International Journal of Advanced Technology and Engineering Exploration, Vol 9(95)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2021.876071

Research Article

A novel weighted approach for automated cardiac arrhythmia beat classification using convolutional neural networks

Ravindar Mogili[1*] and G. Narsimha[2]
Research Scholar, Department of Computer Science, Jawaharlal Nehru Technological University, Hyderabad, India[1]
Professor, Department of Computer Science, Jawaharlal Nehru Technological University, Hyderabad, India[2]

Received: 25-May-2022; Revised: 26-October-2022; Accepted: 28-October-2022

Abstract
Arrhythmia is a cardiac disorder in which the normal blood pumping activity of the heart becomes irregular. This heart malfunction can result in serious heart disease and even death. Therefore, detection and proper treatment of arrhythmia are essential. The abnormal heart behaviour can be recorded using an electrocardiogram (ECG). A one-dimensional convolutional neural network (1D-CNN) with a novel weighted approach was proposed to detect and classify arrhythmia types from ECG signals. The proposed classifier was trained and evaluated using the Massachusetts institute of technology-Beth Israel hospital (MITBIH) arrhythmia database to classify five arrhythmia beat categories (N, S, V, F, and Q), as recommended by the Association for Advancement of Medical Instrumentation (AAMI). The proposed model obtained an overall sensitivity of 94.35%, precision of 94.02%, specificity of 99.5%, and accuracy of 99.65%. The experimental results demonstrate that the proposed CNN model can achieve cutting-edge performance and can be used for arrhythmia diagnosis in real-time.

Keywords
Heart disease, ECG, Arrhythmia, Convolution Neural Network, AAMI.

1.Introduction
According to the World Heart Federation, there will be 23 million deaths due to cardiovascular disease (CVD) annually by 2030 [1]. This is 28.5% higher than the CVD deaths that occurred in 2019. This indicates the need to take appropriate measures to reduce the mortality rate. One broad group of CVDs is arrhythmia, which manifests as an abnormal electrical behaviour of the heart [2]. This electrical activity can be measured using an electrocardiogram (ECG) [3]. The morphological pattern changes in the ECG can be used to identify the arrhythmia type [4]. Some arrhythmia can be potentially life-threatening and some others are not. There are five groups of arrhythmias in the association for advancement of medical instrumentation (AAMI) dataset: non-ectopic (N), ventricular ectopic (V), supraventricular ectopic (S), fusion (F), and unknown (Q) [5]. Although these arrhythmia types are not fatal, they can lead to life-threatening complications if not treated in a timely manner [6].

*Author for correspondence

Therefore the early detection and classification of these arrhythmia types are essential to provide proper treatment and decrease the CVD death rate [7].

To accurately distinguish and classify abnormal heartbeats from ECG signals, remarkable expertise in this area is required. This criterion imposes a number of limitations on professional ECG data analysis: i) it is prone to human error and a time-consuming process. ii) The number of cardiologists available to diagnose heart problems in a large population is limited. iii) The cost of diagnosis is high. These limitations highlight the need for reliable and inexpensive methods to accurately identify and classify abnormal beats to ensure a person's heart health [8]. Therefore, automated arrhythmia identification and classification models are required.

Over the past decade, researchers have developed various pattern recognition models for the automated identification and classification of the arrhythmia beats. The already available arrhythmia identification and classification models have proven to be quite useful for cardiologists and medical institutions [9]. However, ECG beat classification performance can be further improved. The classification of ECG beats is typically performed mainly in three phases [10]. They are: (i) filter noise from raw ECG signals (ii) extract and select the optimal set of features, and (iii) perform classification to identify the beat type. The classification performance depends upon quality and distinctive features extracted from the ECG signal to understand and analyze arrhythmia characteristics. The traditional machine learning algorithms can only perform classification tasks. Noise removal, feature extraction, and selection were performed using other methods. In addition, human interaction may be required between the above stages. However, the convolutional neural network (CNN) can handle all three phases on its own. That is, it can extract and choose the optimal features from acquired raw ECG signals and perform classification. That was the reason, deep neural networks, such as CNN have recently attracted a lot of attention in the applications of biological signal processing.

Even though deep learning models produced promising outcomes in arrhythmia classification, they were seriously influenced by imbalanced data. When a dataset contains extremely imbalance data, especially in deep learning, training models become more biased towards majority class samples [11]. In case of arrhythmia datasets, most of the abnormal heart beats are very rare, which restricts the development of automated models. There exist different methods for dealing with imbalanced data, including affine transformation, random translation, random flipping, and random cropping, and so on. These techniques, however, are not recommended for augmenting because morphological data in ECG signals has more meaning. Beside these, under-sampling and over-sampling can handle imbalanced data classification; each has its own problems. Oversampling not only increases training time but may also lead to model over-fitting. In the case of under-sampling, the model may fail to learn a few properties of the majority classes owing to the dropping of some samples. Therefore, a novel weighted method is proposed to handle imbalanced data. The novel weighted mechanism assigns weights to AAMI classes. The purpose is to pay more attention to minority class beats while training. Additionally, number of CNN layers, optimal number of neurons to be used in the fully connected layers, batch size, loss function, activation function, optimizer, mechanism to prevent over-fitting, and other factors to be determined and optimized. This study enhanced arrhythmia classification performance when compared to state-of-the-art approaches.

The remainder of this paper is organized as follows. Section 2 provides the literature survey. Section 3 discusses the structure of the proposed one-dimensional convolutional neural network (1D-CNN) model, the novel weighted approach and also details of database used in this study. Section 4 provides the details of training and evaluation of proposed 1D-CNN model, as well as the evaluation metrics, and experimental results. Section 5 presents a discussion and comparison of proposed work with the state-of-the-art methodologies, and limitations of the proposed work. The future enhancements and conclusions are presented in section 6.

2.Literature survey

In existing literature, researchers have classified arrhythmia beats using both traditional machine learning algorithms (random forest, support vector machine (SVM), etc.) and deep learning algorithms (CNN, recurrent neural network (RNN), etc). In traditional approach, separate methods are required to extract important features from ECG and then appropriate classifier to be selected in the classification stage. Mian and Fawad [12] extracted statistical features from sub-band coefficients of wavelet decomposition of ECG signal. They classified five arrhythmia types using random forest and achieved 97% accuracy. Sharma et al. [6] applied discrete wavelet transform (DWT) on ECG signal and extracted fractal dimension, Renyi and fuzzy entropy features. Later, k-nearest neighbours (KNN) as used to classify AAMI five classes and attained 98% accuracy. Sahoo et al. [13] extracted Hilbert and wavelet transform based features from the ECG signals. Later, they applied principal component analysis (PCA) to reduce the feature set and used SVM classifier and obtained 98.5% accuracy for classifying five arrhythmia types. Sultan and Ghorbani [14], proposed time–frequency (TF) representation-based algorithm to extract ECG features. TF features along with higher order statistical features and R-R interval of ECG beats are given to ensemble-based decision trees to classify them into 5 AAMI classes. The results demonstrate that N, V, and S classes have accuracy over 99%. But F class beats achieved very less sensitivity of 12.11% and positive predictive value of 51.09%. However, all of these approaches used traditional machine learning algorithms and required hand-crafted features.

The adequacy of the features extracted from ECG signals has substantial effect on the overall performance and reliability of the classification algorithm [15]. The auto-extracted features by CNN have higher quality than hand-crafted features, which enhances classification performance. But different CNN structures exhibit different feature extraction capabilities [16]. Therefore, determining the optimal CNN structure is a design issues. The training time of a CNN model increase when the number of layers increases. However, graphics processing units (GPUs) can be used to efficiently train convolutional based models.

Various noise removal methods have been utilized to improve the classification results of arrhythmia [5, 6, 16, 17]. Acharya et al. [5] proposed a nine-layer CNN model for classifying five categories of ECG beats. They experimented with the original and noise-free ECG data. They found that the average sensitivity and accuracy increased for noise-free ECG signals compared with the original. Yao et al. [16] pre-processed raw ECG signals to filter noise using wavelet decomposition and balanced the classes using data augmentation techniques. Their results proved that data augmentation and the addition of the gated recurrent units (GRUs) to the CNN model can improve the sensitivity of each AAMI beat category. Khan et al. [17] classified five different types of cardiac arrhythmia using a ten-layer 1D-CNN model. They used segmented noise-free ECG data after balancing the classes using the augmentation technique as the input. Their model attained 95.2% overall accuracy, 95.4% and 95.2% average recall and precision respectively. Sharma et al. [6] proved that the average precision, specificity, sensitivity, and accuracy of noise-free ECG data for five classes were better than those of original noisy data.

CNN models can also be deployed in wearable devices to monitor the heart activity in real -time. It can also be trained to detect the abnormal beats of a particular individual. Xiolin et al. [18] proposed a ten-layer CNN model suitable for wearable devices to detect arrhythmia beat types. They reduced and optimized the complexity of the CNN using the multistage pruning technique. Kiranyaz et al. [4] developed a dedicated CNN model for each patient, separately from a generalized model. They claimed that such a model could classify ECG beats of that patient quickly and more accurately. Sarvan and Özkurt [15] applied a nine-layer 1D-CNN model to the original unbalanced AAMI classes of ECG data to detect five categories of the ECG beats. They claimed that when the number of epochs was increased to 300, their model performed better.

Although CNN can auto-extract features and perform classification, researchers used CNN to implement classification with hand-crafted features. Al et al. [19] applied continuous wavelet transform on ECG data to convert it into input that is suitable to pre-trained CNN model popularly known as visual geometry group network (VGGNet) to generate the features. These obtained features are given to fully connected layers to classify the beats. Li and Boulanger [7] proposed to combine hand-crafted features such as statistical and morphological with features extracted from short-time Fourier transform spectrogram of ECG using a CNN to detect abnormal beats. Yu [20] used the wavelet transform to preprocess the ECG signal to detect R peak and find R-R interval. Then ECG segments formed from R peaks are given to an eight-layer CNN model for classifying them into four AAMI classes (N, S, V, F). Their model attained sensitivity of 93.0% and 81.3%, respectively for V and S class beats. Huang et al. [3] transformed the ECG signals into ECG spectrograms using short-time Fourier transform. These spectrograms are given as input to the two-dimensional convolution neural network (2D-CNN) model and classified five different arrhythmia types and achieved 99% average accuracy. Even though 2D-CNN model accuracy improved over 1D-CNN, computation cost increased significantly.

Innovative approaches to extract features from ECG signals and use of attention layer to enhance quality of features have introduced in the recent literature. Mousavi et al. [21] developed a novel approach called ECG language processing (ELP), which relies on variations in a series of different wave morphologies for interpreting ECG beats. This is analogous to word sequences in sentences and inspired by natural language processing. It allows a computer-aided system to interpret ECG beats in the same way doctors do. The proposed ELP approach with RNN attention-based model achieved 97% accuracy when tested on Massachusetts institute of technology-Beth Israel hospital (MITBIH) arrhythmia database according to AAMI criteria. Ma et al. [22] proposed a residual network (ResNet) and bidirectional long short-term memory (BiLSTM) model with an attention mechanism to extract and classify arrhythmia beats according to AAMI criteria. Initially, the ECG signal is divided into segments containing arrhythmia beats after denoising it with DWT. Following that, generative adversarial networks (GAN) was used to balance the beats in minority classes. Later, the spatial characteristics obtained from the segmented ECG beats using ResNet are fused with the temporal characteristics obtained from Bi-LSTM, and then feature enhancement is performed using the attention layer. The proposed model performs classification based on these enhanced features and achieves 99.4% accuracy. Lu et al. [23] proposed an end-to-end classification approach based on CNN and long short-term memory (LSTM) for classifying arrhythmia beats. CNN first extract local morphological features from raw ECG signal, followed by LSTM mining temporal correlated morphological features and perform classification. When tested against the MIT-BIH arrhythmia database, this model has an accuracy of 96.16%. Shoughi and Dowlatshahi [24] designed a hybrid model by merging CNN and BiLSTM in order to accurate classify arrhythmia in ECG heartbeats. In pre-processing stage, they utilized wavelet transform to denoise the ECG data and synthetic minority over-sampling technique (SMOTE) to balance the arrhythmia beat types. In the classification step, CNN is used to extract local features and BLSTM is used to extract correlated high-level features in order to acquire the deep features and conduct classification in accordance with the AAMI standard. The accuracy of this model was 98.71%. Gai [25] transformed a one-dimensional ECG waveform into two-dimensional in order to extract rich informative features using 2D-CNN. They addressed the issue of class imbalance by augmenting artificially created minority beats with up to 10% Gaussian noise in signal amplitude. This model obtained an accuracy of 98.65% when tested against MIT-BIH arrhythmia database. Liu and Zhang [26] developed a novel arrhythmia classification method by including an attention layer into CNN. This model extracted the most informative features straight from raw ECG using an attention mechanism and increased the efficacy of the arrhythmia identification procedure. When evaluated on the MIT-BIH arrhythmia database, it scored average precision of 98.65% and average recall of 98.68%. Zubair and Yoon [27] proposed a novel deep-learning framework that includes temporal transition module composed of numerous convolutional layers with different size kernels. This approach extract both short-term and long-term morphological variations from ECG beats. They handled class imbalance problem by introducing a cost-sensitive loss function that assigns adaptive weights to class samples based on data distribution in the training batch. The proposed model obtained accuracy of 99.81%, sensitivity of 88.82%, and precision of 95.68%. Even though, the accuracy of this model is high, it failed to obtain better sensitivity.

The AAMI dataset formed from the MITBIH arrhythmia database was highly unbalanced. To deal with this problem, various augmentation techniques are used. This involves creating synthetic data through SMOTE, replicating the data, and using transformation operations such as reflection and rotation [17]. Pandey and Janghel [2] designed an 11-layer 1D-CNN model to classify MIT-BIH arrhythmia beats into five categories according to AAMI standard. They used the SMOTE technique to address class imbalance problems. Their model achieved 98.30% accuracy when using a train-test split of 70:30. Jiang et al. [28] used an over-sampling strategy to deal with imbalance data. The auto-encoder was used to extract features, which were then given to CNN after minority classes were over-sampled. This approach achieves a total of 96.6% accuracy in classifying N, V, S, and F category beats according to the AAMI recommendation. However, over-sampling method may result in over-fitting. The authors [5, 16, 17, 22] also used augmented data to balance the AAMI classes. The augmentation not only increases the computing costs significantly, but also brings in new training data that might be quite different from the real data to be tested [29].

The following findings are drawn from the study and analysis of existing works. Since the ECG signal varies widely between individuals and under different physical situations, hand-crafted features are not generalized enough to classify ECG arrhythmia. However, auto-extracted features can bypass these constraints. Second, as CNN can handle noisy data, raw ECG signals can be directly fed to it without any noise filtering effects to reduce computational cost. Finally, the dataset formed according to AAMI criteria is highly unbalanced which has an impact on minority class performance. To address this issue, minority classes were either over-sampled or various data augmentation techniques were used. Due to the negative consequences of augmentation and over-sampling, it is not always appropriate to balance datasets using them. Therefore, dealing with highly unbalanced data requires the development of innovative approaches.

3.Proposed model

The proposed model consists of a 9-layer 1D-CNN model with novel weighted mechanism to classify arrhythmia beats according to AAMI standard by taking raw ECG signal as input. The proposed model architecture is shown in *Figure 1*. The openly available MIT-BIH arrhythmia database [30] was used in this study. It is collected from the following link: https://physionet.org/content/mitdb/1.0.0/. This dataset contains raw ECG signal recordings that have been divided into multiple segments such that each segment with a fixed length and a R-peak in the middle. These extracted ECG segments are fed into a 9-layer 1D-CNN model to predict the class label associated with it. As the classes in the dataset are highly imbalanced, a novel weighted mechanism is introduced to improve the performance of the minority classes as well as overall classification. The role of novel weighted mechanism is to increase the importance of minority classes. Finally, the proposed architecture able to predict the class label associated with input ECG segment as per AAMI standard.

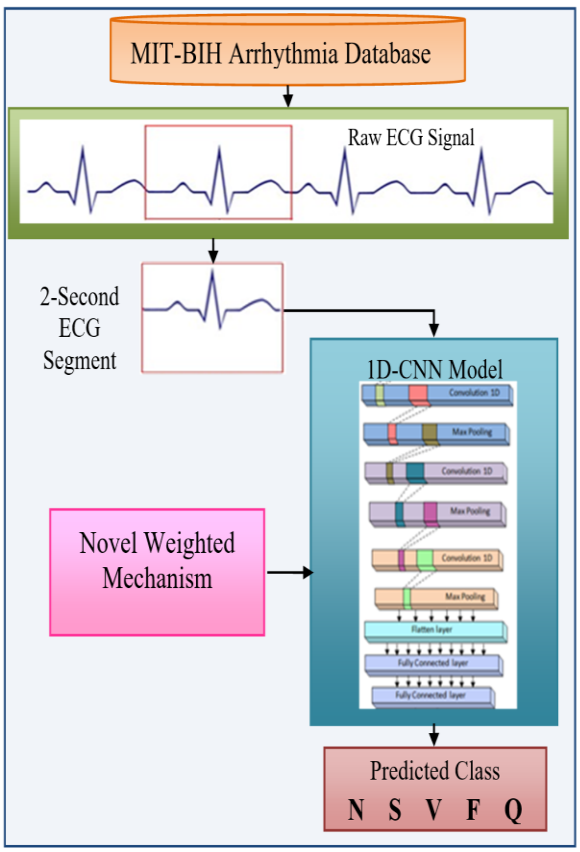

Figure 1 Overview of proposed model

3.1Database used

The MIT-BIH arrhythmia database included 48 ECG records collected from different patients. Each record contains 30-minutes of ECG data obtained using Lead II with a sampling rate of 360 Hz. In each ECG recording, all R-peaks were annotated with the relevant arrhythmia beat type [31]. A Python package called waveform database (WFDB) was used to read the ECG waveform data and the corresponding annotations from these recordings. From each R peak, a 2-seconds ECG segment was produced with a window size of 1-second or 360 samples to the left and 359 samples to the right, with reference to the R-peak position. In other words, the 720 samples represented an ECG segment of 2-seconds. The annotated beat label at the centre R-peak of an ECG segment serves as a beat class. Sample 2-second ECG segments from the generated dataset are presented in *Figure 2*.

A dataset was formed with each generated ECG segment as an input and the beat type associated with it as an output. *Table 1* provides details of the dataset. It contains 15 arrhythmia types and their ECG segment numbers from the MIT-BIH arrhythmia database. According to the AAMI standard, these 15 arrhythmia types were clustered into five groups and their details are also listed in *Table 1*. The generated AAMI dataset was randomly shuffled before splitting. This enabled the training model to conduct a thorough and unbiased exploratory data analysis. *Table 1* clearly demonstrates that the AAMI beat classes were highly imbalanced. Therefore, a stratified train-test split strategy of 70:30 was used. This strategy splits the dataset into two sets such that the training set contains randomly selected 70% of each class data and the remaining 30% of the data is placed in the testing set. This split approach ensured to spread of an equal proportion of beats from all five classes to both sets. The details of the beats obtained according to this split strategy for each class in the testing and training sets are presented in *Table 1*. The training set was further divided into two more sets such that a set with 70% of the data was utilized for training the model and the other 30% was utilized to validate the trained model. This division is clearly illustrated in *Figure 3*.

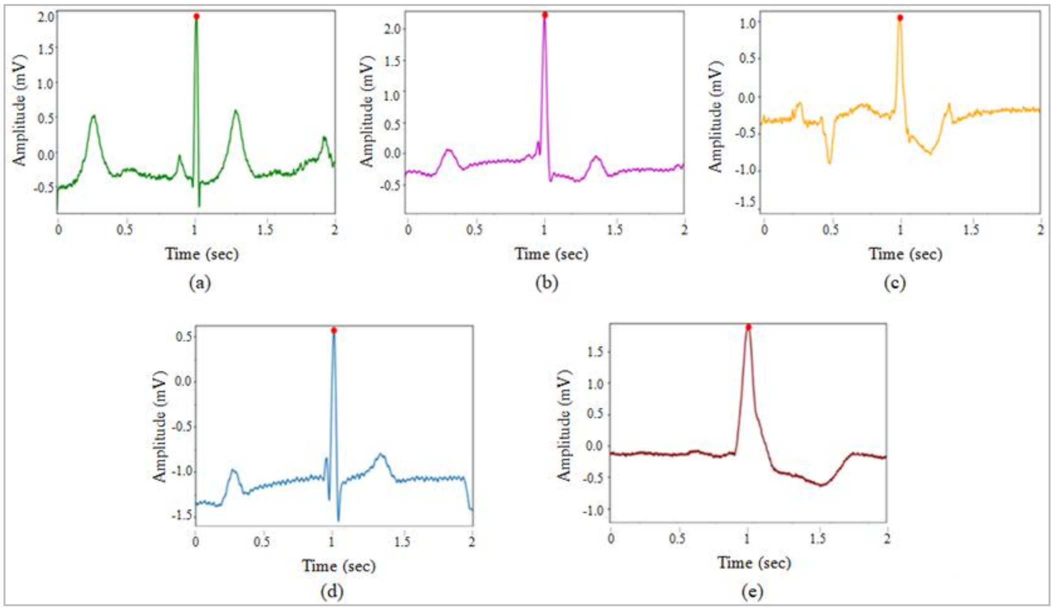

Figure 2 Sample extracted ECG segment beats of (a) Non Ectopic (N) (b) Supraventricular Ectopic (S)(c) Ventricular Ectopic (V) (d) Fusion (F) (e) Unknown (Q)

Table 1 Extracted ECG segments

MIT BIH Arrhythmia Beats		**AAMI Arrhythmia Beats**			
Name of the Beat	**Total Beats**	**Name of the Beat**	**Total Beats**	**Training Beats (70%)**	**Testing Beats (30%)**
Normal beat	74926	Non ectopic beat (N)	90488	63212	27276
Nodal Escape beat	229				
Left bundle branch block	8066				
Right bundle branch block	7251				
Atrial escape beat	16				
Supraventricular premature beat	2	Supraventricular ectopic beat (S)	2780	1982	798
Nodal (junctional) premature beat	83				
Aberrated atrial premature beat	150				
Atrial premature beat	2545				
Premature ventricular contraction	7126	Ventricular ectopic beat (V)	7232	5103	2129
Ventricular escape beat	106				
Fusion of ventricular and normal beat	802	Fusion beat (F)	802	530	272
Unclassified beat	33	Unknown beat (Q)	8033	5695	2338
Fusion of paced and normal beat	982				
Paced beat	7018				

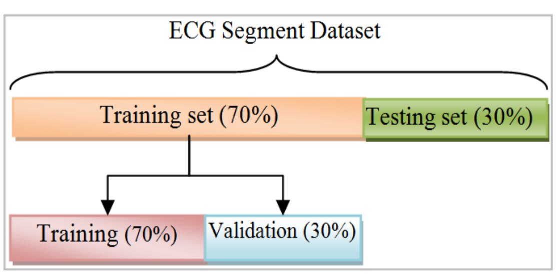

Figure 3 Division of ECG segments as training, testing and validation sets

3.2CNN model architecture

The proposed CNN model contains nine layers: three sets of 1D convolution layers and max-pooling layers, then follows a flattened layer, and finally two sets of fully connected layers. This model uses 720 samples that represent a 2-second ECG segment as input and classifies them into five categories of arrhythmia beats. The CNN analyzes these samples from a segmented ECG image to identify the type of arrhythmia present. The proposed CNN model is illustrated in *Figure 4*.

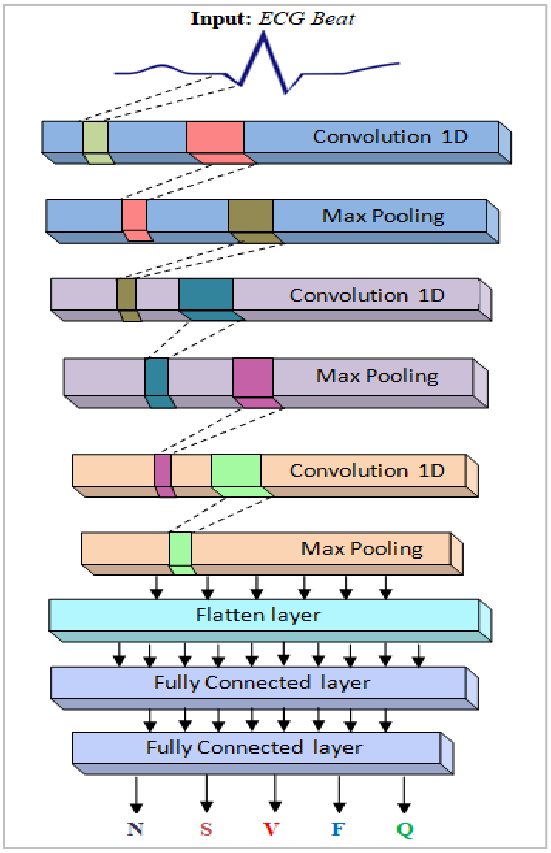

Figure 4 Proposed CNN architecture

The first convolution layer collects preliminary features from the segmented raw ECG beats. The subsequent convolution layers further improve and detect other deeper features. The max-pooling layers role is to decrease the dimensionality of their input to summarize the features detected. This activity minimizes computational complexity of the model by reducing the number of parameters required [16]. The flattened layer flattens the multi-dimensional input array into a single dimension tensor; therefore, it can be suitable as an input to a dense layer. At the end, the classification task carried out using two fully-connected dense layers.

The first 1D convolution layer takes the ECG segments of 720 samples as an input. Then it applies six filters, and the kernel size of five on an input sample of 720×1 to produce a 716×6 matrix as the output. The max-pooling layer after the first 1D convolution layer uses this 716×6 matrix as input to reduce the dimensionality and produce an output of 238×6 matrix. The next 1D convoluted layer uses twelve filters, and the kernel size of five on the input received from its preceding layer to obtain a 234×12 matrix. The second max-pooling layer decreased the dimensionality of the input received to generate 77×12 feature maps as the output. The third and the last convolution layer produce an output of 73×24 matrix by using kernel size of five and thirty-six filters. The third and the last max-pooling layer summarize the extracted features as an output matrix of 24×24. The first fully connected dense layer and also all the three convolution layers used the "relu" as an activation function. The pool size and stride were set as 4 and 3, respectively, in all max-pooling layers. The flattened layer transforms the input received from the final max-pooling layer to obtain 576 intermediate values which were given as inputs to the first fully connected dense layer to obtain 128 outputs. Finally, the second and the last fully connected dense layer use the “softmax” as an activation function on its input to detect the ECG segment class label.

3.3Novel weighted approach

The loss of a multi-class classification problem is typically calculated using the categorical cross entropy loss function [32]. While training, it simply neglects the appearance frequency of the beat class, producing in lower prediction results for classes with fewer samples or minority classes. Wang et al. [33] and Mahajan et al. [34] addressed this issue by assigning appropriate weight to each class type by inferring inverse of the sample number. The weight assigned to a class is inversely proportional to the class frequency. If the number of instances in a class is high, then the weight of that class is low and vice versa. In other words, the samples from the majority class were given less weight, while the samples from the minority class were given more weight. It is given by Equation 1.

$$\mathrm{weight_i\_old} = \frac{1}{\mathrm{no\_of\_samples\_in\_class_i}} \qquad (1)$$

It was observed that using appearance frequency or inverse number of an object class directly leads to highly unstable training, since loss exponentially increases when the training batch includes objects from minority classes. To overcome this problem, we used a linearly scaled form of class weights. That is, Equation 1 multiplied by sum of samples of all the classes and total number of classes represented by n. It is given by Equation 2.

$$\mathrm{weight_i\_mod} = \left(\frac{\left(\sum_{j=1}^{n} \mathrm{no\_of\_class_j\_samples}\right) \times n}{\mathrm{no\_of\_samples\_in\_class_i}}\right) \quad (2)$$

The class weight value can be reduced by taking square root value of Equation 2 and it is given by Equation 3.

$$\mathrm{weight_i\_new} = \sqrt{\left(\frac{\left(\sum_{j=1}^{n} \mathrm{no\_of\_class_j\_samples}\right) \times n}{\mathrm{no\_of\_samples\_in\_class_i}}\right)} \quad (3)$$

When the class weights derived by Equation 3 used, the importance of minority classes is preserved and also loss does not rise exponentially when the training batch includes objects from minority classes.

4.Experimental setup and results

The proposed CNN model was run on an 11th generation Core i7-based system with 16 GB of RAM and a NVIDIA 3070 GPU with 8 GB of memory. The Python programming language was used, together with Keras, and a Google open-source framework called Tensor Flow developed for handling deep learning challenges.

4.1Selecting and tuning hyper-parameters

The training set, described in section 2 and presented in *Table 1,* was utilized to train the model proposed in this study. The "categorical cross entropy" loss function is used because the dataset used in this study has multiple classes (five). When the number of epochs during the model training increased, the characteristics of training data were deeply learned. However, after completion of certain epochs, if the model is trained further, it leads to overfitting instead of generalization. Therefore, it is crucial to select the appropriate number of epochs. Different numbers of trials were conducted on AAMI dataset, and it was found that the number of epochs should not exceed 200. Therefore, the epochs in the model training were set at 200, with "Early Stopping" enabling the monitoring of up to 50 epochs. That is, training terminates after a particular epoch within 200 epochs, such that performance of training model not at all enhances on the validation data for the previous 50 epochs.

Small batches converge faster than large batches, but large batches can reach the optimal minima that small batches cannot [35]. In addition, the selected batch size should be a power of two to get maximum output of GPU processing. The training and validation accuracies, along with the training time for different batch sizes of 64, 128, 256, and 512 were compared. The performance details of different batch sizes were presented in *Table 2.* From the observation of these details, it is determined that 256 was the optimal batch size. The role of the optimizer is to adjust the weights in the network to reduce the losses. The "Adam" is used as an optimizer because it converges faster and is computationally efficient. The author [36] evaluated the effect of different learning rates on a larger dataset and determined that when the learning rate was less than 0.05, it delivered better accuracy than the higher one. The training and validation accuracies and training times for various learning rates (0.05, 0.01, 0.005, 0.001, and 0.0001) were compared keeping batch size as 256. The performance details of different learning rates along with training time were presented in *Table 3.*

Table 2 Comparison of the effect of different batch sizes on training

Batch Size	Training Time	Average Accuracy
64	36.04 min	99.32%
128	34.53 min	99.53%
256	33.12 min	99.58%
512	32.03 min	99.26%
1024	30.42 min	99.19%

Table 3 Comparison of the effect of different learning rates on training

Learning rate	Batch size	Training time	Average accuracy
0.05	256	32.38 min	99.12%
0.01	256	33.22 min	99.38%
0.005	256	35.56 min	99.41%
0.001	256	37.14 min	99.65%
0.0001	256	38.44 min	99.43%

It was found that the training process converged faster at a learning rate of 0.01, but yielded slightly higher accuracy at 0.001. As a result, initially 0.01 was set as the learning rate but was updated to 0.001 after 100 epochs to fine-tune the model.

Using the novel weighted approach proposed in this study, the classes are balanced by applying different weights to the calculated loss for different class samples, according to Equation (3). This brings balance to the dataset and makes the CNN model classify beats more accurately. Moreover, in addition to accuracy as an evaluation metric, precision and recall were also considered to calculate training and validation losses. In *Figure 5*, the green and red curves indicate the validation and training data loss, respectively. The orange and blue curves indicate the accuracies of the validation and training data, respectively. From *Figure 5 (a)*, it can be observed that the validation data loss did not further decrease after approximately 100 epochs. Thus, according to the early stopping with monitoring up to 50 criteria, the training process progressed for a further 50 epochs and thereafter stopped. The accuracy of the validation and training data is shown in *Figure 5 (b).*

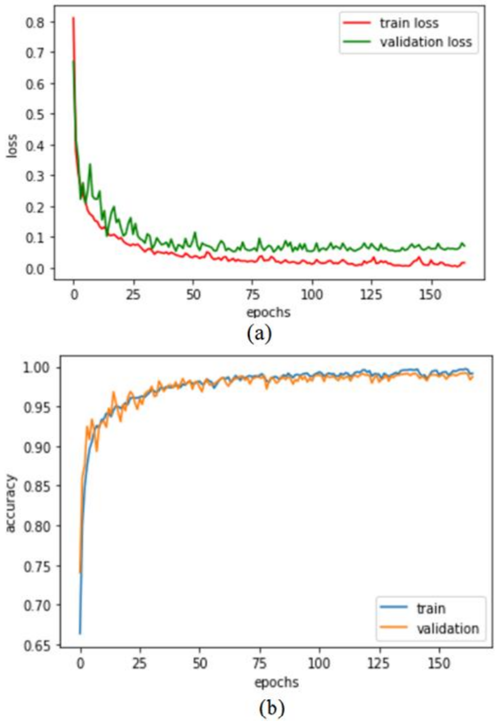

Figure 5 Plotting of 1D-CNN Training and Validation (a) loss curves (b) accuracy curves

When the training process stops, the model in the last epoch may not be optimal. Therefore, the best-trained model is chosen from previous epochs using an application programming interface (API) from Keras called "*ModelCheckpoint*".

4.2Performance assessment metrics

The performance of the CNN model was evaluated by computing the following four basic metrics: precision, sensitivity, specificity, and accuracy. These metrics are computed with given below Equation 4 to Equation 7 by considering true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) obtained from the confusion matrix.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (6)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

As the used dataset is extremely unbalanced, the samples from the majority class may influence the overall performance when the results are computed on the whole dataset, and analysis may lead to the wrong perception. To avoid it, class-wise metrics calculated in addition to the overall or average metrics.

4.3Results

The testing set, described in section 2 and presented in *Table 1*, was utilized to evaluate this trained model. The generated confusion matrix form this trained model using the test dataset is shown in *Figure 6.*

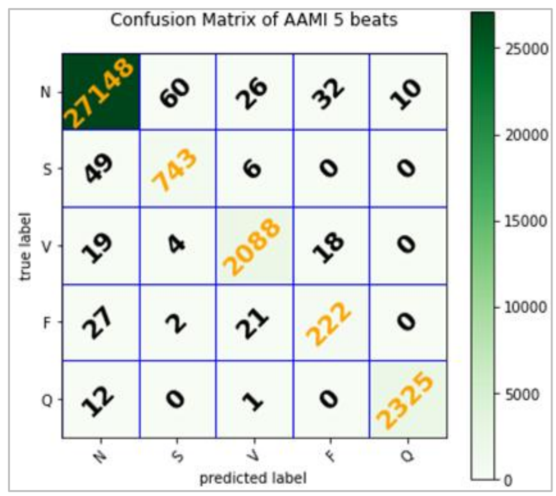

Figure 6 AAMI class arrhythmia confusion matrix

In the existing literature, ECG beat classification is performed using either a subject-oriented scheme or a class-oriented scheme [31]. The proposed model and all compared methods followed the same class-oriented beat classification and used the AAMI standard ECG dataset. The comparison of the proposed 1D-CNN model performance with other state-of-the-art methods are presented in *Table 4*. The highest metric-wise result for each AAMI beat type is also highlighted in *Table 4.*

Table 4 Comparison of 5-types of arrhythmia beat results with state-of-the-art models

Author	Method	Evaluation metric	AAMI beat categories					Average metric
			N	S	V	F	Q	
Acharya et al. [5] (2017)	9 layer 1D-CNN on Denoised ECG	Sensitivity	91.64	89.04	94.07	95.21	97.39	93.47
		Specificity	96.01	98.77	98.74	98.67	99.61	98.36

Author	Method	Evaluation metric	AAMI beat categories					Average metric
			N	S	V	F	Q	
		Precision	85.17	94.76	95.08	**94.69**	98.40	93.62
		Accuracy	95.14	96.82	97.84	97.97	99.16	97.39
Sarvan and Özkurt [15] (2019)	9 layer 1D-CNN on raw ECG	Sensitivity	98.40	82.22	95.38	85.04	99.30	92.07
		Specificity	96.63	99.09	99.60	99.77	99.90	99.00
		Precision	99.33	64.92	94.27	68.38	98.78	85.14
		Accuracy	98.10	98.75	99.32	99.68	99.86	99.14
Pandey and Janghel [2] (2019)	12 layer 1D-CNN on balanced data using SMOTE	Sensitivity	98.61	93.24	96.49	89.61	99.58	95.51
		Specificity	97.89	99.33	99.81	99.46	99.96	99.29
		Precision	99.65	78.28	97.44	55.95	99.00	86.06
		Accuracy	98.51	99.18	99.57	99.39	99.94	99.32
Xiaolin et al. [18] (2020)	10 layer 1D-CNN on balanced data using SMOTE	Sensitivity	98.37	93.60	98.03	88.00	99.60	95.52
		Specificity	98.36	99.06	99.75	99.74	99.93	99.37
		Precision	**99.65**	72.48	96.41	71.90	99.20	87.93
		Accuracy	98.37	98.92	99.63	99.65	99.91	99.30
Yang et al. [32] (2020)	11 layer 1D-CNN on unbalanced Data	Sensitivity	**99.73**	41.46	87.15	59.10	95.49	76.59
		Specificity	82.36	**99.96**	99.79	**99.94**	99.89	96.39
		Precision	96.83	**96.08**	96.72	88.10	98.00	**95.15**
		Accuracy	97.02	98.44	98.94	99.63	99.64	98.73
Mousavi et al. [21] (2021)	ELP + RNN-Attention	Sensitivity	99.24	70.04	89.28	65.71	91.12	83.07
		Specificity	88.30	99.75	**99.94**	99.92	99.84	97.55
		Precision	97.60	87.89	91.88	86.39	97.89	92.33
		Accuracy	97.35	98.99	98.77	99.67	99.20	98.79
Ma et al. [22] (2022)	ResNet + BiLSTM with Attention layer	Sensitivity	99.31	**100**	**98.25**	**99.42**	**99.67**	**99.33**
		Specificity	**99.97**	99.09	98.16	98.87	99.42	99.10
		Precision	NA	NA	NA	NA	NA	NA
		Accuracy	**99.96**	99.55	99.06	99.21	99.32	99.42
Proposed	9 layer 1D-CNN on raw ECG	Sensitivity	99.53	93.11	98.07	81.62	99.44	94.35
		Specificity	98.07	99.79	99.82	99.85	**99.97**	**99.50**
		Precision	99.61	91.84	**97.48**	81.62	**99.57**	94.02
		Accuracy	99.28	**99.63**	**99.71**	**99.70**	**99.93**	**99.65**

* NA: Not Available

5.Discussions

The length of the chosen ECG segment plays an important role in classification performance. The authors [5, 18] used 260 samples or a 0.72-second duration ECG segment, and Pandey and Janghel [2] used 360 samples or a 1-second duration ECG segment as input in their models. One of the reasons why their results are less than ours is their chosen ECG segment length. Doctors usually investigate ECG segment of short duration for diagnosis, but not just a single ECG beat [37]. The ECG segment length is directly proportional to the processing time. Therefore, the shorter the ECG segment length, the lower the processing time. However, when the ECG segment length is less than 2-second, it may not include important details, such as the RR interval. As a result, providing ECG segments of two seconds or longer to the 1D-CNN structure is quite realistic for better arrhythmia beat classification performance. Hence, in this study, we used two-second ECG segments as inputs. We also compared the classification results of ECG segments with 1-second and 2-seconds on the proposed model, and found that the latter produced better results.

The kernel size determines level of feature detail captured from the input. That is, smaller kernels capture lower-level details as features while ignoring higher level details, and larger kernels do the inverse. In this study, a smaller kernel was employed in the first convolution layer and its size was increased in the subsequent convolution layers. As a result, lower level details are initially acquired as features, and next layers captures higher features which are composed of several small features from the previous layers. As both lower and higher-level features are considered, the classification performance is improved significantly. Although [5, 15] followed the same approach used in this work in terms of kernel size, the overall classification performance of the proposed model improved compared to [5,15] due to the use of the novel weighted mechanism.

Observation from *Table 4* reveals that, despite having higher accuracy in normal (N) beats, the model from Ma et al. [22] failed to attain better accuracy in all abnormal arrhythmia beats when compared to the proposed model. But, in case of Q type abnormal beats from Pandey and Janghel [2] has an accuracy of just 0.01 higher than our model. However, the class-wise accuracy of the remaining four beats, as well as the overall accuracy of Pandey and Janghel [2], is lower than model proposed in this study. This demonstrates that the novel weighted mechanism enhances or achieve comparable cardiac classification accuracy than existing methods especially in case of abnormal beats.

Yang et al. [32] classified AAMI beats without using any class-balancing mechanisms. They achieved very low sensitivities of 59.10% and 41.46% for the two minority classes F and S, respectively. But, Ma et al. [22] obtained high sensitivities of 99.42% and 100% for the two minority classes F and S, respectively. They used GAN to artificially generate almost identical beats from existing minority beats and added them to the data set to address the issue of class imbalance. The training and testing sets derived from this enlarged data set may contain nearly identical or duplicate beats, implying that the trained model is already aware of minority class test beats, and thus the results may be biased. This suggests that it might not be an effective strategy as compared to our approach. The authors [2, 5, 18] used the SMOTE technique to produce synthetic beats and balance the classes in the AAMI dataset before performing the classification. However, augmenting synthetic data is not the right approach because it not only increases the training time but also produces data that may not be realistic. Therefore, we performed ECG beat classification using a novel weighted approach to balance the classes in the AAMI dataset without creating synthetic beats. The results in *Table 4* proved that our approach achieved better or very close results than existing methods. Even though overall results of this work are comparable, it is not possible to obtain better sensitivity as compared to the Ma et al. [22].

Finally, when the four evaluation metric values for all five categories of beats are considered, it can be stated that the 1D-CNN model proposed in this study matches or outperforms previous approaches. Furthermore, the model can be used directly on raw ECG signals without using a noise filter. The performance of the proposed model in real-time depends on accurate detection of the R-peak position. As the Pan-Tompkins algorithm [38] can detect the R-peak position with an accuracy of 99.3%, with the help of it, the R peak can be detected to prepare 2-second ECG segments which act as input to the proposed model. Therefore, the approach proposed in this study can be implemented to diagnose cardiac arrhythmia in real-time.

The strengths of our approach in obtaining better results compared to those of state-of-the-art methods are summarized below.

- A novel weighted approach was used to balance the classes in the dataset instead of using synthetic data or oversampling technique.
- The proposed model is fully automatic; hence it does not need separate feature extraction and classification methods.
- Early stopping technique and an API called "*ModelCheckpoint*" from Keras are used to choose the best trained model to evaluate the test dataset.

5.1Limitations

Even though the proposed 1D-CNN model results were better than those of existing methods, a few of the limitations of our approach are as follows:

- The classification work was carried out only according to class-oriented beats. However, subject-oriented beat classification was not performed.

- An ECG segment of duration 2-seconds was chosen to perform beat classification. The effect of segments with duration of 3 or 5-seconds or more was not investigated.
- The raw ECG signals were used in this study without filtering any noise. The effect of noise-filtered signals on the proposed weighted CNN model has not yet been investigated.

In future work, the aforementioned tasks will be carried out, and their results will be compared with the results of this work. Moreover, we can investigate the combination of CNN with other traditional machine learning algorithms for classifying the arrhythmia beats.

A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion

In this work, a 9-layer 1D-CNN model is developed with a novel weighted approach to improve the AAMI recommended arrhythmia beat classification performance compared to existing 1D-CNN models. A set of 2-second ECG segments created from raw ECG signals collected from the MITBIH arrhythmia database was used as the input. The proposed model achieved an overall sensitivity of 94.35%, precision of 94.02%, specificity of 99.5%, and accuracy of 99.65%. Except for the F-category beat sensitivity and precision, the proposed approach enhanced or reached very close class-wise and overall sensitivity, precision, specificity, and accuracy when compared to existing 1D-CNN models. In future work, a hybrid model that combines CNN with other classic machine learning algorithms, such as decision tree, random forest, and SVM needs to be investigated to further enhance the classification performance of arrhythmia beats according to the AAMI recommendations.

Acknowledgment

The authors collected the data set from the link: https://physionet.org/content/mitdb/1.0.0/

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contribution statement

Ravindar Mogili: The paper concept and methodology, software and resources used, investigation and formal analysis, preparing the original draft has been done. **G. Narsimha:** The project supervision, investigation, analysis, proof reading, validation and administration has been done.

References

[1] https://www.nhp.gov.in/world-heart-day-2020_pg. Accessed 15 December 2021.
[2] Pandey SK, Janghel RR. Automatic detection of arrhythmia from imbalanced ECG database using CNN model with SMOTE. Australasian Physical & Engineering Sciences in Medicine. 2019; 42(4):1129-39.
[3] Huang J, Chen B, Yao B, He W. ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network. IEEE Access. 2019; 7:92871-80.
[4] Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering. 2015; 63(3):664-75.
[5] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine. 2017; 89:389-96.
[6] Sharma M, Tan RS, Acharya UR. Automated heartbeat classification and detection of arrhythmia using optimal orthogonal wavelet filters. Informatics in Medicine Unlocked. 2019.
[7] Li H, Boulanger P. Structural anomalies detection from electrocardiogram (ECG) with spectrogram and handcrafted features. Sensors. 2022; 22(7):1-22.
[8] Kumar G, Pawar U, O'reilly R. Arrhythmia detection in ECG signals using a multilayer perceptron network. AICS 2019(pp.353-64).
[9] Golrizkhatami Z, Taheri S, Acan A. Multi-scale features for heartbeat classification using directed acyclic graph CNN. Applied Artificial Intelligence. 2018; 32(7-8):613-28.
[10] Zhou S, Tan B. Electrocardiogram soft computing using hybrid deep learning CNN-ELM. Applied Soft Computing. 2020.
[11] Yang H, Liu J, Zhang L, Li Y, Zhang H. Proegan-ms: a progressive growing generative adversarial networks for electrocardiogram generation. IEEE Access. 2021; 9:52089-100.
[12] Mian QS, Fawad HS. Arrhythmia diagnosis by using level-crossing ECG sampling and sub-bands features extraction for mobile healthcare. Sensors. 2020; 20(8):1-19.
[13] Sahoo S, Mohanty M, Sabut S. Automated ECG beat classification using DWT and Hilbert transform-based PCA-SVM classifier. International Journal of Biomedical Engineering and Technology. 2020; 32(3):287-303.
[14] Sultan QS, Ghorbani AR. ECG arrhythmia classification using time frequency distribution techniques. Biomedical Engineering Letters. 2017; 7(4):325-32.
[15] Sarvan Ç, Özkurt N. ECG beat arrhythmia classification by using 1-D CNN in case of class imbalance. In medical technologies congress 2019 (pp. 1-4). IEEE.
[16] Yao G, Mao X, Li N, Xu H, Xu X, Jiao Y, et al. Interpretation of electrocardiogram heartbeat by CNN

and GRU. Computational and Mathematical Methods in Medicine. 2021.
[17] Khan MM, Siddique MA, Sakib S, Aziz A, Tanzeem AK, Hossain Z. Electrocardiogram heartbeat classification using convolutional neural networks for the detection of cardiac Arrhythmia. In fourth international conference on I-SMAC 2020 (pp. 915-20). IEEE.
[18] Xiaolin L, Cardiff B, John D. A 1d convolutional neural network for heartbeat classification from single lead ECG. In IEEE international conference on electronics, circuits and systems 2020 (pp. 1-2). IEEE.
[19] Al RMM, Bazi Y, Al ZM, Othman E, Benjdira B. Convolutional neural networks for electrocardiogram classification. Journal of Medical and Biological Engineering. 2018; 38(6):1014-25.
[20] Yu X. An ECG arrhythmia image classification system based on convolutional neural network. In journal of physics: conference series 2020 (pp. 1-8). IOP Publishing.
[21] Mousavi S, Afghah F, Khadem F, Acharya UR. ECG language processing (ELP): a new technique to analyze ECG signals. Computer Methods and Programs in Biomedicine. 2021.
[22] Ma S, Cui J, Xiao W, Liu L. Deep learning-based data augmentation and model fusion for automatic arrhythmia identification and classification algorithms. Computational Intelligence and Neuroscience. 2022.
[23] Lu W, Jiang J, Ma L, Chen H, Wu H, Gong M, et al. An arrhythmia classification algorithm using C-LSTM in physiological parameters monitoring system under internet of health things environment. Journal of Ambient Intelligence and Humanized Computing. 2021:1-11.
[24] Shoughi A, Dowlatshahi MB. A practical system based on CNN-BLSTM network for accurate classification of ECG heartbeats of MIT-BIH imbalanced dataset. In international computer conference, computer society of Iran 2021 (pp. 1-6). IEEE.
[25] Gai ND. ECG beat classification using machine learning and pre-trained convolutional neural networks. arXiv preprint arXiv:2207.06408. 2022.
[26] Liu Z, Zhang X. ECG-based heart arrhythmia diagnosis through attentional convolutional neural networks. In international conference on internet of things and intelligence systems 2021 (pp. 156-62). IEEE.
[27] Zubair M, Yoon C. Cost-sensitive learning for anomaly detection in imbalanced ECG data using convolutional neural networks. Sensors. 2022; 22(11):1-15.
[28] Jiang J, Zhang H, Pi D, Dai C. A novel multi-module neural network system for imbalanced heartbeats classification. Expert Systems with Applications: X. 2019; 1:1-15.
[29] Romdhane TF, Pr MA. Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss. Computers in Biology and Medicine. 2020.
[30] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Engineering in Medicine and Biology Magazine. 2001; 20(3):45-50.
[31] Afkhami RG, Azarnia G, Tinati MA. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. Pattern Recognition Letters. 2016; 70:45-51.
[32] Yang F, Zhang X, Zhu Y. PDNet: a convolutional neural network has potential to be deployed on small intelligent devices for arrhythmia diagnosis. Computer Modeling in Engineering & Sciences. 2020; 125(1):365-82.
[33] Wang YX, Ramanan D, Hebert M. Learning to model the tail. Advances in Neural Information Processing Systems 2017.
[34] Mahajan D, Girshick R, Ramanathan V, He K, Paluri M, Li Y, et al. Exploring the limits of weakly supervised pretraining. In proceedings of the European conference on computer vision 2018 (pp. 181-96).
[35] Kandel I, Castelli M. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Express. 2020; 6(4):312-5.
[36] Wilson DR, Martinez TR. The need for small learning rates on large problems. In IJCNN'01. international joint conference on neural networks. proceedings (Cat. No. 01CH37222) 2001 (pp. 115-9). IEEE.
[37] 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. Information Sciences. 2017; 405:81-90.
[38] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering. 1985:230-6.

Ravindar Mogili is a Ph.D. candidate in the Department of Computer Science and Engineering at the Jawaharlal Nehru Technological University, Hyderabad, India. He has 18 years of experience in industry and teaching. He is member of many professional bodies. His research interests include Data Mining, Image Processing, Machine Learning and Deep Learning with applications to the health sector.
Email: ravindarmogili@live.com

Dr. G. Narsimha received B.E (ECE), M. Tech (CSE) and Ph.D in CSE from Osmania University Hyderabad, India. He is currently working as Principal & Professor of Department of CSE at JNTUH College of Engineering Sultanpur, Hyderabad, India. His research interests include Data warehousing and Data Mining, Big Data, Mobile Computing, Cloud Computing, Computer Networks, and Network Security. To date, he has published 153 quality research articles in international conferences and renowned peer reviewed international journals. Clarivate Web of Science indexed 60 of these research publications, while Elsevier SCOPUS indexed 65. He is holding 4 copyrights granted by ACM, USA for his research publications. He has successfully supervised 26 PhD students so far. As of now, 15 Ph.D scholars are pursuing their research work under his ample guidance. He was honoured with outstanding reviewer award from Elsevier Computer Science and Electrical.
Email: narsimha06@jntuh.ac.in

Appendix I

S. No.	Abbreviation	Description
1	1D-CNN	One-Dimensional Convolutional Neural Network
2	2D-CNN	Two-Dimensional Convolutional Neural Network
3	AAMI	Association for Advancement of Medical Instrumentation
4	API	Application Programming Interface
5	BiLSTM	Bidirectional Long Short-Term Memory
6	CNN	Convolutional Neural Network
7	CVD	Cardiovascular Disease
8	DWT	Discrete Wavelet Transform
9	ECG	Electrocardiogram
10	ELP	ECG Language Processing
11	F	Fusion
12	FN	False Negative
13	FP	False Positive
14	GAN	Generative Adversarial Networks
15	GB	Gigabyte
16	GPU	Graphics Processing Unit
17	GRU	Gated Recurrent Unit
18	KNN	k-nearest neighbours
19	LSTM	Long Short-Term Memory
20	MITBIH	Massachusetts Institute of Technology-Beth Israel Hospital
21	N	Non-ectopic
22	PCA	Principal Component Analysis
23	Q	Unknown
24	RAM	Random Access Memory
25	ResNet	Residual Network
26	RNN	Recurrent Neural Network
27	S	Supraventricular Ectopic
28	SMOTE	Synthetic Minority Over-Sampling Technique
29	SVM	Support Vector Machine
30	TF	Time–Frequency
31	TN	True Negative
32	TP	True Positive
33	V	Ventricular Ectopic
34	VGGNet	Visual Geometry Group Network
35	WFDB	Waveform Database