A Novel Approach to Detect Cardiac Arrhythmia Based on Continuous Wavelet Transform and Convolutional Neural Network

Shadhon Chandra Mohonta1, and Md. Firoj Ali2

1Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
2Department of Mechatronics Engineering, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

emails: *shadhonmohonta@gmail.com; and @eeemfa07@gmail.com

1. INTRODUCTION

The amount of cardiovascular diseases upsurges incredibly and 17.3 million people decease every year conveyed by World Health Organization (WHO) (Mendis, Pska, Norving, & Organization, 2011). The arrhythmias are most serious among them, and may cause sudden cardiac arrest or stroke (Huikuri, Castellanos, & Myerburg, 2001). Indeed, proper diagnosis of arrhythmias can considerably preclude such sudden cardiac death. ECG comprises complete details about the normality or abnormality of human heart, and various classes of arrhythmias can be diagnosed from it. In (De Bie, Martignani, Massaro, & Diemberger, 2020), short-term samples of 10 s obtained from pediatric and adult patient’s ECG signals for atrial fibrillation or other abnormal rhythms detection. Furthermore, various feature extraction approaches such as time-domain and frequency-domain, statistical, and time-frequency analysis have been implemented for arrhythmia detection. Here, the features like heart rate, RR interval, R amplitude, PR interval, QRS duration, P-wave, and T-wave duration (Chen, Wang, & Wang, 2018; De Chazal & Reilly, 2006; Mitra, Mitra, & Chaudhuri, 2006) have been extracted from the time-domain analysis. Moreover, statistical features have been extracted in terms of variance, kurtosis, and skewness (Queiroz, Azoubel, & Barros, 2019; Queiroz, Junior, Lucena, & Barros, 2018). Besides, many literatures implemented wavelet transform to classify ECG arrhythmias in time-frequency analysis (Banerjee & Mitra, 2013; Khorrami & Moavenian, 2010).

In addition, to assist clinicians for detecting and classifying arrhythmias automatically a CNN based approached has been emerged in recent years. The CWT and CNN based approach has been proposed jointly for automatic arrhythmia classification with an overall accuracy of 98.74% (Wang et al., 2021). Here, the 2D-scalograms obtained from CWT while features extraction conducted using CNN from those scalograms. Moreover, different cardiac arrhythmias have been detected after the preprocessing of ECG signals using a 34-layer CNN (Brisk et al., 2019), 1D-CNN of 31-layers (Li, Zhou, Wan, Li, & Mou, 2020), CNN with long short-term memory (LSTM) (Chen et al., 2018), and 11-layer deep CNN (Acharya et al., 2017). Besides, a combination of signal quality index (SQI) algorithm and densed-CNN has been proposed to distinguish atrial fibrillation from short ECG segments (9–60 s) (Rubin, Parvaneh, Rahman, Conroy, & Babaeizadeh, 2018). Additionally, for automatic classification of cardiac arrhythmias various new models such as wavelet transform with 2D-CNN (Mohonta, Motin,
 &, Kumar, 2022), deep neural network (DNN) models constitutive of residual convolutional modules and bidirectional LSTM (He et al., 2019), and Deep Multi-Scale Convolutional neural network Ensemble (DMSCE) (Prabhakaran & Dandapat, 2021) have been proposed.

In this work, a new CNN model has been proposed to recognize the subjects with arrhythmias, and discriminate them from the healthy one. To expedite this, ECG segments of 60 s have been transformed into RBG scalograms based on CWT method for pattern recognition. Then, these scalograms of ECG segments have been fed into our proposed CNN model for automatic detection of cardiac arrhythmias. The remainder of the paper is depicted as follows: section 2 describes the materials and methods part. In section 3, a comprehensive results and discussion has been presented based on proposed classifier performance. Finally, the conclusion is in fourth section.

# 2. MATERIALS AND METHODS

Our proposed method for arrhythmia detection has been divided into six stages containing data acquisition, pre-processing, segmentation, R peak detection, time-frequency transformation, and CNN classifier. The overall layout of our proposed method is shown in Figure 1.

![Overall layout for arrhythmia detection](image)

## A. Data Acquisition

The ECG data have been separated into two groups: healthy subjects, and the subjects with arrhythmia. The data is considered from the MIT-BIH database, and the duration of ECG recordings of each subject is 30 min. Here, 13 subjects have been grouped as healthy considered from normal sinus rhythm (NSR) database, whether 52 subjects have been considered from arrhythmia (ARR), atrial fibrillation (AFI), supraventricular arrhythmia (SVA) and malignant ventricular ectopy (MVE) database to form arrhythmic group.

## B. Pre-processing and Segmentation

In pre-processing, the discrete wavelet transform (DWT) has been considered for eliminating the noise, baseline wander and artefacts from ECG signals. The low and high-frequency components of a signal has been decomposed using DWT, and then the signal has been reconstructed. Here, the ECG signal has been decomposed considering ‘sym4’ wavelet where low-frequency and high-frequency information are carried out by the approximation coefficient, and the first and second level detail coefficients respectively. Therefore, the denoised ECG signals have been reconstructed based on inverse wavelet transform only considering the third and fourth level detail coefficients. After that, each ECG recording has been divided into segments with 60 s duration. The segmented raw and filtered ECG signal of 60 s are shown in Figure 2 and Figure 3 respectively. So, the healthy and arrhythmic group have total 390 and 1560 segments respectively.

![Original ECG signal](image)

**Figure 2: Original ECG signal**

![Filtered ECG signal](image)

**Figure 3: Filtered ECG signal**

## C. R Peak Detection

In this paper, the R peaks of ECG signal have been detected in three steps: (i) squaring the magnitude of the signal, (ii) finding the average value, and (iii) detecting R peaks. Here, the threshold value for the minimum peak height is the average value, and 0.14 s is considered as minimum peak height. The detected R peaks of a ECG signal is shown in Figure 4.

![Detection of R peaks](image)

**Figure 4: Detection of R peaks**

There are a number of segments in healthy, and arrhythmic group which contain abnormal, and normal rhythm respectively. To exclude these undesirable segments, the detected R peaks have been used to compute the average heart rate (AHR) for each segment. After that, the segments which are in normal range (60 bpm ≤ AHR ≤ 100 bpm ), and abnormal range
where, $\psi^*(t), r, a$, and $b$ are the complex conjugate of the basic wavelet, time shift, scale factor, and location parameter respectively.

**E. CNN Model**

In this analysis, a CNN network has been proposed to detect arrhythmia from RGB scalogram. It consists of several common layers: convolutional (Conv), Rectified Linear Unit (ReLU), batch normalization (BN), max pooling (MaxPool), fully connected (FC) and softmax. The basic model of proposed CNN is shown in Figure 5.

![Figure 5: The basic model of proposed CNN](image)

The input of the network is RGB time-frequency scalogram of size $227 \times 227 \times 3$. This model has four convolutional blocks. Every convolutional block contains convolutional layer, batch normalization layer, and a pooling layer. A non-linear activation function, ReLU, is used after the convolutional layer to enhance the approximation ability between each layer of the network. Besides, the function of the batch normalization layer is to batch normalize the activation output of the layer. The parameters of the max pooling layer of every convolutional block are same, and have the Kernel size $3 \times 3$ and the stride $2$. But, the parameters of convolutional layer in each block are different. Moreover, the fully connected layer shows the classification results and the softmax function implies the probability that the input belongs to each class. Finally, 70% of images have been considered to train our model, and rest are used for testing. Table 1 represents the summarization of the proposed CNN model.

**Table 1: Summaries the parameters of proposed CNN model**

<table>
<thead>
<tr>
<th>Layer Type</th>
<th>Kernel Size</th>
<th>Stride</th>
<th>Kernel</th>
<th>Padding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv_1</td>
<td>11×11</td>
<td>4</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>MaxPool_1</td>
<td>3×3</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conv_2</td>
<td>5×5</td>
<td>1</td>
<td>256</td>
<td>2</td>
</tr>
<tr>
<td>MaxPool_2</td>
<td>3×3</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conv_3</td>
<td>3×3</td>
<td>1</td>
<td>394</td>
<td>1</td>
</tr>
<tr>
<td>MaxPool_3</td>
<td>3×3</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Conv_4</td>
<td>3×3</td>
<td>1</td>
<td>256</td>
<td>1</td>
</tr>
<tr>
<td>MaxPool_4</td>
<td>3×3</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FC</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

**3. RESULTS AND DISCUSSION**

In this work, our proposed CNN model is a simple one which consists of only four convolutional layers, and one fully connected layer. Here, RGB scalograms of size $227 \times 227 \times 3$ of all ECG segments have been obtained from CWT method for arrhythmia detection. After that, the designed CNN model took these scalograms as an input, and differentiate arrhythmic subjects from the healthy ones automatically. Here, total 348 scalograms are used to train the model, while 150 images are considered for testing purpose. In this approach, 6 epochs and 204 iterations have been considered, and it took around 13 min to compare the classification process. The confusion matrix of the proposed CNN model is shown in Figure 6.

![Figure 6: Confusion matrix for proposed CNN model](image)

Here, three statistical parameters have been computed to evaluate the performance of our classifier which are as follows (Queiroz et al., 2019):

$$Se(\%) = \frac{TP}{TP+FN} \times 100$$  \hspace{1cm} (2)

$$Sp(\%) = \frac{TN}{TN+FP} \times 100$$  \hspace{1cm} (3)
\[
\text{Acc(\%)} = \frac{TP+TN}{TP+FP+TN+FN} \times 100
\]  
(4)

Where, \( Se \) is the sensitivity which represents how the technique is effective for detecting arrhythmia, \( Sp \) is the specificity which denotes how the technique is effective for detecting healthy subjects, and \( Acc \) is the accuracy which measures the effectiveness of the technique regarding diagnosis. Also, \( TP \) is the true positive, \( TN \) is the true negative, \( FP \) is the false positive, and \( FN \) is the false negative.

The accuracy and loss curves of our proposed model are shown in Figure 7 and Figure 8 respectively. It is obvious that the network has reached its stable position after approximately 30 iterations.

The comparison between the performance of our work and other existing methods has been summarized in Table 3. In (Wang et al., 2021), the arrhythmia detection had been performed with a combination of CWT and CNN-based approaches. A deep learning model named ResNet-31 had been used to detect cardiac arrhythmia from single lead ECG signals which showed 99.06\% accuracy (Li et al., 2020). Also, a combination of CNN and long short-term memory (LSTM) based model had been proposed for automatic arrhythmia classification, and achieved 98.10\% accuracy (Chen et al., 2018). Moreover, accuracy of 94.90\% had been achieved for arrhythmia detection using CNN-based approach from ECG segments of 5 s. (Acharya et al., 2017). Arrhythmia has been classified from ECG signal using STFT-based spectrogram and proposed 2D-CNN model, and achieved an overall accuracy of 99.00\%.

It is apparent from Figure 7 and Figure 8 that the accuracy, and loss are 99.39\%, and 0.015 respectively. Besides, the accuracy of the proposed CNN model is better than using a pretrained AlexNet model (Mohonta & Ali, 2021). Besides, the statistical indices such as sensitivity (\( Se \)), specificity (\( Sp \)), and accuracy (\( Acc \)) of this classification are portrayed in Table 2.

In sum, the deep learning model extracts features automatically from CWT scalogram which conveys the time-frequency information of arrhythmia. Actually, the wavelet-based technique decomposes complex information and patterns of an image into elementary forms. Hence, the propounded CNN model has outperformed the other related works, and effectively diagnosed and classified arrhythmic subjects from normal one with an overall accuracy of 99.39\%. However, small number of ECG segments has been considered in this analysis. Further analysis could be done with the same approach using large number of segments.

4. CONCLUSIONS

The information of electrical activity of the heart can be optimized through ECG signal which also helps to diagnose cardiac disorder non-invasively. Here, the
detection of arrhythmia from ECG segments has been conducted based on wavelet transform and CNN classifier. In this diagnosis, a novel CNN model has been incorporated for image recognition that investigates the CWT scalograms, and delivers better performance with an overall accuracy of 99.39%. So, the results demonstrate that the proposed approach automatically extracted features from the ECG signal to distinguish arrhythmic subjects from healthy one successfully. Therefore, this approach can be implemented clinically to guide academicians and medical professionals for efficient arrhythmia diagnosis.

ACKNOWLEDGEMENTS

The authors would like to thank PhysioNet for making the dataset available.

REFERENCES


