Computing (2024) 106:1227–1248
https://doi.org/10.1007/s00607-023-01243-0

SPECIAL ISSUE ARTICLE

Automatic ECG classification using discrete wavelet transform and one-dimensional convolutional neural network

Armin Shoughi[1] · Mohammad Bagher Dowlatshahi[1] · Arefeh Amiri[2] · Marjan Kuchaki Rafsanjani[3] · Ranbir Singh Batth[4]

Received: 15 November 2021 / Accepted: 21 November 2023 / Published online: 23 December 2023

Abstract

This paper presents an approach based on deep learning for accurate Electrocardiogram signal classification. The electrocardiogram is a significant signal in the realm of medical affairs, which gives vital information about the cardiovascular status of patients to heart specialists. Manually meticulous analysis of signals needs high and specific skills, and it is a time-consuming job too. The existence of noise, the inflexibility of signals, and the irregularity of heartbeats keep heart specialists in trouble. Cardiovascular diseases (CVDs) are the most important factor of fatality globally, which annually caused the deaths of 17.9 million people. Totally 31% of all death in the world are related to CVDs, which the age of 1/3 of patients that died because of CVDs is below 70 Because of the high percentage of mortality in cardiovascular patients, accurate diagnosis of this disease is an important matter. We present an approach to the analysis of electrocardiogram signals based on the convolutional neural network, discrete wavelet transformation with db2 mother wavelet, and synthetic minority over-sampling technique (SMOTE) on the MIT-BIH dataset according to the association for the advancement of medical instrumentation (AAMI) standards to increase the accuracy in electrocardiogram signal classifications. The evaluation results show this approach with 50 epoch training that the time of each epoch is 39 s, achieved 99.71% accuracy for category F, 98.69% accuracy for category N, 99.45% accuracy for category S, 99.33% accuracy for category V and 99.82% accuracy for category Q. It is worth mentioning that it can potentially be used as a clinical auxiliary diagnostic tool. The source code is available at https://gitlab.com/arminshoughi/ecg-classification-cnn.

Keywords Cardiovascular diseases · Convolutional neural network · Deep learning · Electrocardiogram signals · Physio bank MIT-BIH arrhythmia database

Mathematics Subject Classification 68T05

Extended author information available on the last page of the article

1 Introduction

Lifestyle changes, job stresses, rush in life and unhealthy nutrition are essential factors in heart diseases leading to heart failure and heart attacks [1]. According to the world health organization (WHO) statistics, approximately 17.9 million people die annually due to cardiovascular diseases. So cardiovascular diseases (CVDs) are essential factors of mortality in the world [2]. So, early diagnosis of CVDs is a necessary and worthwhile job. The electrocardiogram is one of the most powerful tools for cardiovascular disease diagnosis, which consists of the sequence of heartbeats, and it registers electrical fluctuations that result from heart operations by putting electrodes on the skin surface. Each heartbeat consists of a P wave, QRS complex, and T wave, shown in Fig. 1. The P wave is representative of the contraction and depolarization of atriums. Stimulation of them, which occurred by focal automation above the left atrium, starts from the left ventricle to perform blood transformation. The transient delaine appeared when blood was entered, arising from the time of guidance atrioventricular node that starts after each P wave and ends at the starting point of the QRS complex. Then electrical operations of the ventricle are registered as QRS complex. The horizontal line, which is seen after the QRS complex and before the T wave, is named ST segment, which shows the threshold level of ventricle polarizing. The T wave shows the final stage of ventricular depolarization. According to the Association for the Advancement of Medical Instrumentation (AAMI), necessary arrhythmia signals that put patients at risk and need treatment for avoiding potential risks are classified into five categories (N, S, V, F, and Q). Table 1 lists the detailed heartbeat types.

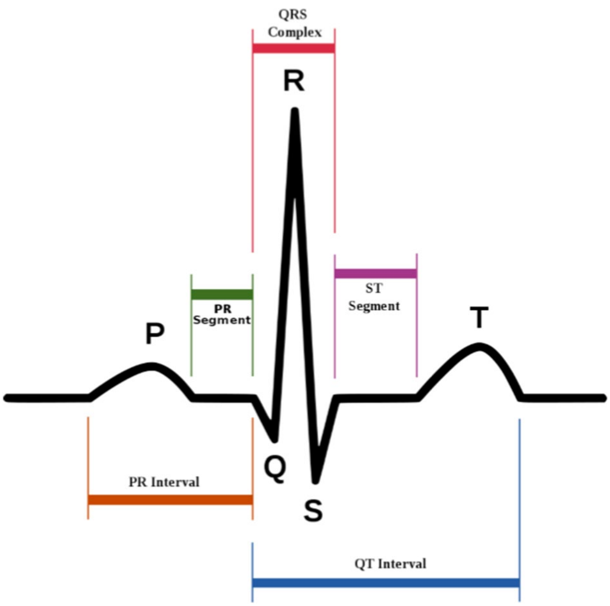

Fig. 1 The normal ECG signal and different parts of it [3]

Table 1 Classification of different ECG arrhythmia into five categories of N, S, V, F, A [4]

AAMI classes	Heartbeat types
Normal beats (N)	Normal beats (N), Left bundle branch block (L), Right bundle branch block (R), Atrial escape beat (e), Nodal (junctional) escape beat (j)
Supraventricular ectopic Beats (S)	Atrial premature contraction (A), Aberrated atrial premature beat (a), Nodal (junctional)premature beat (J), Supraventricular premature beat (S)
Ventricular ectopic beats (V)	Ventricular premature contraction (V), Ventricular escape beat (E)
Fusion beats (F)	A fusion of ventricular, normal beat (F)
Unclassifiable beats (Q)	Paced beat (/), A fusion of paced and normal beat (f), Unclassified beat (Q)

In this study, we focus on the classification of ECG signals according to AAMI standards. For that, at first, we extract ECG beats from ECG signals of the MIT-BIH dataset. Then we classify them into five classes according to AAMI standards. MIT-BIH dataset is an unbalanced dataset; for handling this problem, we use the SMOTE algorithm to create new artificial data according to existing data. Finally, we design a 15-layer CNN, and train and evaluate it with these data.

The rest of this paper is organized as follows. Section 2 describes the related works. Section 3 introduces the database that is used for training, testing, and evaluating the proposed method in this study. Section 4 describes the proposed method. Section 5 presents the experimental setup and shows the achieved result by the proposed method and performance compared to the state-of-the-art approaches. Finally, Sects. 6 and 7 draw the discussion and conclusion of the paper.

2 Related work

In the last decade, different approaches for heart disease classification by analyzing ECG signals were presented. Most of these approaches consist of three levels, which include pre-processing, feature extraction, feature selection, and classification algorithm. At the pre-processing level, because of existing noise originating from different factors such as power cables, incidental body movement, etc. different technologies such as median filter, discrete wavelet transform (DWT), adaptive filters, band-pass filters, low-pass filters, and high-pass filters are used. Mathews et al. [5], to remove baseline wander, passed the signal through median filters with window sizes of 200 ms and 600 ms. Elhaj et al. [6] used DWT with db6 mother wavelet for denoising ECG data at the preprocessing level. Venkatesan et al. [7] used four steps to reduce noise. First, the DC component is eliminated by subtracting the average value. Then, the baseline drift is reduced with a median filter. The low-pass filter removes power frequency interference and electromyogram (EMG) noise. Finally, the high-pass filter can eliminate some other low-frequency noise. Azariadi et al. [8] used band-pass filtered at 1-50 Hz to remove the power line interference and the baseline wandering noises. Chazal et al. [9] used the 12-tap low-pass filter to remove unwanted power-line and high-frequency noise from the baseline-corrected ECG. Li et al. [10] used a high-pass filter with a 1-Hz cutoff frequency to suppress residual baseline wander, a second-order 30 Hz Butterworth low-pass filter to reduce high-frequency noise, and a notch filter to eliminate power line interference. After filtering the signals, important features considering different time and frequency ranges and morphological features are extracted. Technologies used in this level consist of a discrete wavelet transform [6, 8, 11] along with high and finite impulse ratio [12], high order statics [6, 7, 13], and 1D convolutional neural network [14]. On the other hand, to decrease the complexity and time of computing, some researchers utilized feature extraction and feature selection approaches. Some technologies used in this case can be named principal component analysis [6, 15, 16], independent component analysis [6, 12], and fast, independent component analysis [11]. The final level is classification which used different classification algorithms such as support vector machines [6, 12, 15, 17], neural networks [18, 19], deep neural networks [20], convolutional neural networks [14], k-nearest networks [18, 21–23], random forest [16, 18], logistic regression [16] and ensemble learning [13]. Deep learning is one of the most practical techniques in machine learning that is used successfully too much in image verification [24], classification [25], and speech recognition [3], and increases speed and accuracy in pattern recognition tools in the last decade. Convolutional neural networks are the most successful networks in the deep neural networks that are used too much in automatic cardiovascular disease diagnosis by ECG signals. Zhai et al. [24], with a combination of two couple of neighbor heartbeats, created a two-dimensional matrix and made use of it as input for a seven-layer 2D convolutional neural network. They achieved 98.6% accuracy in the classification of the V category and 97.5% accuracy in the classification of the S category. Xiang et al. [25] considered a combination of morphological and time features, as the input of two convolutional neural networks to extract new features. Finally, with extracted features, they could achieve 97.8% accuracy in V and F signal classification. Zhang et al. [26] proposed a model of a six-layer convolutional neural network which, consists of two convolution layers, two pooling layers, and two fully-connected layers. This model classified five classes (N, S, V, F, and Q) from the MIT-BIH dataset with an accuracy of 97.5%. Acharya et al. [27] proposed a nine-layer convolutional neural network to classify five different classes from the MIT-BIH arrhythmia dataset. To overcome the unbalanced number of heartbeat classes in the dataset, they created artificial heartbeats by changing the average, standard division, and z-score of basic data. For adaption with the number of heartbeats from the N class, the number of heartbeats from all classes except the N class was increased, and they were used for training and testing models. They achieved 94.03% accuracy using artificial data, but if this model only trained with original data, total accuracy decreased to 89.07%. Oh et al. [28] proposed combining CNN and LSTM to classify five classes from the MIT-BIH arrhythmia dataset by using segments with variable lengths. This model's structure consists of three convolutional layers and three pooling layers, followed by LSTM layers and two fully-connected layers. LSTM is used for extracting time information of features concluded from convolutional layers. They finally achieved 98.1% accuracy in heartbeat signal classifications. Park et al. [29] applied two morphological feature extraction methods: higher-order statistics and Hermite basis functions. also, they assumed that the QRS complexes of class N and S might have a morphological similarity, and those of class V and F may also have their own similarity. Therefore, they employed a hierarchical classification method using support vector machines, considering those similarities in the architecture. The results showed that their hierarchical classification method performs better than the conventional multi-class classification method. In addition, the Hermite basis functions gave more accurate results than the higher-order statistics. Zubair et al. [30] used CNN with 44 recordings of ECG signals obtained from the MIT-BIH database. They extracted R-peak ECG beat patterns for the training of the three-layer CNN. They achieved 92.70% accuracy in detecting the ECG beats in their respective classes. Li et al. [31] implemented a parallel general regression neural network (GRNN) to classify the heartbeat and achieved a 95% accuracy according to the AAMI standards. they designed an online learning program to form a personalized classification model for patients. The achieved accuracy of the model is 88% compared to the specific ECG data of the patients. Liu et al. [32] proposed several kinds of classification methods for ECG beats and used convolution neural networks and some improved algorithms. In the method which combined the CNN model and SVM classifier, the classification accuracy was up to 91.29%. From the result, we can see that sometimes an uncomplicated method can also have a good result. In addition, we do not need to extract complex features of ECG beats so it is more convenient for real-time diagnosis. Jiang et al. [33] presented an evolutionary optimization of the feed-forward implementation of BbNNs and the application of the BbNN approach to personalized ECG heartbeat classification. Network structure and connection weights were optimized using an EA that utilizes evolutionary and gradient-based search operators. The performance evaluation using the MIT-BIH arrhythmia database showed a sensitivity of 86.6% and overall accuracy of 98.1% for VEB detection. For SVEB detection, the sensitivity was 50.6% and the overall accuracy was 96.6%. These results were a significant improvement over other major techniques compared to ECG signal classification.

3 ECG database

The ECG signals which are used for training, testing, and evaluating this approach are extracted from the open-source physio bank MIT-BIH arrhythmia database [34] proposed by American accurate medical tools [4]. The MIT-BIH Arrhythmia Database was the first generally available set of standard test material for the evaluation of arrhythmia detectors, and it has been used for that purpose as well as for basic research into cardiac dynamics at about 500 sites worldwide since 1980. It has lived a far longer life than any of its creators ever expected. Together with the American Heart Association (AHA) Database, it played an interesting role in stimulating manufacturers of arrhythmia analyzers to compete based on objectively measurable performance, and much of the current appreciation of the value of common databases, both for basic research and for medical device development and evaluation, can be attributed to this experience [35]. This database consists of 48 sequences of half-hour two-channel signals (lead II and lead V) registered from 47 different subjects.

Table 2 The number of extracted beats from the MIT-BIH dataset according to related classification

Type	Number of beats
N	90502
S	2777
V	7226
F	802
Q	8031

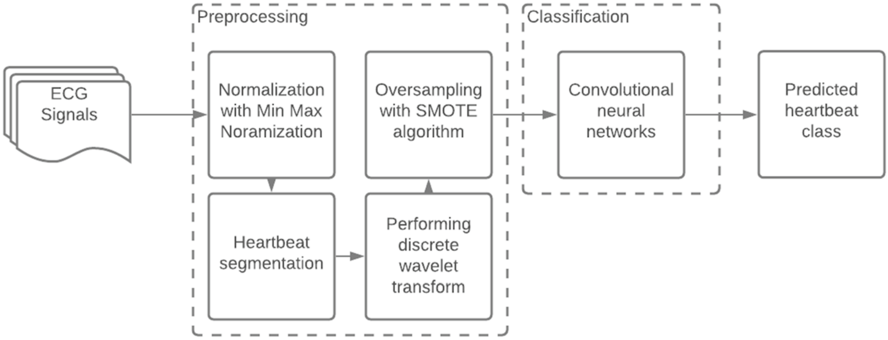

Fig. 2 Flowchart of our proposed method

The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database. This database also includes the class type and R pick location of every beat that was confirmed and reviewed by two or more cardiologists. In this study, 109338 heartbeats by using lead II are recognized and extracted. The number of samples from each separated class is visible in Table 2.

4 Proposed approach

In this part, we present the proposed method for ECG classification. Our proposed method consists of three main parts, including preprocessing, feature extraction, and classification, as shown in Fig. 2. A detailed description of each part is introduced in the subsections below.

4.1 Pre-processing

The proposed approach input is a sequence of ECG beats extracted from ECG signals. To extract the ECG beats from a given ECG signal, we first normalized received ECG signals in 0 to 1 intervals by Min-Max Normalization according to

Eq. 1. Then we found the set of T waves regarding the ECG R-peaks of its corresponding annotation file in the MIT-BIH Arrhythmia database and split the continuous ECG signal to a sequence of heartbeats based on the extracted T waves, and assigned a label to each heartbeat based on the annotation file. In the end, we resized each heartbeat to a predefined fixed length (280 samples).

$$X_{normalized} = \left(\frac{X - X_{minimum}}{X_{maximum} - X_{minimum}} \right) \tag{1}$$

where X is the value of each sample, $X_{minimum}$ is minimum, $X_{maximum}$ is the maximum value of X among all samples, and $X_{normalized}$ is the normalized sample.

4.2 Feature extraction

The ECG signal originates as a continuous waveform, comprising multiple points that convey crucial information about the behavior of the signal. These points become discernible after the process of sampling. The identification of disease patterns relies on these significant points. To achieve this, spectral analysis of the ECG signals and the application of a wavelet transform approach can be employed. This method allows for efficient operation on devices with limited processing power, enabling us to extract the essential points. By eliminating insignificant points, we can effectively reduce the complexity of the proposed approach. The basic idea of the wavelet transformation approach is to overcome the weakness and limitations in the current Fourier transform. Fourier transform works best when the spectral frequency of a signal is static in Statistical terms [36]. On the other hand, if a signal includes X_{HZ} frequency, then this signal will exist equally throughout the entire length of the signal. Thus if the signals are more dynamic, then the results will be worse than ever. The key advantage of the Wavelet Transform compared to the Fourier Transform is the ability to extract both local spectral and temporal information. A practical application of the Wavelet Transform is analyzing ECG signals which contain periodic transient signals of interest. Also, Fourier transforms approximately a function by decomposing it into sums of sinusoidal functions, while wavelet analysis makes use of mother wavelets. Both methods are capable of detecting dominant frequencies in the signals; however, wavelets are more efficient in dealing with time-frequency analysis. The wavelet transform approach is designed for removing unstable signal problems such as ECG signals. This approach is implemented by using a generator function named mother wavelet transform and performing Dilation and Translation [37]. Wavelet transformation uses on large scales for signal processing, and its fundamental merit is the size of the variable window, which is wide in low frequencies and narrow in high frequency and causes clarity of time frequencies in all frequency domains. To access comprehensive information in this field, refer to Mallat studies [38]. A discrete form of wavelet transformation is indicated as the following:

$$\varphi_{jk}(t) = \frac{1}{\|\sqrt{s_0^j}\|}\varphi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right) \quad (2)$$

where φ_{jk} is wavelet function for the specific amounts of k and j, t time, $s_0 > 1$ constant dilation step, τ_0 is time transformation constant and depends on s_0. In this study, the effect of different mother wavelets including db2, db4, db8, db10, sym2, sym4, sym10, coif1, coif3, coif5, and dmey were reviewed and evaluated. The db2 mother wavelet, due to its complexity and its similarity to ECG signals, caused increased efficiency of the proposed approach. By applying the discrete transformation, preliminary data become a wave and is divided into two categories. The first category is named approximation, which includes low frequencies and an indication of the total trend of current data. It plays an essential role in calculating, and we classified ECG signals according to this category. The second category name is details that include high frequencies, and it is an indication of limited changes in data. Feature extraction is performed for 0.25 s. The output of this section is a sequence of 141 extracted features from heartbeats with 280 features. In Fig. 3, five heartbeat waveforms from five different classes with extracted features of them are shown.

4.3 Oversampling

The distribution of heartbeat disease classes in the MIT-BIH dataset is unbalanced, and the number of heartbeats of a normal class is approximately nine times more than the number of samples in other classes. Machine learning approaches when the number of samples in different classes is not equally faced with trouble. In this study, sampling is performed using the SMOTE oversampling algorithm to remove unbalanced data problems in the dataset. This algorithm creates new artificial data by considering similarities among the current data samples. It aims to balance class distribution by randomly increasing minority class examples by replicating them. SMOTE synthesizes new minority instances between existing minority instances. It generates the virtual training records by linear interpolation for the minority class. These synthetic training records are generated by randomly selecting one or more of the k-nearest neighbors for each example in the minority class. The SMOTE algorithm works as follows: at first, a random sample from the minority class is drawn. Then, for the observations in this sample, the k nearest neighbors are identified. One of those neighbors will be taken and identified as the vector between the current data point and the selected neighbor. Then the vector is multiplied by a random number between 0 and 1. Finally, to obtain the synthetic data point, this point is added to the current data points [39].

Following the feature extraction stage, we proceeded to partition the data into 10 subsets of equal size, maintaining a 9:1 ratio. Subsequently, we employed 9 sets as the training set during the network training process, reserving the remaining subset for the final step, namely the test process. This allowed us to assess the trained network and calculate evaluation metrics using independent data that had no influence on the training procedure. To ensure the integrity of the evaluation results and the

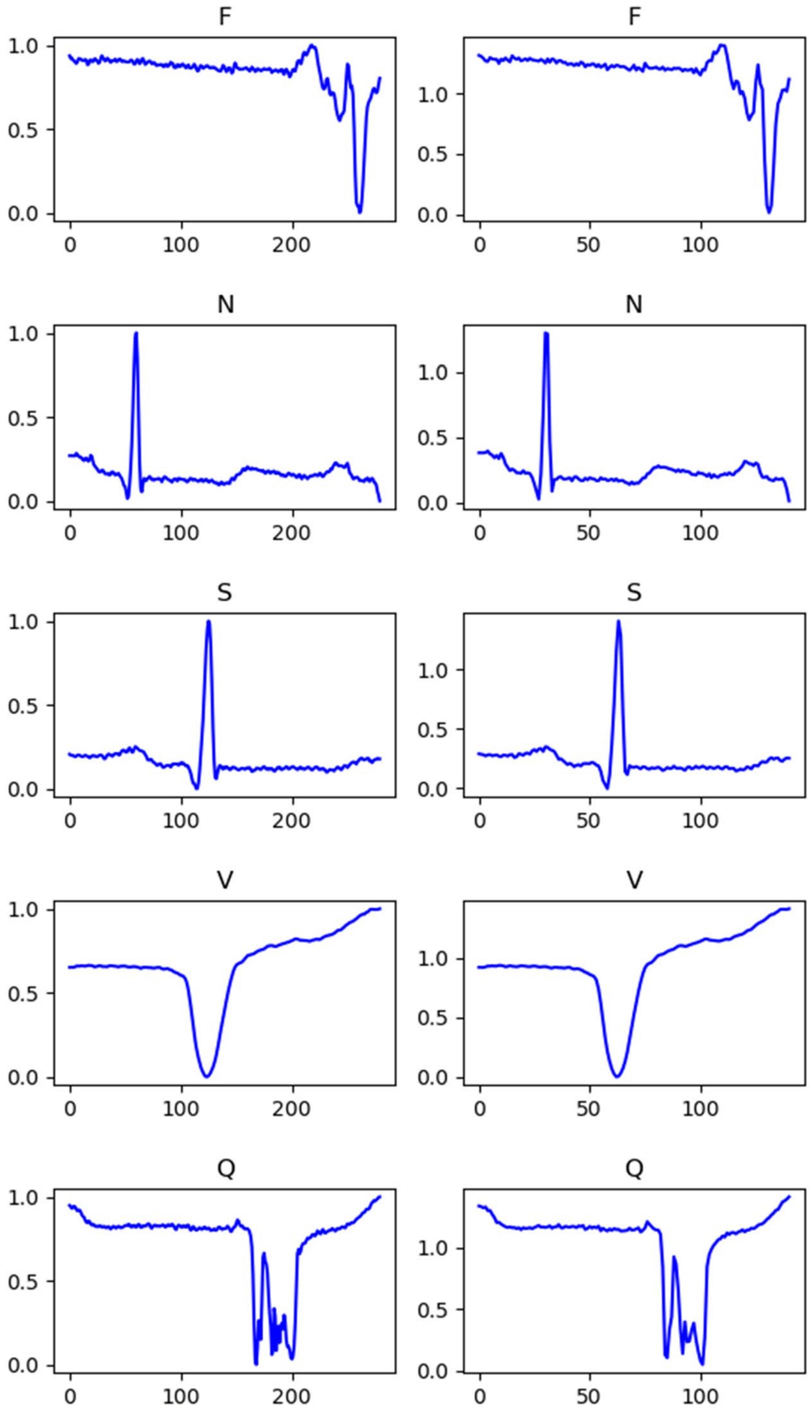

Fig. 3 Preliminary signals waveform of five different classes with extracted feature waveforms of them

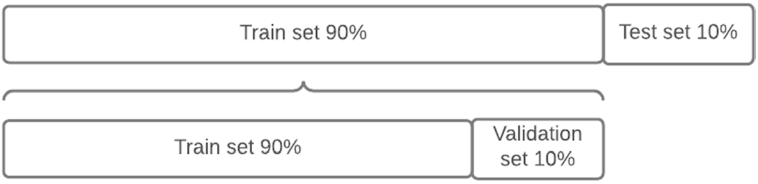

Fig. 4 Distribution of ECG segments used for classifier training, testing, and evaluating

Table 3 The number of presented data samples in different classes in testing and training steps before and after oversampling using SMOTE

	Before oversampling			After oversampling		
	Train data	Test data	Total	Train data	Test data	Total
F	719	83	802	2000	83	2083
N	81464	9038	90502	81464	9038	90502
V	2488	289	2777	5000	289	5289
F	6478	748	7226	6478	748	7226
Q	7255	776	8031	7255	776	8031

accuracy of the final model, we exclusively applied oversampling techniques solely on the training set. This precaution aimed to prevent any inadvertent impact on the outcomes when assessing the model's performance on the testing set. Furthermore, we implemented a 10-fold cross-validation approach by dividing the training set into training and testing subsets at a 9:1 ratio. In each training stage, nine portions were utilized for training purposes, while the remaining segment was reserved for evaluating the classification performance (Fig. 4).

According to the low number of samples in F and S classes and due to avoiding the focus of the model on the classification of classes with more samples by using the SMOTE algorithm, we increased the number of samples in F class from 719 to 2000 samples and the number of samples in S class from 2491 to 5000 samples. Thus the number of each class after performing oversampling with SMOTE is shown in Table 3.

4.4 Classifier architecture

Table 4 shows the structure summary of the proposed classifier network. This network consists of 15 layers, including three sections, each section consists of one convolutional layer with ReLU activation function, one batch-normalized layer, one ReLU layer, and one max-pooling layer. These sections are followed by one flatten layer for flatting and sending inputs into the last two fully-connected layers that are used respectively with ReLU and Softmax activation functions for classifying data into five classes (see Fig. 5).

In each convolutional layer (1, 5, 9 layers), data were convolved using equation (3) with layers kernel (5, 5, 3).

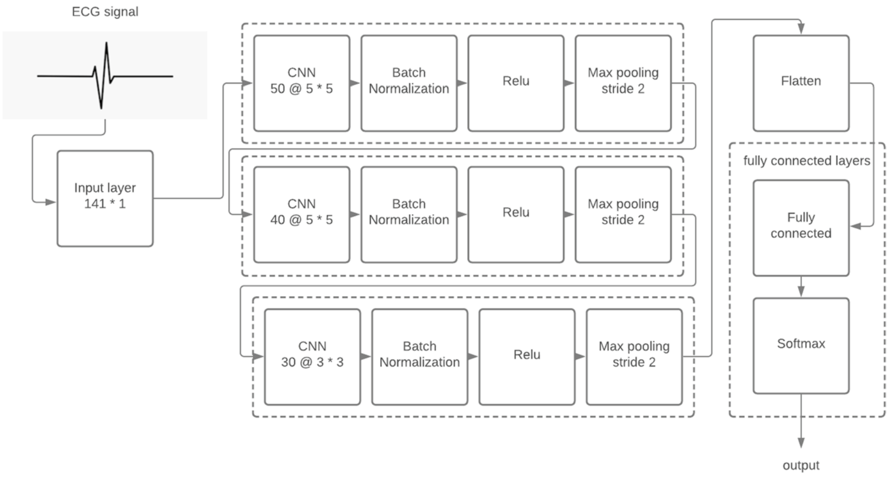

Fig. 5 The architecture of the proposed CNN model

$$x_n = \sum_{k=0}^{N-1} y_k f_{n-k} \tag{3}$$

where y is signal, f is filter, n is the number of signal samples of y, and output vector is shown by x. The result of convolutional layers is considered as batch-normalized layer inputs. Batch normalization indeed is an approach for normalizing inputs to deal with a change in internal variables. During the network training in each iteration, batch-normalization by subtracting the data of every set from the average of the set and dividing by standard division, which was calculated by equations 4 and 5, was performed.

$$\mu_B = \frac{1}{B} \sum_{i=1}^{B} x_i \tag{4}$$

where μ_B is the average of the set, m is the number of set samples, x_i is the i_{th} sample in the set, $x \in B$, and B is the normalizing set's input.

$$\sigma_B^2 = \frac{1}{B} \sum_{i=1}^{m} \left(x_i - \mu_B\right)^2 + \epsilon \tag{5}$$

where σ_B^2 is an indication of set variance, m is the number of samples in each set, x_i is an i_{th} sample of each set, $x \in B$ that B is the input of batch-normalized layer and μ_B is average set. Then by using the result of equations 4 and 5, input samples of x are normalized by equation (6).

$$BN(x) = \curlyvee \circledcirc \frac{x - \mu_B}{\sigma_B} + \beta \tag{6}$$

where σ_B is the average, μ_B is the standard division of B set, γ is the scaling parameter, and β is the displacement parameter it should be trained in common with the other model parameters. The output of batch-normalization to introduce a nonlinear section in the convolutional network and nonlinear training of it was sent as an input to the ReLU layer. In the ReLU function, the gradient was not saturating in a positive area, and it performs more quickly than the Sigmoid and Tangent hyperbolic function and reduces the training error rate. The output of this layer is considered as the input of the max-pooling layer. The aim of using the max-pooling layer here is due to two reasons: (1) decreasing of the following layer calculations, (2) extraction of the local dependency, and keeping the prominent information, which causes a reduction in the volume of calculated data and increment of the processing speed too. The output of the third max-pooling layer or, in other words, the 12^{th} layer, is sent to the flatten layer for flatting the data. Finally, the result as an input of 14^{th} and 15^{th} layers, which are fully connected, are sent for classification and acquiring final results. Thus a total of 71035 parameters exist in the proposed network that 70495 parameters are trainable, and 540 parameters are untrainable.

5 Results

For reviewing the proposed approach, Python programming language, Keras platform, and GPU-based Tensorflow backend in Ubuntu operation system version 16.04 with core i7 2.5 GH CPU, 8 GB RAM, and NVidia GeForce GTX 840 were used. This system is selected due to the aim of this study in performing classifier quickly with a minimum processing capacity for use on low processing power devices. The proposed network is trained during 50 epochs each epoch lasts at 39 s by using Sparse Categorical Cross Entropy as loss function, Adam as an optimizer algorithm, and 113131 data samples from 5 different classes. Overall training of the network was performed for 32 min. Then, the trained model tests using 10934 samples from 5 different classes. The result indicated that the model acquired 98.53% accuracy in heartbeat classification. The classifier accuracy diagram in training and evaluating phases according to the number of epochs in the training stage are shown in Fig. 6. After the training of the network, the classifying operation of each heartbeat is done in 0.9 s.

We compared and analyzed the accuracy of the proposed approach with the state-of-the-art approaches in this field, especially with the studies that observed AAMI standards and classified data according to Table 1. The results in which comparison is performed are shown in Table 5. The proposed approach improved the accuracy of heartbeat signal classifications compared to the other presented approaches in this field, along with the reduction of complexity and, consequently, the increment of classification speed. It is not only related to the idea of using discrete wavelet transform on the db2 mother wavelet for the classification of heartbeat signals and the power of it in locating the placement of different heartbeat section incidents but also shows that the optimized convolutional neural network is convenient for 1D heartbeat classification.

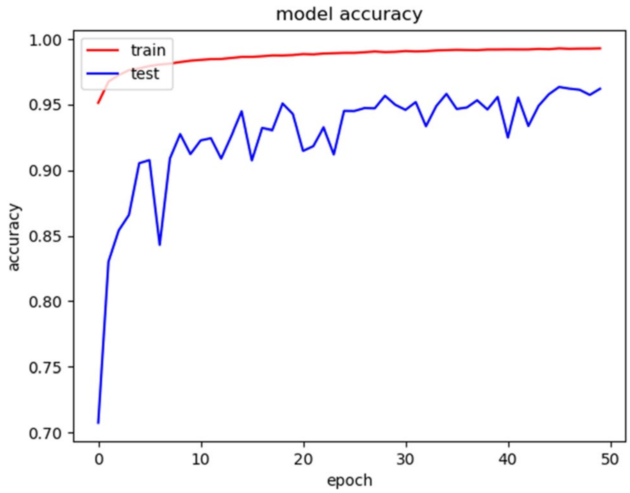

Fig. 6 The diagram of training and evaluating accuracy on MIT-BIH dataset according to the number of epochs

Table 4 A summary of the proposed network structure

No	Layer name	Output shape	Kernel size	Stride	parameter
1	Conv1d	69*100	5	2	600
2	Batch normalization	69*100	3	–	400
3	ReLU	69*100	–	–	0
4	Max-pooling1d	23*100	–	–	0
5	Conv1d	10*90	5	2	45090
6	Batch normalization	10*90	4	–	360
7	ReLU	10*90	–	–	0
8	Max-pooling1d	3*90	–	–	0
9	Conv1d	1*80	3	2	21680
10	Batch normalization	1*80	2	–	320
11	ReLU	1*80	–	–	0
12	Max-pooling1d	1*80	–	–	0
13	Flatten	80	–	–	0
14	Dense	30	–	–	2430
15	Dense	50	–	–	155

5.1 Evaluation metrics

Evaluation metrics that are used include sensitivity, positive predictive value, accuracy, and specificity according to AAMI instructions and, by using equations (7), (8), (9) and (11), are calculated.

Table 5 The summary of the proposed approach compared with other state-of-the-art studies

Article	Preprocessing	Feature exctraction	Classification	Accuracy	Sensitivity	Specificity	PPV
Joshi et al. [40]	Wavelet	Wavelet+PCA+ICA	SVM	86.4%	–	–	–
Zubair et al. [30]	Band pass filter	CNN	Softmax	92.7%	–	–	–
Acharya [27]	Daubechies wavelet six filters	–	CNN	94.3%	89.68%	97.05%	62.73%
Liu et al. [32]	–	LDA	CNN	96.47%	55.11%	86.81%	58.12%
Li et al. [26]	–	–	CNN	97.5%	–	–	–
Yang et al. [18]	–	PCA Net	SVM	97.77%	86.34%	97.74%	93.55
Oh et al. [28]	–	CNN and LSTM	Deep learning	98.10%	97.5%	98.7%	–
Proposed method	MinMax and SMOTE	DWT (db2)	CNN	98.53%	92.94%	99.15%	92.24%

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

$$Positive\ Productivity = \frac{TP}{TP + FP} \tag{8}$$

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

$$Specificity = \frac{TN}{TN + FP} \tag{10}$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative value.

In addition to the evaluation metrics outlined in the AAMI instructions, we further incorporated the Type II error (false negative) to provide a comprehensive assessment of the performance of our classification model in the context of ECG analysis. The Type II error measures the rate at which positive instances are inaccurately classified as negative by the model. Accurate detection of positive instances is of great importance in ECG analysis, particularly when the consequences of missing positive cases, such as critical cardiovascular abnormalities, are significant. The Type II error becomes particularly relevant in such scenarios, as minimizing false negatives is crucial to ensure accurate and timely diagnoses. The calculation of the Type II error rate involves considering the number of false negatives (FN) and the sum of false negatives and true positives (TP) according to the following formula:

$$Type\ II\ Error\ Rate = \frac{FN}{FN + TP} \tag{11}$$

By incorporating the Type II error into our evaluation metrics, we gain insights into the model's ability to correctly identify positive instances and minimize the rate of false negatives. It complements other evaluation metrics, such as sensitivity and positive predictive value, by specifically addressing the potential misclassification of positive cases as negative. The inclusion of the Type II error aligns our evaluation framework with the AAMI instructions, providing a comprehensive assessment of the model's performance. It allows for a deeper understanding of the model's sensitivity to positive instances and highlights the importance of minimizing false negatives for accurate diagnosis and patient care.

These metrics were calculated according to the separation of classes that are shown in Table 6.

Table 6 The confusion matrix of ECG heartbeats on test samples and evaluation criteria calculated from the results acquired from the execution of the model on test data

Class	F	N	S	V	Q	Acc (%)	Sen (%)	Spec (%)	PPV (%)	Type II
F	68	10	0	7	0	99.71	80.00	99.86	81.93	0.14
N	8	9006	32	29	15	98.69	99.08	96.80	99.35	0.01
S	0	23	250	1	0	99.45	91.24	99.66	87.41	0.16
V	7	23	4	651	1	99.33	94.90	99.63	94.48	0.07
Q	0	3	0	1	795	99.82	99.50	99.84	98.03	0.01

6 Discussion

As mentioned before, heartbeat classification approaches are the most useful and practical ways of preventing and recognizing cardiovascular diseases; it is an important research topic in combination with the processes of medicine and technology. This study aims to present a new heartbeat classifications approach based on Deep learning with minimum needed processes and high accuracy compared to other state-of-the-art approaches. Our main goal was to increase the accuracy in the classification of cardiovascular diseases. Table 6 presents a confusion matrix of ECG heartbeats across all folds. According to Table 6, less than 1.4% of the ECG heartbeats are wrongly classified. The minimal PPVs recorded are attributed to the detection of classes F is 81.93% and S is 87.41%. The overall average classification performance (Accuracy, PPV, Sensitivity, and Specificity) is collected in Table 5.

The analysis revealed varying False Negative Error rates across different ECG classes. Notably, for class F, the model exhibited a False Negative Error rate of 0.14. Although this indicates a relatively acceptable rate of misclassification, further optimization may be required to enhance its performance in this class. The model demonstrated exceptional precision in classifying class N, with a False Negative Error rate of only 0.01. This outcome suggests that our proposed method, effectively learned the distinguishing features of class N, resulting in minimal false negatives. For class S, the False Negative Error rate was recorded at 0.16. While this is relatively higher than desired, it could be attributed to the inherent complexity of class S patterns, necessitating more comprehensive feature extraction and data augmentation techniques. Moreover, class V showcased a False Negative Error rate of 0.07. While this indicates a reasonable performance, additional fine-tuning of the model's parameters may yield further improvements in the classification of this class. Finally, the model demonstrated outstanding performance in class Q, with an impressively low False Negative Error rate of 0.01. This suggests that our proposed method, successfully captured the distinctive attributes of class Q, resulting in highly accurate predictions.

In this section, we discuss the highlights of the proposed approach and the effect of each feature on the proposed approach's function. In recent years the function of many heart signal classifications has been evaluated using the MIT-BIH dataset and AAMI standard. In Table 5, the summary of several heartbeat classifications along with the accuracy of which in classification is described. In these approaches, a dataset used for training and evaluation is identical, and data are divided according to AAMI into five classes. In some research, such as [18, 26, 28, 31] researchers ignore the preprocessing levels to reduce the complexity of calculation and focus on increasing classification power.

In some studies, researchers against this opinion and believe that performing proper preprocesses on data not only does not cause an increase in complexity of calculation and classification time but also by processing raw data, causing a decrease in calculation volume and next level of classification inputs and then increased speed and accuracy of classification. For example, Zubair et al. [30], and Jiang et al. [33], and Zadeh et al. [41] used Bandpass filter, Joshi et al. [40] used Wavelet, and Acharya [27] used Daubechies wavelet six filters as a preprocess in their proposed method. At the preprocessing level, we used Min-Max Normalization to normalize heartbeats by focusing on the second approach and accelerating the operation routine, and decreasing calculation volumes. Then by using SMOTE oversampling algorithm, we try to remove the unbalanced problems in a sample number of each class in the MIT-BIH unbalanced dataset. Imbalanced datasets can negatively affect the overall performance of conventional and CNN-based classification systems. In order to alleviate this issue, synthetic data was generated to ensure that the number of samples in each class is proportional. The CNN model trained with well-balanced class data outperformed the model that was trained with a dataset that had a class imbalance as large as 9-fold. After the preprocessing level, there is an extraction level that in some studies such as [26, 30] to accelerate systematic classification of data and avoid additional processing is ignored. On the other hand, some other studies used different feature extraction approaches such as PCA [18], CNN [30], CWT [41], a combination of wavelet and PCA and ICA [29], a combination of CNN and LSTM [28] with the aim of essential feature extraction and ignoring unessential features and decrease processing volume and calculations complexity is used. At this level, we used a discrete wavelet transform using the db2 mother wavelet due to the short calculation time for extracting related features and sending them as input data of the proposed model. An important section of each classification approach is structures and the kind of classifier which is used. The researchers must concentrate on creating accurate classifiers with high accuracy and high speed with training and changing parameters and hyper-parameters and a combination of classifier structures. Some classifiers used in this section are SVM [18, 29], combination of SVM and GA [41], Neural network [31], Block based NN [33], Softmax [30], Deep learning [28], Convolutional neural networks [26, 27]. According to the structure and features of CNN, these networks can improve ECG analysis with a quick process of signals and learning the filters that are resistant to noises. The optimized CNN model in this study receives important extracted features and performs classification automatically. The proposed network finally achieves high classification accuracy (98.53%) compared with other studies. So, this experiment confirmed that balancing classes in the dataset can improve the classification of ECG events through CNN. This study shows that distribution in heart function is directly reflected in ECG heart signal morphological changes. Thus we conclude that the efficient use of CADs systems can help an accurate diagnosis of ECG signal changes. On the other hand, it is approved that a convolutional neural network, which is basically functional for 2D and 3D data processing, is functional for 1D data processing. Regarding the drawback of the proposed method, we can mention that it requires long training hours, specialized hardware to efficiently train (GPU), and the training is computationally expensive. Also, a very large number of signals is required to train the model that can reliably recognize multiple patterns. Nevertheless, after the training of the classification system has been successfully completed, the disease diagnosis operation is carried out very fast. In this way, the proposed system can be deployed in clinical applications to help cardiologists objectively detect ECG heartbeat signals. It can also be used in intensive care units (ICUs) and in rural areas where medical care is not easily available. The pitfalls of the proposed algorithm are as follows: (1) It is assumed that each ECG beat can include one type of arrhythmia which is not always true, and ECG beats can be including more than one arrhythmia type. (2) The system is improved by using an imbalanced dataset.

The overall purpose of this study was to perform the decreased needed calculation processes and increase classification accuracy by using lead II heartbeats MIT-BIH. Our work makes a notable contribution to the field of ECG classification through the effective combination of various techniques and methodological enhancements. By integrating heartbeat segmentation, oversampling, discrete wavelet transform, and a CNN architecture, we have created a comprehensive framework that leverages the strengths of each component. This combination allows for improved accuracy and performance in ECG classification tasks. Despite the presence of previous studies utilizing similar methods and combinations, our work has demonstrated superior results and accuracy, as evidenced by our comparative analysis. Through rigorous evaluation and comparison with existing works, we have consistently observed enhanced performance achieved by our proposed approach. The superiority of our findings indicates the effectiveness of our methodology in accurately classifying ECG signals. Our comprehensive evaluation of the MIT-BIH database highlights the advantages of our approach in achieving higher classification accuracy, particularly in distinguishing between different cardiac abnormalities. The significant improvements obtained by our work emphasize the added value of the integrated techniques and methodological enhancements we have employed. The superior results achieved by our work are attributed to the careful consideration of each component's impact and their synergistic integration. The effective combination of these techniques has resulted in a classification framework that surpasses previous approaches, providing more accurate and reliable ECG analysis. In conclusion, our work not only contributes to the field of ECG classification by combining various techniques and methodological enhancements but also outperforms previous studies that have utilized similar methods and combinations. Our findings reinforce the importance of comprehensive integration and optimization of multiple components in ECG classification algorithms. We believe that our work provides valuable insights and sets a new benchmark for accurate and reliable ECG classification.

Thus next step in addition to finding out a simple and effective classification approach for optimizing convolutional neural networks, and getting better results, is using the proposed approach for testing a combination of other leads for more enrichment of empirical content. It's worth mentioning that frequency-based analysis of ECG signals has limitations that need to be acknowledged, and further investigation is necessary to explore specific frequency ranges associated with different cardiovascular diseases. These limitations include the complexity of frequency content in ECG signals, inter-subject variability in frequency characteristics, the dynamic nature of frequency content, unexplored frequency bands with potential diagnostic value, and dataset specificity. Addressing these limitations requires advanced signal processing techniques, large-scale studies, and exploration of alternative frequency bands. Further research into specific frequency ranges associated with different cardiovascular diseases will enhance the accuracy and clinical relevance of ECG classification, leading to improved diagnosis and management of these conditions.

7 Conclusion

This study addresses the significance of accurate and efficient classification of ECG signals in cardiovascular disease diagnosis. The primary objective is to propose an approach that combines a 15-layer CNN architecture with the utilization of FFT and the DB2 mother wavelet transform for feature extraction. To streamline the computational process, ECG beats are extracted from the raw ECG signals, followed by the removal of irrelevant data features. The preprocessing steps include signal normalization and the application of the discrete wavelet transform using the DB2 mother wavelet. This enables the extraction of essential features and key points crucial for recognizing different types of cardiovascular diseases. The dataset is divided into a 1:9 ratio for the training and testing sets, respectively. The proposed 15-layer CNN architecture encompasses a 1D convolutional layer, a Batch Normalization layer, a Rectified Linear Unit (ReLU) layer, a Max-pooling layer, a flattening layer, a fully-connected layer, an Adam optimizer algorithm, and sparse categorical cross-entropy as the loss function. This architecture demonstrates its effectiveness in accurately classifying ECG signals into five distinct classes. The proposed approach utilizes appropriate preprocessing techniques, such as signal normalization and the discrete wavelet transform, to enhance the accuracy and efficiency of ECG signal classification. Additionally, the integration of a powerful CNN architecture plays a crucial role in significantly improving the overall efficiency of the classification process. This combination of preprocessing techniques and the CNN network results in enhanced accuracy and efficiency, making the proposed approach a valuable tool for precise ECG signal analysis. Comparisons between the proposed approach and state-of-the-art methods, as shown in Table 5, support its efficacy as an accurate classifier for ECG signal classification. Furthermore, the utilization of the FFT and the DB2 mother wavelet transform provides insights into the frequency characteristics of ECG signals. This enables the extraction of relevant frequency-based features, which play a vital role in detecting specific cardiac abnormalities. The selection of the DB2 mother wavelet and the justification for its suitability in capturing the frequency content of ECG signals contribute to the robustness and scientific rigor of the proposed approach. Also, the investigation into the False Negative Error for each class allowed us to identify areas of strength and areas requiring improvement. The model's robust performance in classes N and Q highlights its ability to effectively discriminate between these classes with minimal false negatives. However, further refinement and feature engineering might be necessary to enhance the model's sensitivity to class F, S, and V patterns.

In conclusion, the proposed approach integrates a 15-layer CNN architecture with the use of FFT and the DB2 mother wavelet transform for accurate and efficient classification of ECG signals. The achieved results demonstrate its efficacy in cardiovascular disease diagnosis. The combination of computational efficiency, accurate classification, and consideration of frequency characteristics substantiate the potential of the proposed approach in practical applications within the medical field. As this study lays the groundwork for future research, it is hoped that continued advancements in deep learning models will lead to even more accurate and reliable ECG classification systems for medical diagnosis and patient care.

References

1. Marmot M, Theorell T (1988) Social class and cardiovascular disease: the contribution of work. Int J Health Services 18(4):659–674
2. World Health Organization, Cardiovascular Diseases (CVDs). https://www.who.int/health-topics/cardiovascular-diseases/. Accessed: 15-09-2020
3. Balaskas K, Siozios K (2019) Ecg analysis and heartbeat classification based on shallow neural networks. In: 2019 8th International conference on modern circuits and systems technologies (MOCAST), pp 1–4. https://doi.org/10.1109/MOCAST.2019.8742072
4. ANSI-AAMI (2008) Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms. American National Standards Institute, Inc. (ANSI), Association for the Advancement of Medical Instrumentation (AAMI). ANSI/AAMI/ISO (1998,2008)
5. Mathews SM, Kambhamettu C, Barner KE (2018) A novel application of deep learning for single-lead ecg classification. Comput Biol Med 99:53–62
6. Elhaj FA, Salim N, Harris AR, Swee TT, Ahmed T (2016) Arrhythmia recognition and classification using combined linear and nonlinear features of ecg signals. Comput Methods Prog Biomed 127:52–63
7. Venkatesan C, Karthigaikumar P, Paul A, Satheeskumaran S, Kumar R (2018) Ecg signal preprocessing and svm classifier-based abnormality detection in remote healthcare applications. IEEE Access 6:9767–9773. https://doi.org/10.1109/ACCESS.2018.2794346
8. Azariadi D, Tsoutsouras V, Xydis S, Soudris D (2016) Ecg signal analysis and arrhythmia detection on iot wearable medical devices. In: 2016 5th International Conference on Modern Circuits and Systems Technologies (MOCAST), pp 1–4. IEEE
9. de Chazal P, O'Dwyer M, Reilly RB (2004) Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE Trans Biomed Eng 51(7):1196–1206. https://doi.org/10.1109/TBME.2004.827359
10. Li Q, Rajagopalan C, Clifford GD (2013) Ventricular fibrillation and tachycardia classification using a machine learning approach. IEEE Trans Biomed Eng 61(6):1607–1613
11. Rajagopal R, Ranganathan V (2017) Evaluation of effect of unsupervised dimensionality reduction techniques on automated arrhythmia classification. Biomed Signal Process Control 34:1–8
12. Martis RJ, Acharya UR, Min LC (2013) Ecg beat classification using pca, lda, ica and discrete wavelet transform. Biomed Signal Process Control 8(5):437–448
13. Ghorbani Afkhami R, Azarnia G, Tinati MA (2016) Cardiac arrhythmia classification using statistical and mixture modeling features of ecg signals. Pattern Recogn Lett 70:45–51. https://doi.org/10.1016/j.patrec.2015.11.018
14. Kiranyaz S, Ince T, Gabbouj M (2015) Real-time patient-specific ecg classification by 1-d convolutional neural networks. IEEE Trans Biomed Eng 63(3):664–675
15. Dohare AK, Kumar V, Kumar R (2018) Detection of myocardial infarction in 12 lead ecg using support vector machine. Appl Soft Comput 64:138–147

16. Shimpi P, Shah S, Shroff, M, Godbole A (2017) A machine learning approach for the classification of cardiac arrhythmia. In: 2017 International conference on computing methodologies and communication (ICCMC), pp 603–607 . IEEE
17. Alonso-Atienza F, Morgado E, Fernandez-Martinez L, García-Alberola A, Rojo-Alvarez JL (2013) Detection of life-threatening arrhythmias using feature selection and support vector machines. IEEE Trans Biomed Eng 61(3):832–840
18. Yang W, Si Y, Wang D, Guo B (2018) Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine. Comput Biol Med 101:22–32
19. Ochoa A, Mena LJ, Felix VG (2017) Noise-tolerant neural network approach for electrocardiogram signal classification. In: Proceedings of the international conference on compute and data analysis, pp 277–282
20. Xia Y, Zhang H, Xu L, Gao Z, Zhang H, Liu H, Li S (2018) An automatic cardiac arrhythmia classification system with wearable electrocardiogram. IEEE Access 6:16529–16538. https://doi.org/10.1109/ACCESS.2018.2807700
21. Desai U, Martis RJ, Nayak CG, Sarika K, Nayak SG, Shirva A, Nayak V, Mudassir S (2015) Discrete cosine transform features in automated classification of cardiac arrhythmia beats. In: Emerging research in computing, information, communication and applications, pp 153–162. Springer, Berlin
22. Raj S, Ray KC (2018) Sparse representation of ecg signals for automated recognition of cardiac arrhythmias. Exp Syst Appl 105:49–64
23. Hammad M, Maher A, Wang K, Jiang F, Amrani M (2018) Detection of abnormal heart conditions based on characteristics of ecg signals. Measurement 125:634–644
24. Zhai X, Tin C (2018) Automated ecg classification using dual heartbeat coupling based on convolutional neural network. IEEE Access 6:27465–27472
25. Xiang Y, Luo J, Zhu T, Wang S, Xiang X, Meng J (2018) Ecg-based heartbeat classification using two-level convolutional neural network and rr interval difference. IEICE Trans Inf Syst 101(4):1189–1198
26. Li D, Zhang J, Zhang Q, Wei X (2017) Classification of ecg signals based on 1d convolution neural network. In: 2017 IEEE 19th international conference on e-health networking, applications and services (Healthcom), pp 1–6. IEEE
27. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, San Tan R (2017) A deep convolutional neural network model to classify heartbeats. Comput Biol Med 89:389–396
28. Oh SL, Ng EY, San Tan R, Acharya UR (2018) Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats. Comput Biol Med 102:278–287
29. Park K, Cho B, Lee D, Song S, Lee J, Chee Y, Kim I, Kim S (2008) Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function. In: 2008 Computers in cardiology, pp 229–232. IEEE
30. Zubair M, Kim J, Yoon C (2016) An automated ecg beat classification system using convolutional neural networks. In: 2016 6th International conference on IT convergence and security (ICITCS), pp 1–5. IEEE
31. Li P, Wang Y, He J, Wang L, Tian Y, Zhou T-S, Li T, Li J-s (2016) High-performance personalized heartbeat classification model for long-term ecg signal. IEEE Trans Biomed Eng 64(1):78–86
32. Liu J, Song S, Sun G, Fu Y (2019) Classification of ecg arrhythmia using cnn, svm and lda. In: International conference on artificial intelligence and security, pp 191–201. Springer
33. Jiang W, Kong SG (2007) Block-based neural networks for personalized ecg signal classification. IEEE Trans Neural Netw 18(6):1750–1761
34. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE (2000) Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation 101(23):215–220
35. Moody GB, Mark RG (2001) The impact of the mit-bih arrhythmia database. IEEE Eng Med Biol Mag 20(3):45–50
36. Dibal PY, Onwuka EN, Agajo J, Alenoghena CO (2018) Application of wavelet transform in spectrum sensing for cognitive radio: a survey. Phys Commun 28:45–57. https://doi.org/10.1016/j.phycom.2018.03.004
37. Chakrabarti C, Vishwanath M, Owens RM (1996) Architectures for wavelet transforms: a survey. J VLSI Signal Process Syst Signal Image Video Technol 14(2):171–192
38. Stephane M (1999) A wavelet tour of signal processing. Elsevier

39. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
40. Joshi N, Topannavar P (2014) Support vector machine based heartbeat classification. In: Proc. of 4th IRF Int. Conf, pp 140–144
41. Zadeh AE, Khazaee A (2011) High efficient system for automatic classification of the electrocardiogram beats. Ann Biomed Eng 39(3):996–1011

Authors and Affiliations

Armin Shoughi[1] · Mohammad Bagher Dowlatshahi[1] · Arefeh Amiri[2] · Marjan Kuchaki Rafsanjani[3] · Ranbir Singh Batth[4]

✉ Marjan Kuchaki Rafsanjani
kuchaki@uk.ac.ir

Armin Shoughi
shoughi.ar@fe.lu.ac.ir

Mohammad Bagher Dowlatshahi
dowlatshahi.mb@lu.ac.ir

Ranbir Singh Batth
ranbir.21123@lpu.co.in

1 Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran

2 Shahid Madani Hospital, Lorestan University of Medical Sciences, Khorramabad, Iran

3 Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

4 School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India