| Team Id | PNT2022TMID48615 |
| --- | --- |
| Title | Classification of Arrhythmia by using deep learning with 2-D ECG spectral image representation |

**PROPOSED SCHEME**

The method consists of five steps, i.e., signal pre-processing, generation of 2-D images (spectrograms), augmentation of data, extraction of features from the data (using the CNN model), and its classification based on the extracted features. The details of these steps are presented in the following subsections.

**1. Pre-Processing**

The three primary forms of noise in the ECG signal are power line interference, baseline drift, and electromyographic noise [46]. The noise from the original ECG signal must be removed to ensure that a denoised ECG signal is obtained for further processing. We combined wave let based thresholding and the reconstruction algorithm of wavelet decomposition to remove noise from the original ECG signal [47]. The wavelet thresholding was performed using,

Where, w,x,y represents the wavelet coefficients, ω,x,y represents the estimated wavelet coefficients after threshold, x represents the scale and y represents the shift, λ represents the threshold, and α is a parameter whose value can be set arbitrarily. The wavelet thresholding reduced the electromyographic noise and power line noise interference. Moreover, the reconstruction algorithm of wavelet decomposition was used to remove the baseline drift noise from the noisy ECG signal

**2. Generation of 2-D Images**

While 1-D CNN can be used for time series signals, the flexibility of such models is limited due to the use of 1-D kernels. On the other hand, 3-D CNNs require a large amount of training data and computational resources. In comparison, 2-D CNNs are more versatile since they use 2-D kernels and, hence, could provide representative features for time series data. Hence, for certain applications where sufficient data is available and for

1-D signals that can be represented in a 2-D format, using a 2-D CNN could be beneficial. Herein, for generating 2-D images to be used with the 2-D CNN model, the ECG signal was transformed into a 2-D representation. The 2-D time-frequency spectrograms were generated using the short-time Fourier transform. The ECG signal represents non-stationary data where the instantaneous frequency varies with time. Hence, such changes cannot be fully represented by just using information in the frequency domain. The STFT is a method derived from the discrete Fourier transform to analyze instantaneous frequency as well as the instantaneous amplitude of a localized wave with time-varying characteristics. In the analysis of a non-stationary signal, it is assumed that the signal is approximately stationary within the span of a temporal window of finite support. The 1-D ECG signals were converted into 2-D spectrogram images by applying STFT as follows,

$$XSTFT[m, n] = \sum_{k=0}^{L-1} x[k]g[k - m]e^{-j2\pi nk/L}$$

where L is the window length, and x[k] is the input ECG signal. The log values of XSTFT[m, n] are represented as spectrogram (256 × 256) images

**3. Data Augmentation**

Another significant advantage of using 2-D CNN models is the flexibility it provides in terms of data augmentation. For 1-D ECG signals, data augmentation could change the meaning of the data and hence is not beneficial. However, with 2-D spectrograms, the CNN model can learn the data variations, and augmentation helps in increasing the amount of data available for training. The ECG data is highly imbalanced, where most of the instances represent the normal class. In this scenario, data augmentation can help when those classes that are underrepresented are augmented. For arrhythmia classification using ECG signals, augmenting training data manually could degrade the performance. Moreover, classification algorithms such as SVM, fast Fourier neural network, and tree-based algorithms, assume that the classification of a single image based representation of an ECG signal is always the same [48]. The proposed CNN model works on 2-D images of ECG signals as input data, which allows changing the image size with operations such as cropping. Such augmentation methods would add to the training data and hence would allow better training of the CNN model. Another important issue that arises when using small data with CNN based architectures is overfitting. Data augmentation is a way to deal with overfitting and allows better training of a CNN model. For imbalanced data, data augmentation can help in maintaining a balance between different classes. We have used the cropping method for the augmentation of seven classes of ECG beats; namely, premature ventricular contraction beat (PVC), paced beat (PAB), right bundle branch block beat (RBB), left bundle branch block beat (LBB), atrial premature contraction beat (APC), ventricular flutter wave (VFW), and ventricular escape beat (VEB). These are common types of cardiac arrhythmias and are considered in studies we have used for comparison (refer to the Discussion section). While other methods of augmentation are used, such as warping in image processing applications, the aim here is to augment classes that are under-represented. Towards this end, eight different cropping operations (left top, center top, right top, left center, center, right center, left bottom, center bottom, right bottom) were applied. As a result of cropping, we obtain multiple ECG spectrograms of reduced size ($200 \times 200$), which are then resized to $256 \times 256$ images (using linear interpolation) before being fed into the CNN. This resulted in an eight times increase in the training data, which benefited the training process.

## 4. Deep Neural Network

In this study, a CNN-based model is proposed for an automatic classification of arrhythmia using the ECG signal in a supervised manner. The ECG data used in the study have corresponding labels (ground truth) identifying the type of arrhythmia present. These labels were assigned by expert cardiologists and are used for supervised training. For each heartbeat segment, the arrhythmia class label was transferred to the corresponding spectrogram image representation. The first CNN-based algorithm, introduced in 1989 [49], was developed and used for the recognition of handwritten zip codes. Since then, multiple CNN models have been proposed for the classification of images, among which AlexNet [39] has achieved significant performance for a variety of images. The existing neural networks with the feed-forward process for the automatic classification of the 2-D image was not feasible since these methods do not take into account the local spatial information. However, with the development of CNN architectures and using nonlinear filters, spatially adjacent pixels can be correlated to extract local features from the 2-D image. In the 2-D convolution algorithm, the downsampling layer is highly desirable for extracting and filtering the spatial vicinity of the 2-D ECG images. For these reasons, the ECG signal was transformed into a 2-D representation, and a 2-D CNN algorithm was used for classification. Consequently, high accuracy was obtained in the automatic taxonomy of the ECG waves. The details of the proposed CNN model is presented in Section.