2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

ECG-based Biometric Recognition without QRS Segmentation: A Deep Learning-Based Approach

Jui-Kun Chiu, Chun-Shun Chang and Shun-Chi Wu* *Member, IEEE*

***Abstract* — Electrocardiogram (ECG)-based identification systems have been widely studied in the literature. Usually, an ECG trace needs to be segmented according to the detected R peaks to enable feature extraction from the ECGs of duration equal to nearly one cardiac cycle. Beat averaging should also be applied to reduce the influence of inter-beat variation on the extracted features and identification accuracy. Either detecting R peaks or collecting extra heartbeats for averaging will inevitably lead to a delay in the identification process. This paper proposes a deep learning-based ECG biometric identification scheme that allows identity recognition using a random ECG segment without needing R-peak detection and beat averaging. Moreover, the problem of being vulnerable to unregistered subjects in an identification system is also addressed. Experimental results demonstrated that an identification rate of 99.1% for an identification system having 235 enrollees with an equal error rate of 8.08% was achieved.**

I. Introduction

With the growing popularity of Internet of Things (IoT) devices and their accompanying services in improving an individual's quality of life, designing an effective identity management solution becomes essential to prevent the rising risks of data security and privacy breaches [1]. This task can be realized based on cryptographic methods that require individuals to remember items such as passwords to prove their identities. Biometric technologies that directly utilize individuals' physiological or behavioral attributes for identity recognition can be another category of methods. Compared to passwords, biometric attributes are always with individuals, enabling secure access to IoT devices without needing to remember anything. Example biometrics currently in use include fingerprints, faces, and irises. However, these extrinsic biometric attributes are susceptible to falsification and replay attacks. A contact lens can spoof an iris recognition system with a photographed iris image printed on it [2], [3]. A fake finger with the fingerprint ridges of an individual imprinted on the surface can be fabricated using materials such as glue and gelatin [4]. Therefore, the search for new biometric attributes never stops, and electrocardiograms (ECGs), the recordings of the heart's electrical activity, have been gaining interest [5–10].

Similar to identification based on other attributes, for ECG data to be applicable in an identification task, a feature extraction step is required for the subsequent identification process to be effective. Several feature extraction methods have been proposed in the literature, including fiducial feature-based [11], principal component analysis (PCA)-based [12], and wavelet transform-based [13] approaches. The common characteristic of these methods is that their features are extracted from the ECGs of duration equal to approximately one cardiac cycle. Consequently, an ECG trace needs to be segmented according to the detected R peaks before feature extraction, with the R peaks being detected by, for example, the Pan–Tompkins algorithm [14]. In addition, to reduce the influence of inter-beat variation on the extracted features and thus the ultimate identification accuracy, beat averaging is applied [15]. Either detecting R peaks or collecting extra heartbeats for averaging will delay the identification process. Finally, once the features are obtained, they are used to form biometric templates that are stored in the database or presented to the system for identity recognition with a selected classifier.

This paper presents a convolutional neural network (CNN)-based scheme for ECG biometric identification. Although applying CNNs to ECG biometric identification can also be found [16], we consider a scheme that requires no R-peak detection and beat averaging. A CNN is comprised of alternating layers of three types [17]: convolution, pooling, and fully connected layers having impressive prediction results in computer vision [18], [19]. With the several kernels in the convolution layers of the proposed CNN, the discriminant features of different characteristics from the input ECGs can be extracted for identification. The noise-resistant property of kernels allows us to determine a subject's identity using one single heartbeat without beat averaging to mitigate inter-beat variations [20]. Moreover, the "translation invariance" property of a pooling layer makes the down-sampled feature maps less sensitive to the content shifts in the ECG segments such that no R peak detection is required to align heartbeats before feature extraction. Finally, the method for excluding a subject that is not registered in the system is also considered to work with the proposed identification scheme.

II. Proposed ECG Biometric Identification Scheme

A. Data Preprocessing

Baseline wander and power-line interference are common artifacts that accompany the acquired ECGs. To extract the wander signal to be subtracted from the acquired ECGs for baseline wander removal, two median filters with window sizes of 200 ms and 600 ms are applied. To suppress power-line interference, a second-order IIR notch filter centered at the power line frequency is utilized. Then, we cut the artifact-free ECGs into the segments of a fixed duration T_d. Depending on the sampling rate f_s, the size of the

This work was supported by the Ministry of Science and Technology of Taiwan, R.O.C., under the contracts MOST 109-2221-E-007-012-MY3.

J.-K. Chiu (chiu.steve98@gmail.com), C.-S. Chang (chunshun001@gmail.com), and S.-C. Wu (*corresponding author; +886 3 5715131 x34304; shunchi.wu@mx.nthu.edu.tw) are with the Department of Engineering and System Science, National Tsing Hua University, Hsinchu, Taiwan 30013.

data vector to be fed into the proposed network model is $T_d f_s \times 1$. Notably, these segments were obtained randomly without referencing the detected R peaks, as required in most studies.

B. Network Architecture

The proposed network model for identity determination is depicted in the identification stage of Figure 1. This model is intended to operate in identification mode, that is, an unknown subject's identity is determined without an identity claim. The proposed model's architecture comprises five convolution, three max-pooling, and three fully connected layers. Each of the first two convolution layers is followed by a max-pooling layer. The third, fourth, and fifth convolution layers are connected in series without any intervening layer. The fifth convolution layer is followed by the last max-pooling layer, whose output is applied to three successive fully connected layers. The number of filters in the five convolutional layers is 96, 256, 384, 384, and 256. The filter size for the last three convolution layers and the three max-pooling layers is 3×1, and those for the first and second convolution layers are 11×1 and 5×1, respectively. All filters in the convolution and pooling layers are applied with a stride of one. The number of neurons in the first two fully connected layers is 512, and that of the last fully connected layer is the same as the number of subjects to identify (i.e., the number of enrollees). The rectified linear unit nonlinearity is applied after every convolution and the first two fully connected layers. Owing to the need for multi-label classification (i.e., identity identification), the "softmax" activation function is chosen for the last fully connected layer. The entire network was trained by iterating the cost function of cross-entropy between the actual class labels and the predicted probabilities from the softmax layer in the following experiments. A gradient descent optimizer was used during training. The learning rate was 10^{-4}, the batch size was 50, and the epoch was 100.

C. Exclusion of Unregistered Subjects

A deficiency associated with identity recognition using the proposed network model and shared by the existing methods is that irrespective of whether a subject is registered or not, he/she will be forced to link to one of the enrollees when trying to access the system. The security of the biometric system is thus threatened, and being able to exclude an unregistered subject is essential for the proposed scheme to be applicable in practice. To achieve this latter task, the identity given by the previous network model is only treated as a plausible one whose genuineness needs to be further verified.

Recall that the output vector of the last fully connected layer with the softmax activation function in the proposed model will generate a probability distribution comprising n_s probabilities for a given ECG segment, where n_s represents the number of enrollees. Because each enrollee of the system provides n_t ECG segments for model training, an enrollee will have n_t such probability vectors. Then, we average these probability vectors to obtain his/her own feature vector, which will be stored in the system to represent him/her. Later, when an unknown ECG segment is presented to the identification system and a plausible identity is issued, the user-specific feature vector of that plausible enrollee will be retrieved from the system to evaluate for the "similarity" with the just generated probability vector of the unknown segment.

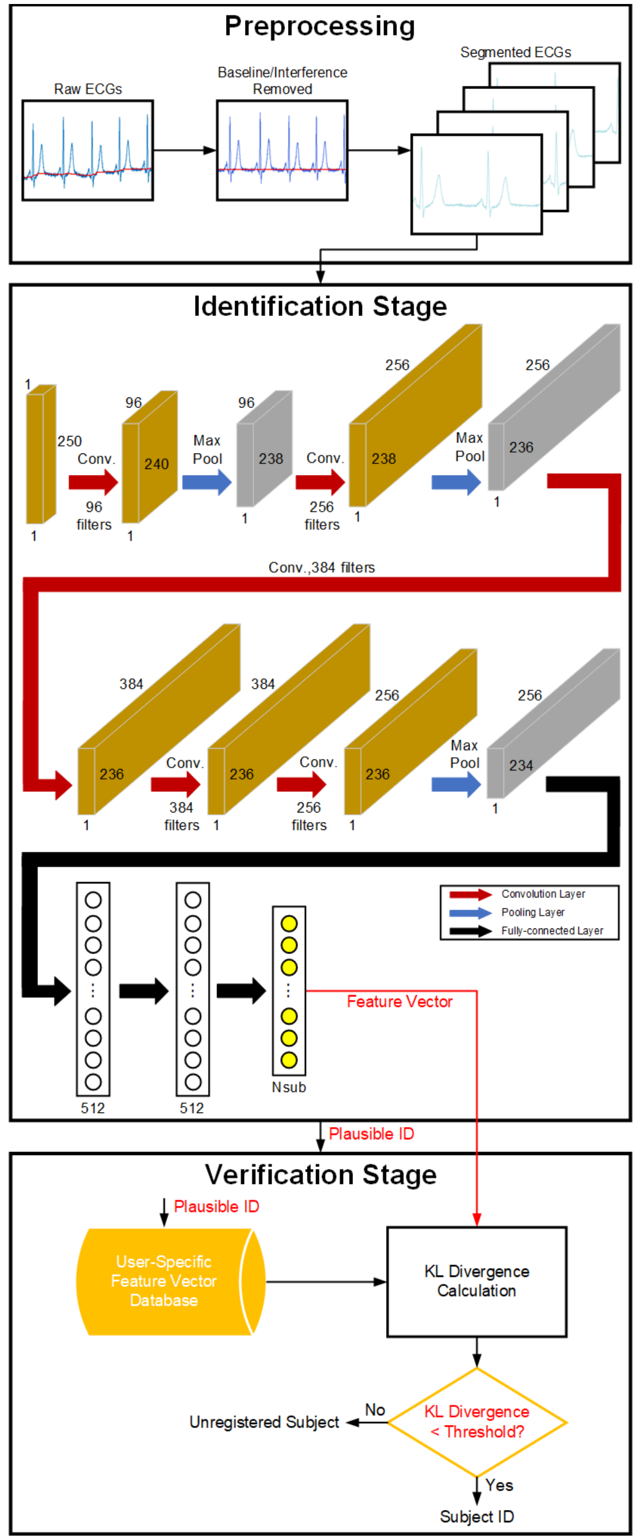

Fig. 1. Algorithm flowchart for the proposed scheme for unknown subject identification and unregistered subject exclusion.

The unknown segment is regarded as belonging to the plausible enrollee if the similarity lies in a preset fence region.

To calculate the similarity, the Kullback–Leibler (KL) divergence is used [21]:

$$D_{KL}(\mathbf{p} \parallel \mathbf{q}) = \sum_{i=1}^{n_s} p_i \log(\frac{p_i}{q_i}) , \quad (1)$$

where $\mathbf{p}$ and $\mathbf{q}$ denote two probability vectors, with p_i and q_i being their i-th elements. To determine the required fence region, the interquartile range (IQR)-based method can be used. Notably, for any enrollee, there are still n_v ECG segments for network model validation, whose probability vectors can be used to calculate the KL divergences with respect to the user-specific feature vector of that enrollee. These n_v values are not the same because of the inter-beat variation and random beat segmentation. IQR is the difference between the first quartile Q_1 and the third quartile Q_3, which are defined as the values in the n_v KL divergences holding 25% and 75% of the data below them, respectively [22]. The fence region is then defined to be from $Q_1 - k_{iqr} \times$ IQR to $Q_3 + k_{iqr} \times$ IQR, where k_{iqr} denotes the control parameter for the region size. However, because it is unnecessary to consider a subject whose KL divergence is lower than the fence region to be unregistered, we only need to check if the calculated KL divergence is above the higher value $Q_3 + k_{iqr} \times$ IQR. Later, if an unknown ECG segment is presented to the identification system and a plausible identity is issued, we will check if the obtained KL divergence is lower than its corresponding $Q_3 + k_{iqr} \times$ IQR. If so, the subject is deemed to be a plausible enrollee; otherwise, he/she is excluded as an unregistered one.

III. Results and Discussion

A. Dataset

The proposed scheme's performance was evaluated using the Lead-I recordings of 285 subjects from the Pysikalisch-Technische Bundesanstalt (PTB) database [23]. Each ECG record was sampled at 1 kHz with a 12-bit resolution. Before proceeding, each ECG record was subjected to baseline wander removal and 50 Hz power-line interference suppression, as discussed in Section II.A. Among these 285 subjects, we randomly selected 235 subjects as the enrollees to implement the CNN-based identification system, and the remaining 50 subjects were used as potential imposters to examine the proposed scheme for unregistered subject exclusion.

B. Results and Discussion

In the first example, we studied the influence of the sampling rate and the duration of an ECG segment on identification accuracy. ECG recordings of different sampling rates were obtained by downsampling the original recordings, whose duration varied from 1 to 3 s for them to cover at least one complete heartbeat waveform. A total of 70,500 ECG data segments were used, of which 47,000, 11,750, and 11,750 were used for training, validation, and testing, respectively. The proposed model was trained on TensorFlow 1.0 with the CUDA 9.0 toolkit and cuDNN v7.3.1 on the computing platform with Intel core I7-9700F CPU, GeForce GTX 2070, and 32 GB RAM. For comparison, we also implemented the fiducial-based approach using 21 morphological attributes [11] and the PCA-based approach [12], whose implementation details can be observed in [15]. During identification, only one heartbeat/segment per subject was used in each trial, and the proportion of correctly identified subjects, *i.e.*, the identification rate (IR), was adopted as the performance measure.

As shown in Figure 2, as f_s decreases, the IRs of the fiducial and PCA-based approaches decrease, whereas that of the proposed CNN-based method increases. This demonstrated that although the morphological details of heartbeats were eliminated after a reduction in f_s, discriminant information could be extracted to enable valid identification through CNN's superior feature extraction capability. Furthermore, the IR of the proposed method increased as T_d increased when f_s was sufficiently small. Note that as f_s or T_d increases, the size of the input data vector increases. This large input size increases the number of parameters of the CNN-based model. For example, when T_d = 2 s, the number of parameters is 32,212,075, 64,980,075, 130,516,075, and 261,588,075 when f_s = 125, 250, 500, and 1,000 Hz, respectively. Because the data segments available for model training were limited, a network model with satisfactory IR might not be obtainable, especially for a large model. Lastly, an IR of 99.1% was obtained when f_s = 125 Hz and T_d = 2 s.

In the next experiment, the efficacy of the proposed scheme for unregistered subject exclusion was evaluated. Two types of errors were considered. A false acceptance occurred if an unregistered subject (or an impostor) was identified as belonging to the database. A false rejection occurred when a registered subject was not identified. The false-positive identification-error rate (FPIR) and false-negative identification-error rate (FNIR) were the metrics used to quantify these two errors. To evaluate the FPIR, the remaining 50 subjects of the PTB dataset served as imposters, and each impostor presented 50 ECG segments of 2 s duration to spoof the system. Further, each of the 235 enrollees presented 50 ECG segments different from those used for training and validation to determine the correct identification rate. Afterward, the FNIR was calculated as the ratio of the number

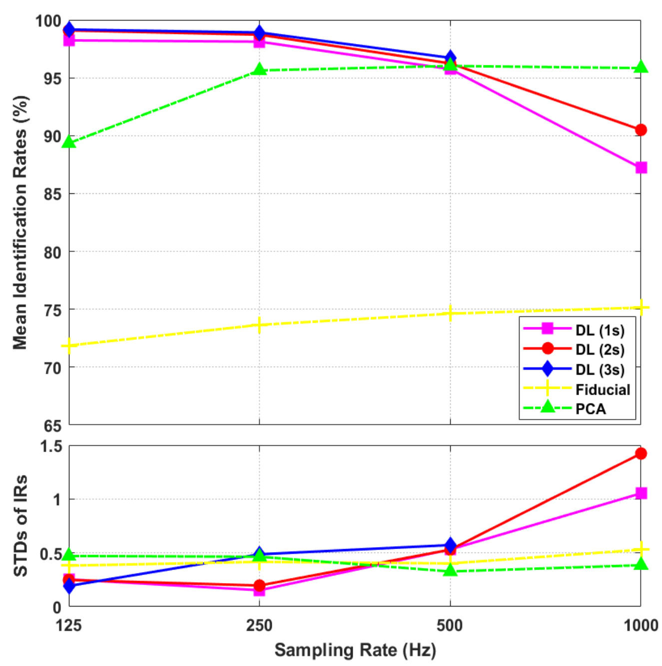

Fig. 2. IRs of various algorithms under different sampling rates.

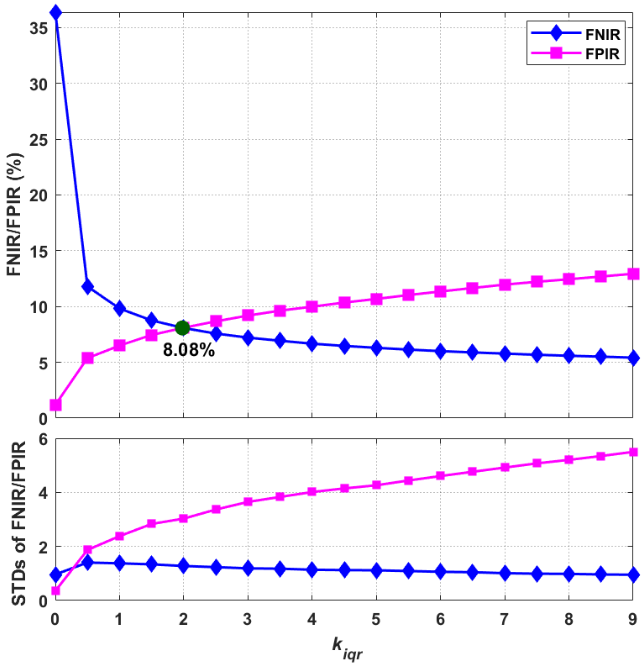

Fig. 3. FPIRs and FNIRs of the proposed scheme under different k_{iqr}.

of incorrect rejections to the number of total access attempts. We repeated the evaluation ten times. The mean FPIRs and FNIRs and their standard deviations under different k_{iqr} values are depicted in Figure 3. The lowest equal error rate achieved by the proposed scheme was 8.08% when k_{iqr} = 2.16. As observed, the value of FNIR was more than 8% when k_{iqr} < 2. This demonstrated that the data variation owing to random segmentation was severe, and incorporating a small fence region increased the likelihood of falsely rejecting a genuine subject. Although the FNIR can be reduced by enlarging the fence region, it will instead boost the FPIR, and other sophisticated approaches to mitigate either the FNIR or FPIR should be resorted to.

IV. Conclusions

This paper has presented a biometric scheme that allows identity recognition using a random ECG segment. Initially, an ECG segment from an unknown subject is presented to the proposed DL-based identification system to determine a plausible identity. The generated probability vector is then sent to calculate the KL divergence with respect to that plausible enrollee's user-specific feature vector to check whether the unknown subject is deemed to be the plausible enrollee with the proposed IQR-based unregistered subject exclusion scheme. The performance of the proposed scheme was evaluated using ECGs from the PTB database. Compared to the two existing approaches, our proposed scheme achieved relatively high recognition rates of more than 99% for an identification system with 235 enrollees. Although incorporating the exclusion scheme lowered the IR because a genuine access attempt might falsely be rejected, access attempts of impostors to the system could be stopped. Because the resulting equal error rate was still as high as 8.08%, we will focus our future studies on proposing a way to address this issue.

References

[1] X. Zhu and Y. Badr, "Identity management systems for the internet of things: a survey towards blockchain solutions," *Sensors,* vol. 18, p. 4215, 2018.

[2] A. Morales, J. Fierrez, J. Galbally, and M. Gomez-Barrero, "Introduction to iris presentation attack detection," *Handbook of Biometric Anti-Spoofing*: Springer, 2019, pp. 135-150.

[3] Z. Wei, X. Qiu, Z. Sun, and T. Tan, "Counterfeit iris detection based on texture analysis," *International Conference on Pattern Recognition*, 2008: IEEE, pp. 1-4.

[4] A. Rattani, W. J. Scheirer, and A. Ross, "Open set fingerprint spoof detection across novel fabrication materials," *IEEE Transactions on Information Forensics and Security,* vol. 10, pp. 2447-2460, 2015.

[5] I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, and J. W. Rohrbaugh, "ECG biometric recognition: A comparative analysis," *IEEE Transactions on Information Forensics and Security,* vol. 7, pp. 1812-1824, 2012.

[6] J. S. Arteaga-Falconi, H. A. Osman, and A. E. Saddik, "ECG Authentication for Mobile Devices," *IEEE Transactions on Instrumentation and Measurement,* vol. 65, pp. 591-600, 2016.

[7] R. Tan and M. Perkowski, "Toward improving electrocardiogram (ECG) biometric verification using mobile sensors: A two-stage classifier approach," *Sensors,* vol. 17, p. 410, 2017.

[8] Y. Li, Y. Pang, K. Wang, and X. Li, "Toward improving ECG biometric identification using cascaded convolutional neural networks," *Neurocomputing,* 2020.

[9] M. Ingale, R. Cordeiro, S. Thentu, Y. Park, and N. Karimian, "ECG biometric authentication: A comparative analysis," *IEEE Access,* vol. 8, pp. 117853-117866, 2020.

[10] S.-C. Wu, P.-L. Hung, and A. L. Swindlehurst, "ECG Biometric Recognition: Unlinkability, Irreversibility and Security," *IEEE Internet of Things Journal*, vol.8, pp. 487-500, 2021.

[11] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing,* vol. 2008, pp. 1-11, 2007.

[12] J. M. Irvine, S. A. Israel, W. T. Scruggs, and W. J. Worek, "eigenPulse: Robust human identification from cardiovascular function," *Pattern Recognition,* vol. 41, pp. 3427-3435, 2008.

[13] A. D. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Transactions on Instrumentation and Measurement,* vol. 57, pp. 248-253, 2008.

[14] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering,* pp. 230-236, 1985.

[15] S.-C. Wu, P.-T. Chen, and J.-H. Hsieh, "Spatiotemporal features of electrocardiogram for biometric recognition," *Multidimensional Systems and Signal Processing,* pp. 1-19, 2018.

[16] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional neural networks for ECG biometric recognition," *Pattern Recognition Letters,* vol. 126, pp. 78-85, 2019.

[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, pp. 436-444, 2015.

[18] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation,* vol. 1, pp. 541-551, 1989.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, pp. 1097-1105, 2012.

[20] H. Phan, L. Hertel, M. Maass, and A. Mertins, "Robust audio event recognition with 1-max pooling convolutional neural networks," *arXiv preprint arXiv:1604.06338,* 2016.

[21] J. H. Oh, J. Gao, and K. Rosenblatt, "Biological data outlier detection based on Kullback-Leibler divergence," *IEEE International Conference on Bioinformatics and Biomedicine*, 2008: IEEE, pp. 249-254.

[22] H. Vinutha, B. Poornima, and B. Sagar, "Detection of outliers using interquartile range technique from intrusion dataset," in *Information and Decision Sciences*: Springer, 2018, pp. 511-518.

[23] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. H. 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*, vol. 101, pp. 215-220, 2000.