

ECG Heartbeat Classification based on Signal-to-Image Transformation Using Convolutional Neural Networks

Murside Degirmenci¹, Mehmet Akif Ozdemir^{2*}, Elif Izci¹, and Aydin Akan³

Abstract

Background: Electrocardiogram (ECG) is a method of recording the electrical activity of the heart and provides a diagnostic mean for heart-related diseases. An arrhythmia is any irregularity of heartbeat that causes an abnormality in one's heart rhythm. Early detection of arrhythmia has great importance to prevent many diseases. Manual analysis of ECG signal is not sufficient for quickly identifying arrhythmias that may cause sudden deaths. Hence, many studies have been presented to developed 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 signals are obtained from MIT-BIH arrhythmia database and are categorized according to five arrhythmia types. The proposed approach identifies arrhythmia classes by using Convolutional Neural Network (CNN) architecture trained by two-dimensional (2D) ECG beat images. CNN architecture is selected due to high image recognition performance. ECG signals are segmented into heartbeats, then each heartbeat is transformed into a 2D grayscale image. The heartbeat images are used as input for the CNN.

Results: The proposed CNN model is compared to other common CNN architectures such as LeNet and ResNet-50 to evaluate the performance of our study. Overall, the proposed study achieved 99.7% test accuracy in the classification of five different ECG arrhythmias.

Conclusions: Testing results demonstrate that CNN trained by ECG image representations provide outstanding classification performance of arrhythmic ECG signals and outperforms similar network architectures. Hence, the proposed approach provides a robust method for the classification of ECG arrhythmias.

Keywords: Arrhythmia, Classification, Convolutional Neural Networks, Deep Learning, ECG Beat Image, Electrocardiogram.

Background

A heartbeat is an event that occurs when the heart contracts and relaxes rhythmically. Electrical system of the heart is made up of two main parts: Sinoatrial (SA) node which is located in the right atrium of the heart and atrioventricular (AV) node that is located in the middle of the heart, between the atrium and the ventricle. The electrical trigger starts in the SA node travels from the atria to the AV node. The signal is delayed in the AV node and then spreads through the lower chambers. The ventricles contract, sending blood throughout the body, and heartbeat occurs. ECG is a tool used to observe electrical cardiac activity. Each heartbeat has a P wave, QRS complex, and a T wave that represents repolarization and depolarization of the atria and ventricles of the heart. Heart rate for a healthy person ranges from 60 to 100 beats per minute [1]. It is controlled by the sinus node which is the natural pacemaker of the heart. The heartbeat depends on one's instant activity that it may beat slower or faster. For example, our heart beats faster when we exercise because the frequency of the electrical impulses goes up, causing an increase in the heart rate to supply a sufficient amount of oxygen to the organs. During resting or sleeping, the heart beats slower than active conditions.

Arrhythmia is any abnormality in the cardiac cycle which 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 threat to daily life. Since cardiac arrhythmias are a major threat to human health, its early and accurate detection is essential in medical sciences [3]. Manual analysis of the ECG signal is not sufficient to quickly identify abnormalities in the heart rhythm. Analyzing long-duration ECG signals by physicians is a time-consuming task that may yield inaccurate results. Developing an automatic cardiac arrhythmia detection algorithm reduces the physician's workload, decreases arrhythmia detection time, and also improves diagnostic efficiency and accuracy. In order to accurately detect abnormalities in the ECG signal, many studies in the literature proposed computer-aided systems by using different feature extraction and classification techniques.

There have been several methods for automatically detecting arrhythmias based on signal processing, feature extraction, and learning algorithms [4, 5]. 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 by applying different filtering operations [6, 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 forms in the time- or the frequency-domain [6]. 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 have been utilized in previous studies for the classification of different arrhythmia types. Support Vector Machine (SVM) [13], Random Forest (RF) [14], Artificial Neural Networks (ANN) [15] have been used. In [16], time-frequency analysis of ECG signal is applied in the detection of cardiac arrhythmias. Pseudo Wigner-Ville Distribution used to obtain time-frequency representation of ECG signals obtained from AHA and MIT-BIH databases. Four different classifiers; Logistic Regression with L2 Regularization (L2-RLR), Adaptive Neural Network Classifier (ANNC), SVM, and Bagging classifier are used to classify ventricular fibrillation, ventricular tachycardia, normal sinus, and other rhythms. Fakhheraldin *et al.* [17] proposed complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) which is a decomposition method for non-linear and non-stationary signals. Average power, coefficients of dispersion, sample entropy, and singular values are calculated as features from Intrinsic Mode Functions (IMFs) extracted by the CEEMDAN algorithm. Features are used to train an ANN and classify five different heartbeats obtained from MIT-BIH database with 99.9% accuracy. 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, selecting not optimal features for classification step, low classification performance are examples of these limitations that directly affect success of the studies [18].

The architecture of conventional neural network algorithms contains input, output, and hidden layers. Deep Learning (DL) is a recent approach that contains more than three layers and has become more favorable compared to conventional techniques [19]. In DL, feature extraction and classification parts are embedded in the model which automatically identifies the optimal features from the input data [20]. DL has become 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.* [21] 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.* [22] proposed a DNN based method to classify ECG signals using three different databases. In [23], the temporal features of ECG heartbeats are detected with DNN on the MIT-BIH database with 99.09% accuracy. Yildirim *et al.* designed 16-layer deep CNN to classify 17 different arrhythmias on ECG signals taken from the MIT-BIH dataset. In [24], 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 architecture for automatic classification of ECG signals [25-27]. Kiranyaz *et al.* [28] introduced one-dimensional (1D) 3-layer CNN with R-peak wave for ECG arrhythmia classification. Baloglu *et al.* [29] used CNN algorithms with the end-to-end structure on 12-lead ECG signals for automatic detection of myocardial infarction on ECG and the proposed structure yielded significant performance with over 99% accuracy. Savalia *et al.* [30] proposed multilayer perception (MLP) and CNN based methods to identify first degree AV block (FAV) and ventricular bigeminy diseases. In [31], authors proposed an 11-layer CNN structure to detect different ECG segments with accuracy of

92.50% using 2 seconds and 94.90% using 5 seconds ECG segments. Yao *et al.* [32] proposed an attention-based time-incremental convolutional neural network (ATI-CNN), that preserve spatial and temporal characteristics of ECG signals with the integration of CNN architecture and recurrent cells. Their results attained to 81.2% accuracy. Besides the use of 1D CNN for ECG signal classification, several methods are proposed based on the 2D CNN algorithm for ECG image classification. The hidden structure of CNN can extract various local features from 2D input samples. The spatially adjacent pixels can be provided by utilizing nonlinear and multiple filters. Hence, the recent state-of-the-art studies proposed 2D CNNs for ECG arrhythmia classification problems, and these approaches have motivated researchers to implement CNNs with 2D image-based estimation process. In these studies, ECG signals are converted into 2D images and provided as input data for CNN [18, 33, 34]. The CNN architecture is considered to be suitable for the analysis and classification of 2D data. It achieves better results compared to other classical techniques in image processing [35]. In [18], a method is presented for heartbeat classification by using 2D CNN and 2D grayscale ECG images that are transformed from ECG beats of the MIT-BIH database. The proposed model achieves accuracy of 99.05%. Huang *et al.* [36] used time-frequency 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%. Jun *et al.* [18] proposed deep 2D CNN based 7 different arrhythmia detection using 2D grayscale ECG beat images and compare their proposed architecture performance with well-known structures such as AlexNet and VGGNet.

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, CNN model is selected for DL model and ECG signals are converted into ECG images for use as an input to the CNN model. Examining ECG images for arrhythmia detection by CNN is much similar to expert examination. CNN model directly processes the input images while noisy ECG beats do not affect the performance of classification. The model automatically extracts features from the raw ECG images.

These properties provide to elimination of noise filtering and feature extraction steps. ECG recordings are taken from MIT-BIH database [37] which includes different arrhythmia types categorized according to irregularity of heart by cardiologists. In this study, five different arrhythmia types; normal (N), left bundle branch block (LBBB), premature ventricular contraction (PVC), premature atrial beat (PAB), and right bundle branch block (RBBB) 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 CNN architecture. Classification performance of the proposed DL approach is compared to LeNet and ResNet-50 architectures. Experimental results show that the proposed image transformation method and network architecture provided 99.7% accuracy for the classification of five different arrhythmic heartbeat types. Compared with other popular DL approaches such as ResNet, the proposed method provides superior performance for classifying ECG images with fewer layers and less complex architectures. Thus, it is aimed to create a deep learning model for real-time wearable ECG arrhythmia systems to make quicker and more accurate predictions.

Methods

The proposed ECG arrhythmia classification algorithm based on 2D deep learning approaches consists of following steps: Obtaining ECG signals from public database, heartbeat segmentation, image transformation, and ECG arrhythmia classification by using 2D CNN architecture. Schematic diagram of the methodology is shown in Fig. 1. ECG arrhythmic signals are taken from the MIT-BIH arrhythmia database [37]. Five different arrhythmia types are selected from the database. In order to generate 2D images that will be applied as input to CNN architecture, ECG signals are segmented into heartbeats, then converted into ECG heartbeat images.

ECG Database

MIT-BIH arrhythmia database [37] 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 [38], MLII lead recordings are used in this study. The duration of each signal is 30-min. with 360 Hz sampling frequency and filtered by 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 [39]. The deep learning algorithms may tend to be biased for the type of arrhythmia classes that include a large number of samples as the number of heartbeats are not equal for each class of arrhythmia in the dataset. Arrhythmia classes with approximately same number of heartbeats are utilized to train the model to eliminate the imbalance effect of the database. Five different arrhythmia types are selected from the database as follows: normal (N), left bundle branch block (LBBB), premature ventricular contraction (PVC), premature atrial beat (PAB), and right bundle branch block (RBBB). ECG signals collected from the database are segmented into beats to generate single heartbeat images for CNN structure. The sample sizes of ECG beats for the considered classes are given in Table 1. Notice that the number of beats in each arrhythmia class is very close.

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 peak of each beat on the signal, separates heartbeats from the signal, and categorizes them according to their arrhythmia types. Each heartbeat is segmented according to its QRS peak location and categorized arrhythmia types. Examples of the segmented

arrhythmic beats are demonstrated in Fig. 2. ECG records and 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.

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

Beat Type	Beat Count	Record Numbers
N	8017	102 104 105 106 114 116
LBBB	8072	109 111 207 214
RBBB	7256	118 124 212 231 232 207
PVC	6970	105 106 107 114 116 119 203 207 208 210 213 214 215 221 228 233
PAB	7024	102 104 107 217

2D Signal-to-Image Transformation

CNN is mainly used for analyzing 2D data since it automatically learns the optimal features from raw image data [40]. 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 as an image in Python. Through this transformation, noise filtering and feature extraction steps are eliminated. CNN model directly processes the input images while noisy ECG beats do not affect the performance of classification. Furthermore, the model automatically extracts features that eliminate 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 used 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. Thus, a large amount of ECG beats is required in this study. 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 grayscale image. All images are generated with the same color and converted into grayscale, hence there is no need to process the RGB color degree in the CNN model. Thus, grayscale image conversion provides a decreases dimension of images and makes the analysis easier. Besides, grayscale image transformation decreases classification time due to reduced dimension and computational complexity.

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 output layer [41]. An artificial neural network consists of more than three layers that have several hidden layers in its structure. The structure is inspired by brain working system that includes hidden layers, many artificial neurons in the multiple 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. Mathematical representation of artificial neuron can define as;

$$y_j = f_j(x) = \varphi(w_j, x + b_j) \quad (1)$$

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 is shown in Fig. 3.

Artificial neural networks have been used in many different areas as computer vision, speech recognition, natural language processing, bioinformatics, drug design, medical image analysis. CNN is a part of artificial neural networks especially designed for analyzing two-dimensional data like images or videos [42]. In contrast conventional machine learning techniques, CNN architecture does 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 [20]. CNN has three characteristic layers; convolution, pooling, and fully connected layer. Convolution and pooling layers are used as feature extraction, fully connected is used as a classifier.

In the convolution layer, input samples are convolved with a specific kernel. Many features are selected by changing the kernel size [33]. Equation of discrete convolution function can define as;

$$(f * g)(x) = \sum_t f(t) g(x+t) \quad (2)$$

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

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

where \mathbf{K} is a convolution kernel, \mathbf{G} is a 2-D signal. Convolution process provides to extract effective features from the input. The pooling layer is applied for reducing the dimension of the input sample by keeping the optimal features. In the fully connected layer, final layer which is convolution or pooling layer flattened and neurons of previous layer are connected to other neurons that are found current layer [43]. After the finishing process, the CNN model is created.

In this study, a novel CNN model is designed to inspired LeNet model which is a well-known CNN model. The novel model has three convolution layer, three pooling layers, and a fully connected layer. Maximum pooling is implemented in the pooling layer. It selects only maximum value within the convolution filter and this provides reducing the number of output neurons. The model structure can be seen in Fig. 4. To specify the output values of the kernel in the layer, activation function is used. In the experiment, rectified linear units (ReLU) are applied as activation function and its equation is shown in (4).

$$f(x) = \max(0, x) \quad (4)$$

where \mathbf{x} is an element of the output kernel after the convolution or pooling layer.

Training and Testing

The cross-validation is usually not processed because of training costs and time duration for deep learning approaches. The random split of train and test size is selected instead of the cross-validation. 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. The dataset is separated into two parts as 80% of total data is used for training and 20% is used for testing with using random validation split. Hence, the test size is included 7519 ECG beat images and label distribution of ECG beats in test sample size found as 1423, 1664, 1604, 1440, 1388 for PAB (0), Normal (1), LBBB (2), RBBB (3), PVC (4) classes respectively. These ECG images are randomly split as train or test samples to achieve balanced distribution. In the learning or training, a proposed model train sample set is utilized and the test sample set is utilized to validate the proposed model. In the training phase, standard backpropagation [44] with a batch size of 64 is implemented for stochastic learning. The weights are updated according to the following equation;

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

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

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

In the proposed algorithm, the learning rate is defined as 0.001. Adam optimizer is used to optimization and cross-entropy is selected to loss function calculation. After every training epoch is completed, the proposed algorithm performs a test on the CNN model. The summary of CNN architecture 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. The summary of LeNet architecture is demonstrated in Table 3.

Table 2 Summary of proposed CNN architecture

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

Table 3 Summary of LeNet architecture

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

Performance Evaluation Metrics

The performance of the proposed model is evaluated utilizing various metrics. These metrics are accuracy, specificity, recall, precision, and F1-Score. The accuracy 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 sensitivity (Recall) metric is referred to as a true positive rate which is predicted as positive, while the specificity metric is referred to as a true negative rate which is predicted as negatives. The precision is referred to the proportion of classified positive cases that are

correctly real positives. The F1-Score metric is calculated by evaluating the harmonic mean of precision and recall values. These performance metrics are calculated as follows;

$$ACC(Accuracy) = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$SP(Specificity) = \frac{TN}{TN + FP} \quad (8)$$

$$RE(Recall) = \frac{TP}{TP + FN} \quad (9)$$

$$PR(Precision) = \frac{TP}{TP + FP} \quad (10)$$

$$F1-Score = \frac{2 \times RE \times PR}{RE + PR} \quad (11)$$

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 is incorrectly assigned for actual class, while **FNs** referred to the number of test samples for actual class, but is incorrectly assigned for different classes. **TPs** are the number of test samples that actually belong to a class and 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 is indicated by X-axis and the percentage of false negative is indicated by Y-axis on the representation of the ROC curve. The ideal ROC curve can be pointed to ranging from 0 to 100. The area between ROC and the axes calculated as the Area Under the ROC Curve (AUC). The AUC is an appropriate metric to examine algorithm performance as it is independent of the prediction

criterion selected. The performance of the proposed model has been evaluated thanks to the performance evaluations mentioned above.

Results

In this paper, a 2D CNN model is proposed to classify five different arrhythmia types. The proposed algorithm consists of three steps; heartbeat segmentation, image transformation, and deep learning for classification. *Keras* and *TensorFlow* libraries are used 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*, *cuda64_101.dll*, and *Tensor Flow 2.0*. Novel deep learning model is used as a CNN architecture which includes three convolutions, three pooling layers, and a fully connected layer. Using an ECG image formation as the input to CNN, preprocessing and feature extraction steps are eliminated.

The proposed 2D CNN based study is separated into two phases; training 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 7519 sample test set. Fig. 5 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. And also, Fig. 5 represents the confusion matrix and ROC curves for the test phase of the study. The confusion matrix consists of ground truth and predicted labels, where the column is the true label, and row is 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. Standard performance metrics such as SPE, PR, RE, F1-Score, and ACC are used to evaluate the performance of the 2D CNN model. Different epoch number and batch size are tested to obtain the highest evaluation parameter results. 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. Accuracy reached the highest value by increasing the epoch number and decreasing the batch size. Finally, 150

epochs and 64 batch sizes are found to be optimum for the proposed experiment. Detailed evaluation metrics of the study are demonstrated in Table 4. Using 150 epoch and 64 batch sizes, accuracy is calculated as 99.7% for the classification of five different ECG arrhythmias.

Table 4 Classification performances of the trained models

Metrics	ResNet-50	LeNet	Proposed
Classifier Type	5 different multi-classes		
SPE (%)	98.30	97.54	99.22
Recall (%)	98.12	98.07	99.7
Precision (%)	98.26	97.88	99.5
Mean Square Error	0.107	0.131	0.049
Training Loss	~0	0.073	~0
Training ACC (%)	100	99.9	100
Loss	0.0548	0.0281	0.0273
F1-Score (%)	98.11	97.78	99.24
Test ACC (%)	98.28	97.89	99.70

Discussion

This study analyzes the 2D ECG Signal-to-Image representations with the CNN approach on arrhythmia identification issues. The performance evaluations of the proposed model 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 in the dataset causes the imbalanced effect 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 our outstanding classification performance.

The deep learning structures can be converted to deeper networks with the implementation of different types of layers. The goal in this situation is to achieve higher classification performances. The deeper networks have complex structures hence their training time duration may be increased to days. The

training time is a crucial point in the design of automatic classification systems. The performances of the complex multi-layered structures can be evaluated considering above parameters. This study proposed a LeNet-based novel CNN structure for classifying arrhythmias and the performance is evaluated with accuracy of 99.7%. Moreover, well-known CNN structures ResNet-50 which includes 70 layers, and LeNet that consists of only 8 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. Our experimental results show that the proposed model outperforms the accuracy of LeNet by utilizing a less complex structure than it. The confusion matrices of the ResNet and LeNet based classification are represented in Fig. 6. Thus, we observed 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, filter dimensions, and other training parameters. The developed automatic cardiac arrhythmia detection algorithm improves the diagnostic efficiency and accuracy while reducing the training time. The innovative contribution of our proposed study can be highlighted as follows:

- This paper proposes a novel deep learning approach for identifying different arrhythmia types utilizing 2D ECG beats that are provided from 1D ECG signals by signal-to-image transformation procedure.
- We demonstrate the benefits of using 2D gray-scale images on the proposed CNN structure. We investigate the performance improvements by using a 2D CNN-based model in the classification of arrhythmia types as compared to well-known CNN approaches such as LeNet, and ResNet-50.

Conclusions

This paper proposed a method for accurate classification of ECG arrhythmias based on 2D CNN architecture. CNN is mainly used for image recognition in the literature. As such we propose

representation of ECG heartbeats as 2D gray-scale images and train a CNN to classify these images. Each ECG signal is segmented into heartbeats according to arrhythmia types. Then heartbeats are transformed into gray-scale ECG images and used as input for CNN. Advantage of transformation from signal to image is the elimination of preprocessing for removing noise. The proposed novel 2D CNN architecture has achieved high performance with the already known CNN structures like LeNet and ResNet, Furthermore, the architecture includes fewer layers compare to ResNet and demonstrate the highest performance for classification of arrhythmic signals. Comparison of arrhythmia classification results of the proposed study with previous studies may be seen in Table 5.

Traditional pattern recognition algorithms consist of preprocessing, feature extraction, feature selection, and classification steps. Through CNN architecture, preprocessing, feature extraction, feature selection steps may be eliminated. These steps are embedded in the CNN model. Furthermore, deep learning approaches are proper to find optimal features due to inherently hidden layers in the structures. This study demonstrated that detection of arrhythmias using ECG time-amplitude images and CNN architecture is a successful method and may help experts for diagnosing cardiovascular diseases quickly and precisely. In future studies, the proposed algorithm may be developed for home healthcare monitoring systems, by combining mobile applications and portable ECG devices, that will automatically detect arrhythmias and send them to physicians.

Table 5 Performance comparison table with recent studies

Author	Year	Database	Types	Classifier	Performance
Chen <i>et al.</i> [45]	2020	MITDB	Normal, Atrial Fibrillation, Ventricular Bigeminy, Pacing Rhythm, Atrial Flutter, Sinus Bradycardia	CNN LSTM	ACC = 99.32% SEN = 97.75% SPE = 99.51%
Huang <i>et al.</i> [36]	2019	MITDB	Normal, PVC LBBB, RBBB, PAB	CNN	ACC: 99%
Oh <i>et al.</i> [25]	2018	MITDB	Normal, PVC LBBB, RBBB, PAB	CNN LSTM	ACC: 98.1% SEN: 97.5% SPE: 98.7%
Yildirim <i>et al.</i> [46]	2018	MITDB	Normal, PVC LBBB, RBBB, Paced Beat	Bi-LSTM	ACC: 99.39% N/A N/A
Acharya <i>et al.</i> [31]	2017	MITDB AFDB CUDB	Normal, Ventricular Fibrillation, Atrial Fibrillation, Atrial Flutter	CNN	ACC: 92.50% SEN: 98.09% SPE: 93.13%
Sahoo <i>et al.</i> [8]	2017	MITDB	Normal, LBBB RBBB, Paced Beat	SVM	ACC: 98.32% SEN: 96.86% SPE: 98.92%
Acharya <i>et al.</i> [47]	2017	MITDB	Normal, Fusion, Unknown Supraventricular ectopic Ventricular ectopic	CNN	ACC: 94.03% SEN: 96.71% SPE: 91.54%
Zubair <i>et al.</i> [48]	2016	MITDB	Normal, Unknown Supraventricular ectopic Ventricular ectopic Fusion	CNN	ACC: 92.70% N/A N/A
Kiranyaz <i>et al.</i> [28]	2015	MITDB	Normal, Unknown, Fusion Supraventricular ectopic Ventricular ectopic	CNN	ACC: 99% SEN: 93.9% SPE: 98.9%
Proposed Study	2020	MITDB	Normal, PVC LBBB, RBBB, PAB	CNN	ACC: 99.7% SEN: 99.7% SPE: 99.22%

Abbreviations

ACC, Accuracy; ANN, Artificial Neural Network; ANNC, Adaptive Neural Network Classifier; ATI-CNN, Attention-based Time-Incremental Convolutional Neural Network; AUC, Area Under the ROC Curve; AV, Atrioventricular Node; CAD, Computer-Aided Diagnosis; CEEMDAN, Complete

Ensemble Empirical Mode Decomposition with Adaptive Noise; CNN, Convolutional Neural Network, DL, Deep Learning; DNN, Deep Neural Network, ECG, Electrocardiogram, FAV, First Degree AV Block; FN, False Negative; FP, False Positive; IMFs, Intrinsic Mode Functions; LBBB, Left Bundle Branch Block; LDA, Linear Discriminant Analysis; LSTM, Long Short-Term Memory; MLII, Modified Limb Lead II; MLP, Multilayer Perception; N, Normal; PAB, Premature Atrial Beat; PR, Precision; PVC, Premature Ventricular Contraction; RBBB, Right Bundle Branch Block; RE, 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, TN, True Negative; TP, True Positive; VMD, Variational Mode Decomposition.

Acknowledgements

The authors declare no conflict of interest.

Founding

Not applicable.

Authors' contributions

Conceptualization, MAO, EI, MD; Methodology, MAO, AA; Data Collection, EI, MD; Writing-Original Draft Preparation, MAO, EI, MD; Writing-Review & Editing, MAO, EI, MD, AA; Funding Acquisition, MAO, AA. 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/>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Biomedical Technologies, Izmir Katip Celebi University, Cigli, Izmir Turkey.

²Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, Izmir Turkey.

³Department of Electrical and Electronics Engineering, Izmir University of Economics, Balçova, Izmir, Turkey.

REFERENCES

1. Agrafioti F: **ECG in biometric recognition: Time dependency and application challenges**: Citeseer; 2011.
2. Sahoo JP: **Analysis of ECG signal for Detection of Cardiac Arrhythmias**. 2011.
3. Luz EJS, Schwartz WR, Cámara-Chávez G, Menotti D: **ECG-based heartbeat classification for arrhythmia detection: A survey**. *Comput Methods Programs Biomed* 2016, **127**:144-164.
4. Nanjundegowda R, Meshram VA: **Arrhythmia Detection Based on Hybrid Features of T-wave in Electrocardiogram**. *Int J Intell Syst* 2018, **11**(1):153-162.
5. Li H, Feng X, Cao L, Li E, Liang H, Chen X: **A new ECG signal classification based on WPD and ApEn feature extraction**. *Circ Syst Signal Pr* 2016, **35**(1):339-352.
6. Houssein EH, Kilany M, Hassanien AE: **ECG signals classification: a review**. *Int J Intell Eng Inform* 2017, **5**(4):376-396.
7. Chang K-M: **Ensemble empirical mode decomposition for high frequency ECG noise reduction**. *Biomed Tech (Berl)* 2010, **55**(4):193-201.
8. Sahoo S, Kanungo B, Behera S, Sabut S: **Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities**. *Measurement* 2017, **108**:55-66.
9. Lin C-C, Yang C-M: **Heartbeat classification using normalized RR intervals and morphological features**. *Math Probl Eng* 2014, **2014**.
10. Mironovova M, Bíla J: **Fast fourier transform for feature extraction and neural network for classification of electrocardiogram signals**. In: *2015 Fourth International Conference on Future Generation Communication Technology (FGCT): 2015*: IEEE; 2015: 1-6.
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**(21):8361-8368.
12. Mert A: **ECG feature extraction based on the bandwidth properties of variational mode decomposition**. *Physiol Meas* 2016, **37**(4):530.
13. Li H, Yuan D, Wang Y, Cui D, Cao L: **Arrhythmia classification based on multi-domain feature extraction for an ECG recognition system**. *Sensors* 2016, **16**(10):1744.
14. Li T, Zhou M: **ECG classification using wavelet packet entropy and random forests**. *Entropy* 2016, **18**(8):285.
15. Dewangan NK, Shukla S: **ECG arrhythmia classification using discrete wavelet transform and artificial neural network**. In: *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT): 2016*: IEEE; 2016: 1892-1896.

16. Mjahad A, Rosado-Muñoz A, Bataller-Mompeán M, Francés-Villora J, Guerrero-Martínez J: **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**:119-127.
17. Abdalla FY, Wu L, Ullah H, Ren G, Noor A, Zhao Y: **ECG arrhythmia classification using artificial intelligence and nonlinear and nonstationary decomposition.** *Signal Image Video P* 2019, **13**(7):1283-1291.
18. Jun TJ, Nguyen HM, Kang D, Kim D, Kim D, Kim Y-H: **ECG arrhythmia classification using a 2-D convolutional neural network.** *arXiv preprint arXiv:180406812* 2018.
19. LeCun Y, Bengio Y, Hinton G: **Deep learning.** *nature* 2015, **521**(7553):436.
20. Yıldırım Ö, Pławiak P, Tan R-S, Acharya UR: **Arrhythmia detection using deep convolutional neural network with long duration ECG signals.** *Comput Biol Med* 2018, **102**:411-420.
21. Zhang C, Wang G, Zhao J, Gao P, Lin J, Yang H: **Patient-specific ECG classification based on recurrent neural networks and clustering technique.** In: *2017 13th IASTED International Conference on Biomedical Engineering (BioMed): 2017*: IEEE; 2017: 63-67.
22. Al Rahhal MM, Bazi Y, AlHichri H, Alajlan N, Melgani F, Yager RR: **Deep learning approach for active classification of electrocardiogram signals.** *Inf Sci* 2016, **345**:340-354.
23. Sannino G, De Pietro G: **A deep learning approach for ECG-based heartbeat classification for arrhythmia detection.** *Future Gener Comput Syst* 2018, **86**:446-455.
24. Pławiak P, Acharya UR: **Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals.** *Neural Comput Appl* 2019:1-25.
25. Oh SL, Ng EY, San Tan R, Acharya UR: **Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats.** *Comput Biol Med* 2018, **102**:278-287.
26. Lee K-S, Jung S, Gil Y, Son HS: **Atrial fibrillation classification based on convolutional neural networks.** *BMC Med Inform Decis Mak* 2019, **19**(1):206.
27. Zhao Z, Deng Y, Zhang Y, Zhang Y, Zhang X, Shao L: **DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network.** *BMC Med Inform Decis Mak* 2019, **19**(1):286.
28. Kiranyaz S, Ince T, Gabbouj M: **Real-time patient-specific ECG classification by 1-D convolutional neural networks.** *IEEE Trans Biomed Eng* 2015, **63**(3):664-675.
29. Baloglu UB, Talo M, Yildirim O, San Tan R, Acharya UR: **Classification of myocardial infarction with multi-lead ECG signals and deep CNN.** *Pattern Recognit Lett* 2019, **122**:23-30.
30. Savalia S, Emamian V: **Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks.** *Bioengineering* 2018, **5**(2):35.
31. 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* 2017, **405**:81-90.
32. 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**:174-182.
33. Kim M-G, Ko H, Pan SB: **A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks.** *J Ambient Intell Humaniz Comput* 2019:1-9.
34. Rajput K, Wibowo S, Hao C, Majmudar M: **On Arrhythmia Detection by Deep Learning and Multidimensional Representation.** *arXiv preprint arXiv:190400138* 2019.
35. Hatami N, Gavet Y, Debayle J: **Classification of time-series images using deep convolutional neural networks.** In: *Tenth International Conference on Machine Vision (ICMV 2017): 2018*: International Society for Optics and Photonics; 2018: 106960Y.
36. Huang J, Chen B, Yao B, He W: **ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network.** *IEEE Access* 2019, **7**:92871-92880.
37. Moody GB, Mark RG: **The impact of the MIT-BIH arrhythmia database.** *IEEE Eng Med Biol Mag* 2001, **20**(3):45-50.
38. Barrella T, McCandlish S: **Identifying Arrhythmia from Electrocardiogram Data.** 2014.
39. Shaker AM, Tantawi M, Shedeed HA, Tolba MF: **Generalization of Convolutional Neural Networks for ECG Classification Using Generative Adversarial Networks.** *IEEE Access* 2020, **8**:35592-35605.
40. Chen J-F, Chen W-L, Huang C-P, Huang S-H, Chen A-P: **Financial time-series data analysis using deep convolutional neural networks.** In: *2016 7th International Conference on Cloud Computing and Big Data (CCBD): 16-18 Nov 2016*: IEEE; 2016: 87-92.
41. Zhao B, Lu H, Chen S, Liu J, Wu D: **Convolutional neural networks for time series classification.** *J Syst Eng Electron* 2017, **28**(1):162-169.
42. Zhang J, Wu Y: **Automatic sleep stage classification of single-channel EEG by using complex-valued convolutional neural network.** *Biomed Tech (Berl)* 2018, **63**(2):177-190.
43. 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.

44. Bouvrie J: **Notes on convolutional neural networks**. 2006.
45. 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.
46. Yildirim Ö: **A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification**. *Comput Biol Med* 2018, **96**:189-202.
47. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, Tan RS: **A deep convolutional neural network model to classify heartbeats**. *Comput Biol Med* 2017, **89**:389-396.
48. Zubair M, Kim J, Yoon C: **An automated ECG beat classification system using convolutional neural networks**. In: *2016 6th international conference on IT convergence and security (ICITCS): 2016*: IEEE; 2016: 1-5.