Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks

Murside Degirmenci
Izmir Katip Celebi Universitesi

Mehmet Akif Ozdemir (✉ makif.ozdemir@ikcu.edu.tr)
Izmir Katip Celebi Universitesi   https://orcid.org/0000-0002-8758-113X

Elif Izci
Izmir Katip Celebi Universitesi

Aydin Akan
Izmir Ekonomi Universitesi

Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks

M. Degirmenci[a], M.A. Ozdemir[a,b,*], E. Izci[a], A. Akan[c]

[a] *Department of Biomedical Technologies, Izmir Katip Celebi University, Cigli, Izmir 35620, Turkey*

[b] *Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, Izmir 35620, Turkey*

[c] *Department of Electrical and Electronics Engineering, Izmir University of Economics, Balcova, Izmir 35330, Turkey*

Abstract

Background: Electrocardiogram (ECG) is a method of recording the electrical activity of the heart and it provides a diagnostic means for heart-related diseases. Arrhythmia is any irregularity of the heartbeat that causes an abnormality in the heart rhythm. Early detection of arrhythmia has great importance to prevent many diseases. Manual analysis of ECG recordings is not practical for quickly identifying arrhythmias that may cause sudden deaths. Hence, many studies have been presented to develop computer-aided-diagnosis (CAD) systems to automatically identify arrhythmias.

Methods: This paper proposes a novel deep learning approach to identify arrhythmias in ECG signals. The proposed approach identifies arrhythmia classes using Convolutional Neural Network (CNN) trained by two-dimensional (2D) ECG beat images. Firstly, ECG signals, which consist of 5 different arrhythmias, are segmented into heartbeats which are transformed into 2D grayscale images. Afterward, the images are used as input for training a new CNN architecture to classify heartbeats.

Results: The experimental results show that the classification performance of the proposed approach reaches an overall accuracy of 99.7%, sensitivity of 99.7%, and specificity of 99.22% in the classification of five different ECG arrhythmias. Further, the proposed CNN architecture is compared to other popular CNN architectures such as LeNet and ResNet-50 to evaluate the performance of the study.

Conclusions: Test results demonstrate that the deep network trained by ECG images provides outstanding classification performance of arrhythmic ECG signals and outperforms similar network architectures. Moreover, the proposed method has lower computational costs compared to existing methods and is more suitable for mobile device-based diagnosis systems as it does not involve any complex preprocessing process. Hence, the proposed approach provides a simple and robust automatic cardiac arrhythmia detection scheme for the classification of ECG arrhythmias.

Keywords: Arrhythmia; Classification; Convolutional Neural Networks; Deep Learning; Electrocardiogram.

1. Introduction

A heartbeat is an event that occurs when the heart contracts and relaxes rhythmically. Electrocardiogram (ECG) is a tool used for observing the electrical activity of the heart. Each heartbeat has a P wave, QRS complex, and T wave that represent repolarization and depolarization of the atria and ventricles of the heart. The heart rate for a healthy person ranges from 60 to 100 beats per minute [1]. The heartbeat depends on one's instant activity that it may beat slower or faster. The heart beats faster when exercising, and it beats slower than active conditions during resting or sleeping.

Arrhythmia is any abnormality in the cardiac cycle that can be considered as an irregular heart rate or irregular waveform [2]. A heart that has an arrhythmic heartbeat cannot pump enough blood throughout the body as well as it should. This condition may damage many organs and pose a threat to daily life. Since cardiac arrhythmias are a major threat to human health, their early and accurate detection is essential in medical practice [3]. Manual analysis of the ECG signal recordings is not efficient to correctly detect abnormalities in the heart rhythm [4,5]. Analysis of long-duration ECG signals by physicians is a burdensome and time-consuming task that may yield inaccurate results. Developing automatic cardiac arrhythmia detection algorithms reduce the physician's workload, decreases arrhythmia detection time, and also improves diagnostic efficiency and accuracy. Many studies in the literature presented some forms of computer-aided systems by using different feature extraction and classification techniques to accurately detect abnormalities in the ECG signals.

There have been several methods for automatically detecting arrhythmias based on signal processing, feature extraction, and machine learning algorithms [6]. Recorded ECG signals are generally contaminated by different noise types or artifacts which may change the characteristics of the ECG signal. In the preprocessing stage, contaminants are removed from ECG signals applying different filtering operations [7]. The feature extraction stage is crucial for the discrimination of arrhythmic signals from regular ones. Features are extracted from the ECG signals by using various methods in

the time or the frequency domain [7]. Among the time-domain ECG morphology and heart rate features [8], R-R interval and linear discriminant analysis (LDA) [9] have been widely used. In the frequency-domain, features based on Fourier transform [10], spectral correlation [11], and variational mode decomposition (VMD) [12] have been used. The preprocessing and feature extraction stages construct an analysis system for the final learning algorithms. Conventional machine learning algorithms such as Support Vector Machine (SVM) [13], Random Forest (RF) [14], and Artificial Neural Networks (ANN) [15] have been utilized in previous studies for the classification of different arrhythmia types. In [16], time-frequency (TF) analysis of ECG signals is applied in the detection of cardiac arrhythmias. Pseudo Wigner-Ville Distribution is utilized to obtain a TF representation of ECG signals obtained from the American Heart Association (AHA) and Massachusetts Institute of Technology (MIT) - Boston's Beth Israel Hospital (BIH) databases. Four different classifiers; Logistic Regression with L2 Regularization (L2-RLR), Adaptive Neural Network Classifier (ANNC), SVM, and Bagging are used to classify ventricular fibrillation, ventricular tachycardia, normal sinus, and other rhythms. Although many studies have developed arrhythmia detection algorithms by using preprocessing, feature extraction, and machine learning techniques, they have limitations for accurately classifying arrhythmias. Loss of ECG beat characteristics in noise filtering, not selecting optimal features for classification, low classification performance is examples of these limitations that directly affect the success of the studies [17].

The architecture of conventional neural network algorithms contains input, output, and hidden layers. Deep Learning (DL) is a novel neural network structure that contains more than three layers and has become more favorable in detection and classification studies compared to conventional techniques [18]. In DL, feature extraction and classification parts are embedded in the model which automatically identifies the optimal features from the input data [19]. DL has become very popular in recent studies since it provides improved performance of ECG arrhythmia classification. DL may be categorized into different types based on the training methods such as recurrent neural networks (RNNs), deep neural

networks (DNNs), convolutional neural networks (CNNs), and Long short-term memory (LSTM). Zhang *et al.* [20] proposed an RNN and clustering-based method to find patient-specific ECG classification algorithms by using the MIT-BIH arrhythmia database. Al Rahhal *et al.* [21] proposed a DNN based method to classify ECG signals using three different databases. In [22], the temporal features of ECG heartbeats are detected with DNN on the MIT-BIH database with 99.09% accuracy. Yildirim *et al.* designed a 16-layer deep CNN to classify 17 different arrhythmias on ECG signals taken from the MIT-BIH dataset. In [23], authors proposed a novel 3-layer deep genetic ensemble of classifiers to detect 17 different arrhythmias which achieved 99.37% classification accuracy.

CNN is a popular deep learning architecture for the automatic classification of ECG signals [24]. Kiranyaz *et al.* [25] introduced a one-dimensional (1D) 3-layer CNN with an R-peak wave for ECG arrhythmia classification. Baloglu *et al.* [26] used CNN algorithms with the end-to-end structure on 12-lead ECG signals for automatic detection of myocardial infarction on ECG signals with over 99% accuracy. Savalia *et al.* [27] proposed multilayer perception (MLP) and CNN-based methods to identify first-degree AV block (FAV) and ventricular bigeminy diseases. In [28], the authors proposed an 11-layer CNN structure to detect different ECG segments with an accuracy of 92.50% using 2 seconds, and 94.90% using 5 seconds ECG segments. Yao *et al.* [29] proposed an attention-based time-incremental CNN (ATI-CNN), that preserves spatial and temporal characteristics of ECG signals with the integration of a CNN architecture and recurrent cells. Their results attained 81.2% accuracy. Besides using a 1D CNN for ECG arrhythmia classification, there are several studies in the literature based on 2D CNNs. The hidden structure of CNN can extract various local features from 2D input samples. The spatially adjacent pixels may be represented by utilizing nonlinear and multiple filters. Hence, the recent state-of-the-art studies proposed 2D CNNs for ECG arrhythmia classification, these approaches motivated researchers to implement CNNs with 2D image-based input data. In these studies, ECG signals are converted into 2D images and provided input for CNN [17,30]. The CNN architecture is considered to be more suitable for the analysis and classification of 2D data. It achieves

better results compared to other classical techniques in image processing [31]. Jun *et al*. [17] proposed deep 2D CNN based on 7 different arrhythmia detection methods using 2D grayscale ECG beat images, and compare the performance of their proposed architecture with well-known structures such as AlexNet and VGGNet. Huang *et al*. [32] used TF spectrograms that are obtained from five different ECG beats by short-time Fourier transform (STFT). Spectrogram images are utilized as input data to the 2D deep CNN which yielded an average accuracy of 99.00%.

Although there are many studies for arrhythmia detection in the literature, most of them experience various problems such as excessive depth in the network, training cost, and computational complexity. Considering the benefits and drawbacks of the existing techniques, this paper proposes a novel DL approach for identifying different arrhythmia types in ECG signals. In this approach, the CNN model is selected as the network structure, and ECG signals are converted into ECG images to be used as input to a new CNN architecture. Five different arrhythmia types are considered for classification. Before converting 1D ECG signals into 2D images, segmentation is applied to the ECG signals to extract the ECG beats. Then each ECG beat is converted into a 2D grayscale image and used as input data for the CNN architecture. The classification performance of the proposed DL approach is compared to LeNet and ResNet-50 architectures.

2. Material and methods

The proposed ECG arrhythmia classification algorithm consists of the following steps; heartbeat segmentation, image transformation, and ECG arrhythmia classification by using 2D CNN architecture. The schematic diagram of the methodology is shown in Fig. 1. The ECG signals are taken from the MIT-BIH arrhythmia database [33]. Five different arrhythmia types are selected from the database. ECG signals are segmented into heartbeats and converted into ECG heartbeat images which are then used to train the network.

MIT-BIH arrhythmia database [33] includes different arrhythmic signals which are independently annotated by two or more cardiologists according to their arrhythmia types. Each record includes two-channel ECG signals which are the modified limb lead II (MLII) and one of the modified leads V1, V2, or V5. Due to the deformation of the second channel, MLII lead recordings are used in this study. The duration of each signal is 30-min with 360 Hz sampling frequency and is filtered by a 0.1-100 Hz bandpass filter. The MIT-BIH database is well-known to be imbalanced by the non-equal number of ECG beats for each arrhythmia which deteriorates the accuracy of DNN and CNN models [34]. The deep learning algorithms may tend to be biased for the type of arrhythmia classes that include many samples as the number of heartbeats are not equal for each class of arrhythmia in the dataset. Recently, some approaches have been proposed to eliminate the imbalance effect in the MIT-BIH database. A novel study proposed a data augmentation technique using 15 different classes from MIT-BIH to generate a balanced database. In this approach, the proposed model trained with the same techniques and hyperparameters using the original imbalanced database and balanced database that is created with the augmentation technique to observe the effect of the balanced dataset. Their results revealed that augmenting the imbalanced original dataset with generated heartbeats outperforms the performance of arrhythmia classification than using the proposed techniques trained with the original dataset [35]. Oh *et al*. [36] utilized a combination of CNN and LSTM for classifying five classes of the MIT-BIH dataset which are normal (N), left bundle branch block (LBBB), right bundle branch block (RBBB), premature ventricular contraction (PVC), atrial premature beat (APB) and normalization technique is applied to standardize the input data. Huang *et al*. [32] proposed 2D CNN using spectrograms of the five different classes of arrhythmias as input for ECG arrhythmia detection. The nearly equal numbers of five different arrhythmias which are N, LBBB, RBBB, PVC, and APB are selected to balance the dataset and classified to achieve the highest accuracy. Recent studies have revealed that data

augmentation techniques and providing an equal number of beat samples among classes of the dataset can be used as approaches to stabilize the imbalance ratio (IR) [37].

Considering the limitations of the MIT-BIH dataset and approaches that are applied to overcome them, arrhythmia classes with an approximately equal number of heartbeats are utilized to train the model and to eliminate the imbalance effect of the database. Besides, the processing of a large amount of heartbeat data plays a vital role in DL-based approaches for successfully analyzing and classifying them. As a result, the process of determining the types of arrhythmia was performed based on the two criteria mentioned above. The specific 5 arrhythmia type that includes many heartbeat data and an approximately equal number of beat samples among them are determined to provide the normalized and big dataset. Consequently, five different arrhythmia types which are N, LBBB, RBBB, PVC, and Paced beat (PB) are selected from the database. Another advantage of selecting the 5 specific types of arrhythmia is that it provides a fair comparison with recent studies that make a certain number of arrhythmia classifications. The sample sizes of ECG beats for the considered classes are given in Table 1.

Table 1 The number of beats for each arrhythmia type and ECG records.

Beat Type	Classes	Annotation	Beat Count	Record Numbers
Normal	N	N	8017	102 104 105 106 114 116
Left Bundle Branch Block	LBBB	L	8072	109 111 207 214
Right Bundle Branch Block	RBBB	R	7256	118 124 212 231 232 207
Premature Ventricular Contraction	PVC	V	6970	105 106 107 114 116 119 203 207 208 210 213 214 215 221 228 233
Paced Beat	PB	/	7024	102 104 107 217

2.2. Heartbeat segmentation

Python programming language is utilized for analyzing and classifying ECG arrhythmic signals. In the MIT-BIH database, each heartbeat of a signal is annotated by cardiologists based on the QRS structure and type of the heartbeats. To identify annotated heartbeats, *WFDB Toolbox* for Python is applied to

the ECG signals. This toolbox finds the QRS complex of each beat on the signal, separates heartbeats from the signal, and categorizes them according to their arrhythmia types. An example drawing of the ECG signals in the MIT-BIH database, and the segmented 2D heartbeat images are illustrated in Fig. 2. ECG records and the number of ECG beats for each arrhythmia type are shown in Table 1. After completing the segmentation, each ECG beat is converted into ECG image formation.

2.3. 2D signal-to-image transformation

CNN is mainly used for analyzing 2D data since it automatically learns the optimal features from raw image data. Based on this popular advantage, a 2D CNN model is proposed in this study for ECG arrhythmia classification. After the segmentation of ECG signals into heartbeats, each beat is transformed into image formation by plotting the time-amplitude waveform and saving it as an image in Python. CNN model directly processes the input images while noisy ECG beats do not affect the performance of classification. Furthermore, the model automatically extracts robust features that eliminate the manual long-duration feature extraction step. Examining ECG images for arrhythmia detection by CNN is much similar to expert examination. Experts analyze ECG signals according to their observation of the ECG graph and diagnose arrhythmia type. They use the time amplitude behavior of ECG signals. The other advantage of image formation is increasing data size for CNN architectures. In general, deep learning methods need a large amount of data for successfully analyzing and classifying them. Through beat segmentation and image transformation steps, the total size of the dataset is increased. After completing the image transformation, each heartbeat image is converted into a *64×64-pixel* grayscale image. There are two main reasons for the image size tune to *64x64*. The first is to choose the lowest resolution that can represent the lines of the ECG signal drawing without any degeneration, while at the same time avoid increasing the training cost for the deep learning architecture where thousands of ECG images are trained. The other is to avoid negative dimensions in the pooling layers of the deep network. Moreover, during the training of the proposed model, various

image sizes are tested to yield the best performance. Further, all images are generated with the same color and converted into grayscale since there is no need to process the RGB color information in the CNN model. The RGB color depth of ECG images is not a significant feature to differentiate arrhythmia types [38]. Thus, grayscale image conversion provides a decreased dimension of images and reduces the computational complexity, and the classification time.

2.4. Convolutional neural network

CNN is a deep learning algorithm based on artificial neural network structures. Conventional machine learning techniques include three layers which are the input layer, one hidden layer, and the output layer. An artificial neural network consists of more than three layers that have several hidden layers in its structure. The structure is inspired by the brain working system that includes many hidden layers. In the hidden layer which includes many neurons, the input is transformed into something that the output layer can use. Neurons provide feature detection from the input data. The mathematical representation of artificial neuron is defined as;

$$y_j = f_j(x) = \varphi\big(w_j, x + b_j\big) \tag{1}$$

where f_j is a function of the input x weighted by a vector of connection weights w_j completed by a neuron bias b_j, and associated with an activation function φ. The schematic diagram of an artificial neuron is visualized in Fig. 3.

Artificial neural networks have been used in many different areas as computer vision, speech recognition, natural language processing, bioinformatics, drug design, and medical image analysis. CNN is a part of artificial neural networks especially designed for analyzing 2D data like images or videos. In contrast to conventional machine learning algorithms, CNN architectures do not need to extract hand-crafted features from the raw data. Both feature extraction and classification parts are embedded in the architecture and so automatically identify the robust features from the input data [19]. CNN has three characteristic layers; convolution, pooling, and fully connected layer.

In the convolution layer, input samples are convolved with a specific kernel. Many features are provided by moving the specific kernel [30]. The equation of discrete convolution function is defined as;

$$(f * g)(x) = \sum_t f(t)g(x + t) \tag{2}$$

where f and g are two functions. For 2D signals like images, the equation is changed as;

$$K * G(i, j) = \sum_{m,n} K(m, n)G(i + m, j + n) \tag{3}$$

where **K** is a convolution kernel, **G** is a 2D signal. The convolution process provides to extract effective features from the input. The pooling layer is used for reducing the dimension of the input sample by keeping the optimal features. In the fully connected layer, all neurons of the current layer are inter-connected to the neurons in the next layer. As such, the results of the convolution and pooling layers are used for classification. Between convolution and fully connected layers, there is a flattened layer where multidimensional feature vectors are transformed into 1D output vectors [39]. Also, in the fully connected layer, data is provided from the flattened layer and the learning process is realized through the neural network. The SoftMax function is used in the last layer to classify each ECG arrhythmia class. When the training of CNN is complete, the model is created for the classification.

In this study, a novel CNN model is designed, inspired by the LeNet model which is a well-known CNN architecture. The proposed model has three convolution layers, three pooling layers, and a fully connected layer. Maximum pooling is implemented as the pooling layer. It selects only the maximum value within the feature matrix obtained by convolution filters which provide reducing the number of output neurons. To specify the output values of the kernel in the layer, the activation function is used. The rectified linear unit (ReLU) is utilized in the proposed model as an activation function which is defined in (4).

$$f(x) = \max(0, x) \tag{4}$$

Here **x** is an element of the output kernel after the convolution or pooling layer. The proposed model structure is illustrated in Fig. 4.

2.5. Training and testing

The cross-validation is usually not processed because of training costs and time duration in deep learning approaches. The deep learning-based studies need to process huge datasets so the validation split technique can be able to provide the function of n-fold-cross-validation. Despite this, the 5-fold cross-validation method was used in addition to the validation split process to observe the performance of the study. In the random validation split process, the dataset is separated into two parts; 80% of total input data is using for training and remains is using for testing. Then, the part reserved for training was applied to a 5-fold cross-validation process. Hence, the test data is included 7519 ECG beat images and label distribution of ECG beats in test sample size found as 1664, 1423, 1604, 1440, and 1388 for Normal (0), PB (1), LBBB (2), RBBB (3), and PVC (4) classes, respectively. These ECG images are randomly split as training or test samples to achieve balanced distribution. In the learning or training, a proposed model training sample set is utilized, and the test sample set is utilized to validate the proposed model. In the training phase, standard backpropagation with a batch size of 64 is implemented for stochastic learning. The weights are updated according to the following equation;

$$w_l = \left(1 - \frac{m_\lambda}{s}\right) w_{l-1} - \frac{m}{x} \frac{\partial c}{\partial w} \tag{5}$$

where **w** is weight, **l** is layer number, **n** is learning rate, **λ** is the regulation parameter, **s** is the total number of training samples, **x** is the batch size, and **c** is the cost function. In addition, the biases are updated according to the following equation;

$$b_l = b_{l-1} - \frac{n}{x} \frac{\partial c}{\partial w} \tag{6}$$

In the proposed algorithm, the learning rate is chosen as 0.001. Adam optimizer is used for optimization, and cross-entropy is selected for the loss function calculation. After all the training

epochs are completed, the proposed algorithm performs a test on the CNN model. The summary of the proposed CNN architecture with layer parameters is shown in Table 2.

The proposed CNN model is compared with LeNet and ResNet-50 architectures to evaluate the accuracy rate of the study. LeNet architecture contains two convolution layers, two pooling layers, and a fully connected layer. As the proposed CNN architecture mimics the LeNet, the summary of LeNet architecture with layer parameters is also demonstrated in Table 3.

Table 2 Summary of proposed CNN architecture.

Layers	Type	Layer Parameters	Input Size
Layer 2	Conv2D	Kernel size=2×2 Kernel= 20 Stride= 1	64×64×1
Layer 3	Max-pooling	Kernel size=2×2 Kernel= 20 Stride= 2	64×64×20
Layer 4	Conv2D	Kernel size=2×2 Kernel= 50 Stride= 1	32×32×20
Layer 5	Max-pooling	Kernel size=2×2 Kernel= 50 Stride= 2	32×32×50
Layer 6	Conv2D	Kernel size=2×2 Kernel= 100 Stride= 1	16×16×50
Layer 7	Max-pooling	Kernel size=2×2 Kernel= 100 Stride= 2	16×16×100
Layer 8	FC	Kernel=500	8×8×100
Layer 9	Output	Kernel=5	500

Table 3 Summary of LeNet architecture.

Layers	Type	Layer Parameters	Input Size
Layer 2	Conv2D	Kernel size=2×2 Kernel= 16 Stride= 1	64×64×1
Layer 3	Max-pooling	Kernel size=2×2 Kernel= 16 Stride= 2	64×64×16
Layer 4	Conv2D	Kernel size=2×2 Kernel= 32 Stride= 1	32×32×16
Layer 5	Max-pooling	Kernel size=2×2 Kernel= 32 Stride= 2	32×32×32
Layer 6	FC	Kernel=500	16×16×32
Layer 7	Output	Kernel=5	500

2.6. Performance evaluation metrics

The performance of the proposed model is evaluated utilizing various metrics which are accuracy (ACC), specificity (SPE), recall (REC), precision (PRE), and F1-Score [40]. The ACC is indicated as the total number of correctly classified ECG beat images divided by the total number of test images. The accuracy-based performance evaluation is performed for machine and deep learning algorithms,

which is not sufficient in the case of imbalanced labeled testing and training sets. Hence, additional performance metrics are included in the evaluations. The REC metric is referred to as a true positive rate which is predicted as positive, while the SPE metric is referred to as a true negative rate which is predicted as negatives. The PRE is referred to as the proportion of classified positive cases that are correctly real positives. The F1-Score metric is calculated by evaluating the harmonic mean of PRE and REC values. These performance metrics are calculated as follows;

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$SPE = \frac{TN}{TN+FP} \tag{8}$$

$$REC = \frac{TP}{TP+FN} \tag{9}$$

$$PRE = \frac{TP}{TP+FP} \tag{10}$$

$$F1\ Score = \frac{2\ x\ REC\ x\ PRE}{REC+PRE} \tag{11}$$

The confusion matrix is provided to express the number of true positives (**TPs**), true negatives (**TNs**), false positives (**FPs**), false negatives (**FNs**), and test sample size. The above five evaluation metrics are generated after the calculation of the confusion matrix process. Here, the **FPs** referred to the number of test samples for different classes but are incorrectly assigned for actual class, while **FNs** referred to the number of test samples for actual class, but are incorrectly assigned for different classes. **TPs** are the number of test samples that belong to a class and are correctly indicated as the same class and **TNs** are the number of test samples for different classes that are correctly predicted by the proposed model.

The Receiver Operating Characteristic (ROC) curve serves as a graphic presentation of the trade-off between sensitivity (the false-positive rates) and specificity (the false-negative rates). The percentage of false-positive (1-specificity) is indicated by the x-axis, goes from 0 to 1 (0-100%) and the percentage of false negatives is indicated by the y-axis, goes from 0 to 1 (0-100%) on the representation of the

ROC curve. When the ideal values (100% sensitivity and 100% specificity) are provided, the point on the ROC curve would be at the upper left-hand corner (0,1) which means that the better the test is at discriminating between cases and non-cases. The area between ROC and the axes calculated as the Area Under the ROC Curve (AUC). It can get any value between 0 and 1 as both x- and y-axis have values goes from 0 to 1. The fact that AUC is closer to 1 indicates that the better the overall performance of the test. The AUC is an appropriate metric to examine algorithm performance as it is independent of the prediction criterion selected. Also, in order to observe the training result clearly, the average of 5 training ACC values obtained after 5-fold cross-validation was calculated. The performance of the proposed model has been evaluated thanks to the performance evaluation metrics mentioned above.

3. Results

In this study, *Keras* and *TensorFlow* libraries are used as the backend for deep learning algorithms. All experiments are carried out on a computer with Nvidia GeForce RTX 2080 Ti GPU and 64 GB RAM by using *Cuda 10.1*, *cudart64_101.dll*, and *Tensor Flow 2.0*.

The proposed 2D CNN-based study is separated into two phases; training (the training data is divided into 5 parts in itself for cross-validation) and testing. The data in the testing phase are those never used in the training phase. The proposed network is trained by the training dataset which contains *30072* samples and validated by a *7519*-sample test set. In Fig. 5, the training and testing performance of the proposed CNN architecture is given in comparison with LeNet and ResNet-50 architectures. Fig. 5 (c) demonstrates the training and validation performance of the proposed architecture about zero training loss, 100% training accuracy, 0.0273 validation loss, and 99.7% validation accuracy. The fact that the training accuracy and validation accuracy are being parallel to each other and converge to 100%, and the training loss and validation loss are being parallel to each other, converges to zero are indicators that the model is not exposed to overfitting during the testing and training phase. Furthermore, Fig. 5

(f) represents the confusion matrix and Fig. 5 (i) represents the ROC curves for the testing phase of the study. The confusion matrix consists of ground truth and predicted labels, where the column is the true label, and the row is the predicted label. The classification results of the proposed CNN model are satisfactory, as observed by the high number of correct answers in the blue cells and the low number of incorrect answers in the white cells. According to the confusion matrix of the proposed study, the correctly estimated number of arrhythmias (TPs), the number of arrhythmias estimated in the false class (FNs), and the success of the model in each class can be calculated. The TPs, FNs, and accuracy for the PB class are obtained as 1417, 6, and 99.58%, respectively. These values are calculated as 1646, 18, and 98.92% for the N class, respectively. Also, these values obtained as 1595, 9, and 99.44% for LBBB class; 1432, 8, and 99.44% for RBBB class and 1371, 17, and 98.78%, for PVC class, respectively. Standard performance metrics such as SPE, PRE, REC, F1-Score, and ACC are used to evaluate the performance of the 2D CNN model. Additionally, various epoch and batch sizes are examined to obtain the highest classification performance. The proposed model is first created with 90 epochs and 256 batch sizes and resulted in 92% accuracy. 100 epochs and 128 batch sizes yielded 96% accuracy. The model yielded the highest accuracy by increasing the epoch number and decreasing the batch size. Nevertheless, increasing the epoch size beyond 150 did not affect accuracy. Thus, the epoch was chosen as 150 to avoid increasing the training cost. Moreover, decreasing the batch size to less than 64 also reduced accuracy. For this reason, the batch size is tuned to 64.

In addition to examining the effect of epoch and batch size changes on the performance of the model, the effect of changing input image sizes, changing learning rate hyperparameter, and the number of model layers on the accuracy was also examined. A visualization of the effect of some parameter changes on accuracy is given in Fig. 6. Note that, in order to tune the optimum parameter, the relevant parameter is changed while the other parameters are fixed to the values where the best accuracy is obtained. Image sizes are chosen to square sizes and a power of 2 in all trials. The main reason for this is to avoid the padding process in the convolutional and pooling layers. Thus, the image size was

chosen as *256x256* and an accuracy of 99.7 was obtained. Afterward, the model was trained with input image sizes at *128x128*, *64x64*, *32x32*, and *16x16*, respectively. The highest accuracy has been achieved with *64x64*, *128x128*, and *256x256* image sizes. Similar to the epoch size selection, the image size was chosen as *64x64* in order to avoid increasing the training cost. Also, the reason why image size lower than 16 was not selected; to keep the resolution at a sufficient level to avoid distortion of the ECG images and to avoid negative dimension when applying the pooling process in the training phase. Additionally, for the ResNet-50 architecture, the lowest image size that can be used as an input without any padding is *64x64*. Another reason why image size was chosen as *64x64* is to compare the trained models as equal conditions as possible. In order to observe the effect of the learning rate hyperparameter on training performance, the learning rate was chosen as 0.01, 0.001, and 0.0001, respectively. Experimental results show that when the learning rate is tuned to 0.001, the highest accuracy is yielded.

Finally, in addition to conducting experimental studies to tune the number of layers in the proposed architecture, the obtained results by training ResNet-50 (more layers) and LeNet (fewer layers) architectures were considered. Therefore, it is adopted to yield higher accuracy with fewer layers (lower training cost). In the proposed model layers containing convolution and subsequent max-pooling with 20, 50, 100, 150, and 200 kernels, respectively, were tested by freezing some layers. The highest performance was yielded in 3-layer architecture. In the parameter tuning process, the principle of providing the lowest cost has been adopted while creating the lightest system. Finally, 150 epoch sizes, 64 batch sizes, 3-layer architecture, 0.001 learning rate, and 64x64 image sizes are tuned to be optimum for the proposed experiment. With these tuned parameters, an outstanding multiclass arrhythmia classification performance with an accuracy of 99.7 was yielded. Detailed performance evaluation metrics of the study are demonstrated in Table 4.

Table 4 Classification performances of the trained models.

Metrics	ResNet-50	LeNet	Proposed

Classifier Type	5 different multi-class		
The Best Training ACC (%)	100	99.9	100
Avg. Cross-Val. ACC (%)	99.16	98.8	99.9
Test ACC (%)	98.28	97.89	99.7
F1-Score (%)	98.11	97.78	99.24
SPE (%)	98.30	97.54	99.22
REC (%)	98.12	98.07	99.7
PRE (%)	98.26	97.88	99.5
Avg. Multi-AUC	0.98	0.97	1.00
Mean Squared Error	0.107	0.131	0.049
Training Loss	~0	0.073	~0
Validation Loss	0.0548	0.0281	0.0273

4. Discussion

This study presents a new method for the ECG arrhythmia classification using 2D ECG signal-to-image representations with the deep learning approach. In our previous study [38], five different arrhythmia types were differentiated using ECG signal and LeNet architecture with an accuracy of 97.42%. The five different heartbeats are categorized as non-ectopic beats, ventricular ectopic beats, fusion beats, supraventricular ectopic beats, and unclassifiable beats. However, the label distribution between heartbeat categories was not balanced. In this study, almost equally distributed (IR= 1.16) five arrhythmias are classified based on a proposed 2D CNN architecture. In addition, while the previous study was built on a traditional CNN architecture, a new architecture that is not as deep and does not have a high computational complexity is proposed in this study. The performance of the proposed model is evaluated with ECG signals from the MIT-BIH database. The balanced label distribution of datasets is very important for deep learning-based studies. The case where the number of heartbeats is not equal for each class causes an imbalanced dataset that affects the accuracy performance in deep learning approaches. In order to eliminate this problem, the balanced label distribution is achieved by determining arrhythmia classes containing an approximately equal number

of heartbeats from the MIT-BIH dataset. The classification success of the balanced data distribution may be observed in this outstanding classification performance.

The deep learning structures may be converted to deeper networks by appending different types of layers to increase the classification performances. The deeper networks have complex structures hence their training time duration may be increased to days. The training time is a crucial issue in the design of automatic classification systems. The performances of the complex multi-layered structures can be evaluated considering the above parameters. The LeNet architecture consists of 7 layers which are 2 convolutional layers, 2 max-pooling layers, and a fully connected layer. The designation of a deeper version of the LeNet structure is aimed at the implementation of effective layers to obtain higher classification achievements. In this study, a novel CNN structure that mimics LeNet has achieved 99.7% accuracy for classifying different arrhythmias. Moreover, well-known CNN structures ResNet-50 which includes 70 layers, and LeNet that consists of only 7 layers are also employed to classify the 2D ECG beat images to compare their results with the proposed model and evaluate the complexity of these models. Although the complexity of ResNet-50 is higher than the LeNet based model, the proposed model yielded higher accuracy than ResNet-50. Comparison of arrhythmia classification results of the proposed study with previous studies is summarized in Table 5.

Many studies in the literature achieved a lower accuracy rate than the proposed approach, although they involve many deep network layers. Recently, Li *et al*. [24] proposed ResNet-31 based ECG heartbeats classification using five classes of single-lead and 2-lead datasets and achieved 99.06% and 99.38% accuracy for training and testing respectively. In another study [41], ECG fragments, and RR interval of six different arrhythmia types are differentiated based on 5-fold cross-validation using 12 layers CNN architecture and two layers of LSTM network with an accuracy of 99.32%. In other studies, 11 layers CNN [32], 10 layers CNN+LSTM [36], 11 layers CNN [28], 9 layers CNN [42], are proposed to differentiate different arrhythmias with an accuracy of over 90%. The proposed approach improves the classification performance with less network layer depth (9 layers CNN). On the other

18

hand, the average training time for each epoch is increased due to deeper layers in the above studies. This CNN architecture-based classification shortens the average training time per epoch due to less network depth. Compared to existing studies, the proposed method achieved high classification performance with a low-depth network, very fast trainable pattern coefficients that take up less space which makes it less complex, and thus may perform better in portable ECG devices. The faster and higher accuracy classification with fewer layers than other studies was achieved in all classes using the proposed approach. Also, compared to the ResNet-based architectures used in [24], they achieved similar accuracy with the proposed study. Outperforming state-of-the-art DL-based approaches, a five-class classification method is presented using 2D ECG beats in grayscale and alleviating imbalanced data distribution. Hence, the proposed study employed a simple and effective CNN approach to differentiate the heartbeat arrhythmias.

Table 5 Performance comparison table with recent studies.

Work	Year	Database	Arrhythmia Types	Beat Count / IR	Classifier	Performance
Chen *et al*. [41]	2020	MITDB	Normal, Atrial Fibrillation, Ventricular Bigeminy, Pacing Rhythm, Atrial Flutter, Sinus Bradycardia	Total: 3027 Test: 605 IR: 21	CNN LSTM	ACC: 99.32% SEN: 97.75% SPE: 99.51%
Huang *et al*. [32]	2019	MITDB	Normal, PVC LBBB, RBBB, APB	Total: 2520 Test: 420 IR: 1.5	CNN	ACC: 99% N/A N/A
Oh *et al*. [36]	2018	MITDB	Normal, PVC LBBB, RBBB, APB	Total: 16499 Test: 1650 IR: 23.97	CNN LSTM	ACC: 98.10% SEN: 97.50% SPE: 98.70%
Yildirim *et al*. [43]	2018	MITDB	Normal, PVC LBBB, RBBB, PB	Total: 7376 Test: 1476 IR: 4.29	Bi-LSTM	ACC: 99.39% N/A N/A
RA-Mahfuz et al. [45]	2021	MITDB	Normal, PVC, LBBB, RBBB, PB	Total: 38000 Test: 1000 IR: 1.24	CNN	ACC: 99.9% SEN: 99.9% SPE: 99.98% AUC: 0.999
Sahoo *et al*. [8]	2017	MITDB	Normal, LBBB RBBB, PB	Total: 109494 Test: 1071 IR: 1.19	SVM	ACC: 98.39% SEN: 96.86% SPE: 98.92%
Acharya *et al*. [42]	2017	MITDB	Normal, Fusion, Unknown Supraventricular Ectopic Ventricular Ectopic	Total: 109449 Test: 10945 IR: 112.96	CNN	ACC: 94.03% SEN: 96.71% SPE: 91.54%
Zubair *et al*. [44]	2016	MITDB	Normal, Unknown Supraventricular Ectopic,	Total: 100389 Test: 88752	CNN	ACC: 92.70% N/A

			Ventricular Ectopic, Fusion	IR:7927		N/A
Kiranyaz *et al.* [25]	2015	MITDB	Normal, Unknown, Fusion Supraventricular Ectopic Ventricular Ectopic	Total: 83648 Test: N/A IR: 5754.23	CNN	ACC: 99% SEN: 93.90% SPE: 98.90%
Li *et al.* [24]	2020	MITDB	Normal, Supraventricular Ectopic, Ventricular Ectopic, Fusion, Unknown	Total: 94013 Test: 18803 IR: 100.26	CNN	ACC: 99.38% SEN: 94.54% SPE: 98.14%
This Work	**2021**	**MITDB**	**Normal, PVC LBBB, RBBB, PB**	**Total: 37339 Test: 7519 IR: 1.16**	**CNN**	**ACC: 99.70% SEN: 99.70% SPE: 99.22% AUC: 1.00**

The experimental results show that the proposed model outperforms the accuracy of LeNet by utilizing a less complex structure. Accuracy-loss graphs obtained from training ResNet-50 and traditional LeNet architecture are shown in Fig. 5 (a) and Fig. 5 (b), respectively. The ROC curves are shown in Fig. 5 (d) and Fig. 5 (e), respectively. Finally, the confusion matrixes are shown in Fig. 5 (g) and Fig. 5 (h), respectively. It can be observed in Fig. 5 (a) that the ResNet-50 architecture has a high validation loss fluctuation. This is thought to be caused by shortcuts in residual networks. Besides, after the 80th epoch, the loss values converged to zero. In Fig. 5 (b), although the accuracy-loss drawings obtained from the LeNet architecture seem more stable, they could not exceed a certain accuracy value. The most stable accuracy-loss graph is obtained by the proposed architecture in Fig. 5 (c). Similarly, the ROC Curves in Fig. 5 (d, e, f), are examined, it is seen that the best ideal result is obtained with the proposed architecture in Fig. 5 (f). Finally, the confusion matrices in Fig. 5 (g, h, i) prove the robustness of the proposed 2D ECG image transformation method. Ultimately, the results demonstrated that the proposed model, ResNet, and LeNet yielded an accuracy of 99.7%, 98.28%, and 97.89%, respectively. These accuracy values show that ResNet and LeNet achieve lower classification success than the proposed model. Thus, it was that the harmony between the higher performances and the complexity of the structure is related to the selection of effective layer types that are in the correct order, optimal filter dimensions, and other training parameters. The developed automatic cardiac arrhythmia detection algorithm improves diagnostic efficiency and accuracy while reducing the training time. The innovative contribution of the proposed study may be emphasized as follows:

- This paper proposes a novel deep learning approach for identifying different arrhythmia types utilizing 2D ECG beats that are obtained from 1D ECG signals by a signal-to-image transformation procedure.

- The benefits of using 2D gray-scale images on the proposed CNN structure were demonstrated.

The proposed method,

- does not require any complex pre-processing of ECG signals, and QRS complex determination to perform classification,

- does not contain any manual computationally demanding feature extraction steps as in the traditional machine learning methods,

- investigates the performance improvements by using a novel 2D CNN-based model in the classification of arrhythmia types compared to well-known CNN approaches such as LeNet and ResNet-50,

- increases classification performance while decreasing the computational cost compared to the well-known CNN architectures for mobile-based decision-making systems,

- optimizes the deep network hyperparameters to yield the best classification performance.

5. Conclusion

This paper proposes a novel approach for the accurate classification of ECG arrhythmias based on 2D CNN architecture. ECG heartbeats are transformed into 2D time-amplitude images to be used as input data for CNN architecture. This study demonstrates that a combination of simple ECG time-amplitude images, and the image classification capability of the CNN architecture may provide the highest classification performance with the elimination of manual preprocessing steps such as noise filtering,

feature extraction, and feature selection. The proposed novel 2D CNN architecture includes relatively fewer layers compare to other CNNs and has yielded better performance compared to the well-known CNN architectures. Additionally, the proposed model was generated with balanced label distribution to minimize the imbalanced dataset effect. Considering the experimental results, the proposed method is a simple, effective, and useful approach that may be used by experts for quickly and automatically identifying cardiovascular problems on ECG signals. In future studies, the proposed algorithm may be implemented for home healthcare monitoring systems, by combining mobile applications and portable ECG devices, that will automatically detect arrhythmias in real-time and share them with physicians.

Abbreviations

ACC, Accuracy; ANN, Artificial Neural Network; ANNC, Adaptive Neural Network Classifier; ATI-CNN, Attention-based Time-Incremental Convolutional Neural Network; APB, Atrial Premature Beat; AUC, Area Under the ROC Curve; AV, Atrioventricular Node; CAD, Computer-Aided Diagnosis; CNN, Convolutional Neural Network; DL, Deep Learning; DNN, Deep Neural Network, ECG, Electrocardiogram, FAV, First Degree AV Block; FN, False Negative; FP, False Positive; IR, Imbalance Ratio; LBBB, Left Bundle Branch Block; LDA, Linear Discriminant Analysis; LSTM, Long Short-Term Memory; MLII, Modified Limb Lead II; MLP, Multilayer Perception; N, Normal; PB, Paced Beat; PRE, Precision; PVC, Premature Ventricular Contraction; RBBB, Right Bundle Branch Block; REC, Recall; ReLU, Rectified Linear Units; RF, Random Forest; ROC, Receiver Operating Characteristic; SA, Sinoatrial Node; SPE, Specificity; STFT, Short-Time Fourier Transform; SVM, Support Vector Machine; TF, Time Frequency; TN, True Negative; TP, True Positive; VMD, Variational Mode Decomposition.

Human and animal rights

The authors declare that the work described has been carried out in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans.

Informed consent and patient details

The authors declare that this report does not contain any personal information that could lead to the identification of the patient(s).

Authors' contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

CRediT authorship contribution statement

M. Degirmenci: Conceptualization, Data Collection, Writing-Original Draft Preparation, Writing-Review & Editing. **M.A. Ozdemir:** Conceptualization, Methodology, Software, Visualization, Writing-Original Draft Preparation, Writing-Review & Editing, Funding Acquisition. **E. Izci:** Conceptualization, Data Collection, Writing-Original Draft Preparation, Writing-Review & Editing. **A. Akan:** Methodology, Writing-Review & Editing, Funding Acquisition. All authors, with equal contributions, edited the paper and gave final approval for the version to be published.

Availability of data and materials

The data used in this work is publicly available from https://physionet.org/content/mitdb/1.0.0/.

Declaration of competing interests

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

References

[1] Agrafioti F. ECG in Biometric Recognition: Time Dependency and Application Challenges. vol. NR93098. Citeseer; 2011.

[2] Sahoo JP. Analysis of ECG signal for Detection of Cardiac Arrhythmias. 2011.

[3] Luz EJ da S, Schwartz WR, Cámara-Chávez G, Menotti D. ECG-based heartbeat classification for arrhythmia detection: A survey. Comput Methods Programs Biomed 2016;127:144–64. https://doi.org/10.1016/j.cmpb.2015.12.008.

[4] Ozdemir MA, Guren O, Cura OK, Akan A, Onan A. Abnormal ECG Beat Detection Based on Convolutional Neural Networks. 2020 Med. Technol. Congr., Antalya, Turkey: IEEE; 2020, p. 1–4. https://doi.org/10.1109/TIPTEKNO50054.2020.9299260.

[5] Izci E, Ozdemir MA, Sadighzadeh R, Akan A. Arrhythmia Detection on ECG Signals by Using Empirical Mode Decomposition. 2018 Med. Technol. Natl. Congr., Magusa: IEEE; 2018, p. 1–4. https://doi.org/10.1109/TIPTEKNO.2018.8597094.

[6] Nanjundegowda R, Meshram VA. Arrhythmia detection based on Hybrid features of T-wave in Electrocardiogram. Int J Intell Eng Syst 2018;11:153–62. https://doi.org/10.22266/ijies2018.0228.16.

[7] Houssein EH, Kilany M, Hassanien AE. ECG signals classification: a review. Int J Intell Eng Informatics 2017;5:376. https://doi.org/10.1504/ijiei.2017.087944.

[8] Sahoo S, Kanungo B, Behera S, Sabut S. Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. Meas J Int Meas Confed 2017;108:55–66. https://doi.org/10.1016/j.measurement.2017.05.022.

[9] Lin CC, Yang CM. Heartbeat classification using normalized rr intervals and wavelet features. 2014 Int Symp Comput Consum Control IS3C 2014 2014;2014:650–3. https://doi.org/10.1109/IS3C.2014.175.

[10] Sharma RR, Kumar M, Pachori RB. Joint Time-Frequency Domain-Based CAD Disease Sensing System Using ECG Signals. IEEE Sens J 2019;19:3912–20. https://doi.org/10.1109/JSEN.2019.2894706.

[11] Khalaf AF, Owis MI, Yassine IA. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. Expert Syst Appl 2015;42:8361–8. https://doi.org/10.1016/j.eswa.2015.06.046.

[12] Mert A. ECG feature extraction based on the bandwidth properties of variational mode decomposition. Physiol Meas 2016;37:530–43. https://doi.org/10.1088/0967-3334/37/4/530.

[13] Moridani MK, Abdi Zadeh M, Shahiazar Mazraeh Z. An Efficient Automated Algorithm for Distinguishing Normal and Abnormal ECG Signal. IRBM 2019;40:332–40. https://doi.org/10.1016/j.irbm.2019.09.002.

[14] Li T, Zhou M. ECG classification usingwavelet packet entropy and random forests. Entropy 2016;18:285. https://doi.org/10.3390/e18080285.

[15] Pandey SK, Janghel RR. ECG Arrhythmia Classification Using Artificial Neural Networks. Lect. Notes Networks Syst., vol. 46, Springer; 2019, p. 645–52. https://doi.org/10.1007/978-981-13-1217-5_63.

[16] Mjahad A, Rosado-Muñoz A, Bataller-Mompeán M, Francés-Víllora J V., Guerrero-Martínez JF. Ventricular Fibrillation and Tachycardia detection from surface ECG using time-frequency representation images as input dataset for machine learning. Comput Methods Programs Biomed 2017;141:119–27. https://doi.org/10.1016/j.cmpb.2017.02.010.

[17] Jun TJ, Nguyen HM, Kang D, Kim D, Kim D, Kim Y-H. ECG arrhythmia classification using a 2-D convolutional neural network. ArXiv Prepr ArXiv180406812 2018.

[18] Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.

[19] Yıldırım Ö, Pławiak P, Tan RS, Acharya UR. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput Biol Med 2018;102:411–20. https://doi.org/10.1016/j.compbiomed.2018.09.009.

[20] Zhang C, Wang G, Zhao J, Gao P, Lin J, Yang H. Patient-specific ECG classification based on recurrent neural networks and clustering technique. Proc. 13th IASTED Int. Conf. Biomed. Eng. BioMed 2017, Institute of Electrical and Electronics Engineers Inc.; 2017, p. 63–7. https://doi.org/10.2316/P.2017.852-029.

[21] Rahhal MM Al, Bazi Y, AlHichri H, Alajlan N, Melgani F, Yager RR. Deep learning approach for active classification of electrocardiogram signals. Inf Sci (Ny) 2016;345:340–54. https://doi.org/10.1016/j.ins.2016.01.082.

[22] Sannino G, De Pietro G. A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. Futur Gener Comput Syst 2018;86:446–55. https://doi.org/10.1016/j.future.2018.03.057.

[23] Pławiak P, Acharya UR. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. Neural Comput Appl 2020;32:11137–61. https://doi.org/10.1007/s00521-018-03980-2.

[24] Li Z, Zhou D, Wan L, Li J, Mou W. Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram. J Electrocardiol 2020;58:105–12. https://doi.org/10.1016/j.jelectrocard.2019.11.046.

[25] Kiranyaz S, Ince T, Gabbouj M. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. IEEE Trans Biomed Eng 2016;63:664–75. https://doi.org/10.1109/TBME.2015.2468589.

[26] Baloglu UB, Talo M, Yildirim O, Tan RS, Acharya UR. Classification of myocardial infarction with multi-lead ECG signals and deep CNN. Pattern Recognit Lett 2019;122:23–30. https://doi.org/10.1016/j.patrec.2019.02.016.

[27] Savalia S, Emamian V. Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks. Bioengineering 2018;5:35. https://doi.org/10.3390/bioengineering5020035.

[28] Acharya UR, Fujita H, Lih OS, Hagiwara Y, Tan JH, Adam M. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Inf Sci (Ny) 2017;405:81–90. https://doi.org/10.1016/j.ins.2017.04.012.

[29] Yao Q, Wang R, Fan X, Liu J, Li Y. Multi-class Arrhythmia detection from 12-lead varied-length ECG using Attention-based Time-Incremental Convolutional Neural Network. Inf Fusion 2020;53:174–82. https://doi.org/10.1016/j.inffus.2019.06.024.

[30] Kim MG, Ko H, Pan SB, Park K. A study on user recognition using 2D ECG image based on ensemble networks for intelligent vehicles. Wirel Commun Mob Comput 2019;2019:1–9. https://doi.org/10.1155/2019/6458719.

[31] Hatami N, Gavet Y, Debayle J. Classification of time-series images using deep convolutional neural networks. Tenth Int. Conf. Mach. Vis. (ICMV 2017), International Society for Optics and Photonics; 2018, p. 23. https://doi.org/10.1117/12.2309486.

[32] Huang J, Chen B, Yao B, He W. ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. IEEE Access 2019;7:92871–80. https://doi.org/10.1109/ACCESS.2019.2928017.

[33] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag 2001;20:45–50. https://doi.org/10.1109/51.932724.

[34] Shaker AM, Tantawi M, Shedeed HA, Tolba MF. Generalization of Convolutional Neural Networks for ECG Classification Using Generative Adversarial Networks. IEEE Access 2020;8:35592–605. https://doi.org/10.1109/ACCESS.2020.2974712.

[35] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med 2017;89:389–96. https://doi.org/10.1016/j.compbiomed.2017.08.022.

[36] Oh SL, Ng EYK, Tan RS, Acharya UR. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Comput Biol Med 2018;102:278–87. https://doi.org/10.1016/j.compbiomed.2018.06.002.

[37] Zhu R, Guo Y, Xue JH. Adjusting the imbalance ratio by the dimensionality of imbalanced data. Pattern Recognit Lett 2020;133:217–23. https://doi.org/10.1016/j.patrec.2020.03.004.

[38] Izci E, Ozdemir MA, Degirmenci M, Akan A. Cardiac Arrhythmia Detection from 2D ECG Images by Using Deep Learning Technique. 2019 Med. Technol. Congr., Izmir: IEEE; 2019, p. 1–4. https://doi.org/10.1109/TIPTEKNO.2019.8895011.

[39]   Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR. Deep learning for healthcare applications based on physiological signals: A review. Comput Methods Programs Biomed 2018;161:1–13. https://doi.org/10.1016/j.cmpb.2018.04.005.

[40]   Ozdemir MA, Degirmenci M, Izci E, Akan A. EEG-based emotion recognition with deep convolutional neural networks. Biomed Eng / Biomed Tech 2021;66:43–57. https://doi.org/10.1515/bmt-2019-0306.

[41]   Chen C, Hua Z, Zhang R, Liu G, Wen W. Automated arrhythmia classification based on a combination network of CNN and LSTM. Biomed Signal Process Control 2020;57:101819. https://doi.org/10.1016/j.bspc.2019.101819.

[42]   Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med 2017;89:389–96. https://doi.org/10.1016/j.compbiomed.2017.08.022.

[43]   Yildirim Ö. A novel wavelet sequences based on deep bidirectional LSTM network model for ECG signal classification. Comput Biol Med 2018;96:189–202. https://doi.org/10.1016/j.compbiomed.2018.03.016.

[44]   Zubair M, Kim J, Yoon C. An automated ECG beat classification system using convolutional neural networks. 2016 6th Int. Conf. IT Converg. Secur. ICITCS 2016, IEEE; 2016, p. 1–5. https://doi.org/10.1109/ICITCS.2016.7740310.

[45]   Rashed-Al-Mahfuz M, Moni MA, Lio' P, Islam SMS, Berkovsky S, Khushi M, et al. Deep convolutional neural networks based ECG beats classification to diagnose cardiovascular conditions. Biomed Eng Lett 2021;1:3. https://doi.org/10.1007/s13534-021-00185-w.
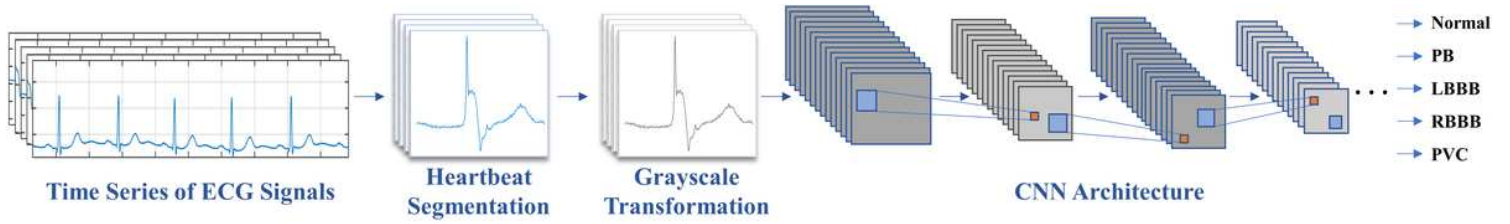
Figures

Figure 1

The graphical representation of the proposed methodology.
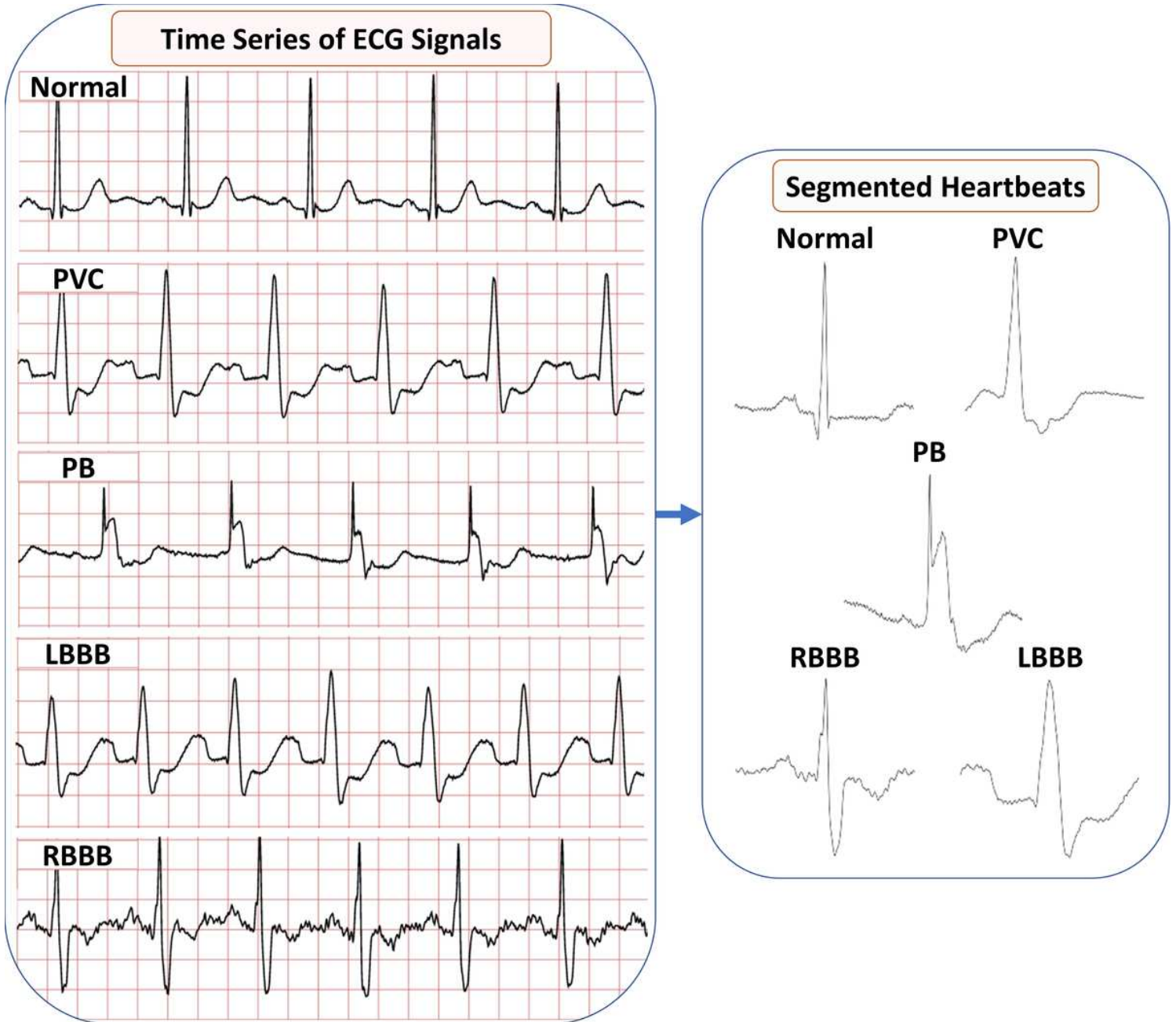

Figure 2

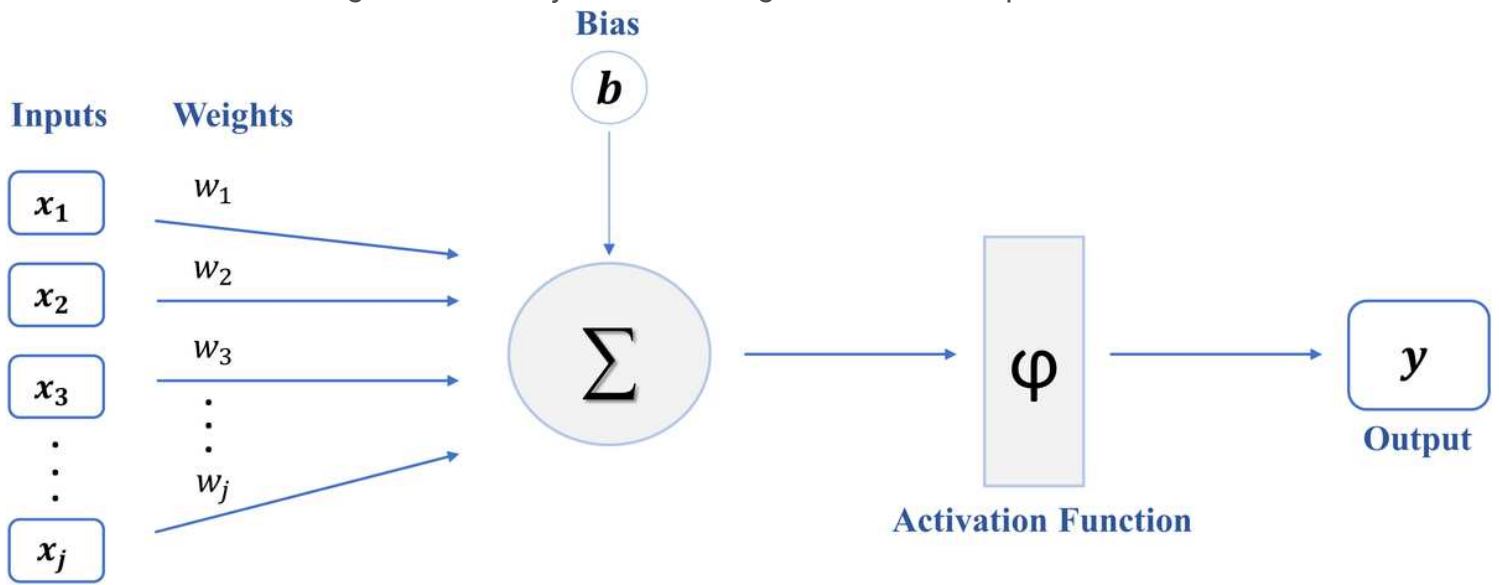Parts of different ECG signals and arrhythmic beat segmentation examples.

Figure 3

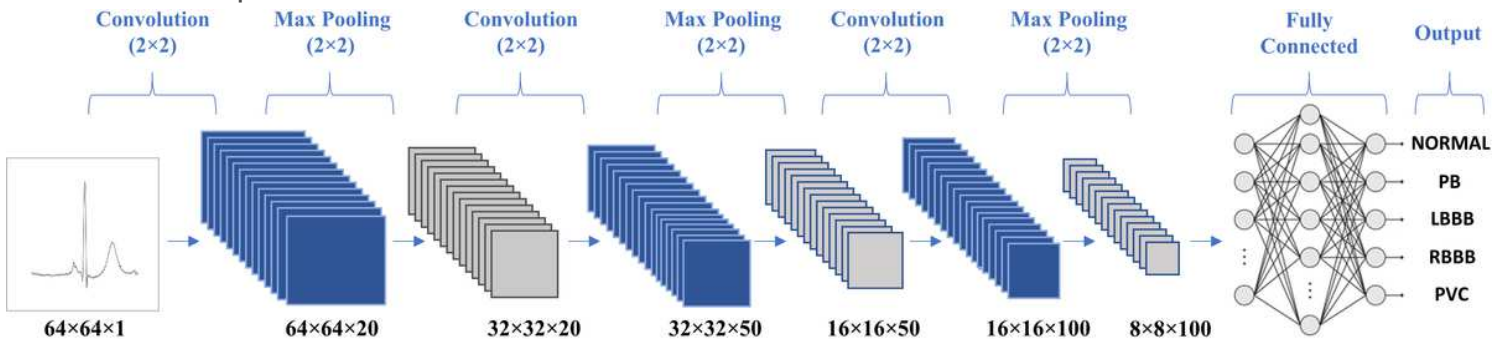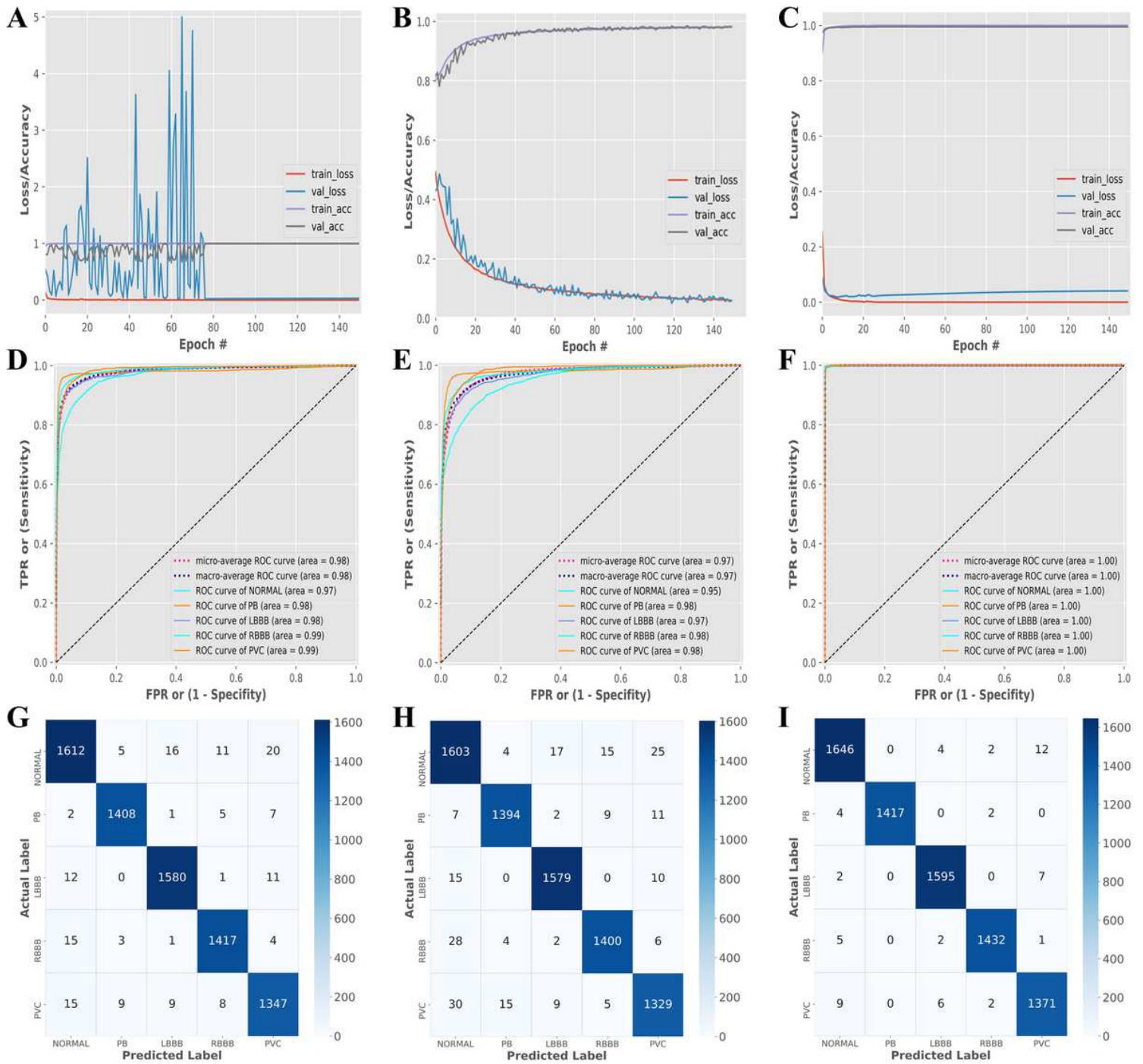Mathematical representation of a standard node in a DNN.

Figure 4

Schematic diagram of the proposed CNN architecture.

Figure 5

Performance metrics of the test phases; (a, b, c) graphs of accuracy and loss functions, (d, e, f) ROC curves, and (g, h, i) confusion matrices. The left column indicates ResNet-50, the mid column indicates LeNet, and the right column indicates the proposed architecture.
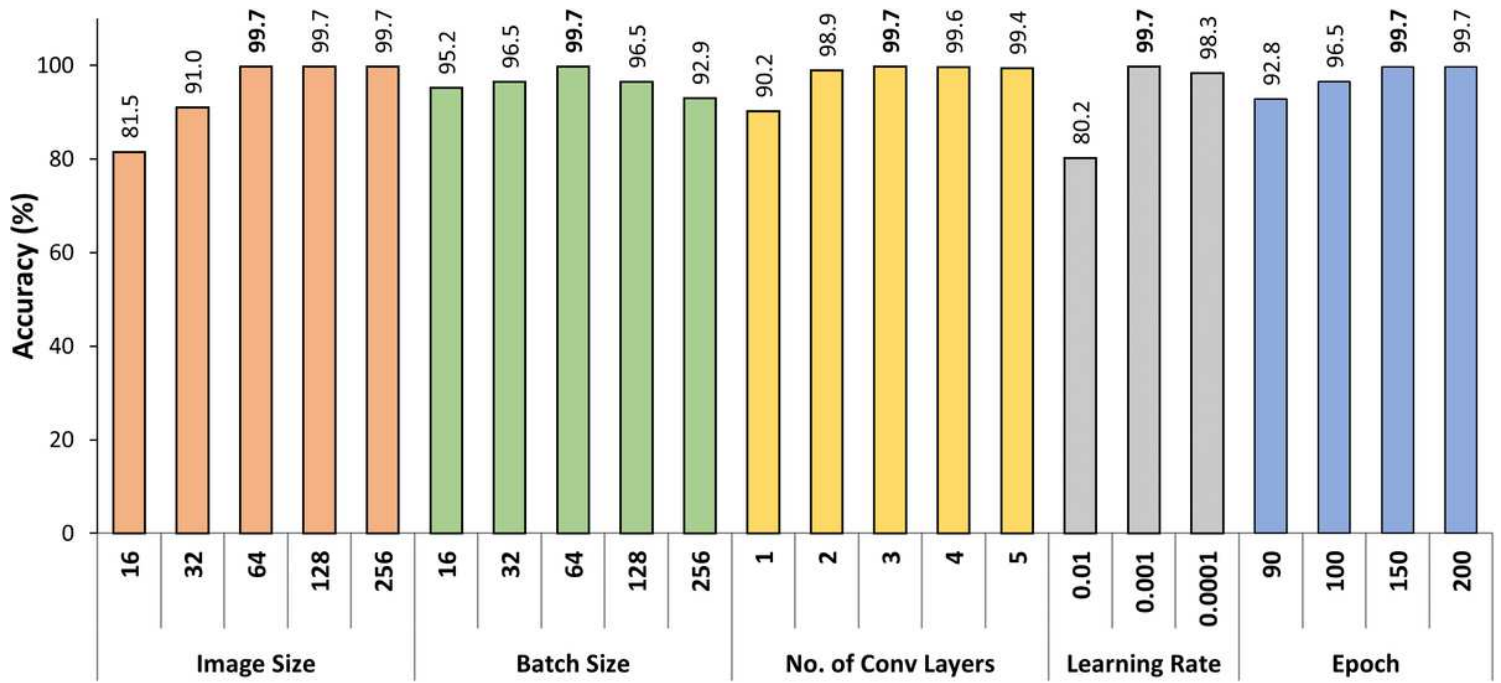
Figure 6

The effect of parameter variations on the performance of the proposed deep network.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- graphicalabstract.png