busch, C. (2012). Ear biometrics: A survey of detection, feature extraction and recognition methods. IET Biometrics, 1(2): 114-129. https://doi.org/10.1049/iet-bmt.2011.0003
[10] Dunn, D. (1993). Using a multi-layer perceptron neural for human voice identification. In Proceedings of the 4th Int. Conf. Signal Process. Applicat. Technol., Newton, MA, USA, 1993.
[11] Murase, H., Sakai, R. (1996). Moving object recognition in eigenspace representation: gait analysis and lip reading. Pattern Recognition Letters, 17(2): 155-162. https://doi.org/10.1016/0167-8655(95)00109-3
[12] Hafemann, L.G., Sabourin, R., Oliveira, L.S. (2017). Offline handwritten signature verification—literature review. In 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, Canada, pp. 1-8. https://doi.org/10.1109/IPTA.2017.8310112
[13] Van Drempt, N., McCluskey, A., Lannin, N.A. (2011). A review of factors that influence adult handwriting performance. Australian Occupational Therapy Journal, 58(5): 321-328. https://doi.org/10.1111/j.1440-1630.2011.00960.x
[14] Monrose, F., Rubin, A.D. (2000). Keystroke dynamics as a biometric for authentication. Future Generation Computer Systems, 16(4): 351-359. https://doi.org/10.1016/S0167-739X(99)00059-X
[15] Karnan, M., Akila, M., Krishnaraj, N. (2011). Biometric personal authentication using keystroke dynamics: A review. Applied Soft Computing, 11(2): 1565-1573. https://doi.org/10.1016/j.asoc.2010.08.003
[16] Forsen, G.E., Nelson, M.R., Staron Jr, R.J. (1977). Personal Attributes Authentication Techniques. Pattern Analysis and Recognition Corp Rome Ny.
[17] Biel, L., Pettersson, O., Philipson, L., Wide, P. (2001). ECG analysis: A new approach in human identification. IEEE Transactions on Instrumentation and Measurement, 50(3): 808-812. https://doi.org/10.1109/19.930458
[18] Wang, J., She, M., Nahavandi, S., Kouzani, A. (2013). Human identification from ECG signals via sparse representation of local segments. IEEE Signal Processing Letters, 20(10): 937-940. 10.1109/LSP.2013.2267593
[19] Tan, W.C., Yeap, H.M., Chee, K.J., Ramli, D.A. (2014). Towards real time implementation of sparse representation classifier (SRC) based heartbeat biometric system. Computational Problems in Engineering, 189-202. https://doi.org/10.1007/978-3-319-03967-1_15
[20] Li, R., Yang, G., Wang, K., Huang, Y., Yuan, F., Yin, Y. (2020). Robust ECG biometrics using GNMF and sparse representation. Pattern Recognition Letters, 129: 70-76. https://doi.org/10.1016/j.patrec.2019.11.005
[21] Goshvarpour, A., Goshvarpour, A. (2019). Human identification using a new matching pursuit-based feature set of ECG. Computer Methods and Programs in Biomedicine, 172: 87-94. https://doi.org/10.1016/j.cmpb.2019.02.009
[22] Huang, Y., Yang, G., Wang, K., Liu, H., Yin, Y. (2020). Learning joint and specific patterns: A unified sparse representation for off-the-person ECG biometric recognition. IEEE Transactions on Information Forensics and Security, 16: 147-160. https://doi.org/10.1109/TIFS.2020.3006384
[23] Zhang, Q., Zhou, D., Zeng, X. (2017). HeartID: A multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications. IEEE Access, 5: 11805-11816. https://doi.org/10.1109/ACCESS.2017.2707460
[24] Salloum, R., Kuo, C.C.J. (2017). ECG-based biometrics using recurrent neural networks. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, p. 2062-2066. https://doi.org/10.1109/ICASSP.2017.7952519
[25] Zhao, Z., Zhang, Y., Deng, Y., Zhang, X. (2018). ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation. Computers in Biology and Medicine, 102: 168-179. https://doi.org/10.1016/j.compbiomed.2018.09.027
[26] Zhou, Z.H., Feng, J. (2019). Deep forest. National Science Review, 6(1): 74-86. https://doi.org/10.1093/nsr/nwy108
[27] Bassiouni, M.M., El-Dahshan, E.S.A., Khalefa, W., Salem, A.M. (2018). Intelligent hybrid approaches for human ECG signals identification. Signal, Image and Video Processing, 12: 941-949. https://doi.org/10.1007/s11760-018-1237-5
[28] Ergin, S., Uysal, A.K., Gunal, E.S., Gunal, S., Gulmezoglu, M.B. (2014). ECG based biometric authentication using ensemble of features. In 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), Barcelona, Spain, pp. 1-6. https://doi.org/10.1109/CISTI.2014.6877089
[29] Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D., Wiederhold, B.K. (2005). ECG to identify individuals. Pattern Recognition, 38(1): 133-142. https://doi.org/10.1016/j.patcog.2004.05.014
[30] Kyoso, M., Uchiyama, A. (2001). Development of an ECG identification system. In 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, pp. 3721-3723. https://doi.org/10.1109/IEMBS.2001.1019645
[31] Plataniotis, K.N., Hatzinakos, D., Lee, J.K. (2006). ECG

biometric recognition without fiducial detection. In 2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, Baltimore, MD, USA, pp. 1-6. https://doi.org/10.1109/BCC.2006.4341628

[32] Jung, W.H., Lee, S.G. (2017). ECG identification based on non-fiducial feature extraction using window removal method. Applied Sciences, 7(11): 1205. https://doi.org/10.3390/app7111205

[33] Page, A., Kulkarni, A., Mohsenin, T. (2015). Utilizing deep neural nets for an embedded ECG-based biometric authentication system. In 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS), Atlanta, GA, USA, pp. 1-4. https://doi.org/10.1109/BioCAS.2015.7348372

[34] da Silva Luz, E.J., Moreira, G.J., Oliveira, L.S., Schwartz, W.R., Menotti, D. (2017). Learning deep off-the-person heart biometrics representations. IEEE Transactions on Information Forensics and Security, 13(5): 1258-1270. https://doi.org/10.1109/TIFS.2017.2784362

[35] Chen, X.Q., Wang, R.P. (2012). The Heartbeat Classification Based on PCA. Applied Mechanics and Materials, 195-196: 402-406.

[36] Zhang, H., Zhang, L. Q. (2005). ECG analysis based on PCA and support vector machines. In 2005 International Conference on Neural Networks and Brain, Beijing, pp. 743-747. https://doi.org/10.1109/ICNNB.2005.1614733

[37] Ojala, T., Pietikäinen, M., Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, 29(1): 51-59. https://doi.org/10.1016/0031-3203(95)00067-4

[38] Houam, L., Hafiane, A., Boukrouche, A., Lespessailles, E., Jennane, R. (2014). One dimensional local binary pattern for bone texture characterization. Pattern Analysis and Applications, 17: 179-193. https://doi.org/10.1007/s10044-012-0288-4

[39] Ojala, T., Pietikainen, M., Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7): 971-987. https://doi.org/10.1109/TPAMI.2002.1017623

[40] Wang, K., Yang, G., Huang, Y., Yin, Y. (2020). Multi-scale differential feature for ECG biometrics with collective matrix factorization. Pattern Recognition, 102: 107211. https://doi.org/10.1016/j.patcog.2020.107211

[41] Su, K., Yang, G., Wu, B., Yang, L., Li, D., Su, P., Yin, Y. (2019). Human identification using finger vein and ECG signals. Neurocomputing, 332: 111-118. https://doi.org/10.1016/j.neucom.2018.12.015

[42] Ding, J., Luo, Q., Jia, L., You, J. (2020). Deep forest-based fault diagnosis method for chemical process. Mathematical Problems in Engineering, 2020: Article ID 5281512. https://doi.org/10.1155/2020/5281512

[43] Utkin, L., Konstantinov, A., Meldo, A., Ryabinin, M., Chukanov, V. (2019). A deep forest improvement by using weighted schemes. In 2019 24th Conference of Open Innovations Association (FRUCT), Moscow, Russia, pp. 451-456. https://doi.org/10.23919/FRUCT.2019.8711886

[44] Liu, F., Yang, A. (2019). Application of gcForest to visual tracking using UAV image sequences. Multimedia Tools and Applications, 78: 27933-27956. https://doi.org/10.1007/s11042-019-07864-y

[45] Louis, W., Komeili, M., Hatzinakos, D. (2016). Continuous authentication using one-dimensional multi-resolution local binary patterns (1DMRLBP) in ECG biometrics. IEEE Transactions on Information Forensics and Security, 11(12): 2818-2832. https://doi.org/10.1109/TIFS.2016.2599270

[46] Pal, A., Singh, Y.N. (2018). ECG biometric recognition. In Mathematics and Computing: 4th International Conference, ICMC 2018, Varanasi, India, pp. 61-73. https://doi.org/10.1007/978-981-13-0023-3_7

[47] Paiva, J.S., Dias, D., Cunha, J.P. (2017). Beat-ID: Towards a computationally low-cost single heartbeat biometric identity check system based on electrocardiogram wave morphology. PloS one, 12(7): e0180942. https://doi.org/10.1371/journal.pone.0180942

[48] Da Silva, H.P., Lourenço, A., Fred, A., Raposo, N., Aires-de-Sousa, M. (2014). Check Your Biosignals Here: A new dataset for off-the-person ECG biometrics. Computer methods and programs in biomedicine, 113(2): 503-514. https://doi.org/10.1016/j.cmpb.2013.11.017

[49] Lourenço, A., Silva, H., Fred, A. (2012). ECG-based biometrics: A real time classification approach. In 2012 IEEE International Workshop on Machine Learning for Signal Processing, Santander, Spain, pp. 1-6. https://doi.org/10.1109/MLSP.2012.6349735

[50] Rehman, U.U., Kamal, K., Iqbal, J., Sheikh, M.F. (2019). Biometric identification through ECG signal using a hybridized approach. In Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence, pp. 226-230. https://doi.org/10.1145/3330482.3330496