Multi-feature Sparse Representations Learning via Collective Matrix Factorization for ECG Biometric Recognition

CHUNYING LIU¹,², JIJIANG YU¹,², YUWEN HUANG¹,² and FUXIAN HUANG¹,²

¹School of computer, Heze University, Heze 274015, China
²Computer Information Processing Lab, Heze University, Heze 274015, China
Corresponding author: Fuxian Huang (e-mail: huangfuxian@126.com).

This work was supported in part by the NSFC-Xinjiang Joint Fund under Grant U1903127 and in part by the Natural Science Foundation of Shandong Province under Grant ZR2020MF052.

ABSTRACT Electrocardiogram (ECG) signal is a promising biometric trait, and many methods have been proposed for ECG biometric recognition. However, it is challenging to design a robust and precise method to improve the recognition performance of ECG signals with noise and signal variation. We present a multi-feature sparse representations learning model via collective matrix factorization for ECG biometric recognition, MSRCMF for short. First, we extract one-dimensional local binary pattern (1D-LBP), shape and wavelet features of ECG signals and then obtain their sparse representations. Second, to extract discriminative information and preserve the intra- and inter-subject similarities, we leverage the collective matrix factorization on multiple sparse representations and the label information to obtain the latent semantic space. At last, we can recognize the ECG signals in the learned semantic space. Extensive experiments on four ECG databases show that MSRCMF can achieve competitive performance compared to state-of-the-art methods.

INDEX TERMS Electrocardiogram biometric recognition, multiple feature, sparse representations learning, collective matrix factorization, label information

I. INTRODUCTION

Electrocardiogram (ECG) signal is a novel biometric trait as the identity of people, and ECG biometric recognition has been widely received attention [1]. Compared to other biometric trains, such as fingerprints, faces, iris, gaits and veins, ECG signals have several distinctive advantages as follows [2], [3]. 1) Liveness detection. Only living persons include ECG signals, which can prove the intrinsic liveness. 2) High security. Due the continuity and irregularity of ECG signals, it is difficult to counterfeit them comparison with the fingerprints, faces, iris and so on. 3) Hybrid information. ECG signals provide non only the information of identity recognition, but also disease detection and heart status. 4) Small data size. ECG signals are one-dimensional data, and most of the other biometric traits are two-dimensional images. In addition, the good-quality ECG signals could be captured from a finger, and it is suitable and acceptable for the user as a biometric modality, which is crucial for improvement of security and privacy [4].

ECG signal is a promising biometric trait, and many methods have been proposed for ECG biometric recognition [5], [6]. However, there are still some issues that need to be further perform this technique robustly and precisely for the following reasons: 1) Noise. There are noises produced by ECG signal-acquisition devices, such as the noises from power line interference, baseline wandering, and electrode motion artifacts [7]. In addition, ECG signals are acquired on the skin’s surface by sensors, and it is inevitable with contact noise. 2) Signal variation. ECG signals are easily influenced by the external factor, and the waveforms of ECG signals change over time [8]. Owing to noise and signal variation, it is a challenging problem to design a robust and precise method.

Sparse representation learning (SRL) can efficiently remove noises, and many methods based on sparse representation have been proposed for ECG biometric recogni-
C. Y. Liu et al.: Multi-feature Sparse Representations Learning via Collective Matrix Factorization for ECG Biometric Recognition

FIGURE 1. Simple Illustration of Our Framework

SRL can represent a test sample by a linear combination of dictionary elements. However, the noise and sample variation can destroy the linear structure of dictionary elements, and the performance of SRL degrades when data are corrupted. Therefore, a challenging problem is how to design a robust method to improve the performance of SRL with noise and sample variation.

Recently, many multi-feature learning methods are proposed for biometric recognition [13], [14], [15], [16]. Different features from the same subject can provide complementary information, and the multi-feature learning can improve the recognition performance. Different features have different meanings, and the concatenated features may not be necessary to improve recognition performance. Therefore, a key problem of multi-feature fusion is how to learn a unified representation from multiple features.

The collective matrix factorization (CMF) can transform the original space into the latent semantics space and remove redundant information, and many CMF-based methods are proposed for biometric recognition [17] [18] [19]. However, the existing CMF-based methods learn semantic representations from cross-domains, without considering multiple features. Hence a challenging problem is to design CMF-based learning method from multiple features.

To tackle the above issues, we design a multi-feature sparse representations learning model via collective matrix factorization for ECG biometric recognition, which can obtain the latent semantic space from multiple sparse representations, and is intrinsically different with current CMF-based frameworks.

2. To further enhance the discrimination of learned features, we integrate label information to preserve the intra- and inter-subject similarities.

3. We use an iterative optimization strategy to solve the optimization problem in MSRCMF, and the time complexity is linear to the database size.

II. RELATED WORK

A. ECG BIOMETRIC RECOGNITION

There are fiducial-based, non-fiducial-based, and hybrid-based approaches for ECG biometrics recognition [3].

Fiducial-based approaches use the feature of time domain, amplitude waveform, and intervals [20]. Venketsh et al. [21] used nine feature of ECG signals in the spatial domain for ECG biometric recognition. The challenge in fiducial-based approaches is that amplitudes and intervals of fiducial points are sensitive to noise. Therefore, these approaches are generally applicable to ECG signals in a relatively noise-free environment. Non-fiducial-based approaches usually extract the entire waveform morphology of ECG features. Hejazi et al. [22] extracted the autocorrelation coefficients for ECG biometric recognition. Tantawi et al. [23] extract the discrete wavelet coefficient of R-R intervals for ECG identification recognition. Louis et al. [6] extracted the multi-resolution local binary patterns of ECG signals for identification recognition. Wang et al. [17] fused the multi-scale differential feature and one-dimensional local binary features of ECG signals. Wang et al. [17] fused the multi-scale differential feature and one-dimensional local binary features of ECG signals. Lim et al. [24] fused the hybridized the discrete wavelet transforms and heart-rate interval for ECG biometric recognition. Bassioni et al. [25] presented a comparative analysis of the authentication performance of ECG biometric systems.

Recently, deep learning for ECG biometric recognition is a hot research area. AlDuwaile et al. [26] designed a small convolutional neural network to achieve better generalization.
capability by entropy enhancement of a short segment of a heartbeat signal for ECG biometric recognition. Labate et al. [27] presented a deep-ECG, a convolutional Neural Networks biometric approach for ECG signals. Srivastva et al. [28] proposed a novel ensemble of the state-of-the-art pre-trained deep neural networks for ECG biometric recognition. Da et al. [55] used two separate convolutional neural network techniques to extract useful representations for ECG biometrics. Hammad et al. [29] developed a secure multimodal ECG biometric system that used convolution neural network (CNN) and Q-Gaussian multi support vector machine (QG-MSVM) based on a different level fusion. Eduardo et al. [30] proposed an ECG-based biometric recognition approach using a deep autoencoder for feature learning.

Deep learning shows good performance based on large-scale training data for massive hyperparameters tuning. However, ECG biometric databases are relatively small, making it difficult to train a deep learning model.

B. SPARSE REPRESENTATION LEARNING

Assume that \( X^t = [X_1^t, X_2^t, ..., X_C^t] \in R^{d_t \times n} \) represents training samples, \( C \) is the total number of classes. \( X_i^t = [x_1^t, x_2^t, ..., x_n^t] \in R^{d_t \times n} \), and \( n = \sum_{i=1}^{C} n_i \).

A test sample \( y \) can be reconstructed by a sparse linear combination of training samples \( X^t \) as:

\[
y = X^t w_p \tag{1}
\]

where \( w_p = [0, ..., 0, x_{i1}, x_{i2}, ..., x_{in}, 0, ..., 0] \in R^n \) is a sparse coefficient vector with only nonzero elements associated with the \( i^{th} \) class. It should be noted that the advantages of representing the test sample as a linear combination of training samples have been explored in [31]–[33].

According to sparse representation learning [33], we can obtain the sparse representation coefficient of sample \( y \) by solving the following optimization problem:

\[
\min_{w_p} \left\| y - X^t w_p \right\|^2_2 + \lambda_1 \left\| w_p \right\|_1 \tag{2}
\]

where \( \lambda_1 \) is a regularization coefficient, \( w_p \in R^n \) is a sparse representation coefficient, and \( \left\| \bullet \right\|_1 \) is the \( L_1 \) norm.

III. PROPOSED METHODOLOGY

A. NOTATIONS

Suppose \( \{X^t\}_{i=1}^{D} \) is the training set of multiple sparse representation matrices, where \( D \) is the number of multiple features, each sparse representation matrix \( X^t = [x_1^t, x_2^t, ..., x_n^t] \in R^{d_t \times n} \) contains \( n \) samples, and \( d_t \) is the dimension of the \( t^{th} \) feature. Let \( L \in R^{p \times n} \) denotes the ground-truth label matrix, where \( p \) is the number of subjects, \( L_{ki} \) is 1 if the \( i^{th} \) base feature belongs to subject \( k \), otherwise \( L_{ki} = 0 \). \( \{Y^t\}_{i=1}^{D} \) represents testing sets with \( D \) different sparse representation, and each sparse representation matrix \( Y^t = [y_1^t, y_2^t, ..., y_n^t] \in R^{q \times n} \) consists of \( q \) samples. A test sample \( y_i = \{y_{i1}^t, y_{i2}^t, ..., y_{in}^t\} \) contains \( D \) features, and our objective is to identify the class that sample \( y_i \) belongs to.

B. PREPROCESSING

First, to remove the noise of ECG signals, we use the wavelet threshold denoising method to obtain the useful signals. Second, the R peaks of ECG signals are located by the Pan-Tompkins algorithm [34], and we can divide ECG signal into heartbeat segmentations by setting each heartbeat fixed length before and after the R peak. Then, we normalized all each heartbeats between 0 and 1. At last, all heartbeats from the same subject are superposed, and ECG templates are obtained by the mean heart value.

C. MULTIPLE FEATURE EXTRACTION

We extracted the shape, wavelet, and 1D-LBP features from ECG segments, respectively. We extract 30 shape features from heartbeat, which include 15 temporal distance features, six amplitude distances, and 10 morphological descriptors [35], [36]. The discrete wavelet coefficients of the heartbeat are obtain as wavelet features, and we chose Daubechies wavelet Db3 with five levels of decomposition to obtain the heartbeat feature value [37]. The 1D-LBP method [38] can effectively extract binary codes from ECG signals by comparing each sampling point with its neighbors. We can obtain the multiple sparse representation matrices from the extracted features by the sparse representation learning.

D. FORMULATION

CMF techniques can remove redundant information and obtain the latent semantic representation. Specifically, a matrix \( X^t \) can represented by two rank-r matrices \( U^t \) and \( V^t \) via CMF as

\[
X^t = U^t V^t, \tag{3}
\]

where \( X^t \in R^{d_t \times n} \), \( U^t \in R^{d_t \times r} \), \( V^t \in R^{r \times n} \), and \( r \) is the dimension of the learned latent semantic space, \( r \ll d_t \), and each column vector \( v_i \) of \( V \) is a latent factor vector of a sample in the latent semantic space.

Intuitively, different sparse representations of a sample all describe the same subject, justifying that they should have the same semantics. Based on this, we can find that different sparse representations share the same latent semantic space learned by CMF. In the light of this, we have the following cost function for each sparse representation:

\[
\min_{U^t, V^t} \lambda \left\| X^t - U^t V^t \right\|^2_F + \gamma \left( \left\| U^t \right\|^2_F + \left\| V^t \right\|^2_F \right) \tag{4}
\]

where \( \lambda \) and \( \gamma \) are the balance parameter, and \( \left\| \bullet \right\|_F \) denotes the Frobenius norm.

In the original space, different sparse representations of one sample have the same semantic labels, so the representations of one sample should also contain the same label in the common latent semantic space. With this in mind, we define

\[
\min_{G, U^t, V^t} \alpha \left\| L - GV \right\|^2_F + \gamma \left\| G \right\|^2_F \tag{5}
\]

where \( L \in R^{p \times n} \) is the ground-truth label matrix, \( G \in R^{p \times r} \) is the label projection matrix, and \( p \) is the number of classes.
Combining Eq. (4) and Eq. (5), we have the following objective function:

$$\min_{U^t, V, G} \lambda \left\| X^t - U^t V \right\|^2_F + \alpha \left\| L - G V \right\|^2_F + \gamma \left( \left\| G \right\|^2_F + \left\| U^t \right\|^2_F + \left\| V \right\|^2_F \right)$$

(6)

By solving Eq. (6), we can obtain the learned semantic matrix $V$ of training samples. Furthermore, to obtain the projection matrices by the learned semantic representation on the training set, we need the specific function to map testing samples from the original space to the semantic space. Thus, a project matrix is needed and can be formulated as follows:

$$\min_{W^t} \beta \left\| V - W^t X^t \right\|^2_F + \gamma \left\| W^t \right\|^2_F$$

(7)

where $W^t$ is the projection matrix, $\gamma$ is the balance parameter. Combining Eqs. (6) and (7), we have the following overall objective function:

$$\min_{U^t, V, G, W^t} \lambda \sum_{t=1}^{D} \left\| X^t - U^t V \right\|^2_F + \alpha \left\| L - G V \right\|^2_F + \gamma \sum_{t=1}^{D} \left( \left\| G \right\|^2_F + \left\| U^t \right\|^2_F + \left\| V \right\|^2_F + \left\| W^t \right\|^2_F \right)$$

(8)

where $D$ is the number of sparse representations.

### E. Optimization

The objective function in Eq. (8) is non-convex with multiple variables $U^t$, $W^t$, $V$, $G$. To tackle the optimization problem, we can solve it by varying any one of these matrix while fixing the others until convergence or the maximum number of iterations is reached as following steps.

**Step 1: Fixing others and solving $U^t$.**

Fix the other variables, we can obtain $U^t$ by solving the following minimization problem:

$$\min_{U^t} \lambda \left\| X^t - U^t V \right\|^2_F + \gamma \left\| U^t \right\|^2_F$$

(9)

To solve the Eq. (9), we set the derivative w.r.t. $U^t$ to zero, and can obtain $U^t$ by a closed-form solution as:

$$U^t = \lambda X^t V^T (\lambda V V^T + \gamma I)^{-1}$$

(10)

where $I \in R^{r \times r}$ is the identity matrix.

**Step 2: Fixing others and solving $G$.**

Fix the other variables, we can obtain $G$ by solving the following minimization problem:

$$\min_{G} \alpha \left\| L - G V \right\|^2_F + \gamma \left\| G \right\|^2_F$$

(11)

Similar to the optimization of solving $U^t$, we can obtain the analytical solution of $G$ as follows:

$$G = \alpha L V^T (\alpha V V^T + \gamma I)^{-1}$$

(12)

After $c$ iterations, the order of magnitude of the overall computational complexity in the training stage is $O(n(r^2 + Dd + d^2 + Cr)c)$, where $C$ is the number of classes, $r$ is the dimension of the latent semantic space, $n$ is the size of the training set, and $d = \sum_{t=1}^{D} d_t$ is the dimension of the $t^{th}$ feature, $D$ is the number of features. It is noted that the time complexity is linear to the database size.

### F. Matching

We divide the testing set into the enroll and probe set, which are used as matching templates and query samples, respectively.

First, as illustrated in Fig. 1, testing set is preprocessed and segmented, we obtain the ECG heartbeats. Second, we extract 1D-LBP, shape and wavelet features of heartbeats...
and then obtain the sparse representations of the enroll and probe heartbeats, denoted as $X_{enroll}^t, X_{probe}^t, X_{enroll} = \{X_{enroll}^t\}_{t=1}^D, X_{probe} = \{X_{probe}^t\}_{t=1}^D$, respectively. Third, we extract the high-level semantic representations of the enroll and probe set, denoted as $W^t X_{enroll}^t$ and $W^t X_{probe}^t$, respectively. At last, we use the Euclidean distance to calculate the similarity between the probe and enrolled samples as

$$d_{ij} = \sum_{t=1}^D \sqrt{\sum_{k=1}^r (W^t X_{enroll}^t(i) - W^t X_{probe}^t(j))^2}$$

(16)

where $X_{enroll}^t(i)$ is the $i^{th}$ enroll sample, $X_{probe}^t(j)$ is the $j^{th}$ probe sample, $d_{ij}$ is the distance between the $i^{th}$ enroll and the $j^{th}$ probe sample. We can recognize the probe sample by comparing $d_{ij}$ to the set threshold.

IV. EXPERIMENTS

A. DATABASES

We evaluate the proposed method on four public databases: MIT-BIH Arrhythmia Database (MITDB) [39], PTB Diagnostic ECG Database (PTBDB) [40], ECG-ID database [41] and University of Toronto Database (UofTDB) [42], and UofTDB is off-the-person ECG databases.

MITDB is widely used in ECG biometrics and heart disease, and includes 48 two-lead recordings of 47 subjects with a 360 Hz sampling frequency and 11 bits per sample resolution. In our experiments, we choose each one recording from 47 subjects.

PTBDB contains 549 recordings from 290 subjects with healthy subjects and heart diseases patients. Each subject from PTBDB has 1 to 5 recordings from 12-lead standard and three Frank leads, ranging between 38.4 and 104.2 seconds. In our experiments, 248 subjects with one recording are chosen, and each recording is longer than 100 seconds.

ECG-ID database contains 310 recordings of 90 volunteers, which has 44 male and 46 females with age from 13 to 75. Each recording of ECG-ID includes ECG lead I, recorded for 20 seconds and digitized at 500 Hz with 12-bit resolution over a nominal ±10 mV range [41]. The number of records from each volunteer varies from 2 to 20, and the acquisition time is from one day to over 6 months. In our experiments, we utilize 89 subjects with two recordings, and the volunteer 74 has only one record, so we discarded it.

UofTDB was specifically created for biometrics, and was captured from the thumbs of both hands with 200 Hz and 12 bit resolution. It is the largest off-the-person public database right now. ECG data in UofTDB covers six sessions, which are named as S1, S2, S3, S4, S5 and S6, and are captured from 1020 subjects between 18 and 52 years. All sessions are acquired under five postures: sitting, standing, exercising, supine and tripod. The first session includes 1012 subjects, from 100 subjects selected to participate in follow-up sessions over a period of six months. In our experiment, we consider 46 subjects in a sitting posture, who participate in all five sessions (S1, S2, S3, S4 and S6).

In MITDB, ECG-ID and PTBDB databases, each subject select one session, and we randomly take 60% of sessions per subject as the training set, 30% for enroll set, and 10% for probe set. We choose two recordings of each subject in ECG-ID, which are called as T1 and T2, respectively. To test our method for multisession data, in ECG-ID, we consider one recording of each subject as the training set, and the other recording is used as the testing set. In UofTDB, we consider S1 as the training set, and S2, S3, S4 and S6 as testing set. In all databases, the heartbeats are segmented by setting 800 milliseconds time length. Experiments are performed on a PC with a 3.60–GHz Core i5-5200U CPU and 8 GB RAM.

B. PERFORMANCE METRICS

To evaluate the performance of the proposed approach, we perform the identification mode and authentication mode. In the identification mode, we use

$$\text{heartbeat recognition rate} = \frac{N_{\text{correct beat}}}{N_{\text{probe beat}}}$$

(17)

where $N_{\text{probe beat}}$ is the total number of probe heartbeats, and $N_{\text{correct beat}}$ is the number of correctly identified probe heartbeats.

Like most work in the literature [3], [5], [43], we also use the subject recognition rate as the identification criterion, which is computed by several continuous heartbeats of voters. For this criterion, each segment includes several continuous heartbeats, and we define

$$\text{subject recognition rate} = \frac{N_{\text{correct}}} {N_{\text{sample}}}$$

(18)

where $N_{\text{sample}}$ is the total number of probe segments, and $N_{\text{correct}}$ is the number of correctly identified probe segments. In all experiments, we use three heartbeats as a segment, which are acquired in about 2-4 seconds, and it is acceptable for practical application.

In the authentication mode, we perform a similarity measure with a segment and all other segments from the same class and different classes. The equal error rate (EER) is the equality of the false acceptance rate (FAR) and false rejection rates rate (FRR).

$$\text{FAR} = \frac{N_{\text{FA}}}{N_{\text{IRA}}} \times 100\%$$

(19)

$$\text{FRR} = \frac{N_{\text{FR}}}{N_{\text{GRA}}} \times 100\%$$

(20)

where $N_{\text{FA}}$ is the number of false acceptance, $N_{\text{IRA}}$ is the number of impostor attempts, $N_{\text{FR}}$ is the number of false rejection, $N_{\text{GRA}}$ is the number of genuine user attempts.

C. EFFECTS OF PARAMETER SETTINGS

First, we vary $\lambda$, $\alpha$, $\beta$, and $\gamma$ in range of [0,10] with step size of 0.01, and obtain the best parameter values at the same
time. When $\lambda = 2.46$, $\alpha = 0.25$, $\beta = 0.12$, and $\gamma = 0.45$, we can achieve the best performance.

Second, we evaluate the performance influence of the number of iterations $c$, and the result is shown as Fig. 3.

**FIGURE 2.** Influence of number of iterations $c$ on performance

From Fig. 2, we can see that the heartbeat recognition performance begins to reach better value when $c$ reaches 60, which means that our approach can converge.

At last, we evaluate the performance influence of the dimension $r$ of the semantic space, and the result is shown as Fig. 3.

**FIGURE 3.** Influence of parameter $r$ on performance (a) MITDB; (b) PTBDB; (c) ECG-ID; (d) UofTDB.

As we can see from Fig. 3, on all databases, the ranges of $r$ are different when it achieves much better heartbeat recognition rate on the four databases.

### D. PERFORMANCE OF PROPOSED METHOD

First, we evaluate the performance of MSRCMF and baseline methods for ECG biometrics. We adapt shape, wavelet, 1D-LBP and single-feature MSRCMF as the baseline methods. Table 1 lists the results of heartbeat recognition rates with different baseline methods.

| Method          | MITDB | PTBDB | ECG-ID | UofTDB |
|-----------------|-------|-------|--------|--------|
| shape           | 65.77% | 62.46% | 63.87% | 62.63% |
| wavelet         | 73.61% | 63.18% | 64.21% | 62.52% |
| 1D-LBP          | 78.24% | 66.27% | 67.78% | 67.35% |
| MSRCMF with only shape | 89.04% | 79.64% | 80.53% | 80.13% |
| MSRCMF with only wavelet | 90.84% | 80.76% | 81.87% | 81.23% |
| MSRCMF with only 1D-LBP | 91.68% | 83.49% | 85.28% | 84.17% |
| MSRCMF          | 95.67% | 86.48% | 87.64% | 87.23% |

As shown in Table 1, the performance of MSRCMF is superior to all baseline methods on the four databases, which demonstrates that the CMF-based multi-feature sparse representation with label information has high discrimination. For example, the heartbeat recognition rates of MSRCMF is 95.67%, 86.48%, 87.64%, and 87.23% on MITDB, PTBDB, ECG-ID, and UofTDB, respectively.

Second, we compare the performance of MSRCMF with multi-feature biometrics methods, such as kerSMBR [44], RMRLRJS-C [45], JSSLCP [3], and RCNMF [46]. Table 2 lists the heartbeat recognition performance of multi-feature methods.

| Method          | MITDB | PTBDB | ECG-ID | UofTDB |
|-----------------|-------|-------|--------|--------|
| kerSMBR         | 82.24% | 72.71% | 74.67% | 73.03% |
| RMRLRJS-C       | 83.17% | 74.15% | 75.82% | 74.26% |
| RCNMF           | 92.69% | 84.26% | 86.34% | 85.75% |
| JSSLCP          | 94.57% | 86.07% | 87.45% | 86.74% |
| MSRCMF          | 95.69% | 86.48% | 87.64% | 87.23% |

From Table 2, we can see that MSRCMF has better performance than the other multi-feature methods, in part because MSRCMF uses multi-feature sparse representations and semantic embedding with label information, which has a better discriminative feature in semantic space.

Furthermore, we compares the heartbeat recognition performance with the multisection records. Table 3 lists the heartbeat recognition performance of our method and other methods in different scenarios.

From Table 3, compared to the baseline methods, our method achieves the best results on ECG-ID databases in different scenarios. Compared with deep methods, the performance of our method is comparable with MobileNet in [47], and is worse than Small CNN in [26]. The multi-section data have more signal variation, and our method is superior to the baseline methods. The multi-feature sparse representations of the multi-section data are complementary, and the collective matrix factorization can remove redundant information, so our method can enhance the discrimination of learned representations for multi-section data. In addition, our method integrates label information to preserve the intra- and inter-subject similarities, which can improve performance for multi-section data.
TABLE 3: Comparison of results with multisession scenarios on ECG-ID

| Method  | Training | Testing | Length of Signal | HR (%) |
|---------|----------|---------|-----------------|--------|
| shape   | S1       | S2      | One heartbeat   | 61.26  |
|         | S2       | S1      | One heartbeat   | 62.94  |
| wavelet | S1       | S2      | One heartbeat   | 62.35  |
|         | S2       | S1      | One heartbeat   | 63.78  |
| 1D-LBP  | S1       | S2      | One heartbeat   | 65.52  |
|         | S2       | S1      | One heartbeat   | 66.84  |
| MobileNet | S1   | S2      | 0.5 s segment   | 86.56  |
|         | S2       | S1      | 0.5 s segment   | 88.47  |
| Small CNN | S1  | S2      | 0.5 s segment   | 92.81  |
|         | S2       | S1      | One heartbeat   | 95.54  |
| MSRCMF  | S1       | S2      | One heartbeat   | 86.12  |
|         | S2       | S1      | One heartbeat   | 87.36  |

At last, we validate the robustness of MSRCMF. Gaussian noise with different variance $\sigma^2$ is added on four databases, and we compare the robustness of MSRCMF, RCNMF, shape, wavelet, and 1D-LBP. The heartbeat recognition rates with different noise levels are shown in Fig. 4.

FIGURE 4. Heartbeat recognition rates with different noise level on (a) MITDB; (b) PTBDB; (c) ECG-ID; (d) UofTDB.

It can be seen that MSRCMF with the influence of Gaussian noise can perform more stably than the other methods.

E. COMPARISON WITH STATE-OF-THE-ART METHODS

To further evaluate the performance of our method, we compare the proposed method with the state-of-the-art methods on four databases, including the methods of [6] [48] [50] [54], which are non-deep ECG biometric methods, and the methods of [49] [51] [52] [53] [55] [56], which are deep neural network-based ECG biometric methods. The performance comparison is shown in Table 4.

From Table 4, compared to the non-deep methods on all four databases, MSRCMF has the best recognition performance. On MITDB, our method achieves the best recognition rate compared to the deep learning methods. On ECG-ID, PTBDB and UofTDB, MSRCMF slightly worse than the methods in [26] [51] and [56]. Therefore, we can see that MSRCMF performs competitively with state-of-the-art methods on all four databases.

It is worth noting that MSRCMF and our previous work in [57] are different. First, our previous work in [57] extracted multiple LBP histograms as feature descriptors from segmented ECG signals, and MSRCMF obtain the sparse representations as features descriptor. Second, our previous work in [57] only constrained the semantic matrix $V$, and MSRCMF used Frobenius norm to constrain all dynamic matrices, including the latent basis factor $U$, semantic representation $V$, label projection matrix $G$, project matrix $W$. Third, all matrix variables of the objective function are non-negative during the computing procedure in our previous work, and all matrices in MSCMF are not necessary non-negative. Therefore, MSRCMF and our previous work are entirely different in the feature extraction, object function and computing procedure.

F. TIME COST ANALYSIS

In this section, we give the time cost analysis of the proposed method. The time costs are showed in Table 5. Due to the unavailable source code of the other methods, we only show the average time of our proposed method in training, preprocessing, feature extraction, and matching of each heartbeat.
From Table 5, we can see that the training feature extraction, and matching of our method is fast, and it is feasible time-consuming in practical application.

V. CONCLUSIONS AND FUTURE WORK

We have proposed a multi-feature sparse representations learning model via collective matrix factorization for ECG biometric recognition, which can extract discriminative information and preserve the intra- and inter-subject similarities in the latent semantic space. MSRCMF can obtain the latent semantic space from multiple sparse Representations, and is intrinsically different with current CMF-based frameworks. Extensive experiments on four ECG databases show that MSRCMF can achieve competitive performance compared to state-of-the-art methods. In future work, we will explore a new ECG recognition approach, which can incorporate the novel loss function into deep learning to further improve the performance.

REFERENCES

[1] P. J. Ribeiro, J. S. Cardoso, L. Andre, “Evolution, Current Challenges, and Future Possibilities in ECG Biometrics,” IEEE Access, vol. 6, pp. 34746-34776, Jun. 2018. 10.1109/ACCESS.2018.2849870.
[2] Kim M. S. B. Pan. Deep learning based on 1-D ensemble networks using ECG for real-time user recognition. IEEE Trans. Industrial Informatics, 2019, 15(10): 5656-5663. DOI: 10.1109/TII.2019.2909730.
[3] Huang Y W, Yang G P, Wang K K, Liu H Y, Yin Y L. Learning Joint and Specific Patterns: A Unified Sparse Representation for Off-the-Person ECG Biometric Recognition. IEEE Trans. on Information Forensics & Security, 2021, 16(1): 147-160. DOI: 10.1109/TIFS.2020.3006384.
[4] M. S. Islam, N. Alajlan. "Biometric template extraction from a heartbeat signal captured from fingers," Multimodal Tools Appl, vol. 76, 2017, pp. 12709–12733.
[5] S. C. Wu et al., “Cancelable biometric recognition with ECGs: Subspace-based approaches,” IEEE Trans. INFORE SEC, vol.14, no.5, pp.1323-1336, Oct. 2018
[6] Louis W., Komeili M., Hatzinakos D. "Continuous authentication using One-Dimensional Multi-Resolution Local Binary Patterns in ECG biometrics," IEEE Transactions on Information Forensics & Security, 2016, 11(12):2818-2832. 10.1109/TIFS.2018.2876838.
[7] H. Limaye and V. Deshmukh, “ECG noise sources and various noise removal techniques: a survey,” International Journal of Application or Innovation in Engineering & Management, vol. 5, no. 2, pp. 86-92, 2016.
[8] J. R. Pinto, J. S. Cardoso, and A. Lourencô, “Evolution, current challenges, and future possibilities in ECG biometrics,” IEEE Access, vol. 6, pp. 34746-34776, 2018.
[9] Wang J, She M, Nahavandi S, Kouzani A Z. Human identification from ECG signals via sparse representation of local segments. IEEE Signal Processing Letters, 2013, 20(10): 937-940. DOI: 10.1109/LSP.2013.2267593.
[10] Jaafar H, Ramli N, Nasir A S A. Implementation of kernel sparse representation classifier for ECG biometric system. Journal of Telecommunication, Electronic and Computer Engineering, 2018, 10(1): 89-94.
[11] Li R, Yang G, Wang K, Huang Y, Yuan F, Yin Y. Robust ECG biometrics using GMMF and sparse representation. Pattern Recognition Letters, 2019, 129: 70-76. DOI: 10.1016/j.patrec.2019.11.005.
[12] Goshvarpour A, Goshvarpour A. Human identification using a new matching pursuit-based feature set of ECG. Comput. Methods Programs biomedicine, 2019, 172: 87-94. DOI: 10.1016/j.cmpb.2019.02.009.
[13] W. Hu, W. Li, X. Zhang, and S. Maybank, “Single and multiple object tracking using a multi-feature joint sparse representation,”IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 4, pp. 816-833, 2014.
[14] R. Primorac, R. Togneri, M. Bennamoun, and F. Sohel, “Generalized joint sparse representation for multimodal biometric fusion of heterogeneous features,” in 2018 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, 2018, pp. 1-6.
[15] H. Zhang, V. M. Patel, and R. Chellappa, “Low-rank and joint sparse representations for multi-modal recognition,” IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4741-4752, 2017.
[16] X. Liu, S. Zhang, P. C. Yuen, and R. Chellappa, “Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker,” IEEE Transactions on Image Processing, vol. 27, no. 4, pp. 2022-2037, 2017.
[17] K. Wang, G. Yang, Y. Huang, and Y. Yin, “Multi-scale differential feature for eeg biometrics with collective matrix factorization,” Pattern Recognition, vol. 102, pp. 211-231, 2020.
[18] G. Ding, Y. Guo, J. Zhou, and Y. Gao, “Large-scale cross-modality search via collective matrix factorization hashing,” IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5427-5440, 2016.
[19] J. Lu, V. E. Liong, and J. Zhou, “Simultaneous local binary feature learning and encoding for homogeneous and heterogeneous face recognition,” IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 8, pp. 1979-1993, 2017.
[20] S. S. Abdeladayem and T. Bouriai, “A novel approach for eeg-based human identification using spectral correlation and deep learning,” IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 2, no. 1, pp. 1-14, 2019.
[21] N. Venkatesh and S. Jayaraman, “Human electrocardiogram for bitmectrics using dtw and fida,” in 2010 20th International Conference on Pattern Recognition. IEEE, 2010, pp. 3838-3841.
[22] 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,” Digital Signal Processing, vol. 52, pp. 72-86, 2016.
[23] M. M. Tantawi, K. Revett, A.-B. Salem, and M. F. Tolba, “A wavelet feature extraction method for electrocardiogram (ecg)-based biometric recognition,” Signal, Image and Video Processing, vol. 9, no. 6, pp. 1271-1280, 2015.
[24] C. L. P. Lim et al., “Heartrate-dependent heartwave biometric identification with thresholding-based GMM-HMM methodology,” IEEE Trans. Industrial Informatics, vol. 15, no. 1, pp. 45-53, 2019.
[25] M. M. Bassioni et al., “Intelligent hybrid approaches for human ECG signals identification,” Signal, Image and Video Processing, vol. 12, no. 5, pp. 941-949, Feb. 2018.
[26] D. A. AliDuwaila,M. S.Islam,”Using Convolutional Neural Network and a Single Heartbeat for ECG Biometric Recognition,”Entropy, 2021, vol. 23, no. 733,pp.1-22.
[27] R. D. Labati, E. Munoz, V. Piuri, R. Sassi, and F. Scotti, “Deep-ecg: convolutional neural networks for eeg biometric recognition,” Pattern Recognition Letters, vol. 126, pp.78-85, 2019.
[28] R. Srivastava, A. Singh, Y. N. Singh, “FlexNet: A fast and robust ECG biometric system for human recognition,” Information Sciences, 2021, vol. 558, no. 3, pp. 208-228.
[29] M. Hammad, Y. Liu, and K. Wang, “Multimodal biometric authentication systems using convolution neural network based on different level fusion of eeg and fingerprint,” IEEE Access, vol.7, 2018, pp. 26527-26542.
[30] A. Eduardo, H. Aidos, and A. Fred, “Ecg-based biometrics using a deep autoencoder for feature learning-an empirical study on transferability,” in International Conference on Pattern Recognition Applications and Methods, vol.2. SCITEPRESS, 2017, pp. 463-470.
[31] Lan X, Zhang S, Yuen P C, Chellappa R. Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker. IEEE Trans. on Image Processing, 2017, 27(4): 2022-2037. DOI: 10.1109/TIP.2017.2777183.
[32] Panagakis Y, Kotropoulos C, Arce G R. Music genre classification via joint sparse low-rank representation of audio features. IEEE Trans. Audio, Speech, and Language Processing, 2014, 12(2): 1905-1917. DOI: 10.1109/TASLP.2014.2355774.
[33] Wright J, Yang A Y, Ganesh A, Sastry S, Ma Y. Robust face recognition via sparse representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227. DOI: 10.1109/TPAMI.2008.79.
Halaman ini tidak dapat dibaca secara alami dan tidak relevan dengan teks yang diberikan.