A Practical Cross-Domain ECG Biometric Identification Method

Huan Sun
school of electronic information engineering, Beijing Jiaotong University
Beijing, China
17120115@bjtu.edu.cn

Yuchun Guo
school of electronic information engineering, Beijing Jiaotong University
Beijing, China
ychguo@bjtu.edu.cn

Bin Chen
school of electronic information engineering, Beijing Jiaotong University
Beijing, China
16120044@bjtu.edu.cn

Yishuai Chen
school of electronic information engineering, Beijing Jiaotong University
Beijing, China
yschen@bjtu.edu.cn

***Abstract*—With the boosting of application, biometric identification with fingerprint or face-image suffers from forging attacks. The electrocardiogram (ECG) as a kind of biometric identification is of higher resistance against such attacks and receives research attention. The state-of-art method has recognition accuracy of about 95%. However, we find that the accuracy will degrade dramatically to 40% if it is applied in a practical context when a significant interval between training period and applying period. The critical reasons for this failure are as follows: 1) the extracted features are temporal sensitive due to that continuous samples being used in training and testing period in the existing schemes; 2) the features highly relevant to the performance are not utilized sufficiently in CNN classifier; 3) the optimal parameter setting for obtaining enough effective samples for individuals has not been investigated. This paper targets on proposing a practical cross-domain ECG biometric identification method to solve the above problems. Specifically, we: 1) determine the best parameters of the non-fiducial random sampling method to obtain enough effective samples for individuals; 2) propose a method to extract deep features across time, frequency and energy domain which are temporal insensitive and individual distinguishable; 3) introduce a channel attention module into the CNN and modify its activation function to optimize the recognition performance. We validate our method on PTBDB and ECG-ID databases. Experiments show that the identification accuracy reaches 56.93% and 85.94% respectively, with an improvement of 41.5% and 20.7% over the existing method.**

***Index Terms*—ECG, identification, interval, MFCC, Attention**

I. Introduction

Fingerprints or face-images are widely used as biometric identifications. But such biometric identifications are facing forgery attacks. Recently, some other kinds of human signs with higher resistance against faking attacks are being investigated for biometric identification. The electrocardiogram (ECG) is among such signs. ECG refers to the time series of various forms of potential changes of heartbeats from the body surface induced by electrocardiograph in each cardiac cycle by pacing points, atria and ventricles, accompanied by changes in bioelectricity. Biel et al. first proposed the application of ECG signal to biometric identification [1]. Pathoumvanh et al. and Zheng et al. considered the ECG identification as a template matching problem in terms of vector distance [6], [8], [15]. Some researchers used convolutional neural network (CNN) to improve the performance of ECG identification, it had been proved that CNN had the best performance in ECG identification among previous works [2], [5], [10], [13].

However, we find that existing models face significant intervals for existing models to be applied in practical application. Existing work neglected that for a biometric identification to be practical, that biometric feature collection and authentication is often departed in time for a long range whereas ECG in a long-time scale is highly dynamic. In the above work, the samples of ECG records used in training and testing periods for designing a model are extracted in continuous time, even some ECG record as short as no more than ten seconds are used in model training. With experiments on PTBDB and ECG-ID, we found that the existing models degraded in performance dramatically in practical context from 96.6% and 94.89% to 40.24% and 71.2%. This is due to the following problems. 1) The extracted features are sensitive to the interval between the training and applying period. 2) The features highly relevant to the performance are not utilized sufficiently in CNN classifier. 3) The optimal parameter setting for obtaining enough effective samples for each individual has not been investigated.

In this paper we propose a cross-domain ECG biometric identification method oriented for practical context. Our contributions as follows: 1) we evaluate the parameter performance of the non-fiducial random sampling method to obtain enough effective samples for each individual; 2) we propose a method to extract individual-distinguishable feature insensitive to timespan between feature collection and recognition period via utilizing deep features across time, frequency and energy domain; 3) we introduce a channel attention module into the CNN and modify the activation function to optimize the recognition performance. Verification results on PTBDB and ECG-ID show that the recognition accuracy of our model can achieve 56.93% and 85.94%.

The rest of this paper is organized as follows. Section 2 reviews previous works. Section 3 elaborates our method in detail. Section 4 verifies our method on real databases. Finally, Section 5 concludes this paper.

978-1-7281-0962-6/19/$31.00

II. Related Work

A. Sample generation

Sample generation methods can be roughly divided into three categories: fiducial fixed-length segmentation, non-fiducial fixed-length segmentation and blind segmentation. The first one is to find the location of R peak first, then take a part of the fixed-length sequence before R peak as the training set and a part of the fixed-length sequence after R peak as the testing set [14]. Non-fiducial fixed-length segmentation first calculates the length of RR interval. The length of RR interval with the largest number of occurrences is defined as the length of the segmented sample, to obtain samples with complete beats [7], [9]. The number of training samples and testing samples acquired by the two methods depends heavily on the length of ECG records, which is not suitable for machine learning and deep learning. To solve the limitation, some researchers proposed blind segmentation [6]. Blind segmentation randomly selects a position contained in the period as the starting point, and the fixed-length sequence is truncated the sample. Blind segmentation has no special requirement for the starting position, so it can obtain a large number of samples to meet the needs of training in neural network.

B. Feature extraction

For ECG signal feature extraction, almost all of the related methods have been applied, individually or combinedly. These methods can be divided into two categories according to the extent of time dependence. The first one, which highly depends on time, includes autocorrelation (AC), short-time Fourier transform (SIFT) and discrete wavelet transform (DWT). These methods are not suitable for feature extraction from sequences with obvious fluctuations over time. The other one, which on the contrary slightly depends on time, includes Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Meier Frequency Cepstrum Coefficients (MFCC) and Statistical Feature Computing (SFC). To the best of our knowledge, no one has integrated these four time-independent feature extraction methods simultaneously, especially combining across different domains.

C. Classification

Classifier is a general term that functions by classifying samples in data mining, including decision tree, logical regression, naive Bayesian, neural network and other algorithms. To optimize the performance of classifier, some researchers proposed attention module to improve image classification performance [4], [11]. Recently, several studies have proved that introducing attention mechanism in network structure can improve the ability of feature expression of network model. Attentions can not only tell the network model what to pay attention to, but also enhance the representation of specific areas. Attention model in deep learning actually simulates the attention model of the human brain. For example, when we look at a picture, although we can see the whole picture, only a small part of it is focused on, when we look deeply and carefully. At this moment, the human brain mainly focuses on a certain proportion of the object. However, due to that attention to different parts of the observed object is not balanced at the same moment, weights for different channels are assigned the same in commonly used classifier. But in fact, there should be different weights in classifier. After introducing attention module, better recognition accuracy can be obtained even using ordinary neural network.

III. Method

Considering time domain, frequency domain and energy domain, this paper proposes a practical cross-domain ECG biometric identification method, including improved fixed-length sample acquisition, cross-domain feature extraction and optimized classifier. The complete framework of the proposed method is shown in Fig 1.

A. Dataset

a) PTB Diagnostic ECG Database: PTB database is a digitized ECG data compiled by the German National Institute of Metrology. It is mainly collected for research, algorithm benchmarking and teaching purposes. The database contains 290 individuals, totaling 549 records. Some individuals have one record, others have two or more records (up to 5). Especially, more than 50 individuals have 2 or more records, this makes it possible for us to obtain samples with distinct intervals of two periods. The database contains 81 women and 209 men, with age ranging from 17 to 87 years, and the average age is 57.2 years. Further, the average age for male is 55.5 years and that for female is 61.6 years. Among 290 individuals, 238 have heart disease, and the remaining 52 are healthy. The duration of each record varies from a dozen seconds to two minutes. Each record contains 15 leads and Lead I. The raw ECG signals are rather noisy and contain both high and low frequency noise components, so PTBDB needs some filtering process.

b) ECG-ID Database [3]: Database ECG-ID was created and contributed by Tatiana Lugovaya, as well, it was used in his masters thesis. This database contains 310 ECG records of 90 people (Person01-Person90). Each record consists of three files: .atr, .dat, .hea. .atr represents an annotation file, .dat represents the data storing Lead I ECG signal, and recorded for 20 seconds, and digitized at 500 Hz with 12-bit resolution over a nominal 10 mV range, and .hea contains some information like gender, age and recording date. The records were collected from volunteers (44 men and 46 women aging from 13 to 75 who were students, colleagues, and friends of the authors). The number of records corresponding to each person varies from 2 (collected during one day) to 20 (collected respectively over 6 months). Except Person 33, every volunteer has at least 2 records. The same as the operation for PTBDB, we can obtain samples with distinct intervals of two periods. Although each record includes both raw and filtered signals, we need filter the data further.

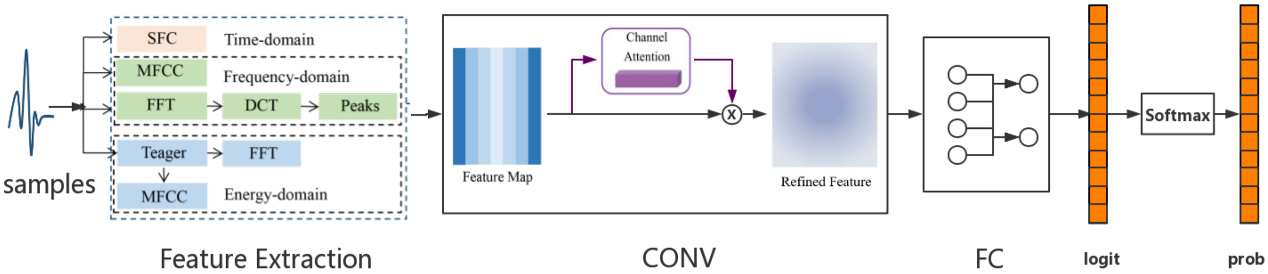
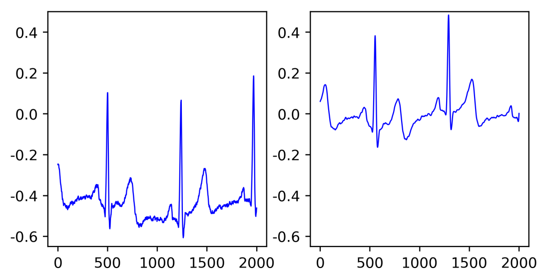
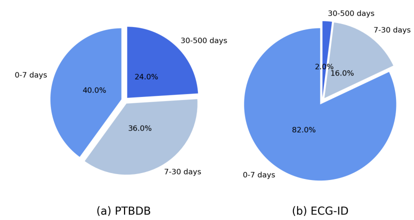


Fig. 1. The block diagram of the proposed method. CONV: Convolutional layer. FC: Fully connected layers.



(a) Raw ECG signal. (b) Filtered ECG signal.

Fig. 2. Waves Before And After Preprocessing.



(a) PTBDB (b) ECG-ID

Fig. 3. The Interval Distribution on Different Databases.

B. preprocessing

The original ECG signal is obtained from electronic acquisition equipment, so there will be three kinds of noise: power frequency interference, random noise and baseline drift, which will seriously interfere with the classifier's judgment. We design a filter bank including Butterworth filter and IIR filter for 50HZ power frequency interference, random noise and baseline drift of electronic components affected by temperature in acquisition equipment. After filtering, we get a relatively pure ECG signal, this process makes feature extraction and classification more accurate. Fig 2 (a) and Fig 2 (b) show the original ECG signal and the filtered ECG signal, respectively.

C. Sample acquisition

The method of sample acquisition proposed in this paper is improved non-fiducial fixed-length sample acquisition with obvious intervals. Firstly, there should be a obvious interval between training period and testing period. The interval distribution on databases PTBDB and ECG-ID is shown in Fig 3.

Then, a sample can be obtained by randomly placing the starting position of the sample at each given period and retrieving a sequence with fixed length in order. How many samples are needed, how many starting points are placed, that is, the number of starting points corresponds to the number of samples. This method random selection makes for the selected training set or testing set have similar statistical characteristics with complete interval ECG records. The following expression represents a sample obtained,

$$X_s^i = R_s^j \left[P_i : E_i \right] \tag{1}$$

$$j = \begin{cases} 0, & \text{record } = \text{visit0} \\ 1, & \text{record } = \text{visit1} \end{cases} \tag{2}$$

where X_s^i denotes the first sample x of the acquired $s-th$ individual, denotes the j record R of the selected $s-th$ individual, denotes the starting position of the first sample, denotes the terminating position of the first sample.

$$E_i = P_i + t * f \tag{3}$$

where, t denotes the duration of a single sample and f denotes the sampling frequency.

D. Feature Extraction

The feature extraction method proposed in this paper is based on time domain, energy domain and frequency domain, which is divided into three channels for feature extraction, and then combined.

Time domain. ECG signal is a non-stationary periodic signal, statistical characteristics in time domain represent the macroscopic description of signals over a period of time. We selecte the characteristics including mean, standard deviation, kurtosis and skewness to represent the samples in time domain.

Frequency domain. In this domain, we obtain two features using Meier Frequency Cepstrum Coefficient (MFCC) , fast Fourier transform (FFT) and Discrete Cosine Transform (D-CT). The heart structure can be displayed in the envelope of the short-term power spectrum of ECG signal, and MFCC is

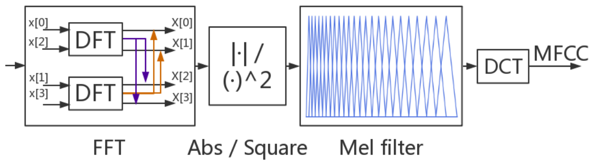


Fig. 4. MFCC Dynamic Feature Extraction Process.

a characteristic of accurately describing this envelope. MFCC (Mel Frequency Cepstral Coefficents) is a feature widely used in automatic speech and speaker recognition. It was first proposed by Davis and Mermelstein in 1980. ECG signal and speech signal have similar generating mechanism, so we apply MFCC to ECG signal feature extraction. The dynamic feature extraction process in MFCC is shown in Fig 4.

MFCC method is proposed for speech signal identification. The extracted features cannot match the ECG signals perfectly. Therefore, we add a group of features which are extracted directly by FFT and DCT cascade without Mel filter. Fourier transform is defined as follows: Let $f(t)$ be the time domain expression of a sample signal and the Fourier transform expression will be:

$$\mathrm{F}(\omega) = \mathrm{F}[f(t)] = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}dt \tag{4}$$

Then, discrete cosine transform (DCT) is performed according to the spectrum of fast Fourier transform (FFT), and a set of peaks is extracted from the sequence obtained by discrete cosine transform (DCT), which is the feature of the second dimension. The definition of one-dimensional discrete cosine transform is as follows: Let $\{f(x)|x = 0, 1, N-1\}$ be a discrete signal sequence, the discrete cosine transform equation is as follow,

$$\mathrm{F(u)} = C(u)\sqrt{\frac{2}{N}}\sum_{x=0}^{N-1} f(x)\cos\frac{(2x+1)u\pi}{2N} \tag{5}$$

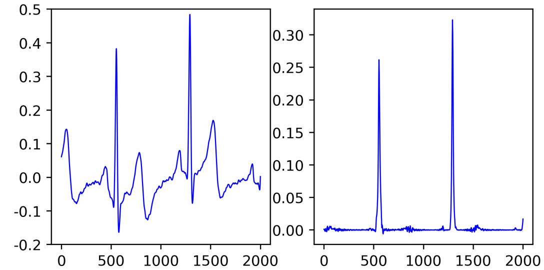
$$(u, x = 0, 1, 2, \ldots\ldots N-1)$$

where C(u) is the kernel function.

Energy domain. After filtering in the pre-processing stage, we get ECG signals enabling to eliminate baseline drift and low-frequency noise, but it is still difficult to extract key features based on such signals. Therefore, we introduce Teager energy operator to eliminate zero-mean noise. Teager energy operator provides a measure for the energy of ECG signal to some extent, representing the modulation state of a single formant energy. In the discrete domain, the Teager energy operator of signal $s(t)$ can be expressed as follow,

$$\psi_D(s) = s^2(n) - s(n+1)s(n-1) \tag{6}$$

where represents discrete Teager energy operators, and n denotes time.



(a) Filtered ECG signal. (b) ECG signal after Teager.

Fig. 5. Teager energy operator.

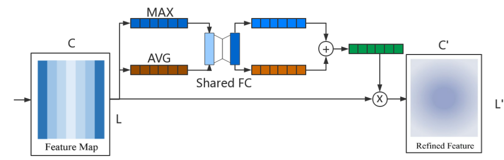


Fig. 6. Channel Attention Module. MAX: Maxpooling. AVG: Average-pooling. Shared FC: Shared fully connected layers.

As shown in Fig 5 (a) and Fig 5 (b), by transforming the energy operator, we can get the FFT spectrum of eliminating spurious noise.

E. Channel attention module

To solve the limitation of CNN in learning channel information, we introduce channel attention module (CAM) to improve learning ability of CNN before classification. As is shown in Fig 6, the size of feature map is $C * L$, the size of refined feature is $C * L$. Similar to previous studies [6], this CAM uses max-pooling (MAX), average-pooling (AVG), and a shared full-connected (FC) module with one hidden layer. The number of neurons for the hidden layer is set to C/r, where r is the reduction ratio. Certainly, we set the number of output neurons as C. Then, we add the two tensors and multiply the new tensor with feature map. Finally, we obtain refined feature. We define the CAM as follow,

$$F_d = F * (\sigma(MLP(AVG(F) + \sigma(MLP(MAX(F))) \tag{7}$$

where F denotes feature map, and F_d denotes refined feature. The improvement lies in that we replace ReLU with sigmoid as the activation function. The experimental results show that this increases the accuracy by one percentage point.

IV. Experimental Results and Discussion

In order to prove the effectiveness of the proposed method when there is an adequately long the interval between training period and testing period, we validate it on PTB and ECG-ID respectively.

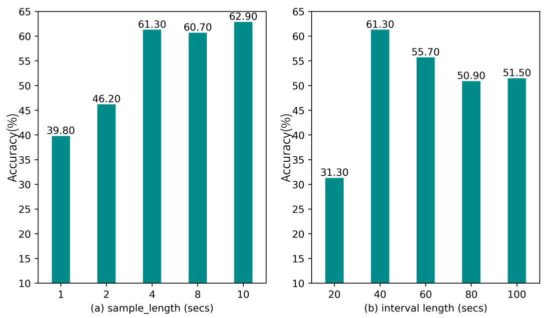


Fig. 7. Recognition accuracy under different sampling settings. (a): Recognition accuracy with different sample lengths. (b): Recognition accuracy with different lengths of training or testing period.

A. Setup

The experimental settings include sample acquisition settings and parameter settings of the network model. Next, they are introduced in detail.

Sample length. Sample length will affect the results of frequency conversion, deficiency in sample length will lead to spectrum leakage, while the contrary will need additional windows for further processing, so we select a single sample length of the following five values: 1, 2, 4, 8, 16, and other parameters remain unchanged. Fig 7(a) is the recognition rates corresponding to different time lengths with the best settings over 10 subjects. From the experimental results, it can be found that the recognition accuracies with sample length set to 4, 8 and 10 seconds are all relatively satisfying, whereas samples with 4 seconds as the length is the final choice, considering both recognition accuracy and computational complexity.

Period length. In order to maintain the consistency of the overall distribution of the training dataset and the testing dataset, the training period and the testing period should be set to have the same length. Taking the training interval as an example, the long period will increase the processing complexity and need to be further divided, while the short period will not be able to express the whole record well. Therefore, the period length is selected from the following values: 20, 40, 80, 100, and 150. As mentioned above, the sample length adopted in our experiment is 4 seconds, with other parameters are unchanged. As is displayed in Fig 7(b), the highest accuracy appears when the period length is 40 seconds, which becomes the final setting in our experiment.

Other Parameters. The interval length refers to the interval length between the training interval and the testing interval. Influenced by the length of a single record in the original database, the maximum intervals available in different databases are also different. Most individuals in PTBDB and ECG-ID have multiple recordings. We select two recordings for training and testing for individuals with multiple recordings. As shown in Fig 3, the average interval length of PTBDB is 63 days and that of ECG-ID is 9 days. After a lot of experiments, we find that the model performed best when the number of training set, validation set and testing set samples were 300, 150 and 300, respectively, and the neurons in the full connective layer are set to 1024 and 256. In addition, we replace ReLU with Leaky ReLU as the activation function.

B. Results with different intervals

DWT has been proved to be the best feature extraction method in previous researches [8], [12], [13]. In order to prove that the statement is reliable when necessarily based on continuous sampling, we compare continuous and discontinuous sampling on PTBDB and ECG-ID, and prove that the effectiveness of current methods or models depends heavily on time continuity. Once sampling is discontinuous, the accuracy will drop sharply. As is shown in Table I, for the results on ECG-ID, when the sampling interval increases from 0 (50-C-O) to around 9 (50-N) days, the recognition rate decreases from 94.89% to 71.2%, which is about 23 percentage points lower. Pathoumvanh et al. used continuous sampling on PTBDB and used neural network to identify 10 individuals (10-C-B), and obtained 93% recognition accuracy [8]. We have also done 10 individual experiments (10-C-O), using CNN with CAM, and we get 94% recognition accuracy, which proves the validity of CAM. For the results on PTBDB, when the sampling interval increases from 0 (50-C-O) to around 63 (50-N) days, the recognition accuracy decreases from 96.6% to 40.24%, which decreases by about 56 percentage points. It can be found that the recognition rate based on DWT features depends heavily on the continuity of time. Once the time is discontinuous, the recognition rate will be greatly reduced, so we need to find a recognition method that can overcome the time dependence.

TABLE I
CONTINUOUS AND DISCONTINUOUS SAMPLING

	PTBDB	ECG-ID
10-C-B	97%	–
10-C-O	99.2%	–
50-C-B	–	93%
50-C-O	96.6%	94.89%
50-N	40.24%	71.2%

C. Results of different methods

In this part, we compare the performance of the previous optimal method with that of our proposed method in the case where there is an obvious interval between two visits, then, PTBDB and ECG-ID are used for verification. As is shown in Table II, In the case of large sampling interval, the recognition accuracy of the best methods proposed in previous studies on PTBDB and ECG-ID is only 40.24% and 71.2%. Using frequency domain features (MFCC + Peaks), the recognition accuracy can reach 50.62% and 82.03%. Adding time domain features (MFCC + Peaks + SFC), the recognition accuracy can be increased by about 3 percent. Finally, the recognition accuracy of our proposed method (MFCC + Peaks + SFC + FFT) can achieve 85.94% and 56.93% on PTBDB and ECG-ID.

TABLE II
RECOGNITION RESULT IN DIFFERENT METHODS

	PTBDB	ECG-ID
DWT	40.24%	71.2%
MFCC+Peaks	50.62%	82.03%
MFCC+Peaks+SFC	53.04%	84.53%
MFCC+Peaks+SFC+FFT	**56.93%**	**85.94%**

V. CONCLUSION

In this paper, we have proposed a practical cross-domain ECG biometric identification method, which can effectively improve the recognition accuracy though the intervals between training period and testing period were large enough. We have: 1) obtained sufficient effective samples by determining the optimal parameters of the non-reference random sampling method; 2) extracted individual-distinguishable feature insensitive to timespan between training and testing period via utilizing deep features across time, frequency and energy domains; 3) introduced a channel attention module into the CNN and modified the activation function to optimize the classifier's performance. A series of experimental results have shown that our method can still achieve high recognition accuracy (56.93% and 85.94% on PTBDB and ECG-ID) even the interval is large enough. Furthermore, in the case where there is a large interval, the recognition system has a space for optimizing the recognition accuracy.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China under Grants 61572071, 61872031, and 61301082.

REFERENCES

[1] L. Biel, O. Pettersson, L. Philipson, and P. Wide. Ecg analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(3):808–812, 2001.

[2] H. Chen, K.-K. Tseng, F. Zeng, H.-N. Huang, and S.-Y. Tu. A new ecg identification with neural network. In *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*, pages 427–430. IEEE, 2012.

[3] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

[4] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[5] V. Khoma, M. Pelc, and Y. Khoma. Artificial neural network capability for human being identification based on ecg. In *2018 23rd International Conference on Methods & Models in Automation & Robotics (MMAR)*, pages 479–482. IEEE, 2018.

[6] J. Liu, L. Yin, C. He, B. Wen, X. Hong, and Y. Li. A multiscale autoregressive model-based electrocardiogram identification method. *IEEE Access*, 6:18251–18263, 2018.

[7] J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Trans. Biomed. Eng*, 32(3):230–236, 1985.

[8] S. Pathoumvanh, S. Airphaiboon, B. Prapochanung, and T. Leauhatong. Ecg analysis for person identification. In *The 6th 2013 Biomedical Engineering International Conference*, pages 1–4. IEEE, 2013.

[9] F. Porée, A. Gallix, and G. Carrault. Biometric identification of individuals based on the ecg. which conditions? In *2011 Computing in Cardiology*, pages 761–764. IEEE, 2011.

[10] M. Tantawi, K. Revett, M. Tolba, and A. Salem. A novel feature set for deployment in ecg based biometrics. In *2012 Seventh International Conference on Computer Engineering & Systems (ICCES)*, pages 186–191. IEEE, 2012.

[11] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–19, 2018.

[12] Y.-T. Wu, H.-T. Wu, and W.-W. Chang. Compressed domain ecg biometric identification using jpeg2000. In *2015 12th International Joint Conference on e-Business and Telecommunications (ICETE)*, volume 5, pages 5–13. IEEE, 2015.

[13] Q. Zhang, D. Zhou, and X. Zeng. Heartid: A multiresolution convolutional neural network for ecg-based biometric human identification in smart health applications. *IEEE Access*, 5:11805–11816, 2017.

[14] Z. Zhang and D. Wei. A new ecg identification method using bayes' teorem. In *TENCON 2006-2006 IEEE Region 10 Conference*, pages 1–4. IEEE, 2006.

[15] G. Zheng, H.-Y. Zhang, and S.-Z. Ji. Ecg based identification under different physical status. In *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*, volume 1, pages 197–203. IEEE, 2018.