This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2771220, IEEE Access

Research and Implementation of ECG-based Biological Recognition Parallelization

Yiming Miao, Yuanwen Tian, Limei Peng,
M. Shamim Hossain, *Senior Member, IEEE,* Ghulam Muhammad

Abstract*—Nowadays, biological recognition technologies attracts more attention in ECG (Electrocardiograph) signals, which vary among different people and are difficult to counterfeit. However, the robustness of recognition can not be well sustained in the case of diversified application scenarios and huge human crowds. In order to tackle this problem, this paper puts forward a fiducial and non-fiducial mixed feature extraction method, which can effectively complete the multidimensional feature modeling of ECG signal. In addition, this paper proposes a LOMF (LDA based on multiple features) algorithm based on ECG mixed feature to solve time-overhead problem of big data training. LOMF includes ECG signal preprocessing, sub-block division, and block training. By combining the MapReduce distributed computing framework and the secondary retrieval method based on the multi-dimensional feature space, LOMF is parallelized to improve recognition rate and computing efficiency at the same time. The experiment results show that, in the diversified scenarios, utilizing ECG mixed feature can return a higher recognition rate than the traditional ECG one-dimensional feature. Moveover, compared with the traditional linear discriminant analysis (LDA) and support vector machine (SVM) algorithms, the precision of LOMF increases by 7%-8%, which depends on the most competitive advantage of using LOMF. LOMF fits MapReduce parallel framework well so it is more effective than traditional algorithms, especially on diversified application scenarios (such as Internet, etc.) where the amount of data grows rapidly.

***Index Terms*—ECG, Biological recognition, ECG mixed feature extraction, Parallel computation, LDA.**

I. INTRODUCTION

In the last decade, the electrocardio activity record of the heart, i.e. ECG (Electrocardiograph) are proved effective on identity recognition [2]. In professional analytical research, ECG-based biological recognition technology are considered better than traditional ways for the following reasons. ECG exists in all living individuals [5] and manifests typical biological features. More importantly, ECG signal is difficult to counterfeit. So far, many different methods have been used for the identity recognition by ECG and have been reported in scientific literature [4], [7].

ECG biological recognition can always obtain a recognition rate as high as 99% where environmental and personal factors remain unchanged. However, a critical problem lies in the complexity of ECG signal which results in the limited application scenarios of conventional algorithm [4]. In other words, once the environmental and personal factors are changed, the conventional ECG recognition method can hardly gain the ideal effect.

In 2015, Antonio Fratini et al. further summarized the ECG biological recognition technology [4]. Several existing ECG identity recognition solutions are only applicable to some specific scenarios. When the cardinal number of crowds becomes larger and the environmental diversity becomes deepened, an universally applicable ECG recognition solution is not available. The challenge of maintaining the accuracy rate of ECG biological recognition in the case of diversified application scenarios and huge amount of crowds [10] was not considered in [4]. Thus, this paper solves above problems from two perspectives.

First, ECG feature extraction. Based on fiducial and non-fiducial features, this research extracts a multidimensional ECG mixed feature so as to incorporate as much ECG feature information as possible. Through this method, the integrity of ECG feature can enable ECG-based biological recognition to adapt to various scenarios.

Second, in order to tackle the problem resulted by huge human crowds, two steps need to be improved. The first step is compression and classification of ECG mixed feature by pattern recognition technology. The second step is to improve the execution efficiency of algorithm. We proposes a LOMF algorithm, and carries out a parallelization based on MapReduce computing framework to promote computing efficiency to a higher level while realize improvement of precision.

II. SYSTEM ARCHITECTURE

LOMF is an algorithm based on ECG mixed feature. The system flow is designed as follows, first conduct the detailed block analysis on ECG mixed feature, then execute the quadratic search based on multi-mode feature space [11], finally the recognition result is returned. The overall system flow diagram is shown in Fig. 1. The detailed operation is elaborated below:

Y. Miao and Y. Tian are with School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China (yimingmiao@hust.edu.cn, yuanwentian1997@gmail.com).

L. Peng is with Department of Integrative Systems Engineering, Ajou University, South Korea (auroraplm@ajou.ac.kr).

M. S. Hossain is with Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia (mshossain@ksu.edu.sa).

G. Muhammad is with Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia (ghulam@ksu.edu.sa).

Corresponding authors: M. Shamim Hossain (mshossain@ksu.edu.sa) and Limei Peng (auroraplm@ajou.ac.kr)

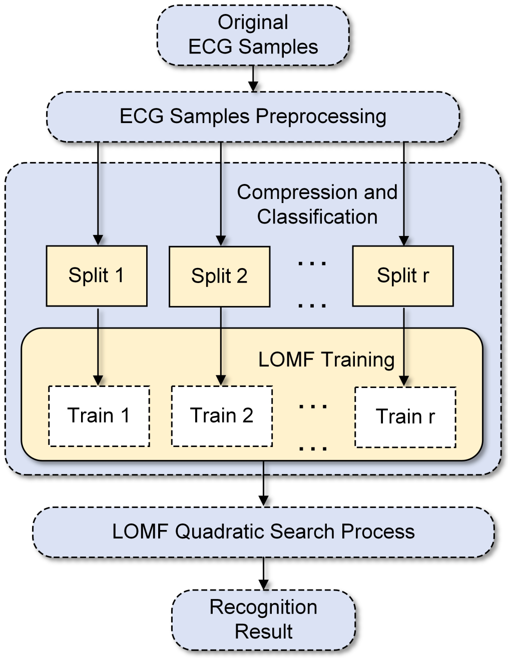


Fig. 1. System Flow Diagram.

*1) **ECG signal preprocessing process**:* Before implementing the LOMF-based distributed algorithm, since ECG signal contains a large number of clutter signals with very low signal-to-noise ratio, the preprocessing process are required for ECG multidimensional feature in order to eliminate the noise interference in ECG signal and enhance the accuracy of electrocardio identity recognition.

*2) **Compression and classification process**:* After obtaining the preprocessed ECG samples, the data redundancy is a big problem. Traditionally, a classification process should be followed immediately. However, different from traditional methods, we first carry out the split sub-block for ECG mixed feature and train them respectively, then we classify each split sub-block independently based on MapReduce framework, so as to finish the compression and classification of multiple ECG mixed features.

*3) **Quadratic search process**:* LOMF algorithm obtains multiple feature space after the sub-block training and classification. Based on the multidimensional feature space, LOMF conducts the quadratic search process. This paper will focus on the design and experiment on quadratic search process in Section V, and analyze the importance of quadratic search process of multidimensional feature space on LOMF in Section IV.

III. ECG Mixed Feature Extraction

ECG mixed feature involves fiducial and non-fiducial features [12]. The main principles are as follows. (1) ECG image feature is deemed as non-fiducial feature. (2) The relative position of PQRST three points is deemed as fiducial feature. (3) The time domain space of ECG signal is deemed as the periodic feature [1]. (4) ECG frequency spectrum is deemed as the variable feature of ECG signal. Before conducting the

TABLE I
Field Table for ECG Mixed Feature Extraction

QS, RS, RT, ⋯.
QS_duration, RS_duration, RT_duration, ⋯.
Shape[1], Shape[2], Shape[3], ⋯.
Spectrum[1], Spectrum[2], Spectrum[3], ⋯.

LOMF algorithm, all ECG samples should finish the unified feature description of ECG signal. The LPC-based frequency spectrum feature and ECG periodic signal feature are shown in Fig. 2(a) and Fig. 2(b) respectively. ECG mixed feature is a data set with 4 different fields, as shown in Table I. The QRS amplitude feature is placed in the first field. QRS time domain feature is stored in the second field. The ECG signal periodic morphological feature is stored in the third field. ECG frequency spectrum feature is stored in the fourth field. These four fields have different lengths.

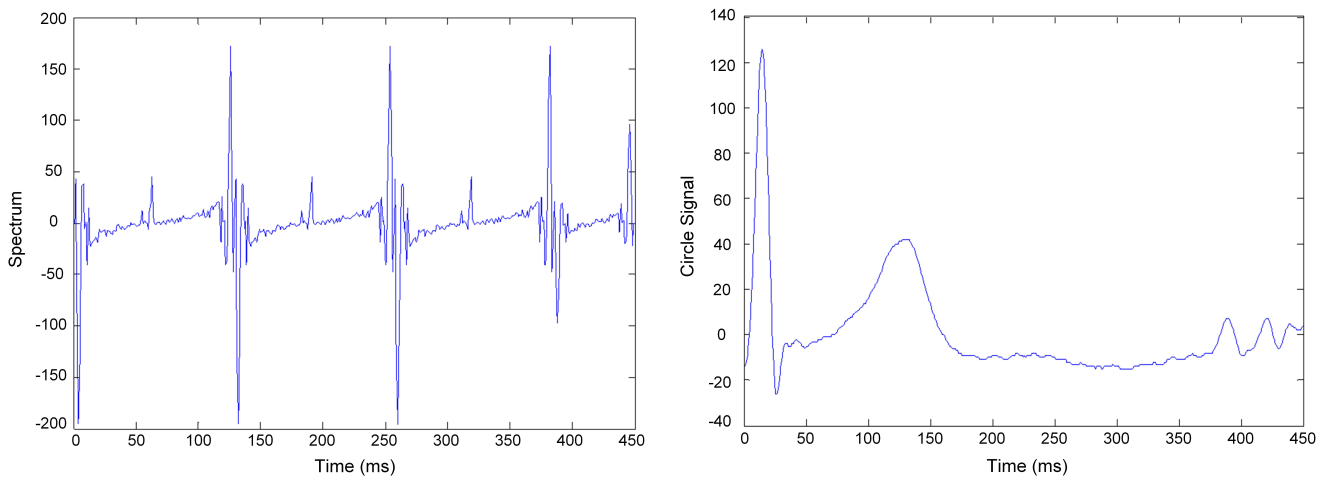
IV. Theoretical Model and Parallelization Implementation of LOMF Algorithm

A. Compression and Classification of ECG Mixed Feature

LOMF algorithm should conduct blocking first. However, before blocking, ECG feature has relatively large data size, involving the data of four dimensions, and the data in each dimension has its own length. Therefore, during searching and matching, the quadratic feature analysis should be made on ECG signal with mixed feature [9]. This paper adopts the algorithm integrating two-dimensional principal components analysis (2DPCA) and LDA [13] for data analysis, aiming to decrease the influence of the linearly independent feature on the recognition effect, and increase the influence of the linearly dependent feature on the recognition effect to the best.

Suppose that X shows a feature matrix in ECG mixed feature, SUM feature matrixes as the samples should be analyzed, among which some samples belong to the same individual while some are from different individuals. Firstly, 2DPCA algorithm computes the mean value of all samples, as shown in Formula 1. Then the difference value between each sample and mean value finishes the normalization processing. Y_i is the result output of normalization, and finally each sample is reset, as shown in Formula 2. Through the covariance diagonalization, each variance of Y_i column vector maintains the maximum, i.e. maintaining the greatest difference between Y_is. In other words, after the original data is converted into this base, the covariance is 0, and the field variance is the maximum. The covariance matrix computation is shown in Formula 3. Then, it computes the feature value and feature vector of G, and selects the corresponding feature vector of the former N feature values to constitute the feature space. The new coordinate space of the Sample Y_i after projection is shown in Formula 4, $Z_i \in R^{m\times t}$.

$$\overline{X} = \frac{1}{SUM}\sum_{i=1}^{SUM} X_i \tag{1}$$



(a) ECG Spectrum

(b) ECG Circle

Fig. 2. ECG Spectrum and Circle

$$Y_i = X_i - \overline{X} \tag{2}$$

$$G_i = \frac{1}{SUM}\sum_{i=1}^{SUM} Y_i^T Y_i \tag{3}$$

$$Z_i = X_i U, i = 1, 2, \cdots, SUM \tag{4}$$

Other dimension features of ECG mixed feature are compressed by the above 2DPCA algorithm as well, and obtain $Y(a)$, $Amplitude(b)$, $Duration(c)$ and $Z(d)$. Then the classification process of LDA algorithm is as follows.

step 1 Suppose that ECG signal features in four dimensions are $Y(a)$, $Amplitude(b)$, $Duration(c)$ and $Z(d)$, respectively.

step 2 Compute the intra-class dispersion matrix S_w and intra-class dispersion S_b of all training samples, as shown in Formula 5 and Formula 6. Thereinto, X represents ECG signal feature, $Y(a)$, $Amplitude(b)$, $Duration(c)$ and $Z(d)$, respectively.

step 3 According to the Fisher Criterion $J_w = argmax\frac{|W^T S_b W|}{|W^T S_w W|}$, four right feature space can be computed $W = [w_1, w_2, \cdots, w_t]$. The corresponding feature vector is shown in Formula 7.

step 4 ECG signal blocks in four dimensions can obtain the following four feature vectors and four feature space after LDA projection, as shown in Table II.

$$S_w = \sum^{N}\sum^{K} P_i (X_j^i - \overline{X^i})^T (X_j^i - \overline{X^i}) \tag{5}$$

$$S_b = \sum^{N}\sum^{K} P_i (\overline{X^i} - \overline{X})^T (\overline{X^i} - \overline{X}) \tag{6}$$

$$Y_i = X_i W \tag{7}$$

TABLE II
ECG MIXED FEATURE

ECG features	Feature Space	Feature Vector
$Y(a)$	$W_y = [w_1, w_2, \cdots, w_t]$	$Y(a)W_y$
$Amplitude(b)$	$W_{amp} = [w_1, w_2, \cdots, w_t]$	$Amplitude(b)W_{amp}$
$Duration(c)$	$W_{dur} = [w_1, w_2, \cdots, w_t]$	$Duration(c)W_{dur}$
$Z(d)$	$W_z = [w_1, w_2, \cdots, w_t]$	$Z(d)W_z$

B. Similarity Measurement of LOMF Algorithm

During the pattern recognition, the specific classifier is needed for the distance computation, and the nearest neighbor classifier is the most widely used classifier. However, in this paper, the similarity of the mixed feature can not be computed by the Euclidean distance, mainly because the feature measurement of ECG mixed feature is different. Therefore, we carry out the blocking LDA feature extraction of ECG mixed feature. For example, Y represents the feature of ECG image and reflects the feature of each pixel point. $Amplitude$ represents the amplitude of ECG waveform, in degree. $Duration$ represents the unit of ms. Thus, in the case of unifying LDA in the traditional method, the link between features will be ignored. To enable the relations among different variables to be independent of the measurement unit, this paper adopts the Mahalanobis distance as the similarity measurement distance between mixed features [14], as shown in Formula 8. In the formula, σ is the covariance matrix of ECG total sample. D is the Mahalanobis distance between the test sample X and the total sample. This paper obtains the feature space and feature vector in four dimensions, as shown in Formula 9 and Formula 10. In the formula, X represents the feature recombination of test sample.

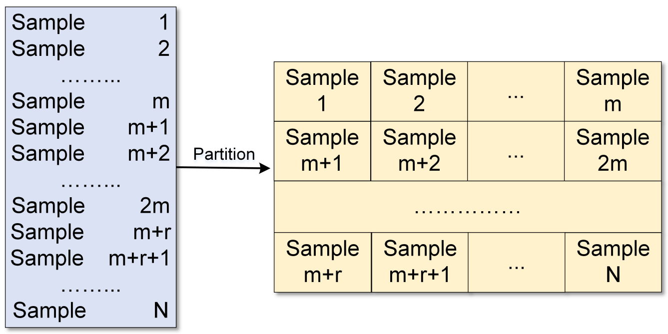


Fig. 3. LOMF Split

$$D^2 = \frac{(X-\vec{u})^T(X-\vec{u})}{\sigma} \tag{8}$$

$$X = [Y_x, Amplitude_x, Duration_x, Z_x] \tag{9}$$

$$\vec{u} = [Y, Amplitude, Duration, Z] \tag{10}$$

C. Process Design of LOMF Algorithm

One characteristic of LOMF algorithm proposed in this paper is based on MapReduce distributed framework [15]. LOMF algorithm firstly splits all sample data by the Mapper node numbers, obtains multiple sub-blocks, and then map each sub-block to distribute to different Mappers for the task of training and recognition. In fact, the sub-blocking in the Step 1 is the foundation for the support of incremental training, as shown in Fig. 3. The so-called incremental training refers to only training the new samples alone, rather than training the original samples, in the case of adding the new samples. LOMF block training process is as follows.

step 1 Each Sample represents ECG mixed feature data pair in four dimensions [16], with the total sample numbers of N, representing as $Sample[N]$. Firstly, N Samples are split into r pieces once, and each piece is composed of m samples. All samples are converted into $SubSamples_i = \{Sample_i, Sample_{i+1}, \cdots, Sample_{i+m}\}$, $i = \{1, 2, 3, \cdots, r\}$.

step 2 The blocking quadratic feature extraction is made for $SubSample[i]$, for the purpose removing the redundant data of ECG mixed feature in four dimensions, and generate four pairs of feature space [$FeatureSpaceA$, $FeatureSpaceD$, $FeatureSpaceS$, $FeatureSpaceZ$] and feature vector set [$Amplitude\ vector, Duration\ vector, Spectrum\ vector, Shape\ vector$]. The ECG mixed feature block processing process by 2DPCA in combination with LDA is shown in Fig. 4.

step 3 The measurement difference between samples is eliminated by the Mahalanobis distance, and each $Result[j] = min\{D_i^j, D_{i+1}^j, \cdots, D_{i+m}^j, \}$. Also, the searching is finished by the quadratic search method based on dual feature space, as shown in Fig. 5.

step 4 The quadratic search should be subject to two search processes. The first one is searching and screening between the similar feature space. The second one is searching and screening between the dissimilar feature space. Moreover, if gaining the result finally, the statistical probability should be the maximum, i.e. finding the sample point occurring for the most times. Firstly, after the projection for the first time, the Sample Test has r projection matrixes different from the traditional method, and each one is a four-dimensional feature vector, to find the sample with the minimum Mahalanobis distance in each projection space, and constitute the result set $Result[m]$. According to the similarity maximum principle of test sample and template sample, the probability of some result from projection is the maximum in all sample space, and the sample can be deemed as the recognition result.

step 5 Incremental training. The quadratic search method based on dual feature space in this paper is firstly to conduct the block training before the feature extraction of the sample. In a similar way, while adding the new samples, the ECG mixed feature is extracted, and LDA quadratic feature analysis is made in the block. Then the newly generated feature space and feature vector are added into the existing $Space[A, D, S, Z]$. The incremental training process is finished.

The feature space of the traditional pattern recognition algorithm [17] is always a extra-large space plane, but the LOMF algorithm is multiple mutually independent feature space. The dissimilar feature space is not influenced each other, and the similar feature space is associated. The quadratic search refers to two search processes of LOMF algorithm as needed during the pattern recognition. The first one is searching and screening between the similar feature space. The second one is searching and screening between the dissimilar feature space. Moreover, the statistical probability should be the maximum finally, i.e. finding the sample point occurring for the most times, to gain the recognition result. The traditional mode is the traversal visit, and finds the extreme value sample point, to gain the result. It is the greatest difference between LOMF algorithm and traditional distributed algorithm.

D. Parallelization Implementation of LOMF Algorithm

This part is aimed at solving the parallelization problem of LOMF algorithm. The traditional distributed algorithm is training based on the overall sample [3], so that the sample-related feature extraction algorithms such as LDA and principal components analysis (PCA) do not support the incremental training. On the contrary, LOMF algorithm proposed in this paper firstly blocks, with mutual independence between pieces. In addition, the traditional algorithm is to compute a public feature space and then search a minimum Mahalanobis distance. However, our LOMF algorithm conducts the sub-block training, sub-block recognition, and comprehensive sequencing of samples, to find the result with the most frequency spectrums [6], i.e. the maximum probability. It is the greatest difference between LOMF algorithm and traditional algorithm.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2771220, IEEE Access

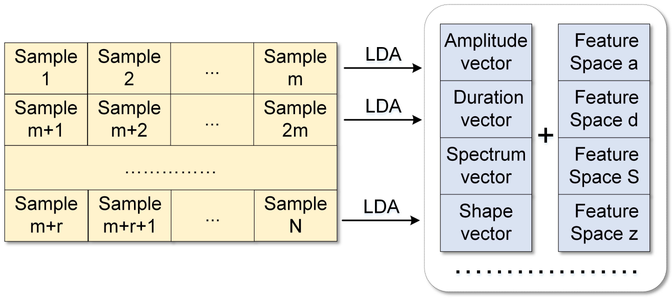


Fig. 4. LOMF Block Processing

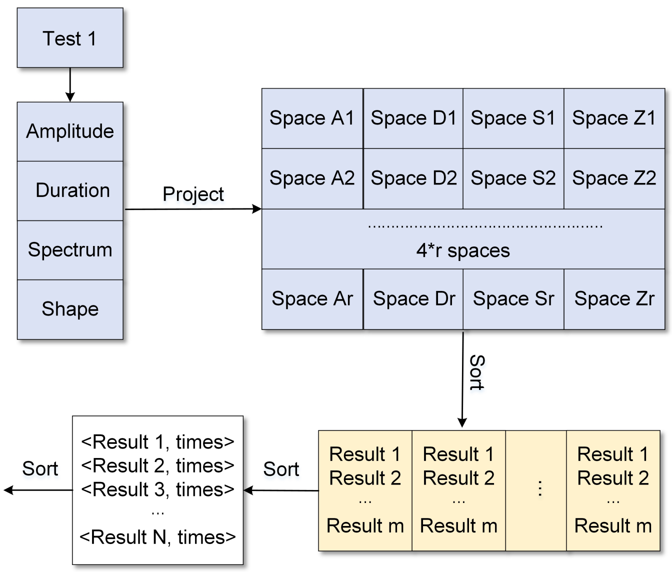


Fig. 5. Search Process of LOMF

Based on the distributed computing platform as shown in Fig. 6, this paper implements the parallelization of LOMF algorithm.

- Find a result closest to Test sample from the database. Firstly, it is required to conduct the feature extraction of Test sample. Test is distributed to the node, and each node stores a private small feature space, called as feature space split. The advantage is to refine the feature space, and also assign the task of feature extraction to all nodes for computation. Moreover, while adding the new data, the algorithm directly trains and computes the new data and gains new feature space split, and then assigns to any node randomly.
- Improve the numbers of Test features. The traditional distributed algorithm has only one Test feature. In this paper, r LOMF parallelization algorithms are proposed, and also maintain consistency with the numbers of feature space split. Before finding the optimum solution, two searches are needed. The first one is to find the nearest Mahalanobis distance in accordance with the nearest neighbor classifier. The second one is to the statistics of maximum probability because LOMF algorithm introduces ECG mixed feature.

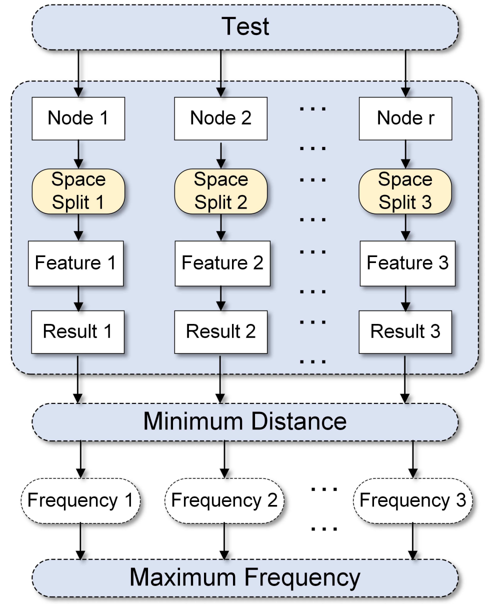


Fig. 6. Parallelization Search of LOMF Algorithm

V. Experimental Process and Result

A. Experimental Parameter Setting

The test platform is 4 computers, and the computer configuration is Intel(R) Xeon(R) i7 CPU, 32-core 16-thread, 128g memory, 16T hard disk, operating system of Ubuntu12.04, and running Hadoop version of 0.23.11. The experiment hardware environment is shown in Table III [8].

TABLE III
Hardware Environment

Parameters	Configuration
CPU	AMD FX-4100 32-core 3.60GHz
memory	128G DDR3
OS	Ubuntu 12.4
OpenCV version	2.4.9
Hadoop version	Hadoop 0.23.11
MapReduce	Yarn (MapReduce v2)

The test data is MIT-BH database, involving 100 experimental samples in total. Each sample has 2*100 ECG signal files, totaling to 20,000 experimental samples. In consideration to the diversity of data sources, this paper adds 100 pathological ECG samples. Each ECG sample includes 10,000 signal values, approximately ECG signal collected by a person within 10s.

B. Experimental Result

In this part, we will test the LOMF algorithm based on ECG mixed feature, involving time overhead and recognition

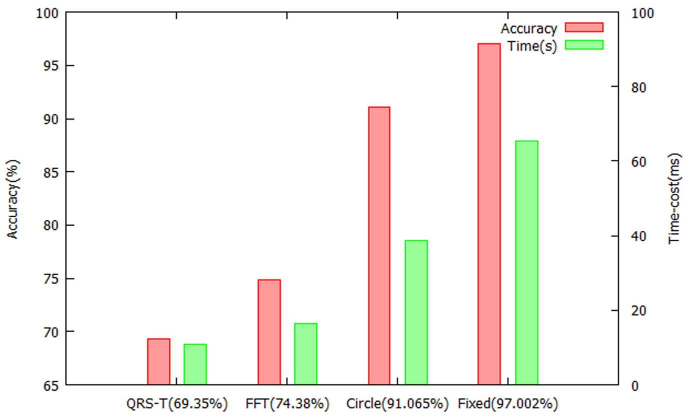


Fig. 7. Contrast of Various Feature Recognition Rates of ECG

accuracy rate of algorithm. During the experimental process, the following aspects are tested. 1. In view of ECG feature selection problem, this paper contrasts the ECG mixed feature, fiducial feature and non-fiducial feature. 2. For the recognition rate of algorithm, this paper contrasts three methods of PCA, LDA, and 2DPCA combining with LDA(2DPCA+LDA). 3. For the training overhead of algorithm, this paper contrasts the LOMF parallelization algorithm and traditional distributed recognition algorithm.

*1) **Contrast of ECG mixed feature performances**:* From Fig. 7, the accuracy rate and time consumption of QRS-T feature is the minimum. Although ECG feature has good invariance, the recognition rate is only about 70%, and the effect is not ideal. The reason analysis is made in the case of a lot of samples, and ECG signal of each individual is collected at different times and contains very large volatility. It can be seen that, the single mode feature is difficult to represent the individuality of ECG in the case of a lot of samples.

TABLE IV
CONTRAST DATA OF VARIOUS FEATURE PERFORMANCES OF ECG

Type	Correct	Incorrect	Proportion	Execution time
QRS-T	13975	6175	0.693548	10.862(s)
FFT	15080	5070	0.748387	16.432(s)
Circle	18066	2 084	0.91065	38.701(s)
Fixed features	19546	604	0.97002	65.389(s)

The recognition result of further detailed analysis is shown in Table IV. It reveals that, QRS-T feature gains 69% of recognition rate from 20,000 samples, but 6,175 samples have the recognition error. The recognition rate of ECG feature based on FFT frequency spectrum is up to about 75%, which has the definite improvement than QRS-T but can not yet meet the ideal standard. The recognition result independently based on ECG periodic pattern is up to 91% of recognition rate, which can be said as one type with the highest recognition rate. LDA algorithm (type of fixed features) based on ECG mixed feature proposed in this paper is up to 97.002% of recognition rate, and only has the error in the recognition of 600 abnormal ECG samples.

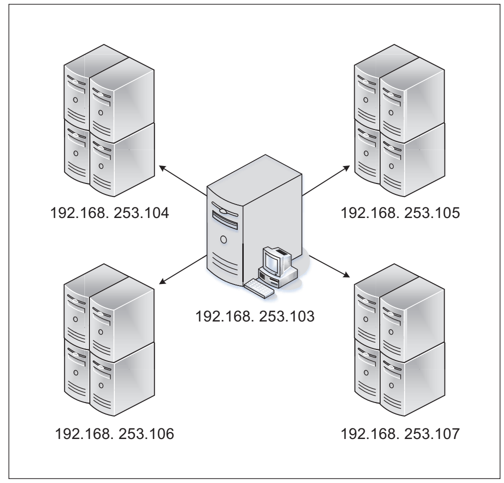


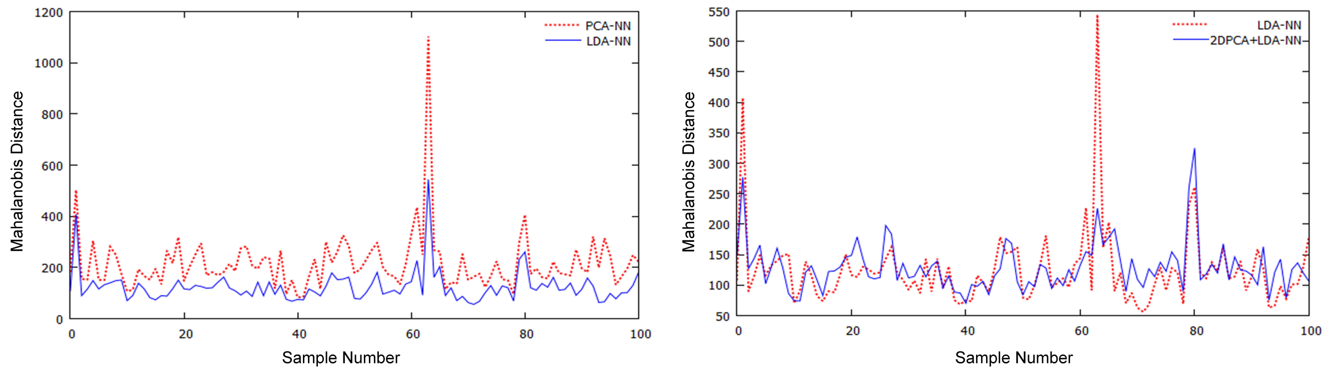
Fig. 9. Distributed System Deployment Diagram

*2) **Classification performance test of LOMF**:* The experiment contrasts the accuracy of different classification algorithms by recognizing the Mahalanobis distance of 100 samples, involving PCA-NN (PCA Nearest Neighbor), LDA-NN (LDA Nearest Neighbor), and 2DPCA+LDA-NN (2DPCA combining with LDA Nearest Neighbor), as shown in Fig. 8. The result of contrast between PCA-NN and LDA-NN algorithms is shown in Fig. 8(a). It can be seen that, during the test of all 100 samples, the performance of LDA-NN algorithm is superior to PCA-NN. Fig. 8(b) shows the performance contrast between LOMF algorithm (i.e. 2DPCA+LDA-NN algorithm) and LDA-NN algorithm. 2DPCA+LDA-NN algorithm is significantly superior to LDA-NN algorithm in two samples of No. 3 and No. 65. Moreover, during the test of other samples, as a whole, the difference between both is not large. But, in No. 65 pathological sample, it can be seen that 2DPCA+LDA still keeps the minimum Mahalanobis distance, i.e. high reliability.

*3) **Parallelization performance test of LOMF**:* In the case of big data of crowds, the experiment tests the LOMF distributed recognition algorithm and conventional LDA algorithm, and analyzes the experimental result. 39,100 ECG samples are selected from MIT-physionet database to incorporate into the training set, involving 1,000 pathological samples and 1,000 external samples, to ensure the diversity of data. There are 6,000 test samples.

It selects 5 servers and deploys 5 computers respectively based on the above hardware parameters. One host is regarded as the master node, and other 4 hosts are regarded as the computational nodes. All test data exist in HDFS (Hadoop Distributed File System), and MapReduce is the second generation of parallel computing framework, called as Yarn. All computational nodes are deployed in the same local area network. The system deployment is shown in Fig. 9.

Fig. 10 shows that the traditional training mode and LOMF parallel training mode respectively train different amount



(a) Performance Contrast between PCA and LDA

(b) Performance Contrast between LDA and 2DPCA+LDA

Fig. 8. Performance Contrast between PCA, LDA and 2DPCA+LDA

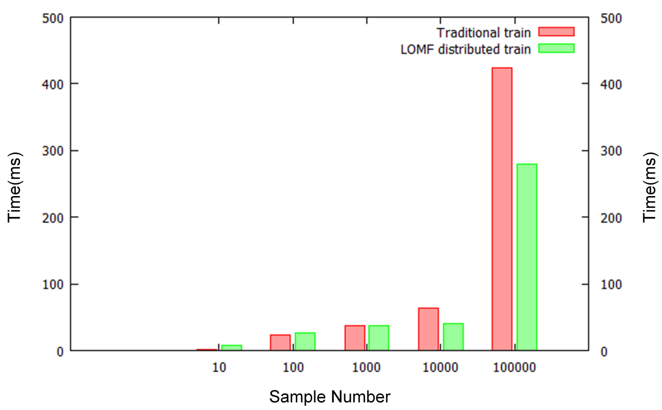


Fig. 10. Contrast of Time Overhead

of samples, and finally gain the result of contrast of time overhead. In the case of the order of magnitude less than 100, the time overhead of LOMF algorithm is higher than that of traditional stand-alone mode. The analysis suggests that, the MapReduce process of LOMF needs the split of sample and the block training before searching, and the overhead of LOMF algorithm is higher than that of conventional method when the data set is smaller. However, with the increase of sample quantity to the level of 10,000, the block training of LOMF multidimensional feature manifests huge advantages. After the parallelization of LOMF algorithm, the samples are split and assigned to multiple computational nodes for the parallel training, so the training process of ECG mixed feature under massive data set is faster, and the precision of LOMF algorithm has higher recognition rate than the traditional single-mode ECG feature.

VI. Conclusion

This paper focuses on keeping the robustness of recognition by utilizing various features of ECG in the case of diversified application scenarios and huge amount of crowds, which a remained problems of ECG identity recognition technology proposed in 2015. As a result, improvements have been made in two aspects. As to selection of ECG feature, we puts forward a method of integrating ECG fiducial and non-fiducial mixed features as the main feature points of ECG. Additionally, we implements the parallelization of LOMF algorithm, which conducts classification process based on LDA algorithm sub-block, and quadratic search process based on multimode feature space. The experiment indicates that LOMF proposed in our paper performs much better than traditional recognition algorithms. By balancing the influence of different feature factors on recognition result, LOMF can improve the recognition accuracy while avoiding the high time overhead. In the future, we will continue to work on new ECG features and pay attention to collect more ECG data sets, in order to optimize the algorithm in a larger data set to achieve better results and improve the feasibility of the algorithm.

Acknowledgement

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia, for funding this work through the research group project no. RGP-228

References

[1] M. Chen, Y. Hao , K. Hwang, L. Wang, L. Wang, "Disease Prediction by Machine Learning over Big Healthcare Data", *IEEE Access*, Vol. 5, No. 1, pp. 8869-8879, 2017.

[2] R. Hoekema, G.J. Uijen, A. Oosterom, "Geometrical aspects of the interindividual variability of multilead ECG recordings", *IEEE Trans Biomed Eng*, vol.48, no.5, pp.551 559, 2001.

[3] M. Chen, Y. Qian, Y. Hao, Y. Li, J. Song, "Data-Driven Computing and Caching in 5G Networks: Architecture and Delay Analysis", *IEEE Wireless Communications*, Vol. 25, No. 1, 2018

[4] A. Fratini, M. Sansone, P. Bifulco, et al., "Individual Identification Via Electrocardiogram Analysis", *Biomedical Engineering Online*, vol.78, no.14, pp.808-812, 2015.

[5] J. Wan, C. Zou, S. Ullah, C. Lai, M. Zhou and X. Wang, "Cloud-enabled Wireless Body Area Networks for Pervasive Healthcare", *IEEE Network*, vol. 27, no. 5, pp. 56 – 61, 2013.

[6] M. Chen, J. Yang , X. Zhu, X. Wang, M. Liu, J. Song, "Smart Home 2.0: Innovative Smart Home System Powered by Botanical IoT and Emotion Detection", *Mobile Networks and Applications*, DOI 10.1007/s11036-017-0866-1, 2017.

[7] D. Sung, J. Kim, M. Koh, K. Park, "ECG authentication in post-exercise situation", *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp.446-449, 2017.

[8] M. Chen, Y. Miao, Y. Hao, K. Hwang, "Narrow Band Internet of Things", *IEEE Access*, DOI: 10.1109/ACCESS.2017.2751586, 2017.

[9] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng and Q. Liang, "Data Fusion Method based on Mutual Dimensionless", *IEEE/ASME Transactions on Mechatronics*, DOI: 10.1109/TMECH.2017.2759791, 2017.

[10] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, C. Youn, "Wearable 2.0: Enable Human-Cloud Integration in Next Generation Healthcare System", *IEEE Communications*, vol.55, no.1, pp.54 61, 2017.

[11] Min Chen, P. Zhou, G. Fortino, "Emotion Communication System", *IEEE Access*, vol.5, pp.326-337, 2017.

[12] E. K. Zaghouani, A. Benzina, R. Attia, "ECG based authentication for e-healthcare systems: Towards a secured ECG features transmission", *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp.1777-1783, 2017.

[13] M. Chen, Y. Hao, L. Hu, K. Huang, V. Lau, "Green and Mobility-aware Caching in 5G Networks", *IEEE Trans. Wireless Communications*, DOI: 10.1109/TWC.2017.2760830, 2017.

[14] J. Wan, S. Tang, D. Li, S. Wang, C. Liu, H. Abbas and A. Vasilakos, "A Manufacturing Big Data Solution for Active Preventive Maintenance", *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2039 – 2047, 2017.

[15] M. Chen, J. Yang, Y. Hao, S. Mao, K. Hwang, "A 5G Cognitive System for Healthcare", *Big Data and Cognitive Computing*, vol.1, no.1, DOI:10.3390/bdcc1010002, 2017.

[16] S. J. Kang, S. Y. Lee; H. I. Cho, H. Park, "ECG Authentication System Design Based on Signal Analysis in Mobile and Wearable Devices", *IEEE Signal Processing Letters*, vol.23, no.6, pp.805-808, 2016.

[17] M. Chen, X. Shi, Y. Zhang, D. Wu, M. Guizani, "Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network", *IEEE Trans. Big Data*, DOI:10.1109/TBDATA.2017.2717439, 2017.