



OPEN

Person identification with arrhythmic ECG signals using deep convolution neural network

Awabed Al-Jibreen¹✉, Saad Al-Ahmadi¹, Saiful Islam² & Abdel Momin Artoli¹

Over the past decade, the use of biometrics in security systems and other applications has grown in popularity. ECG signals in particular are attracting increased attention due to their characteristics, which are required for a trustworthy identification system. The majority of ECG-based person identification systems are evaluated without considering the health-state of the individuals. Few person identification systems consider person-by-person health-state annotation. This paper proposes a person identification system considering the health-state annotated ECG signals where each person's beats overlap among variant arrhythmia classes. This overlapping between the normal class and other arrhythmia classes grants the ability to isolate normal beats in the train set from the Arrhythmic beats in the test set. Therefore, this paper investigates the effect of arrhythmic heartbeats on biometric recognition. An effective lightweight CNN based on depth-wise separable convolution (DWSC) is proposed to enhance the performance of person identification for several common arrhythmia types using the MITBIH dataset. The proposed methodology has been tested on nine arrhythmia types and presents how different types of arrhythmia affect ECG-based biometric systems differently. The experimental results show excellent recognition performance (99.28%) on normal heartbeats and (93.81%) on arrhythmic heartbeats, outperforming other models in terms of mean accuracy.

Information security and privacy are considered one of the most important concerns in many organizations, which need to protect their systems, networks, websites, and other services as well as their data. The biometric identification procedure is often recommended as an effective tool for physical and logical access control, as well as surveillance to ensure information security. Biometrics is the measurement of an individual's physiological or behavioral features^{1,2}. Biometrics have also been used to lessen hacker assaults in hospitals and other healthcare facilities^{3,4}. Due to the widespread use of smart devices like smartphones and smartwatches, biometric recognition has recently gained popularity³. However, traditional biometric modalities such as fingerprint, voice, face, iris, and handwriting have found to be vulnerable, as they can be easily replicated and used fraudulently. For example, a fingerprint can be recreated by latex, the face is vulnerable to artificial masks, an iris can be faked using contact lenses with copied iris features printed on it, and the voice can be mimicked easily. Some of them might be "lost" following severe injuries and trauma or drastically changed by plastic surgery⁵. Additionally, these biometric systems' performance is not always satisfactory in a number of applications⁵. ECG signal also known as heartprint⁶ is an emerging biometric modality employed to improve security. This modality has the inherent liveness characteristic as they can only be obtained from an alive person. As the heart is confined in the structure of the body, ECG signals are difficult to be fabricated or altered unlike face, finger, or iris biometrics⁴. Moreover, good quality ECG signals can be easily be captured from fingers⁶. All these make it a more acceptable and secured biometric modality.

There are numerous existing ECG-based biometric systems with excellent performance on regular ECG signals^{7,8}. These systems either developed their own deep learning-based systems or used the transfer learning method with pre-trained CNNs. These systems perform differently depending on the health state. On arrhythmic ECG signals, the same systems perform less well, than regular ECG signal-based recognition system performance. There has not been much research or analysis done on the arrhythmic ECG signal's structure. Authors in⁹ test their multi-layer perception network model on arrhythmic signals. The most recent advancement of ECG-biometric systems that have been proposed is based on deep-learning techniques; nevertheless, there is no detailed investigation on whether a deep-learning-based model is more suitable for arrhythmic ECG signals.

¹Computer Science Department, College of Computer and Information Sciences, King Saud University, 11543 Riyadh, Saudi Arabia. ²Department of Computer Engineering, Faculty of Engineering, TED University, 06420 Ankara, Türkiye. ✉email: 439204595@student.ksu.edu.sa

To the best of our knowledge, no previous study has tested the arrhythmic versus normal heartbeats of the same patient separated into train and test sets to proposed a complete analysis of the impact of different types of arrhythmia on the ECG-biometric system. As a result, our research suggests a thorough examination of how various types of arrhythmia affect an ECG-based biometric system. A deep-learning, beat-labeled ECG-based biometric system is also proposed in this study. The research's contribution includes the following:

- A deep CNN model is designed based on the depth-wise separable convolution (DWSC) operation with a good trade-off between the low parameter complexity and the high representation capacity for achieving good person identification performance on ECG signals.
- Using the designed CNN model, a robust ECG-based identification system is developed to reduce the adverse impacts of four different types of arrhythmia.
- We isolate the arrhythmic heartbeat set and the normal heartbeat set of the same individuals in train and test sets in order to analyze the effects of arrhythmia on the person identification system. To the best of our knowledge, no person identification system isolates the heartbeats of the same individual based on the heartbeat health-state annotation in the training and test sets.
- We found a relationship between the heartbeat's health state and the performance of the ECG-based person identification system: if the person does not suffer from cardiac disease, the performance is excellent.
- Experimentally, we show a negative impact of arrhythmia on the performance of the person identification system. The severity of the adverse effects varies depending on the type of arrhythmia.

The rest of the paper is structured as follows: Next section presents the related work. Then the proposed model is described. The following section presents results, the evaluation criteria, and the implementation details. In the conclusion, the results are highlighted and future research is underlined.

Related work

ECG identification systems are built based on either hand engineering/machine learning methodologies or deep learning methods. Many ECG-based recognition systems have been proposed since 2001 when ECG-signal was first introduced as a biometric modality¹⁰. Most of the state-of-art studies proposed solutions based on Deep Learning models for ECG Biometric systems. CNN and BLSTM are the most applied models in ECG-based biometric systems that perform the best. Yazhao Li et al.¹¹ suggest a CNN architecture for identification and feature learning process. Eko Ihsanto et al.¹² suggested Residual Depth-wise Separable Convolutional Neural Network (RDSCNN) deep learning approach for ECG-based recognition. Zhidong Zhao et al.¹³ test their suggested system using various dataset scenarios, regular dataset and the Atrial Fibrillation dataset were both used for system evaluation. The results of the suggested method are good, coming in at 99% and 98%, respectively.

There are three algorithms where biometric performance is evaluated separately between healthy and arrhythmic subjects. Authors in¹⁴ suggested a methodology and tested it and compare it versus other two methodologies. They test the three techniques on 112 PTB subjects (98 subjects have arrhythmia beats and 14 healthy subjects.). The first methodology is proposed in¹⁴, the Pulse Active transform (PAT). PAT is a transformation technique; it decomposes a signal into a series of related cyclic triangular waveforms (pulse active feature sets). Matching scores are generated by the Euclidean Distance as the distance measure. The second technique is proposed in¹⁵, where authors proposed a set of extensive ECG features proven to be invariant to anxiety state. These descriptors are the fiducial points; they contain physiological information of the heart. They develop filtering technique that merges heuristic and quantitative information. The third technique is proposed in¹⁰. Authors use about twenty-one fiducial features from single lead. These features include amplitudes, waves and segments. Soft independent modeling of class analogy (SIMSA model) is applied for classification. The first step in SIMCA is applying PCA for each class to express data variance. Next step is building models of different classes using measurement sets. The proposed methods in¹⁰ and in¹⁵ are applied in¹⁴ on same data and the results are shown in Table 1. It can be concluded in the results that in all the three different proposed methodologies, Healthy subjects score better identification performance versus Arrhythmic subjects. It also shows that PAT outperforms other two other methodologies.

Above review shows that Deep learning methods scores higher performance than hand engineering methods. It is noted that the same system performs differently based on the data health state (normal/arrhythmic).

Study	Dataset	The proposed model	Considered health-state annotation	Person health state	Performance
Biel et al. ¹⁰	PTB	21 Fiducial Features, PCA, SIMCA	Person by Person	Healthy Unhealthy	85.4 76.3
Israel et al. ¹⁵	PTB	Fiducial Features, LDA	Person by Person	Healthy Unhealthy	81.4 74
Safie et al. ¹⁴	PTB	Pulse Active Transform PAT	Person by Person	Healthy Unhealthy	95 85
Huang, et al. ¹⁶ Thentu, et al. ¹⁷ LYNN, et al. ¹⁸ Xin Liu, et al. ¹⁹	MITBIH	Deep Learning Techniques	Health-state not considered	Mixed	99.98 EER=0.13 98.4 99.77

Table 1. ECG based biometric systems on normal and arrhythmic subject.

Researches have examined biometric systems using both normal and arrhythmic datasets, but none of them have provided a thorough examination of the various types of arrhythmic signal and how they impact system performance negatively. Moreover, no ECG-based Biometric system has tested Arrhythmic beats versus Normal beats for same person in Test/Training set.

Methods

Deep neural networks make it possible to thoroughly examine data sets without requiring in-depth domain expertise or analytical prowess. Deep learning-based methods automatically learn features, unlike traditional biometric methods that require manual feature extraction. Our goal is to examine and develop the best segmentation and deep learning model for identification system using arrhythmic ECG signals. Figure 1 presents the block diagram of the proposed methodology. Convolution Neural Network (CNN) is proposed in this study as both feature learner and classifier. The details of the proposed CNN architecture are presented in the following subsection.

Change of heartbeat morphology due to arrhythmia

Arrhythmia, or an abnormal rhythm of the heart, changes the morphology of heartbeats. There are various types of arrhythmia, and each type has its own unique pattern²⁰. Therefore, the change in morphology depends on the type of arrhythmia. Due to the morphological change of the ECG signal, it has an impact on biometric systems that use an ECG as a basis. Figure 2a,b depicts the morphological changes of individual subject for different arrhythmias in MIT-BIH dataset²⁰. In Fig. 2c,d show morphological changes collectively for all subjects for variant types of arrhythmias in MIT-BIH dataset.

The statistical summary of the ECG beat samples is represented in Fig. 2, which provides the following features:

- The distance between each box's top and bottom is called the interquartile range.
- The center red line denotes the sample median for each box.
- Box whiskers are lines that extend above and below the box. The whisker length is measured from the end of the interquartile range to the furthest observation.
- The outliers are those that fall outside of the whisker length. An outlier is a value that is more than 1.5 times the interquartile range from either the top or bottom of the box. An outlier is indicated by a red plus sign.

Data preprocessing approach

The ECG signal is a recording of the electrical activity of the heart that represents the regular pattern of heartbeats made up of P-QRS-T waveforms. Some preprocessing is required to get the ECG signal ready for the following stages, including signal segmentation to have the most relevant segments. In this study, beats are segmented based on the locations of the R-peaks recorded the index-annotation file in the MIT-BIH dataset. The R-peak is simpler to detect because of its lengthy duration and high amplitude²¹. We apply the segmentation technique on 181 ms-wide window (0.5 s) following the research in^{22–24} where this signal length gives the best result on the PTB arrhythmic dataset, with R-peak at its central point.

The adopted preprocessing approach in this research is inspired by the information given in^{25–27}. In this research, R-peak position is basically centered in the beats segmentation process. However, R-peak is later re-positioned according to ECG intervals. The authors in²⁷ state that the PQ interval consumes 80 ms, which is roughly 22.22% of the ECG beat. QR-interval consumes 120 ms, roughly 33.33% of the whole beat. T-wave consumes 160 ms, which is roughly 44.44%. Therefore, T-wave consumes about double of the part consumed by P-wave. In²⁵, authors present similar calculations, suggesting that QT-interval consumes about three times what PQ-interval consumes. It is also shown that the PQ interval ranges from 0.6 to 0.79, which is roughly 25.6% of the whole ECG beat. QT-interval consumes from 0.81 to 1.25 (about 0.44) which is 31.43% of the whole ECG beat in²⁶. These calculations suggest that R-peak is not supposed to be in the center of the ECG segment. The

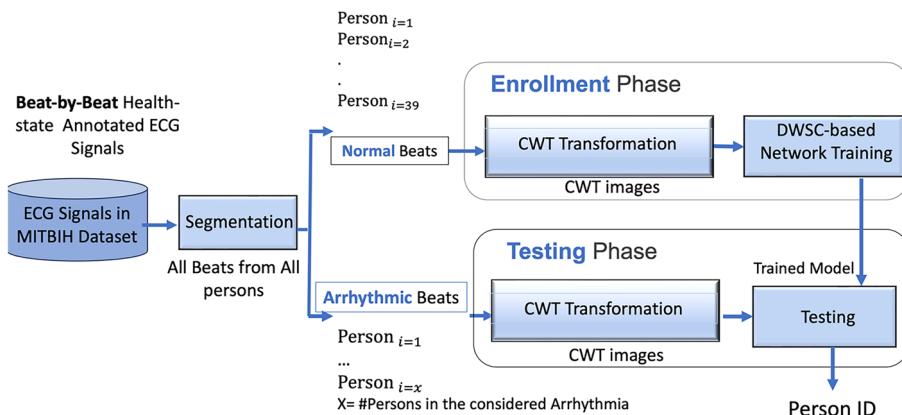


Figure 1. Block diagram of the proposed method.

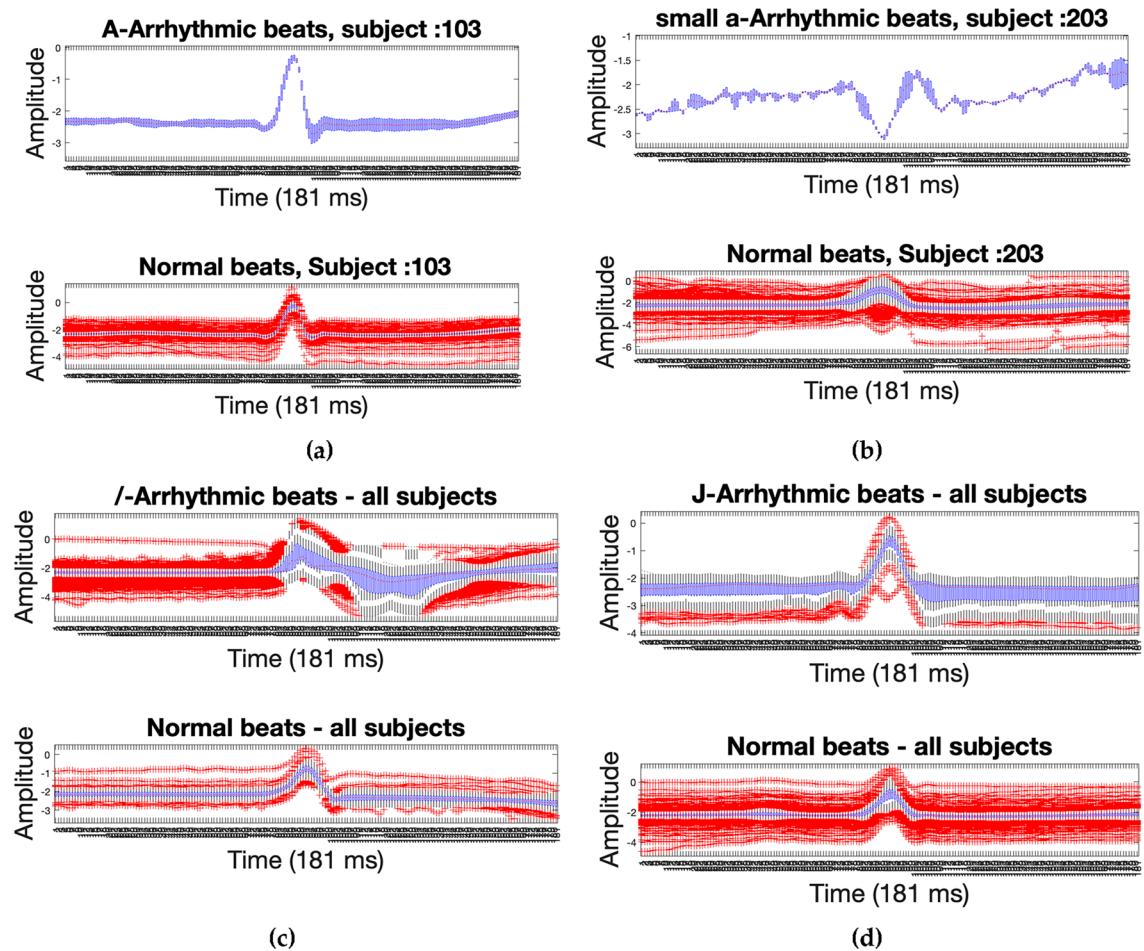


Figure 2. Heartbeat morphology of ECG signal around R-peak for arrhythmic (up) vs normal (down) beats for subjects of MIT-BIH dataset: **(a)** Atrial premature beats (A-arrhythmia) for subjects# 103; **(b)** Aberrated atrial premature beat (small a-arrhythmia) for subjects# 203; **(c)** Paced beat (/—Arrhythmia) for all common Subjects; **(d)** Nodal (junctional) premature (J- Arrhythmia) for all common subjects. **(a)** and **(b)** subfigure shows the intra-individual similarities; inter-individual differences among the individuals can be observed from these two subfigures.

adopted approach is to place the R-peak before the center of the ECG segment. This approach has increased the recognition performance for four types of arrhythmic heartbeats.

After beat segmentation, ECG signal is transformed via a time–frequency transformation technique into a two-dimensional image. Scalogram refers to this representation, which is made up of the absolute value of the signal's CWT coefficients. Figure 3 demonstrates the CWT transformation of ECG normal beats (right) and ECG specific-arrhythmia-type beats (left).

Developing deep learning-based ECG biometric model

Deep learning has become the most widely used model in ECG biometric systems. Different deep learning architectures adopted in ECG biometric systems have been investigated in²⁹. Deep learning-based systems typically outperform traditional hand-engineered and machine-learning-based systems. Deep learning-based system adaptation can be accomplished in two ways. The first involves starting from scratch and training a new network. The second method uses Transfer Learning, which enables us to train an already-built and previously trained CNN before on our dataset. High model efficiency necessitates a straightforward architecture with fewer parameters. By using a sufficient number of training samples and suitable numbers of parameters, over-fitting and under-fitting issues are easy to be avoided. The following subsection present an explanation of the proposed CNN and its components.

Depth-wise separable convolution (DWSC)

The proposed CNN include variant types of layers: convolution layer, batch normalization layer, and activation layer. Convolutional layers are the basic layer where the convolutional process could be standard or depth-wise convolution. In depth-wise convolution, channel-wise and spatial-wise computations are fulfilled in one step. The approach proposed in³⁰ served as the main inspiration for the suggested model. The authors in³⁰

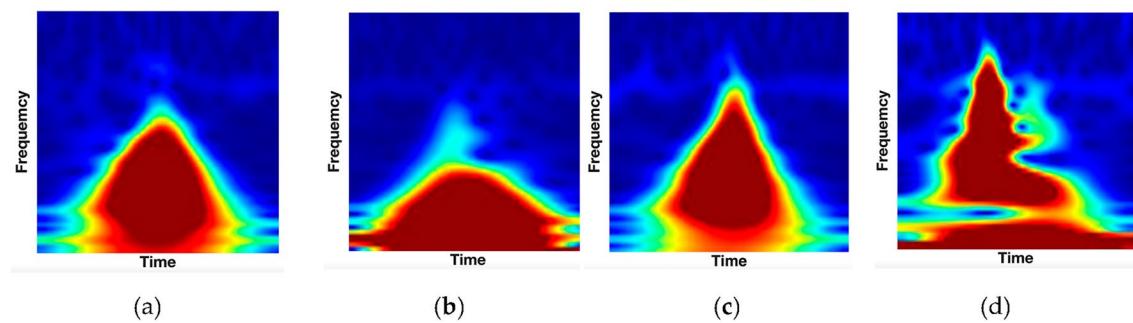


Figure 3. The CWT transformation images of ECG heartbeats from the MITBIH Dataset²⁸ of ECG Beats for: (a) ECG Normal Beats; (b) Premature ventricular Contraction (V-Arrhythmia); (c) Atrial premature beat (A-Arrhythmia); (d) Paced beat (/—Arrhythmia); Heartbeats morphology differences among the three types of Arrhythmias can be observed from the three subfigures.

propose a depth-wise separable convolution architecture. Figure 4 illustrates the basic building block of the proposed model. The basic block of the proposed model contains depth-wise separable convolutions (DWSC) that were proposed in^{30,31}. DWSC consists of two processes: depth-wise and point-wise convolution, Depth-wise convolution convolves each input feature map with a single convolutional filter. Then, point-wise convolution is used to make a linear combination of depth-wise convolution outputs.

The first step in the proposed model is applying a standard convolutional filter K to the input feature map F to produce the output feature map O :

$$O_{k,l,n} = \sum_{i,j,m} K_{i,j,m,n} \cdot F_{k+i-1,l+j-1,m} \quad (1)$$

Then, depth-wise separable convolution is applied by: (1) applying a 3×3 depth-wise convolution \tilde{K} to Each input channel,

$$\hat{O}_{k,l,m} = \sum_{i,j} \tilde{K}_{i,j,m} \cdot F_{k+i-1,l+j-1,m} \quad (2)$$

(2) Combining the depth-wise output by applying 1×1 pointwise convolution \check{K} ,

$$O_{k,l,n} = \sum_m \check{K}_{m,n} \cdot \hat{O}_{k-1,l-1,m} \quad (3)$$

where: m and n is number of the input and output feature maps respectively. Depth-wise Separable Convolution is presented in Fig. 5. The architecture of the proposed CNN model is shown in Table 2.

Activation function

The proposed model is incorporated with nonlinearity through two Activation Functions: (1) Scaled Exponentially Linear Units(SeLU), proposed in³², Where $\lambda = 1.0507$ and $\alpha = 1.6733$

$$SeLU(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{otherwise} \end{cases} \quad (4)$$

SeLU non-linearity is one of ELU's variants. SeLU's important property is introducing self-normalizing³³, it is applied in the first two blocks in the proposed model. (2) Gaussian error linear units (Gelu) activation is applied in the next following blocks to introduce regularizers, such as dropout with the activation function. Gelu procedure is performed by the following formula, where x is the input:

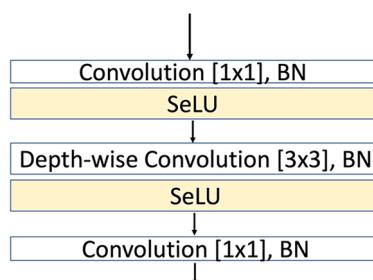
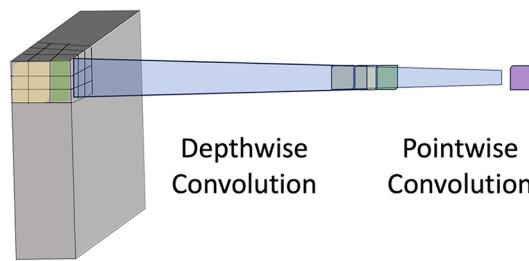


Figure 4. Basic building block (DWSC).

**Figure 5.** Depth-wise Separable Convolution.

Layer/block	Data size × Filter size * #Filters/nodes	Convolution stride
Input layer	$32^2 \times 3$	-
Basic block 1	$32^2 \times 3 \times 32$	2
Basic block 2	$16^2 \times 3 \times 64$	1
Basic block 3	$16^2 \times 3 \times 128$	2
Basic block 4	$8^2 \times 3 \times 256$	1
Basic block 5	$8^2 \times 3 \times 512$	1
Average pooling	$1^2 \times 512$	-
Fully connected	$1^2 \times 200$	-
Elu activation	$1^2 \times 200$	-
Fully connected	$1^2 \times 39$	-
Softmax	$1^2 \times 39$	-
Classification	$1^2 \times 39$	-

Table 2. Architecture of the proposed CNN model.

$$GeLU_{(x)} = \frac{1}{2} \times \left(1 + \tanh \left[\sqrt{\frac{2}{\pi}} (x + 0.044715x^3) \right] \right) \quad (5)$$

Instead of using gate inputs by their sign, like in ReLU, we adopt the GeLU activation function to introduce the nonlinearity weighted inputs computed by the inputs value. The experiments in³⁴ show that GeLU matches or exceeds models with RELUs or ELUs within variant fields such as NLP fields, computer vision, and automatic speech recognition. The Gaussian distribution function computes the cumulative distribution function (CDF)³⁴.

Experiments and Results

Arrhythmic datasets are provided by PhysioNet³⁵ for a patient population. MIT-BIH²⁸ includes 47 participants total—25 men, ages 32 to 89, and 22 women, ages 23 to 89. A detailed statistic is presented in Table 3. This study focuses on this dataset because it is the only one that offers beat-labeling (A or N) for every subject. Every other dataset with an arrhythmia offers a subject-labeling. Therefore, the MIT-BIH dataset is the only one of these

Beats health state	#Persons ^a	#Persons with normal beats
Normal (Healthy)	39	39
Premature Ventricular Contraction (V- Arrhythmia)	36	29
Atrial Premature (A- Arrhythmia)	36	22
Fusion of ventricular and normal (F-Arrhythmia)	16	2
Nodal (junctional) escape (small-j Arrhythmia))	4	3
Aberrated atrial premature (small a)	6	6
Fusion of paced & normal beat (small-f)	3	3
Isolated QRS-like artifact (J-Arrhythmia)	20	16
Nodal (junctional) premature (J-Arrhythmia)	6	5
Right Bundle Branch Block (R-Arrhythmia)	6	2

Table 3. The statistical Information of the MITBIH Dataset. ^aIn The MITBIH Dataset, PersonID may overlap among multiple Arrhythmia classes.

datasets that is beat-annotated rather than person-annotated. As a result, the MIT-BIH dataset is the main focus of this study.

Evaluation

The number of samples TP, FP, TN, and FN is computed by the confusion matrix, where T is true, F is false, P is positive, and N is negative. Evaluation criteria are computed using these sample numbers as follows³³:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (6)$$

$$\text{Specificity (TNR)} = \frac{TN}{FP + TN} \quad (7)$$

$$\text{Recall (TPR)} = \frac{TP}{TP + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

The evaluation of different pre-trained CNNs identification systems for regular beats was the first step in this study. The top results in the research and the complexity of the networks are taken into consideration when selecting the five pretrained networks. Table 4 demonstrates that the proposed model achieves recognition rates better than or competitive to the state-of-the-art networks, including Xception³⁶, Shufflenet³⁷ and MobilenetV2³⁸, which provided the best outcomes of the MITBIH dataset on regular beats. The table also shows whether the network is a Directed Acyclic Graph (DAG) or a linear architecture network.

Exploring and evaluating the various ECG biometric system architectures for both normal and arrhythmic datasets will help us achieve our first goal. Therefore, as a second step, we compare different arrhythmia types' negative impact on ECG biometric systems in Table 5, where Mobilenet⁴¹ is tested on each arrhythmia type separately, isolating the same subject's arrhythmic heartbeats from his/her Normal heartbeats in the test set or train set. ROC curves are shown in Fig. 6 while Fig. 7 shows the cumulative distribution plot (CDF). Figure 8 shows the confusion charts for the four Arrhythmia types. Table 5 also shows the case where both the train and test sets contain only arrhythmic beats, the results are excellent (97.67%). This is because the network trains and learns the characteristics and morphologies of the same type of beats in the test set.

Results

At first, we tested the effect of different components of the proposed model on the performance of recognition using arrhythmic heartbeats. it is observed from Table 4 that the four arrhythmia types (J, A, R, & j) has less impact on the recognition performance. In general, the performance is excellent if the person does not have an arrhythmia as presented in Table 3. However, the performance is worse if the person has arrhythmia. Depending on the type of arrhythmia, the performance varies. For instance, of all arrhythmias, atrial premature (A-arrhythmia) scores very well. This is because, as seen in Figs. 2a and 3b, A-arrhythmia does not significantly alter the form of the heartbeat. In contrast with the occurrence of premature ventricular contraction arrhythmia (V-arrhythmia) which severely changes the morphology of heartbeats, as shown in Fig. 3a.

Depth-wise separable convolution (DWSC) has been applied in several studies related to image processing tasks^{30,42-44}. Authors in³⁰ evaluate their model, which is based on DWSC on ImageNet and CIFAR100 datasets

Model	# Learnable parameters	#Layers	Training time (in hours)	#Epochs for training	Image size	Recognition rates on regular heartbeats	CNN architecture
Inception ³⁹	23.9 M	315	52.4	7	299 ²	96.97	DAG
Xception ³⁶	22.9 M	171	67.3	4	299 ²	99.32	DAG
SqueezeNet ²⁰	1.2 M	68	3.1	5	227 ²	98.68	DAG
Googlenet ⁴⁰	6.9 M	144	11.7	15	224 ²	98.50	DAG
MobilenetV2 ⁴¹	3.5 M	154	5.3	5	224 ²	99.17	DAG
Shufflenet ³⁷	1.4 M	172	2.5	5	224 ²	99.29	DAG
Proposed Model	665 k	46	58 Min	10	32 ²	99.28	Linear

Table 4. Deep network models parameters, time consumed and recognition rate on regular heartbeat. M: Millions, K: Kilos. Significant values are in bold.

Arrhythmia type	Evaluation metrics									
	Train set (N) versus test set (A)					Train set (A) versus test set (N)				
	Acc	Prec	Sens	Rec	F1	Acc	Prec	Sens	Rec	F1
All arrhythmia types	31.79	0.23	31.68	99.9	0.46	48.16	5.36	47.5	96.2	0.10
Premature ventricular Contraction (V-Arrhythmia)	18.5	*	18.5	100	*	19.27	17.3	79.9	100	0.3
Atrial premature (A-Arrhythmia)	93.64	36	93.8	99.5	0.53	64.16	12.2	63.8	96.5	0.22
Fusion of ventricular and normal (F-Arrhythmia)	60.71	0.32	60.8	99.6	0.01	31.29	0.96	31	98.7	0.02
Nodal (junctional) escape (small-j Arrhythmia))	80.7	17.3	79.9	100	0.3	80.7	17.3	79.9	100	0.3
Aberrated atrial premature (small a)	25.3	*	25.3	100	*	52.16	19.8	45.8	99.9	0.33
Fusion of paced & normal beat (small-f)	61.61	17.3	79.9	100	0.3	-	-	-	-	-
Isolated QRS-like artifact (—Arrhythmia)	57.6	*	57.6	*	*	-	-	-	-	-
Nodal (junctional) premature (J-Arrhythmia)	59.26	*	59.26	100	*	49.46	9.95	46.8	98.8	0.18
Right bundle branch block (R-Arrhythmia)	93.57	*	93.57	100	*	-	-	-	-	-

Table 5. ECG-based identification performance of variant types of arrhythmia. N: normal beats. A: arrhythmic beats. *These values are affected by the zero values in confusion matrix.

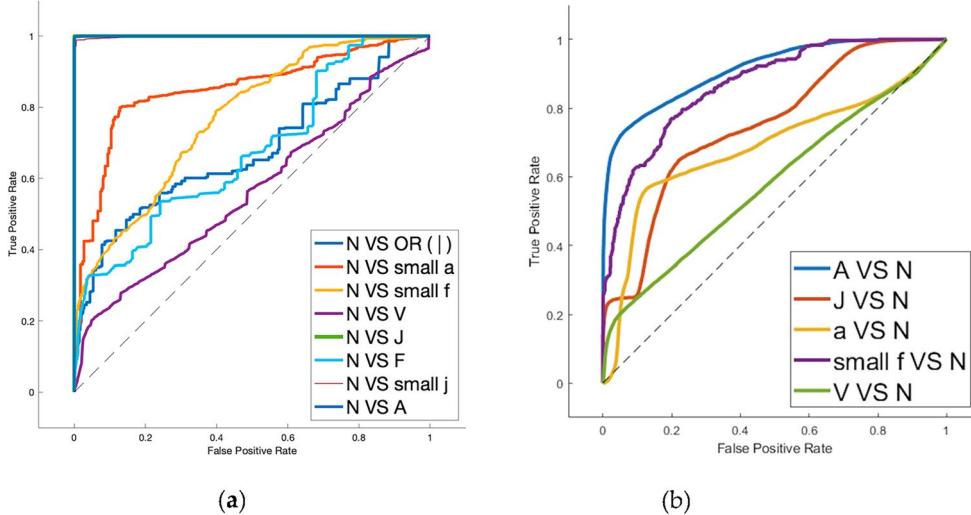


Figure 6. ROC Curves of the validation curve on person identification system evaluated on the MIT-dataset for: (a) Normal beats (training set) versus one arrhythmia type beats (test set); (b) one arrhythmia type (training set) versus normal beats (test set).

and show that the depth-wise approach scored the highest score, requiring much lower parameters compared with the other state-of-the-art approaches.

We evaluated the proposed model and other six pretrained networks on Normal betas (as a Train set) vs. Arrhythmic Beats (J, A, R, & J) as a test set and results are presented in Table 6. In addition, we tested these networks and the proposed model on Normal beats (as a Train set) vs. Arrhythmic Beats (All arrhythmia types) as a test set, and results are presented in Table 6. In addition, to explore DWSC effects, we tested the proposed model with DWSC versus the proposed model with standard convolution (SC) instead. Results in Table 6 shows that the proposed model with DWSC score competitive accuracy with about $5 \times$ lower number of learnable parameters than the proposed model without DWSC operation.

The main contribution in this study is the results shown in Table 7 where the proposed model improves performance of person identification system with the four types of arrhythmias. Table 7 shows that the proposed model outperforms other models in term of the mean accuracy. Considering the proposed model's simplicity (664 k learnable Parameters) and other models' complexity (2 to 20 million learnable parameters, this is another achievement of this study in addition to highest mean accuracy. The proposed model is evaluated on normal heartbeats and achieves a good recognition performance (99.6%) presented in Table 8.

Discussion

The discriminative feature of the proposed model is the shape it takes. It starts from low number of channels, in the first block, then the number of channel increases while we go deeper. The feature map spatial resolution starts from the input image's size then it became smaller while we go deeper in the network for simpler computations and better features learning. The smaller spatial resolution (down-sampling) is achieved through stride = 2 in



Figure 7. Cumulative distribution plot (CDF) in Person identification on the MITBIH dataset for: **(a)** normal versus all arrhythmic beats; **(b)** normal versus normal; **(c)** all-arrhythmic beats versus all-arrhythmic beats.

Confusion Chart (N vs A-Arrhythmia)

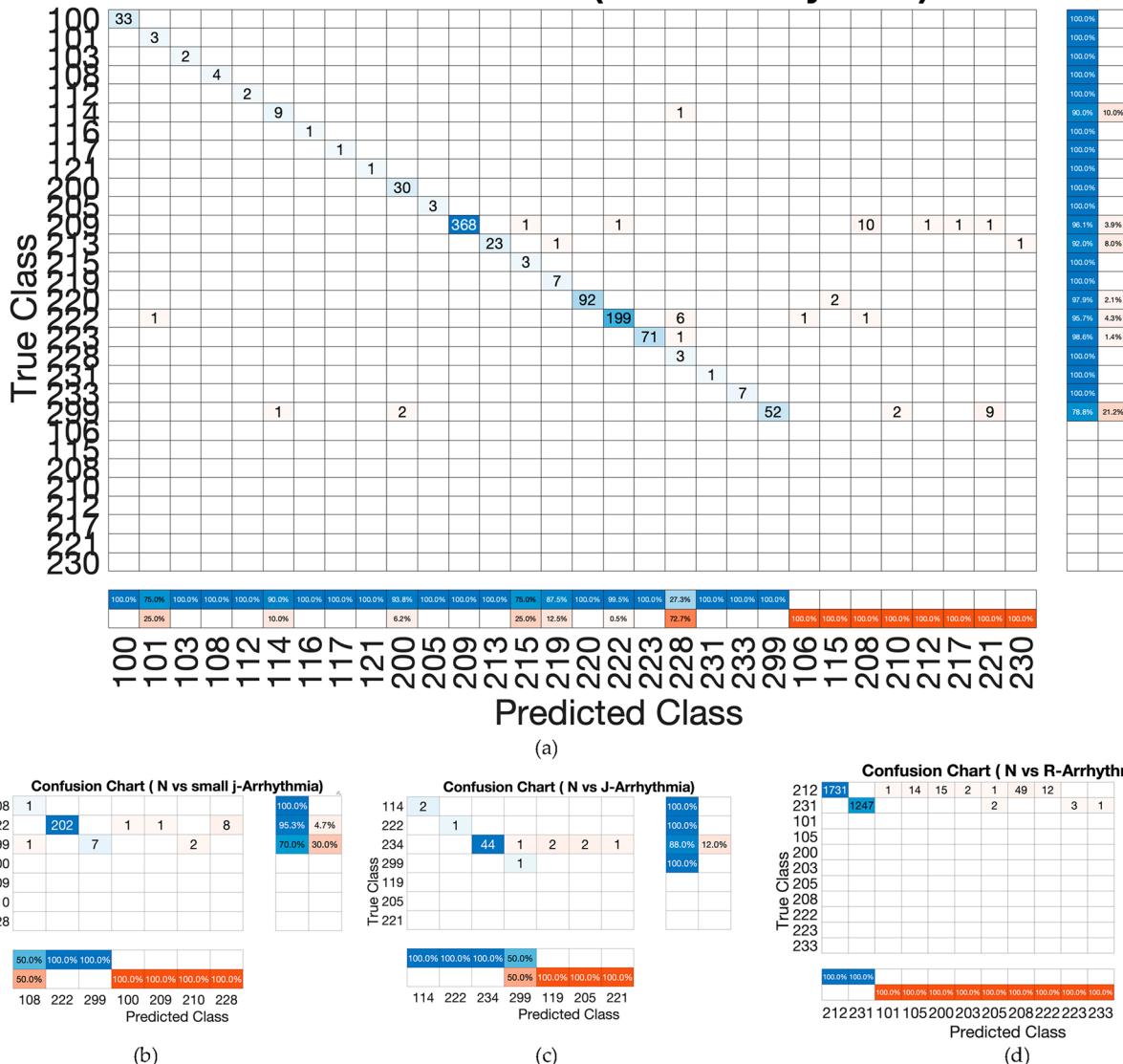


Figure 8. Confusion charts of the proposed methodology results for the four considered arrhythmia. The last two columns are the true positive and false negative rates respectively. The last two rows are the false positive and true negative rates respectively.

Model	Activation	Accuracy (%)			
		W/o DWSC (SC-model)		With DWSC	
		(J, A, R, & j) Arrhythmia	All Arrhythmia	(J, A, R, & j) Arrhythmia	All Arrhythmia
Googlenet ⁴⁰	ReLU ³⁹	57.32	27.30	NA	NA
Inception ³⁹	ReLU ²⁰	78.45	26.65	NA	NA
SqueezeNet ²⁰	ReLU ⁴⁰	69.41	23.09	NA	NA
Xception ³⁶	ReLU ³⁶	NA	NA	95.53	30.28
Shufflenet ³⁷	ReLU ⁴¹	NA	NA	94.41	28.69
MobileneV2 ⁴¹	ReLU ³⁷	NA	NA	95.07	29.56
Proposed Model	SeLU/GeLU	96.08	30.11	96.22	29.17

Table 6. Recognition accuracy of normal (training set) versus arrhythmic beats (test set) with and w/o DWSC. Significant values are in bold.

Arrhythmia Type	Inception	Xception	Googlenet	SqueezeNet	Shufflenet	Mobilenet	Proposed methodology
Atrial Premature (A-Arrhythmia)	90.82	86.86	70.43	81.98	91.34	93.64	95.41 ± 0.2093
Nodal (junctional) escape beat (small-j Arrhythmia)	94.62	91.93	87.36	89.73	95.52	90.84	94.17 ± 0.2348
Nodal (junctional) premature (J-Arrhythmia)	61.11	90.74	46.30	54.54	61.11	59.26	88.89 ± 0.3172
Right Bundle Branch Block (R-Arrhythmia)	59.17	98.57	54.32	64.21	95.87	93.57	96.75 ± 0.1773
Mean accuracy	76.43	92.03	64.6	72.62	85.96	84.33	93.81

Table 7. Identification results of the proposed methodology compared with other models (N vs. A). Significant values are in bold.

Refs	Preprocessing, feature extraction and classification	Acc
⁴⁵ _2019	Segmentation & Convolutional Neural Network	98.9
¹⁸ _2021	Heartbeat Segmentation and RNN	98.4
⁴⁶ _2022	Wavelet transform & Kalman filter & (IPSO-SVM)	95.17
⁴⁷ _2023	Auto Correlation/Discrete Cosine Transform, Feature reduction method & City Block	99.4
Proposed	Heart Beat Segmentation, CWT & CNN	99.58

Table 8. Performance comparison of the proposed model and other models in the MITBIH Dataset.

some convolution operations but not everyone to avoid aggressive feature map spatial size reduction. This model shape is inspired from the networks such as VGG16, VGG19, Mobilenet, Shifflent and Xception networks. These networks are DAG networks, which have more complex architectures. The proposed model is light and simple, its architecture is linear so that each layer in the proposed model has a single input and each output goes to a single layer. Unlike DAG architecture, the proposed model is free of Addition/Concatenation operation and residual connections. Residual connections are not needed in the proposed model due to its simplicity and few number of layers.

The proposed architecture consists of five repeated building blocks based on 3×3 DWSC as in Fig. 4. It is a lightweight model and has roughly 664.6 Kilos learnable parameters. We compare the complexity of same model but based on Standard Convolution (SC) which has 3.8 million and found it is $5.7 \times$ fewer learnable parameters than SC-Model. Assuming number of channels N , DWSC use number of learnable parameters = $filterSize^2 \times 1 \times 1 \times N$, against $filterSize^2 \times N^2$ learnable parameters in case the SC-model. This occurs at the same or small increase in accuracy compared with the same SC -model, as seen in Table 6. Therefore, in the proposed model (DWSC-model) provides competitive accuracy with much lower number of parameters compared with the SC-model and compared with other networks with millions learnable parameters. This is due to the main benefits of DWSC operation which is providing more efficient use of network parameters and an excellent trade-off between the model complexity and its representation capability^{36,37}. The pointwise 1×1 convolution is simply a linear projection of exact same spatial size and the same number of channels; however, it allows the model to include extra non-linearity through the activation function without affecting the receptive field.

Conclusion

The proposed lightweight CNN based on depth-wise separable convolution (DWSC) was tested on normal and different arrhythmic heartbeats and it was found effective to enhance the performance of person identification for several common arrhythmia types. The experimental results show excellent recognition performance on normal heartbeats (99.28%) as well as on arrhythmic heartbeats (93.81%). It outperformed other commonly used models in term of average accuracy as shown in Table 7. We generated ROC curves (Fig. 6) using the FAR and FRR for each class compared to the rest classes (one-versus-rest mode of classification). Although it is possible to balance the error rates for an authentication system by fine tuning the network parameters which we like to leave as a future work. As other possible future works, we also aim to investigate challenges such as automatic scalability due to increase of users, recognition with noisy signals, and effect of ECG signal variability on recognition performances due to different physical conditions such stress or exercise activities.

The model's scalability (i.e., adding more users/classes) is an important requirement of an identification system in real-world situations. Instead of requiring additional training of the entire model from scratch, the system should be able to automatically adopt itself. However, the automatic adaptation of an existing model presents a challenge and we would like to look into a few solutions for this as a future work. First, there have been some recent studies on Automated Machine Learning (AutoML)⁴⁸, which could be incorporated into biometrics. Another alternative strategy is to use transfer-learning to retrain the previously trained model on the new pattern, after fine-tuning the old model by adjusting the final fully connected layer and the classification layer to account for the additional class. This technique could involve optimizing the model's hyperparameters for training on the new class using optimization algorithms such as Bayesian optimization for automated hyperparameter tuning

algorithms. With the use of these methods, the model's performance on the new task can be enhanced by automatically finding the ideal hyperparameters without the need for human intervention. Finally, automated incremental learning (Online learning) for an adaptive deep learning framework⁴⁹ could also be exploited to handle the scalability.

Data availability

The data supporting reported results presented in this study, MIT-BIH Dataset, are openly available in PhysioNet at <https://physionet.org/content/mitdb/1.0.0/>, reference number YYYY²⁸. The analyzed and segmented data generated during this study can be found in: <https://github.com/AwabedM/ECG-Biometrics/tree/main/Data>. Where no new data were created.

Received: 9 December 2023; Accepted: 20 February 2024

Published online: 23 February 2024

References

- Azam, S. N. A. M., Sidek, K. A. & Ismail, A. F. Photoplethysmogram based biometric identification incorporating different age and gender group. *J. Telecommun. Electron. Comput. Eng. (JTEC)* **10**, 101–108 (2018).
- Noor-Alyanis, S. R. & Al-Dhaqm, A. Biometrics authentication techniques: A comparative study. *Int. J. Adv. Appl. Sci.* **7**, 97–103 (2020).
- Camara, C., Peris-Lopez, P. & Tapiador, J. E. Human identification using compressed ECG signals. *J. Med. Syst.* **39**, 148 (2015).
- Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M. & Zhang, D. Biometric recognition using deep learning: A survey. arXiv preprint [arXiv:1912.00271](https://arxiv.org/abs/1912.00271) (2019).
- Beritelli, F. & Serrano, S. Biometric identification based on frequency analysis of cardiac sounds. *IEEE Trans. Inf. Forensics Secur.* **2**, 596–604 (2007).
- Islam, M. S. *et al.* Heartprint: A dataset of multisession ECG signal with long interval captured from fingers for biometric recognition. *Data* **7**, 141 (2022).
- Ammour, N. *et al.* Deep contrastive learning-based model for ECG biometrics. *Appl. Sci.* **13**, 3070 (2023).
- Aslan, H. İ & Choi, C. VisGIN: Visibility graph neural network on one-dimensional data for biometric authentication. *Expert Syst. Appl.* **237**, L121323 (2024).
- Sidek, K. A., Khalil, I. & Jelinek, H. F. ECG biometric with abnormal cardiac conditions in remote monitoring system. *IEEE Trans. Syst. Man Cybern. Syst.* **44**, 1498–1509 (2014).
- Biel, L., Pettersson, O., Philipson, L. & Wide, P. ECG analysis: A new approach in human identification. *IEEE Trans. Instrum. Meas.* **50**, 808–812 (2001).
- Li, Y., Pang, Y., Wang, K. & Li, X. Toward improving ECG biometric identification using cascaded convolutional neural networks. *Neurocomputing* **391**, 83–95 (2020).
- Ihsanto, E., Ramli, K., Sudiana, D. & Gunawan, T. S. Fast and accurate algorithm for ECG authentication using residual depthwise separable convolutional neural networks. *Appl. Sci.* **10**, 3304 (2020).
- Zhao, Z., Zhang, Y., Deng, Y. & Zhang, X. ECG authentication system design incorporating a convolutional neural network and generalized S-transformation. *Comput. Biol. Med.* **102**, 168–179 (2018).
- Safie, S. I., Nurfazira, H., Azavitra, Z., Soraghan, J. J. & Petropoulakis, L. In *2014 IEEE Region 10 Symposium*. 667–671 (IEEE).
- Israel, S. A., Irvine, J. M., Cheng, A., Wiederhold, M. D. & Wiederhold, B. K. ECG to identify individuals. *Pattern Recogn.* **38**, 133–142 (2005).
- Huang, Y.-W., Yang, G.-P., Wang, K.-K., Liu, H.-Y. & Yin, Y.-L. Multi-scale deep cascade bi-forest for electrocardiogram biometric recognition. *J. Comput. Sci. Technol.* **36**, 617–632 (2021).
- Thentu, S., Cordeiro, R., Park, Y. & Karimian, N. In *2021 IEEE International Conference on Consumer Electronics (ICCE)* 1–6 (2021).
- Lynn, H. M., Pan, S. B. & Kim, P. A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks. *IEEE Access* **7**, 145395–145405. <https://doi.org/10.1109/access.2019.2939947> (2019).
- Liu, X., Si, Y. & Yang, W. A novel two-level fusion feature for mixed ECG identity recognition. *Electronics* **10**, 2052 (2021).
- Iandola, F. N. *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint [arXiv:1602.07360](https://arxiv.org/abs/1602.07360) (2016).
- Murawwat, S., Asif, H. M., Ijaz, S., Malik, M. I. & Raahemifar, K. Denoising and classification of Arrhythmia using MEMD and ANN. *Alexand. Eng. J.* **61**, 2807–2823 (2022).
- AlDuwaile, D. A. & Islam, M. S. Using convolutional neural network and a single heartbeat for ECG biometric recognition. *Entropy* **23**, 733 (2021).
- Islam, M. S. & Alajlan, N. In *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 1–5 (IEEE).
- Zihlmann, M., Perekrustenko, D. & Tschanne, M. In *2017 Computing in Cardiology (CinC)*, 1–4 (IEEE).
- Escalona, O., Mukhtar, S., McEneaney, D. & Finlay, D. Armband sensors location assessment for left Arm-ECG bipolar leads waveform components discovery tendencies around the MUAC line. *Sensors* **22**, 7240 (2022).
- Hwang, H. B., Kwon, H., Chung, B., Lee, J. & Kim, I. Y. ECG authentication based on non-linear normalization under various physiological conditions. *Sensors* **21**, 6966 (2021).
- Srivastva, R., Singh, A. & Singh, Y. N. PlexNet: A fast and robust ECG biometric system for human recognition. *Inf. Sci.* **558**, 208–228 (2021).
- Moody, G. B. & Mark, R. G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* **20**, 45–50 (2001).
- Lynn, H. M., Hong, T., Kim, H., Kim, S. H. & Kim, P. In *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, 0065–0070 (IEEE).
- Guo, Y., Li, Y., Wang, L. & Rosing, T. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 8368–8375.
- Enkvetchakul, P. & Surinta, O. Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition. *Appl. Sci. Eng. Progress* **15**, 3810–3810 (2022).
- Klambauer, G., Unterthiner, T., Mayr, A. & Hochreiter, S. Self-normalizing neural networks. In *Advances in Neural Information Processing Systems*, Vol. 30 (2017).
- Yang, Y., Xinyang, S., Wang, Q. & Fang, C. Enhancement of electromagnetic scattering computation acceleration using LSTM neural networks. *Electronics* **12**, 3900 (2023).
- Hendrycks, D. & Gimpel, K. Gaussian error linear units (gelus). arXiv preprint [arXiv:1606.08415](https://arxiv.org/abs/1606.08415) (2016).
- Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **101**, 215–220 (2000).
- Chollet, F. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1251–1258.
- Zhang, X., Zhou, X., Lin, M. & Sun, J. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 6848–6856.

38. Howard, A. G. *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* (2017).
39. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2818–2826.
40. Szegedy, C. *et al.* In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 1–9.
41. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L.-C. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 4510–4520.
42. Senecal, J. J., Sheppard, J. W. & Shaw, J. A. In *2019 International Joint Conference on Neural Networks (IJCNN)* 1–8 (IEEE).
43. Xu, M. & Zhang, X.-L. Depthwise separable resnet with squeeze-and-excitation blocks for small-footprint keyword spotting. *arXiv preprint arXiv:2004.12200* (2020).
44. Liu, J. *et al.* Alzheimer’s disease detection using depthwise separable convolutional neural networks. *Comput. Methods Prog. Biomed.* **203**, 106032 (2021).
45. Kim, M.-G., Ko, H. & Pan, S. B. A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks. *J. Ambient Intell. Humaniz. Comput.* **11**, 1859–1867 (2020).
46. Li, N. *et al.* The Identification of ECG signals using WT-UKF and IPSO-SVM. *Sensors* **22**, 1962 (2022).
47. Meltzer, D. & Luengo, D. Efficient clustering-based electrocardiographic biometric identification. *Expert Syst. Appl.* **219**, 119609 (2023).
48. Salehin, I. *et al.* AutoML: A systematic review on automated machine learning with neural architecture search. *J. Inf. Intell.* **2**, 52–81 (2024).
49. Jameel, S. M., Hashmani, M. A., Rehman, M. & Budiman, A. An adaptive deep learning framework for dynamic image classification in the internet of things environment. *Sensors* **20**, 5811 (2020).

Author contributions

Supervision, S.A. and A.A.; Conceptualization, S.I., S.A. and A.A.; Experiments Design, S.I. S.A., A.A. and A.J.; methodology, A.J. and S.A.; software, S.A. and A.J.; Figures preparation, A.J.; writing—original draft preparation, A.J.; writing—review and editing, S.I., S.A., and A.A.; All authors have read and agreed to the published version of the manuscript.”

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A.-J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024