1DIEN: Cross-session Electrocardiogram Authentication Using 1D Integrated EfficientNet

LIPING ZHANG, School of Computer Science, China University of Geosciences, Wuhan, China
SHUKAI CHEN, School of Computer Science, China University of Geosciences, Wuhan, China and College of Computer Science and Technology, National University of Defense Technology, China
FEI LIN, Wuhan Maritime Communication Research Institute, Wuhan, China
WEI REN, Anhui Engineering Research Center of Intelligent Perception and Elderly Care, Chuzhou University, China, Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet, Anhui University of Technology, China, and School of Computer Science, China University of Geosciences, Wuhan, China
KIM-KWANG RAYMOND CHOO, Department of Information Systems and Cyber Security, University of Texas at San Antonio, USA
GEYONG MIN, Department of Computer Science, University of Exeter, UK

The potential of using electrocardiogram (ECG), an important physiological signal for humans, as a new biometric trait has been demonstrated, and ongoing efforts have focused on utilizing deep learning (e.g., 2D neural networks) to improve authentication accuracy (with some efficiency tradeoffs). In most of the existing ECG-based authentication approaches, the ECG recordings for enrollment and testing are collected within short intervals (e.g., within an hour). However, since ECG biometrics change over time, this design may decrease authentication accuracy when ECG recordings are collected weeks or even months prior. In this article, we propose 1D Integrated EfficientNet (1DIEN) to achieve cross-session ECG authentication. We adopt 1D neural networks as a lightweight alternative to 2D neural networks, and a voting scheme is designed to reduce variance and improve general authentication performance. We use three public ECG databases

The research was financially supported by the National Natural Science Foundation of China (No. 62172303), the Open Research Project of the Hubei Key Laboratory of Intelligent GeoInformation Processing (No. KLIGIP-2019B09), the open Foundation of Anhui Engineering Research Center of Intelligent Perception and Elderly Care, Chuzhou University (No. 2022OPA01), the Knowledge Innovation Program of Wuhan - Basic Research (No. 2022010801010197), the Opening Project of Nanchang Innovation Institute, Peking University (No. NCII2022A02), and the Foundation of Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet, Anhui University of Technology, Ma'anshan, Anhui, 243032, China (IASII22-02). The work of K.-K.R. Choo was supported only by the Cloud Technology Endowed Professorship.
Authors' addresses: L. Zhang, School of Computer Science, China University of Geosciences, Wuhan, China; email: carolyn321@163.com; S. Chen, School of Computer Science, China University of Geosciences, Wuhan, China, and College of Computer Science and Technology, National University of Defense Technology, China; email: chenshukai@cug.edu.cn; F. Lin, Wuhan Maritime Communication Research Institute, Wuhan, China; email: 4419737@qq.com; W. Ren, Anhui Engineering Research Center of Intelligent Perception and Elderly Care, Chuzhou University, China, Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet, Anhui University of Technology, Ma'anshan, China, and School of Computer Science, China University of Geosciences, Wuhan, China; email: weirencs@cug.edu.cn; K.-K. R. Choo, Department of Information Systems and Cyber Security, University of Texas at San Antonio; email: Raymond.Choo@utsa.edu; G. Min, Department of Computer Science, University of Exeter, UK; email: g.min@exeter.ac.uk.

1551-6857/2023/08-ART17 $15.00
https://doi.org/10.1145/3609800

(i.e., an inter-session database, a mixed-session database, and an intra-session database) to evaluate our proposed 1DIEN under different authentication scenarios. The experimental results show that our approach achieves satisfactory performance for ECG authentication at a 3-month interval and is suitable for practical applications.

CCS Concepts: • **Security and privacy → Biometrics;**

Additional Key Words and Phrases: 1D Efficientnet, ECG, authentication

ACM Reference format:
Liping Zhang, Shukai Chen, Fei Lin, Wei Ren, Kim-Kwang Raymond Choo, and Geyong Min. 2023. 1DIEN: Cross-session Electrocardiogram Authentication Using 1D Integrated EfficientNet. *ACM Trans. Multimedia Comput. Commun. Appl.* 20, 1, Article 17 (August 2023), 17 pages.
https://doi.org/10.1145/3609800

1 INTRODUCTION

Biometric-based authentication methods, such as FaceID and TouchID, are widely used in a broad range of consumer technologies. There are, however, limitations in existing solutions. For example, it is known that FaceID will probably be unavailable for users wearing facial masks. There are also security concerns, for example, potential vulnerability that can be exploited to carry out forgery attacks [1, 2].

Electrocardiogram (ECG) is a physiological signal that is generally utilized for heart disease diagnosis and can be an effective alternative to improve both flexibility and security. In addition, it is easy to acquire user ECG data using inexpensive, daily consumer devices (e.g., iWatch [3]). Meanwhile, counterfeiting ECG signals is computationally challenging, partly due to its liveness trait [4]. ECG also represents individuals' physiological and mental status and other information [5], and therefore exhibits great potential as a practical biometric trait. Not surprisingly, many research efforts have been made to design ECG-based authentication solutions. In earlier approaches such as those reported in [6–8], fiducial points (i.e. onset, offset, and peak points of ECG waveforms) of ECG signals are utilized to identify different individuals. However, the accuracy of these approaches is highly dependent on the performance of fiducial detectors and is sensitive to environmental noises [9]. To reduce the reliance on fiducial detectors and further improve authentication accuracy, deep learning technologies have been used in more recently proposed approaches. Of the different deep learning models that have been utilized in ECG authentication, the **2D convolutional neural network (2DCNN)** appears to be one of the most widely used models, partly due to its image processing capability.

There is a tradeoff in using 2D CNN models to improve authentication accuracy [10–12], mainly in computational costs. For example, the number of parameters in 2D CNN models can be large, consequently necessitating more storage costs compared to 1D CNN models. In addition, for a 2-second input ECG signal with a 1,000 sampling frequency, the source data size of 2D CNN (generally $3 \times 224 \times 224$ points) is larger than 1D CNN (1,000 points). Therefore, 1DCNN models might be a better alternative to achieve lightweight ECG authentication. It is, however, not easy to simply apply 1DCNN models to 1D ECG signals. First, few CNN models were initially designed for 1D signals, which limits the development of 1D CNN models. Second, unlike the large image database in image processing tasks, the scale of ECG databases is too small to train a CNN model. Third, the data length of ECG recordings is generally limited. Consequently, a CNN model may learn less from the input. Therefore, developing a 1D CNN model with high authentication accuracy remains a research challenge in the ECG authentication literature.

Furthermore, in most approaches, the enrollment data and the testing data were collected within short intervals (usually within 1 hour). In practice, ECG biometrics may change over time. In other words, when the interval increases, the authentication accuracy of these approaches may decrease. In addition, in real-world applications, users may log in to the authentication system after a long interval since registration. Therefore, this cross-session ECG authentication requirement should be considered in a practical authentication system. This compounds the challenge of designing ECG authentication solutions.

Motivated by the above challenges, we propose **1D Integrated EfficientNet (1DIEN)** to achieve cross-session ECG authentication, which falls under the Authentication of People at a Building Entrance (SCK) use case category [13]. The major contributions of this article are summarized as follows:

- To realize cross-session ECG authentication, a 1DIEN is proposed by considering three ECG authentication scenarios, including cross-session authentication, mixed-session authentication, and intra-session authentication. Furthermore, we conduct multiple experiments using three public ECG databases, including a database with 3-month-interval ECG data. Experimental results show that our proposed approach achieves cross-session ECG authentication and is suitable for practical applications.
- In our proposed ECG authentication scheme, we employ a voting scheme that ensembles three 1D CNN models derived from the base model using the Compound scaling method. The purpose of using this voting scheme is to reduce variance and improve the general performance of our ECG authentication scheme. Specifically, each model in the ensemble makes a prediction for a given sample, and the final prediction is determined by a majority vote among the models. Experimental results demonstrate that our approach can reduce variance and improve general performance.

The rest of this article is organized as follows. Section 2 briefly introduces the state-of-the-art literature, and our proposed approach is described in Section 3. In Section 4, we describe the experimental setup, the databases used, and the training strategy. Then, in Section 5, we present the experimental results. Finally, we conclude this article in Section 6.

2 RELATED WORK

This section will briefly review some recent approaches for single-session ECG authentication, as well as those that support cross-session authentication.

As introduced in Section 1, fiducial points of ECG signals were utilized as biometrics in some earlier approaches. Arteaga-Falconi et al. [6] proposed an authentication algorithm based on ECG. Specifically, they detected multiple fiducial points (Q, R, S, P, T, LP, and TP points) to create feature vectors. To achieve individual authentication, they also proposed a hierarchical algorithm that takes each feature independently. In their approach, four public databases were involved for algorithm evaluation, among which the maximum duration of ECG recordings is within 2 hours.

Huang et al. [7] also proposed an ECG-based authentication scheme incorporating multiple fiducial features. But in their approach, five features were utilized as physiological biometrics of individuals, including average activation time, average QR duration/amplitude, and average RS duration/amplitude. Moreover, they proposed a Kullback–Leibler divergence-based method to realize template matching. To evaluate authentication accuracy in a noisy environment, they utilized three ambulatory databases, but none of them were long enough for cross-session authentication.

Zhang et al. [8] proposed a hybrid authentication scheme that involves both fiducial features and non-fiducial features of ECG. For fiducial features, they collected the amplitudes and duration

features among PQRST points. For non-fiducial features, they computed morphological features and collected spectral features using **Linear Prediction Coefficients (LPCs)**. To achieve ECG authentication, they proposed a **2D Principle Component Analysis (2DPCA)**-based algorithm that could extract biometrics from multiple feature spaces. In their experiments, one database was adopted for evaluation, and the duration of all recordings in this database was within 1 hour.

However, the accuracy of fiducial detectors is sensitive to many factors, such as the acquisition condition and the mental and physical status of users. Therefore, the identification accuracy of fiducial detectors will decrease in a noisy environment, and the accuracy of ECG authentication will also be influenced. To solve this issue, many efforts have been made to reduce the use of fiducial detectors. Meanwhile, deep learning technologies have been used to leverage authentication accuracy.

Abdeldayem and Bourlai [10] proposed an ECG authentication approach based on **Convolutional Neural Network (CNN)**. Compared to other fiducial detectors, R-peak detectors were robust enough against a noisy acquisition environment. Therefore, only R-peak detectors were kept in their scheme to split ECG recordings into segments. After this, these segments were converted to spectral images using **Short-Term Fourier Transform (STFT)** and **Continuous Wavelet Transform (CWT)**. Then, a 2D CNN was adopted to extract biometrics from spectral images and finally to achieve ECG authentication. In Zhao et al.'s design [11], 2D CNN and R-peak-based segmentation were also adopted. But **Generalized S-Transformation (GST)** was utilized to generate spectral images for feature representation.

In [14–16], 1D CNN, a lightweight alternative, was utilized to achieve ECG authentication, and R-peak detectors were also adopted in their approach to segment ECG recordings as the input of 1D CNN. However, although a remarkable authentication accuracy was achieved in the above approaches [8, 10, 11, 14–16], the demand for cross-session authentication was not considered. Therefore, after several months since registration, when a user tries to log in to the authentication system, the approaches above [8, 10, 11, 14–16] will probably lose usability.

To address the above challenges, cross-session authentication was considered in some recent approaches [17, 18]. In da Silva Luz et al.'s design [17], a cross-session database, namely the **Check Your Biosignals Here initiative (CYBHi)** [19], was utilized to evaluate their CNN models. Specifically, they considered two authentication scenarios: (1) the enrollment dataset was collected before the probe dataset, and (2) the enrollment dataset was collected after the probe dataset. Moreover, they built two different CNN-based models respectively for two different inputs: 1D CNN for 1D ECG heartbeats and 2D CNN for spectral images. Experimental results show that the accuracy of cross-session authentication is currently not so satisfactory.

Belo et al. [18] also considered the cross-session scenario in their approach. For comparative analysis, they conducted cross-session authentication experiments using two different 1D neural networks, namely **Recurrent Neural Network (RNN)** and **Temporal Convolutional Neural Network (TCNN)**. Apparently from their experimental results, TCNN works better than RNN, but the accuracy of both models still needs improvement.

In this article, we propose a 1DIEN to achieve cross-session ECG authentication. To improve authentication accuracy, we propose a voting scheme that integrates three neural networks. Moreover, to present a comprehensive analysis, we have conducted experiments for different authentication scenarios (seen in Section 4).

3 METHODOLOGY

This section introduces our methodology for ECG authentication. In our work, the authentication process is simplified as a closed-set classification problem. ECG recordings are first preprocessed before R-peak-based segmentation. Next, the outliers of ECG segments are detected and removed.

After this, the ECG segments are fed to our proposed 1DIEN to train the model. Finally, the probe datasets will pass through the trained model to obtain the classification result, i.e., the authentication result.

3.1 Signal Preprocessing

In the first step, all ECG recordings from different data sources are resampled to a uniform frequency of 500 Hz. Then, the resampled signals are preprocessed through noise removal and R-peak-based segmentation. The details of these methods are listed below:

(1) Noise Removal. In our design, we use a fourth-order zero-phase Butterworth Bandpass filter to denoise ECG recordings, and the cutoff frequency is 0.5 Hz and 40 Hz.

(2) R-peak-based Segmentation. To locate the location of R peaks, we use the well-known Pan-Tompkins algorithm [20] for R-peak detection. Then, the ECG recordings are segmented by a window centered at R peaks. Notably, the window size is set at $2^{\lfloor log_2 freq \rfloor}$, where $freq$ refers to the frequency of input ECG recordings after resampling.

3.2 Outlier Removal

After signal preprocessing, a simple outlier removal algorithm is adopted to avoid possible effects of highly corrupted ECG segments. Specifically, for both the enrollment dataset and probe dataset, we compute the mean values of all ECG segments at each sample point, i.e., the mean ECG segment. Then, the Euclidean distance of each segment to the mean ECG segment is computed. After this, 15% of the ECG segments are discarded to select heartbeats that are more similar to the standard one.

3.3 1D Integrated EfficientNet

To extract biometrics from ECG segments, we follow the concept of EfficientNet [21], one of the most powerful CNN models among recent CNN-based frameworks. Based on the **Neural Architecture Search (NAS)** technology and the MBconv in EfficientNet, we propose a 1DIEN that incorporates three primary 1D EfficientNet models. The details of our proposed models are listed in the context below. The architecture of our proposed 1D EfficientNet is given in Figure 1 and Table 1. As shown in Figure 1, multiple ECG segments are concatenated as the input of our proposed model. Therefore, the length of input data in our model is $N * L$, where N refers to the concatenation number and L refers to the data length of each segment. In our model, the input data are first sent to a 1D convolutional layer to extract the primary features. Then, the output of the first layer will pass through seven MBConv blocks with different scales, seen in Table 1. After this, the output will be fed to a fully connected layer and a softmax layer to obtain the classification results.

In the 1D MBConv block, given in Figure 2, the input data first passes through a $1 * 1$ 1D convolutional layer to achieve dimension raising. The output channel of this layer will be increased to n times the input channel, where n is a hyperparameter to control the zooming scale. In the first MBConv block, n is set at 1, and it is 6 in the other six MBConv blocks. Then, the output data will be sent to a Squeeze-and-Excitation block (seen in Figure 3) after a Depwise Conv block. After this, the output of the Depwise Conv block will pass through another $1 * 1$ 1D convolutional layer to realize dimension reduction. If the shape of the input feature map is equal to the output, a shortcut connection operation will be applied to this MBConv block, and a dropout layer is activated after the second 1D convolutional layer.

To figure out more architectures based on our primary model, we adopt the Compound Scaling method proposed by Tan and Le [21]. In our design, two hyperparameters are used to control the

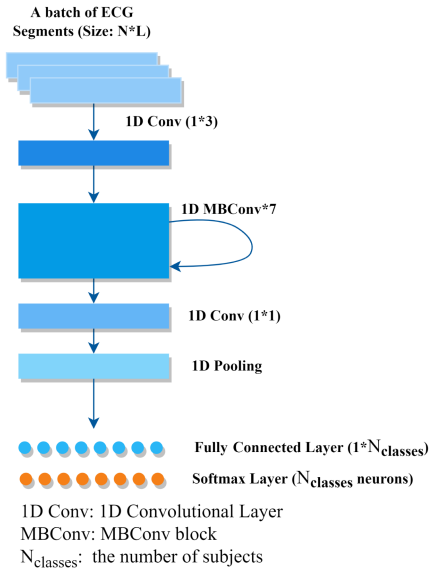


Fig. 1. Proposed 1D Efficient Net.

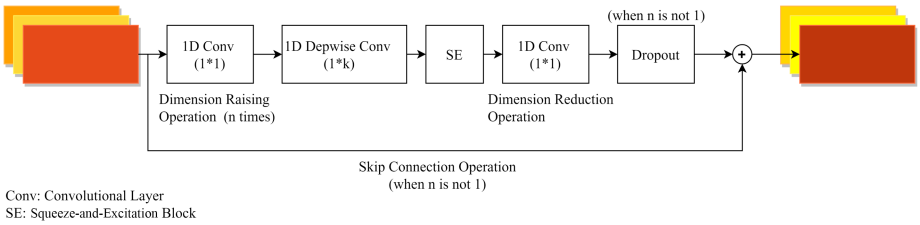


Fig. 2. 1D MBConv block of our proposed model.

Table 1. The Architecture of Our Proposed 1D EfficientNet

Stage	Operation	Scale	Channels	Layers
1	1D Conv (1*1)	$N * L$	32	1
2	MBConv ($n = 1, k = 3$)	$N * L/2$	16	1
3	MBConv ($n = 6, k = 3$)	$N * L/2$	24	2
4	MBConv ($n = 6, k = 5$)	$N * L/2$	40	2
5	MBConv ($n = 6, k = 3$)	$N * L/2$	80	3
6	MBConv ($n = 6, k = 5$)	$N * L/2$	112	3
7	MBConv ($n = 6, k = 5$)	$N * L/2$	192	4
8	MBConv ($n = 6, k = 3$)	$N * L/2$	320	1
9	Conv (1*1) & Pooling	$N_{classes}$	1	1
10	FC	$N_{classes}$	/	1
11	Softmax	$N_{classes}$	/	1

Conv: the convolutional layer, MBConv: the MBConv block, FC: the fully connected layer.
N: ECG segment number, L: ECG segment length, $N_{classes}$: the number of subjects in the ECG database.

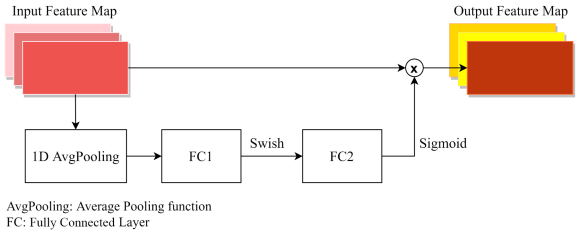


Fig. 3. 1D Squeeze-and-Excitation block of our proposed model.

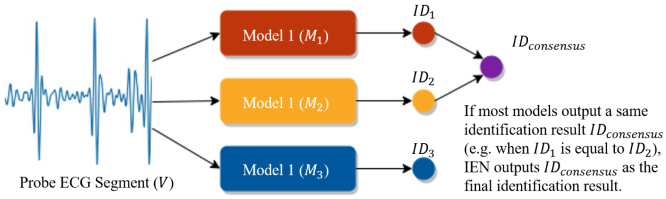


Fig. 4. The proposed voting scheme.

Table 2. The Control Parameters of Our Proposed Model

Model	Depth Factor σ_d	Width Factor σ_w
Model 1	1.0	1.0
Model 2	1.1	1.2
Model 3	1.2	1.4

scale of our model, i.e., the depth factor σ_d and the width factor σ_w. Specifically, the depth factor σ_d controls the depth of our model, i.e., the number of layers N_L in each MBConv block, as in Equation (1):

$$N_L = \lceil N_L \cdot \sigma_d \rceil. \tag{1}$$

The width factor σ_l controls the width of our model, i.e., the number of output channels N_C, as in Equation (2):

$$N_C = 8^{\lfloor log_8^{N_C \cdot \sigma_w} \rfloor}. \tag{2}$$

After exploratory experiments, we have figured out three pairs of (σ_w, σ_l) as in Table 2 that can achieve higher performance.

To reduce variance and improve the general performance of our ECG authentication scheme, a voting scheme was employed to integrate multiple models as in Figure 4. For each authentication process, each fine-grained model M_i outputs an identification result ID_i for a given probe input V. The final prediction is determined by a majority vote among the models; i.e., if most models in the ensemble output the same identification result $ID_{consensus}$, the integrated model will output $ID_{consensus}$ as the identity of V, as in Figure 4. The proposed 1D Integrated EfficientNet is presented in Figure 5.

4 EXPERIMENTS

This section introduces our experiments from the experimental setup, the adopted ECG databases, the training strategy of our proposed model, and the closed-set authentication, i.e., the evaluation experiments.

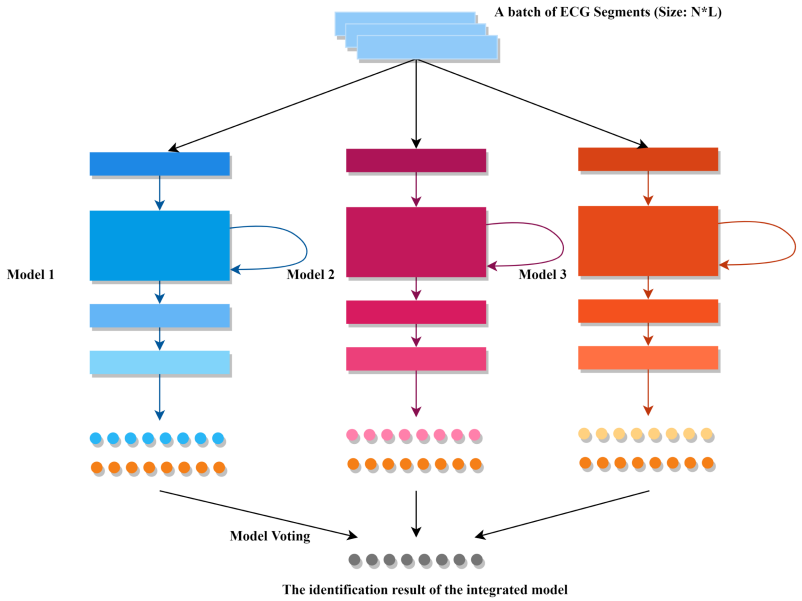


Fig. 5. Our proposed 1D Integrated EfficientNet.

4.1 Experimental Setup

In our work, we conducted our entire experiments on a Lenovo laptop with 2.30 GHz Intel(R) Core(TM) i7-10875H CPU, 16.0 GB RAM, NVIDIA RTX 2060 GPU (6 GB), and a 64-bit Windows 10 operating system. The whole experiment was realized using Python programming language (version: 3.8.5). The proposed 1D Integrated EfficientNet was built using torch 1.7.1 [22] and scikit-learn 0.23.2 [23].

4.2 Data Source

To evaluate our proposed 1D Integrated EfficientNet for cross-session authentication, we adopt CYBHi [17] in our experiments. Meanwhile, for comparative analysis, two other public databases are used, including the **ECG-ID Database (ECGID)** [24, 25] and **PTB Diagnostic ECG Database (PTBDB)** [25, 26]. The details of CYBHi, ECGID, and PTBDB are listed below. The general information of each database is briefly given in Table 3.

(1) CYBHi Database. CYBHi is a public off-the-person ECG database that contains a short-term database and a long-term database. This database was acquired with a Bluetooth wireless ECG acquisition unit (12-bit resolution and 1 kHz sampling frequency) in bioPLUX research.[1] The short-term database involves 65 healthy individuals with an average of 31.1 ± 9.46 years, and the acquisition procedure lasts 2 full days. During acquisition, the participants were asked to go through three phases with an average duration of 5 minutes. The long-term database contains two sessions of ECG data with a 3-month acquisition interval. In this database, 63 healthy participants were involved in both sessions, and the average age of participants is 20.68 ± 2.83 years. In addition, all recordings were collected during a 2-second rest phase.

[1] http://biosignalplux.com

Table 3. Overview of the Data Source

Database	Frequency (Hz)	Subjects	Type
CYBHi	1,000	65&63	N
ECGID	500	90	N
PTBDB	1,000	290	N/AbN

N = normal, AbN = abnormal.

(2) ECGID. ECGID is another off-the-person ECG database with 12-bit resolution and 500 Hz sampling frequency. In this database, 90 subjects aged from 13 to 75 years were involved, and the ECG recordings were acquired periodically over 6 months.

(3) PTBDB. PTBDB is a large-scale database that includes 549 ECG recordings from 290 individuals with an average age of 61.6 years. The sampling frequency and the resolution are 1 kHz and 16 bits, respectively.

4.3 Closed-set Authentication

To evaluate the performance of our proposed IEN for closed-set ECG authentication, we conducted experiments using different concatenation numbers $N \in [1, 4]$ in three different scenarios: cross-session, mixed-session, and intra-session authentication. In our experiments, the ECG recordings were segmented by a window centered at R peaks with a size of $2^{\lfloor log_2 freq \rfloor}$, where $freq$ refers to the frequency of input ECG recordings after resampling. Since all ECG signals were resampled to 500 Hz in our experiments, one heartbeat contains 256 points. Therefore, the length of each ECG sample for different authentication scenarios was $256N$, where $N \in [1, 4]$ is the concatenation number.

(1) Cross-session Authentication. For the cross-session authentication scenario, we utilized the long-term dataset of CYBHi, which includes ECG data from 63 individuals collected from two sessions. To construct the training dataset, we use the ECG data from the first session, and half of the ECG data from the second session are randomly selected to form the validation set. The other half from the second session (not overlapping with the validation set) is used to build the inference dataset. During training, our proposed model is trained on the training dataset and evaluated on the validation dataset. Then, the trained model is further evaluated on the inference dataset. The above process is referred to as the $S1 - S2$ phase in cross-session ECG authentication. To further investigate whether the proposed model learned the changes in ECG biometrics over time, we also construct another dataset by using the ECG data from the second session as the training dataset and half of the ECG data are randomly selected from the first session to form the validation set. The other half from the first session (not overlapping with the validation set) is used to build the inference dataset. We train this model on the training dataset, use the validation dataset to evaluate model performance, and evaluate the trained model on the inference dataset. This process is referred to as the $S2 - S1$ phase in cross-session ECG authentication.

(2) Mixed-session Authentication. For the mixed-session authentication scenario, ECGID is adopted and randomly split into training, validation, and inference datasets with a 60-20-20 ratio. This design allowed the model to learn the changes in biometrics over 6 months for each individual. Then, we trained our proposed model on the training dataset, used the validation dataset to evaluate model performance during training, and evaluated the trained model on the inference dataset.

(3) Intra-session Authentication. For the intra-session authentication scenario, PTBDB is used to simulate the scenario where a user may register in the authentication system and soon attempt to log in. Similarly, we randomly split the data into training, validation, and inference datasets with a 60-20-20 ratio. Then, we trained our model on the training dataset, used the validation dataset to evaluate model performance during training, and evaluated the trained model on the inference dataset.

To simulate an unregistered user attempting to log in, we set aside 10% of the individuals as the "unknown" class. To create this "unknown" class, we randomly select a subset of ECG signals from the original dataset. For example, if the original dataset contains ECG signals from 100 individuals, 10% of the individuals are randomly selected and labeled as "unknown." During the actual implementation of our proposed ECG authentication scheme, the model receives an input ECG signal and outputs a classification label. If the input ECG signal belongs to a registered individual, the output label indicates the specific identity of the corresponding individual. Otherwise, the model outputs the label "unknown" to indicate that the user is not recognized.

To summarize, in all three authentication scenarios, we trained our model on the training dataset, used the validation dataset to evaluate model performance during training, and evaluated the trained model on the inference dataset. We also explored the effects of concatenation number N on model performance for each experimental scenario.

4.4 Training Strategy

For model training, a **Stochastic Gradient Descent (SGD)** optimizer [27] with mini-batches of size 64 and momentum coefficient of 0.9 is employed. To control the learning rate when training loss decreases, we adopt a learning scheduler as in Equation (3):

$$lr = \left(\left(1 + \frac{cos\left(\frac{\pi}{epoch}\right)}{2}\right) \cdot (1 - lrf) + lrf\right) \cdot lr_{init}, \tag{3}$$

where *epoch* refers to the last epoch number, lrf is the learning factor, lr refers to the new learning rate, and lr_{init} refers to the initial learning rate. In our experiments, the learning factor lrf and the initial learning rate are both 0.01. As introduced in Section 3.3, three fine-grained models are selected after 100 training epochs. The closed-set authentication results can be figured out by adopting our proposed voting scheme.

5 DISCUSSIONS

This section presents a detailed discussion of our experimental results and provides a comprehensive comparison between our proposed approach and four related works [10, 11, 16, 18].

5.1 Results and Discussion

In our experiments, we used four evaluation parameters to assess the performance of our proposed IEN: **True Positive (TP)**, **False Positive (FP)**, **True Negative (TN)**, and **False Negative (FN)**. Based on these parameters, we adopted several evaluation metrics, including accuracy, recall, precision, and F1-score. These metrics are defined as follows:

Accuracy measures the ratio of correct predictions to total predictions and is computed using Equation (4):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{4}$$

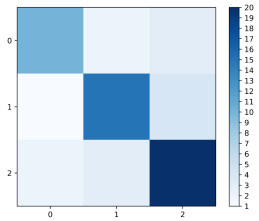


Fig. 6. A sample confusion matrix for a three-class classification problem.

Recall measures the ratio of correct positive predictions to the number of actual positives and is computed using Equation (5):

$$Recall = \frac{TP}{TP + FN}. \tag{5}$$

Precision measures the ratio of correct positive predictions to the number of correct predictions and is computed using Equation (6):

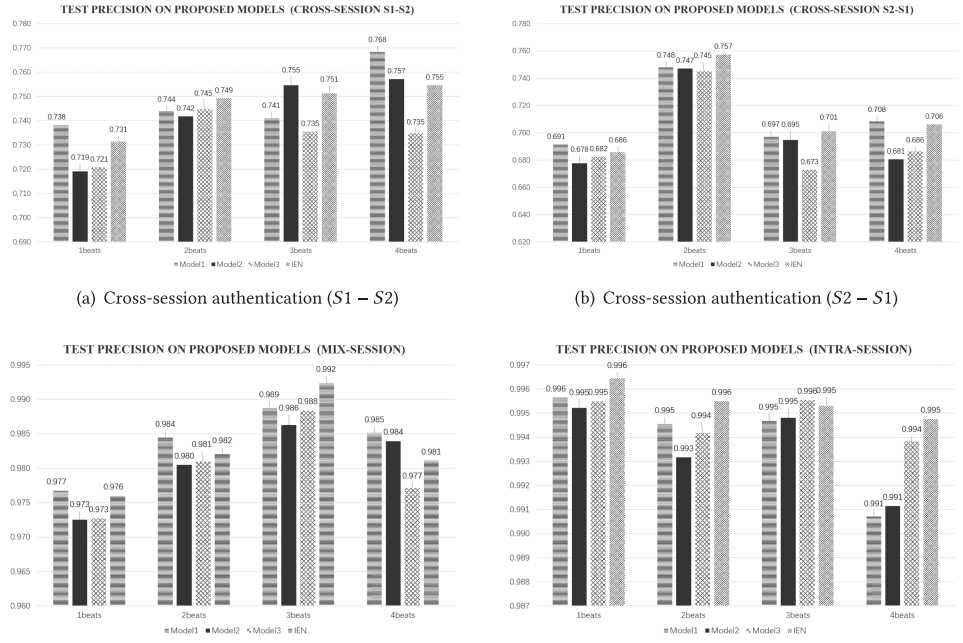
$$Precision = \frac{TP}{TP + FP}. \tag{6}$$

F1-score is the weighted average of precision and recall and is computed using Equation (7):

$$F1 - Score = \frac{2 \cdot (Precision \cdot Recall)}{Precision + Recall}. \tag{7}$$

In addition to the evaluation metrics introduced above, we also use a confusion matrix to evaluate the performance of our proposed IEN. A confusion matrix is a table that shows the number of TP, FP, TN, and FN predictions made by a classification model. For a multi-class classification problem, the confusion matrix has one row and one column for each class. The rows represent the actual classes and the columns represent the predicted classes. Figure 6 shows an example of a confusion matrix for a three-class classification problem. In Figure 6, the color shade represents the size of the value in each cell of the confusion matrix. The darker the color, the larger the value; the lighter the color, the smaller the value. The diagonal cells (from top left to bottom right) show the number of correct predictions for each class. The off-diagonal cells show the number of incorrect predictions. The confusion matrix can be used to calculate various evaluation metrics such as accuracy, recall, precision, and F1-score.

Figure 7 presents the precision of our proposed model with different concatenation numbers n. In the $S1 - S2$ phase of cross-session authentication, as shown in Figure 7(a), the highest precision for each concatenation number n (1–4) is 73.8%, 74.9%, 75.5%, and 76.8%, respectively. In the $S2 - S1$ phase, as shown in Figure 7(b), the highest precision for $n = 1, 2, 3, 4$ is 69.1%, 75.7%, 70.1%, and 70.8%, respectively. Therefore, in the $S1 - S2$ phase and $S2 - S1$ phase of the cross-session authentication scenario, the optimal concatenation number n is 4 and 2, respectively. Figure 7(c) presents the test precision in mixed-session authentication. When the concatenation number n is 3, the test precision reaches 99.2%, and the highest precision for $n = 1, 2, 4$ is 97.7%, 98.4%, and 98.5%, respectively. Therefore, for mixed-session authentication, we collect three heartbeats for each authentication process. In intra-session authentication, as shown in Figure 7(d), the best test precision for each $n \in 1, 2, 3, 4$ is 99.6%, 99.6%, 99.6%, and 99.5%, respectively. In this authentication scenario, when the concatenation number is 1, the test precision is slightly higher. Based on our experimental results, we found that the optimal concatenation numbers for the $S1 - S2$ phase and $S2 - S1$ phase of cross-session authentication, mixed-session authentication, and intra-session



(a) Cross-session authentication ($S1 - S2$)

(b) Cross-session authentication ($S2 - S1$)

(c) Mixed-session authentication

(d) Intra-session authentication

Fig. 7. The precision with different concatenation numbers n.

authentication are 4, 2, 3, and 1, respectively. Therefore, the length of ECG samples for these scenarios is 1,024, 512, 768, and 256 points, respectively. This setting is adopted for the subsequent experiments.

Table 4 summarizes the precision for each authentication scenario on the inference dataset. As shown in Table 4, for the $S1 - S2$ phase in cross-session authentication, the precision on model1, model2, model3, and IEN is 76.8%, 75.7%, 73.5%, and 75.5%, respectively. For the $S2 - S1$ phase, the test precision on four databases is 74.8%, 74.7%, 74.5%, and 75.7%, respectively. In the mixed-session scenario, the precision on IEN reaches 99.2%, and on model1, model2, and model3, the precision is 98.9%, 98.6%, and 98.8%, respectively. In the intra-session scenario, the precision for model2 and model3 is 99.6%, and the precision for model1 and IEN are both 99.6%. The experimental results show that cross-session authentication is still challenging compared to other authentication scenarios. Moreover, from Table 4, our proposed IEN can effectively reduce variance and improve the general performance of authentication.

Next, we evaluated the authentication performance of our proposed IEN on the inference dataset. From Table 5, the accuracy, recall, precision, and F1-score on the $S1 - S2$ phase are 68.1%, 71.0%, 75.5%, and 55.2%, respectively. Moreover, for the $S2 - S1$ phase, they are 69.0%, 71.3%, 75.7%, and 56.2%, respectively. In mixed-session experiments, however, all evaluation metrics (accuracy, recall, precision, and F1-score) of IEN are 99.2%, 99.3%, 99.2%, and 98.7%, respectively. In the intra-session authentication, the accuracy reaches 99.7%, while recall, precision, and F1-score are 99.6%, 99.6%, and 99.3%, respectively. According to the experimental results, our proposed IEN works effectively for both mixed-session authentication and intra-session authentication.

Furthermore, we also evaluated the performance of our proposed IEN on the training dataset. As shown in Table 6, the accuracy, recall, precision, and F1-score for cross-session, mixed-session, and intra-session authentication scenarios all reach 100.0%. These experimental results demonstrate

Table 4. The Precision for Different Authentication Scenarios

	Phase	Model1	Model2	Model3	IEN
CS[1]	$S1-S2$[2]	76.8%	75.7%	73.5%	75.5%
	$S2-S1$	74.8%	74.7%	74.5%	75.7%
MS[3]	/	98.9%	98.6%	98.8%	99.2%
IS[4]	/	99.6%	99.5%	99.5%	99.6%

[1]CS: cross-session authentication.
[2]S1: first session data of long-term CYBHi, S2: second session data of long-term CYBHi.
[3]MS: mixed-session authentication.
[4]IS: intra-session authentication.

Table 5. The Accuracy, Recall, Precision, and F1-score for Different Authentication Scenarios on Inference Dataset

	Phase	**Accuracy**	**Recall**	**Precision**	**F1-score**
CS[1]	$S1-S2$	68.1%	71.0%	75.5%	55.2%
	$S2-S1$	69.0%	71.3%	75.7%	56.2%
MS[2]	/	99.2%	99.3%	99.2%	98.7%
IS[3]	/	99.7%	99.6%	99.6%	99.3%

[1]CS: cross-session authentication.
[2]MS: mixed-session authentication.
[3]IS: intra-session authentication.

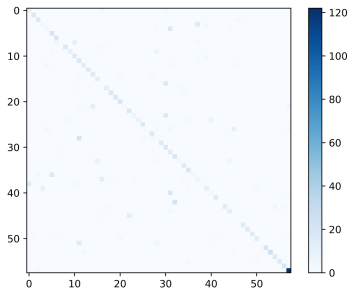
Table 6. The Accuracy, Recall, Precision, and F1-score for Different Authentication Scenarios on Training Dataset

	Phase	**Accuracy**	**Recall**	**Precision**	**F1-Score**
CS[1]	$S1-S2$	100.0%	100.0%	100.0%	100.0%
	$S2-S1$	100.0%	100.0%	100.0%	100.0%
MS[2]	/	100.0%	100.0%	100.0%	100.0%
IS[3]	/	100.0%	100.0%	100.0%	100.0%

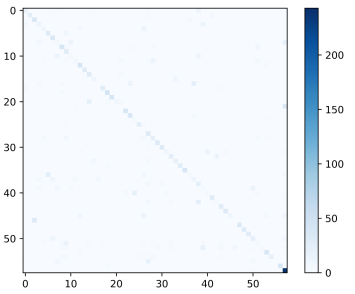
[1]CS: cross-session authentication.
[2]MS: mixed-session authentication.
[3]IS: intra-session authentication.

that our model is fully trained on the training datasets and has correctly learned the biometrics of the training ECG data.

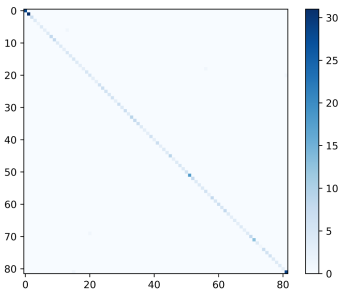
To visualize the distribution of model predictions in ECG authentication, the confusion matrices for different ECG authentication scenarios are presented in Figure 8. Specifically, Figure 8(a) and Figure 8(b) show that the dark cells are generally distributed along the diagonal, indicating a high number of correct predictions, while the few dark cells are distributed in the off-diagonal area. In Figure 8(c), almost all dark cells are distributed along the diagonal, indicating a high level of accuracy in mixed-session ECG authentication. In Figure 8(d), the color shade of the cells is lighter compared to other ECG authentication scenarios, but almost all dark cells are still distributed along the diagonal. This indicates that our proposed IEN model makes correct predictions in most cases for intra-session ECG authentication. Overall, Figure 8 shows that our proposed IEN model seldom makes incorrect predictions in mixed-session and intra-session ECG authentication. Meanwhile, our proposed IEN model also makes correct predictions in most cases for cross-session ECG authentication.



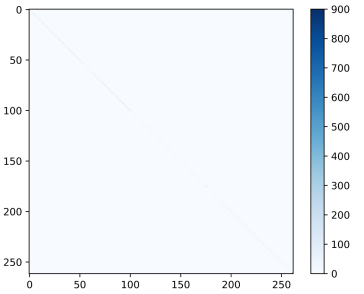
(a) Confusion Matrix for Cross-session authentication ($S1-S2$)



(b) Confusion Matrix for Cross-session authentication ($S2-S1$)



(c) Confusion Matrix for Mixed-session authentication



(d) Confusion Matrix for Intra-session authentication

Fig. 8. The confusion matrix for different ECG authentication scenarios.

5.2 Comparison and Discussion

To provide a comprehensive evaluation of our proposed approach, we compared its authentication results with those of four other approaches [10, 11, 16, 18] under different experimental scenarios.

Belo et al. used two deep learning models, TCNN and RNN, to achieve cross-session ECG authentication. In their design, **Relative Score Threshold Classification (RSTC)** was employed to classify the outputs of their models. Since TCNN and RNN models both focused on the temporal features of ECG signals, these two models could be used to analyze the changes in the ECG signals over time [18]. Ihsanto et al. proposed an ECG authentication approach using **Residual Depthwise Separable Convolutional Neural Networks (RDSCNN)** [16], which utilized 1D CNN to extract ECG biometrics from 1D ECG signals. In contrast, Zhao et al. proposed a 2DCNN-based ECG authentication method that extracted biometrics from a series of ECG trajectory images extracted via the Getframe technique for mixed-session authentication [11]. Abdeldayem and Bourlai also used 2DCNN to extract ECG biometrics in intra-session ECG authentication. They employed STFT and CWT to generate spectro-temporal images of ECG signals for ECG biometrics extraction [10]. Compared with the above approaches [10, 11, 16, 18], our proposed method simply uses 1DIEN to extract ECG biometrics without complex classification methods and signal preprocessing methods, which could be applied to multiple ECG authentication scenarios.

Table 7 illustrates the authentication results of different approaches [10, 11, 16, 18]. As shown in Table 7, for cross-session authentication, Belo et al.'s approach achieves 60.3% and 61.0% precision for the $S1-S2$ and $S2-S1$ phases, respectively, while the precision of our proposed approach for the $S1-S2$ and $S2-S1$ phases is 75.5% and 75.7%, respectively. In mixed-session authentication, Ihsanto

Table 7. Authentication Results under Different Experimental Scenarios

	Phase	Work	Method	Results
CS[1]	$S1-S2$	Belo et al. [18]	TCNN/RNN+RSTC	Precision: 60.3%
		Ours	**1DIEN**	**Precision: 75.5%**
	$S2-S1$	Belo et al. [18]	TCNN/RNN+RSTC	Precision: 61.0%
		Ours	**1DIEN**	**Precision: 75.7%**
MS[2]	/	Ihsanto et al. [16]	RDSCNN	Accuracy: 94.4%
		Zhao et al. [11][3]	2DCNN+Getframe	Accuracy: 96.9%
		Ours	**1DIEN**	**Accuracy: 99.2%**
IS[4]	/	Abdeldayem and Bourlai [10]	2DCNN+STFT/CWT	Accuracy: 95.1%
		Ours	**1DIEN**	**Accuracy: 99.7%**

[1]CS: cross-session authentication.
[2]MS: mixed-session authentication.
[3]Only 50 of 90 subjects in ECGID were involved in Zhao et al.'s approach [11].
[4]IS: intra-session authentication.

et al.'s approach and Zhao et al.'s approach achieve authentication accuracy of 94.4% and 96.9%, respectively, while our proposed approach reaches 99.2% for this scenario. Notably, only 50 of 90 subjects in ECGID were involved in Zhao et al.'s approach [11]. For intra-session authentication, we compared our approach with Abdeldayem and Bourlai's [10], and the accuracy is 95.1% and 99.7%, respectively.

From Table 7, it is clear that our proposed approach exhibits better authentication performance compared to other approaches [10, 11, 16, 18]. Specifically, in cross-session authentication, our proposed approach improves the accuracy by nearly 15.0%. Meanwhile, although in the compared approaches [10, 11, 16] the authentication accuracy of mixed-session and intra-session authentication reaches a relatively high level, our approach achieves even higher accuracy (over 99.0%).

Overall, our proposed approach offers a simpler and more efficient solution for ECG biometric authentication with superior accuracy, making it suitable for various ECG authentication scenarios.

6 CONCLUSION

In this article, we presented a cross-session ECG authentication approach using our proposed 1DIEN model. In our design, only an R-peak detector is utilized for feature extraction, which reduces the reliance on fiducial detectors and is more robust to environmental noises. Meanwhile, 1D neural networks (instead of 2D neural networks) are used in the proposed IEN, which requires fewer computational and storage resources. Our ECG authentication scheme also achieves improved general performance using the proposed voting scheme, and three databases (i.e., a cross-session database (CYBHi), a mixed-session database (ECGID), and an intra-session database (PTBDB)) were utilized for validation experiments. Experimental results showed that our proposed approach outperforms competing approaches in terms of authentication accuracy, while also achieving cross-session ECG authentication.

REFERENCES

[1] Saad Khan, Simon Parkinson, Liam Grant, Na Liu, and Stephen Mcguire. 2020. Biometric systems utilising health data from wearable devices: Applications and future challenges in computer security. *ACM Computing Surveys (CSUR)* 53, 4 (2020), 1–29.

[2] Elakkiya Ellavarason, Richard Guest, Farzin Deravi, Raul Sanchez-Riello, and Barbara Corsetti. 2020. Touch-dynamics based behavioural biometrics on mobile devices—A review from a usability and performance perspective. *ACM Computing Surveys (CSUR)* 53, 6 (2020), 1–36.

[3] Apple Inc. 2022. Take an ECG with the ECG app on Apple Watch. [EB/OL]. Retrieved March 25, 2022, from https://support.apple.com/en-us/HT208955

[4] Majid Komeili, Narges Armanfard, and Dimitrios Hatzinakos. 2018. Liveness detection and automatic template updating using fusion of ECG and fingerprint. *IEEE Transactions on Information Forensics and Security* 13, 7 (2018), 1810–1822.

[5] Ikenna Odinaka, Po-Hsiang Lai, Alan D. Kaplan, Joseph A. O'Sullivan, Erik J. Sirevaag, and John W. Rohrbaugh. 2012. ECG biometric recognition: A comparative analysis. *IEEE Transactions on Information Forensics and Security* 7, 6 (2012), 1812–1824.

[6] Juan Sebastian Arteaga-Falconi, Hussein Al Osman, and Abdulmotaleb El Saddik. 2016. ECG authentication for mobile devices. *IEEE Transactions on Instrumentation and Measurement* 65, 3 (2016), 591–600. DOI: http://dx.doi.org/10.1109/TIM.2015.2503863

[7] Pei Huang, Linke Guo, Ming Li, and Yuguang Fang. 2019. Practical privacy-preserving ECG-based authentication for IoT-based healthcare. *IEEE Internet of Things Journal* 6, 5 (2019), 9200–9210. DOI: http://dx.doi.org/10.1109/JIOT.2019.2929087

[8] Yin Zhang, Raffaele Gravina, Huimin Lu, Massimo Villari, and Giancarlo Fortino. 2018. PEA: Parallel electrocardiogram-based authentication for smart healthcare systems. *Journal of Network and Computer Applications* 117 (2018), 10–16. DOI: http://dx.doi.org/10.1016/j.jnca.2018.05.007

[9] Min-Gu Kim and Sung Bum Pan. 2019. Deep learning based on 1-D ensemble networks using ECG for real-time user recognition. *IEEE Transactions on Industrial Informatics* 15, 10 (2019), 5656–5663.

[10] Sara S. Abdeldayem and Thirimachos Bourlai. 2018. ECG-based human authentication using high-level spectro-temporal signal features. In *2018 IEEE International Conference on Big Data (Big Data'18)*. 4984–4993. DOI: http://dx.doi.org/10.1109/BigData.2018.8622619

[11] Zhidong Zhao, Yefei Zhang, Yanjun Deng, and Xiaohong Zhang. 2018. ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation. *Computers in Biology and Medicine* 102 (2018), 168–179.

[12] Sara S. Abdeldayem and Thirimachos Bourlai. 2019. A novel approach for ECG-based human identification using spectral correlation and deep learning. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 2, 1 (2019), 1–14.

[13] Song-Kyoo Kim, Chan Yeob Yeun, Ernesto Damiani, and Nai-Wei Lo. 2019. A machine learning framework for biometric authentication using electrocardiogram. *IEEE Access* 7 (2019), 94858–94868.

[14] Ruggero Donida Labati, Enrique Muñoz, Vincenzo Piuri, Roberto Sassi, and Fabio Scotti. 2019. Deep-ECG: Convolutional neural networks for ECG biometric recognition. *Pattern Recognition Letters* 126 (2019), 78–85.

[15] Mohamed Hammad, Paweł Pławiak, Kuanquan Wang, and Udyavara Rajendra Acharya. 2021. ResNet-attention model for human authentication using ECG signals. *Expert Systems* 38, 6 (2021), e12547.

[16] Eko Ihsanto, Kalamullah Ramli, Dodi Sudiana, and Teddy Surya Gunawan. 2020. Fast and accurate algorithm for ECG authentication using residual depthwise separable convolutional neural networks. *Applied Sciences* 10, 9 (2020), 3304.

[17] Eduardo José da Silva Luz, Gladston J. P. Moreira, Luiz S. Oliveira, William Robson Schwartz, and David Menotti. 2018. Learning deep off-the-person heart biometrics representations. *IEEE Transactions on Information Forensics and Security* 13, 5 (2018), 1258–1270. DOI: http://dx.doi.org/10.1109/TIFS.2017.2784362

[18] David Belo, Nuno Bento, Hugo Silva, Ana Fred, and Hugo Gamboa. 2020. ECG biometrics using deep learning and relative score threshold classification. *Sensors* 20, 15 (2020), 4078.

[19] Hugo Plácido Da Silva, André Lourenço, Ana Fred, Nuno Raposo, and Marta Aires-de Sousa. 2014. Check your biosignals here: A new dataset for off-the-person ECG biometrics. *Computer Methods and Programs in Biomedicine* 113, 2 (2014), 503–514.

[20] Jiapu Pan and Willis J. Tompkins. 1985. A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering* BME-32, 3 (1985), 230–236. DOI: http://dx.doi.org/10.1109/TBME.1985.325532

[21] Mingxing Tan and Quoc Le. 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.), Vol. 97. PMLR, 6105–6114. https://proceedings.mlr.press/v97/tan19a.html

[22] PyTorch. 2022. From research to productioin. [EB/OL]. Retrieved March 29, 2022, from https://pytorch.org/

[23] scikit learn. 2022. Machine Learning in Python. [EB/OL]. Retrieved March 29, 2022, from https://scikit-learn.org.cn/

[24] Tatiana S. Lugovaya. 2005. Biometric human identification based on electrocardiogram. Master's Thesis, Faculty of Computing Technologies and Informatics, Electrotechnical University "LETI," Saint-Petersburg, Russian Federation.

[25] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220.
[26] R. Bousseljot, D. Kreiseler, and A. Schnabel. 1995. Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet.
[27] Shun-ichi Amari. 1993. Backpropagation and stochastic gradient descent method. *Neurocomputing* 5, 4 (1993), 185–196. https://www.sciencedirect.com/science/article/pii/092523129390006O

Received 8 June 2022; revised 26 May 2023; accepted 12 July 2023