Automated Detection of Arrhythmias Using Different Intervals of Tachycardia ECG Segments with Convolutional Neural Network

Article *in* Information Sciences · April 2017

DOI: 10.1016/j.ins.2017.04.012

CITATIONS
558

READS
5,681

6 authors, including:

U Rajendra Acharya
University of Southern Queensland
948 PUBLICATIONS **57,411** CITATIONS

SEE PROFILE

Shu Lih Oh
Ngee Ann Polytechnic
74 PUBLICATIONS **8,869** CITATIONS

SEE PROFILE

Yuki Hagiwara
Fraunhofer IKS
60 PUBLICATIONS **7,609** CITATIONS

SEE PROFILE

Automated Detection of Arrhythmias Using Different Intervals of Tachycardia ECG Segments with Convolutional Neural Network

U. Rajendra Acharya [a,b,c], Hamido Fujita [d,*], Oh Shu Lih [a], Yuki Hagiwara [a], Jen Hong Tan [a], Muhammad Adam [a]

[a] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

[b] Department of Biomedical Engineering, School of Science and Technology, SIM University, Singapore

[c] Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia

[d] Iwate Prefectural University (IPU), Faculty of Software and Information Science, Iwate 020-0693 Japan

[*]Postal Address: Iwate Prefectural University (IPU), Faculty of Software and Information Science, Iwate 020-0693 Japan

Telephone: +81-19-694-2578; Email Address: HFujita-799@acm.org

ABSTRACT

Our cardiovascular system weakens and is more prone to arrhythmia as we age. An arrhythmia is an abnormal heartbeat rhythm which can be life-threatening. Atrial fibrillation (A-Fib), atrial flutter (AFL), and ventricular fibrillation (V-Fib) are the recurring life-threatening arrhythmias that affect the elderly population. An electrocardiogram (ECG) is the principal diagnostic tool employed to record and interpret ECG signals. These signals contain information about the different types of arrhythmias. However, due to the complexity and non-linearity of ECG signals, it is difficult to manually analyze these signals. Moreover, the interpretation of ECG signals is subjective and might vary between the experts. Hence, a computer-aided diagnosis (CAD) system is proposed. A CAD system will ensure that the assessment of ECG signals is objective and accurate. In this work, we present a convolutional neural network (CNN) technique to automatically detect the different ECG segments. Our algorithm consists of an eleven-layer deep CNN with the output layer of four neurons, each representing the normal (NSR), A-Fib, AFL, and V-Fib ECG class. In this work, we have used ECG signals of two seconds and five seconds' durations without QRS detection. We achieved an accuracy, sensitivity, and specificity of 92.50%, 98.09%, and 93.13% respectively for two seconds of ECG segments. We obtained an accuracy of

94.90%, sensitivity of 99.13%, and specificity of 81.44% for five seconds of ECG duration. This proposed algorithm can serve as an adjunct tool to assist clinicians in confirming their diagnosis.

Keywords – Arrhythmia, atrial fibrillation, atrial flutter, convolution neural network, deep learning, electrocardiogram signals, ventricular fibrillation.

1. Introduction

According to the report by the United Nations in 2015, the world is facing an aging population [31]. It is estimated that the number of people aged 60 years and above will grow by 56.00% from 901 million to 1.4 billion by 2030. Furthermore, the growing population (60 years and older) is expected to be double by 2050, rising to nearly 2.1 billion [31]. The increase in elderly population poses economic [25, 31] and health care issues [20, 31] to the world. Our cardiovascular system grows weaker and becomes more receptive to diseases as we grow older [12]. Moreover, the arteries stiffens and muscle wall of the left ventricle thickens with aging, resulting in decrease in the compliance of blood vessels of the arteries [7]. Consequently, it affects the overall function of the heart which leads to arrhythmia. Hence, arrhythmia is one of the health conditions that the elderly subjects encounter in the society [5, 7]. Arrhythmia is defined as the abnormal rhythm of the heartbeat which can be harmless or critical. The atrial fibrillation (A-Fib), atrial flutter (AFL), and ventricular fibrillation (V-Fib) are the recurrent types of arrhythmias reported in the elderly [7].

The A-Fib is a commonly occurring arrhythmia caused due to various health complications. During A-Fib, the contraction of the atria is asynchronous due to the fast firing of electrical impulses from several parts of cardiac re-entry [30]. Re-entry occurs when an impulse fails to die out after normal activation of the heart and continues to re-excite the heart. In fact, the electrocardiogram (ECG) rhythm of A-Fib is fast and beating at a rate of 150 to 220 beats in a minute. It has an abnormal R-R interval, irregular and fast ventricular contraction, and P wave is absent in the ECG signal [12]. In AFL, the atrial contracts rapidly between 240 and 360 beats per minute and have a replicating saw-tooth waveform, known as flutter wave. AFL occurs when the atria undergoes chaotic electrical signals [30]. V-Fib is usually caused by rapid heartbeat known as ventricular tachycardia (VT). This abnormal heartbeat is due to abnormal electrical impulses in the ventricles. During this, ventricles contract chaotically and haphazardly. It can be seen in

the ECG morphology, which records an unrefined and erratic fluctuation of ECG signal with the absence of QRS complex wave [12]. Typical plots of NSR, A-Fib, AFL, and V-Fib ECG signals are shown in Figure 1 and Figure 2.

Therefore, the morphology of ECG signals contains vital details about the conditions of the heart. Thus, the ECG signal is beneficial in the detection and diagnosis of cardiac health [30]. However, ECG signals are highly nonlinear and any changes in the ECG signals is difficult to observe as its amplitude is in millivolts [2, 30]. Further, the indications of cardiac abnormalities are faithfully indicated in the ECG signals during 24-hour Holter recording. Thus, manual interpretation of the ECG signals can be time-consuming, taxing and subjective due to the long recordings. Moreover, there is a great possibility that important information captured in the ECG morphology may be overlooked. Hence, a computer-aided diagnosis (CAD) system can be employed to reduce subjective variabilities in the diagnosis and reduce the time taken to analyze the ECG signals [22].

Table 8 shows the studies conducted on CAD system to automatically detect arrhythmias and categorize different types of arrhythmia into their respective classes. Wang et al. [36] performed short-time multifractal characterization on three classes of ECG beats (A-Fib, V-Fib, and VT) and recorded an accuracy of 99.40% for A-Fib, 97.20% for V-Fib and 97.80% for VT using fuzzy Kohonen network classifier. Martis et al. [21, 23] have conducted a three-class study to automatically diagnose A-Fib, AFL, and NSR ECG signals. In [21], they have employed higher order spectra methods on 641 NSR, 855 A-Fib, and 887 AFL ECG beats. Then these ECG beats are subjected to independent component analysis (ICA) to select highly significant features. Their method yielded an accuracy, sensitivity, and specificity of 97.65%, 98.16%, and 98.75% with the k-nearest neighbor classifier. In their another study [23], they performed a discrete cosine transform combined with ICA on the ECG beats. Their proposed technique attained an average accuracy of 99.45%.

In addition, Fahim et al. [26] employed a data mining approach with expectation-maximization-based clustering on 50 compressed ECG signals obtained from an open-source database. They used correlation-based feature subset selection technique to reduce the number of features. Then, the selected features are fed into the classifier. They detected A-Fib, premature ventricular contraction, and V-Fib with an average accuracy of 97.00% using the rule-based system. Acharya et al. [29] proposed a CAD system to automatically detect and identify same four ECG classes (NSR, A-Fib, AFL, V-Fib) using the entire database (614,526 ECG beats) obtained from an open-source database [3]. They extracted entropy features from the ECG signals. These extracted

features were subjected to feature reduction and the selected 14 significant features were fed into the decision tree classifier, yielding an accuracy of 96.30%, sensitivity of 99.30%, and specificity of 84.10%. Further, Desai et al. [28] also implemented a CAD system to diagnose the four-class arrhythmia (A-Fib, AFL, NSR, and V-Fib). However, they used a smaller dataset (3,858 ECG beats) obtained from the same open-source database [3] in their work. They applied the recurrence quantification analysis parameters to the ECG beats. Then, the features are arranged according to the F-value index. They achieved an accuracy of 98.37% with the rotation forest classifier.

However, from the literature [21, 23, 26, 28, 29], it can be noted that these CAD systems have a standardized workflow whereby the signals are pre-processed first, then segmented. Then the signals are subjected to features extraction, followed by feature selection to select only significant features for classification. In this study, we did not follow the conventional process of an automated CAD system. This is unlike the previous works recorded in Table 8 as *no features extraction or selection* is implemented in this work. We employed an eleven-layer convolutional neural network (CNN) to automatically classify the four classes of ECG signals (NSR, A-Fib, AFL, and V-Fib). Hence, in this study, there is no need to experiment with different features extraction techniques or determine which classifier performs the best with the extracted features.

CNN has recently been employed in the automated classification of ECG signals. Kiranyaz et al. [27] studied the patient-specific ECG categorization and monitoring system using three-layer CNN with only R-peak wave. They attained an accuracy of 97.60% and 99.00% in the detection of supraventricular ectopic beats and ventricular ectopic beats respectively. Zubair et al. [17] used CNN with 44 recordings of ECG signals obtained from 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 into their respective classes (normal, fusion beat, supraventricular ectopic beat, unknown beat, and ventricular ectopic beat). These works [17, 27] detected QRS wave in their automated classification. Nevertheless, in our study, no detection of QRS wave is implemented.

2. Data Used

In this work, the ECG signals were obtained from a publicly available arrhythmia database. We have obtained V-Fib (Ventricular Fibrillation) ECG signals from Creighton University ventricular

tachyarrhythmia, A-Fib (Atrial Fibrillation) and AFL (Atrial Flutter) ECG signals from MIT-BIH atrial fibrillation, and A-Fib (Atrial Fibrillation), AFL (Atrial Flutter), and NSR (Normal Sinus Rhythm) ECG signals from MIT-BIH arrhythmia database [3]. In this work, we have used lead II ECG signals.

The details of the ECG signals used in this study is shown in Table 1. We have used two different durations of ECG segments (two seconds and five seconds) in this work. The total number of ECG segments used for net A (two seconds) and net B (five seconds) is 21,709 and 8,683 respectively.

Table 1
Overview of the data used in this study.

Database	Data Used
MIT-BIH Atrial Fibrillation (afdb)	A-Fib, AFL
MIT-BIH Arrhythmia (mitdb)	A-Fib, AFL, NSR
Creighton University Ventricular Tachyarrhythmia (cudb)	V-Fib

Table 2
Overview of the ECG segments (two and five seconds) used in this study.

Type	Number of Segments (*Two Seconds*) (Net A)	Number of Segments (*Five Seconds*) (Net B)
NSR	902	361
A-Fib	18,804	7,521
AFL	1,840	736
V-Fib	163	65
Total segments	**21,709**	**8,683**

3. Methodology

3.1 Pre-processing

The ECG signals from MIT-BIH atrial fibrillation and Creighton University ventricular tachyarrhythmia are sampled at a frequency of 250 Hz whereas the ECG signals acquired from

MIT-BIH arrhythmia are sampled at a frequency of 360 Hz. Hence, the ECG signals from MIT-BIH arrhythmia database are downsampled from 360 to 250 Hz. Then, all the ECG signals are denoised and the baseline is removed with Daubechies wavelet 6 [4].

Further, the ECG signals are segmented and sorted according to the cardiac conditions with the prescribed annotations retrieved from the public database. In this study, we segmented the ECG signals of *four* classes into net A and net B without any wave detection. Each segment is normalized with Z-score normalization to address the problem of amplitude scaling and to eliminate the offset effect before we feed the ECG segments into the 1-dimensional deep learning CNN for training and testing. An illustration of two seconds (net A) and five seconds (net B) ECG segments used in this work are shown in Figure 1 and Figure 2 respectively.

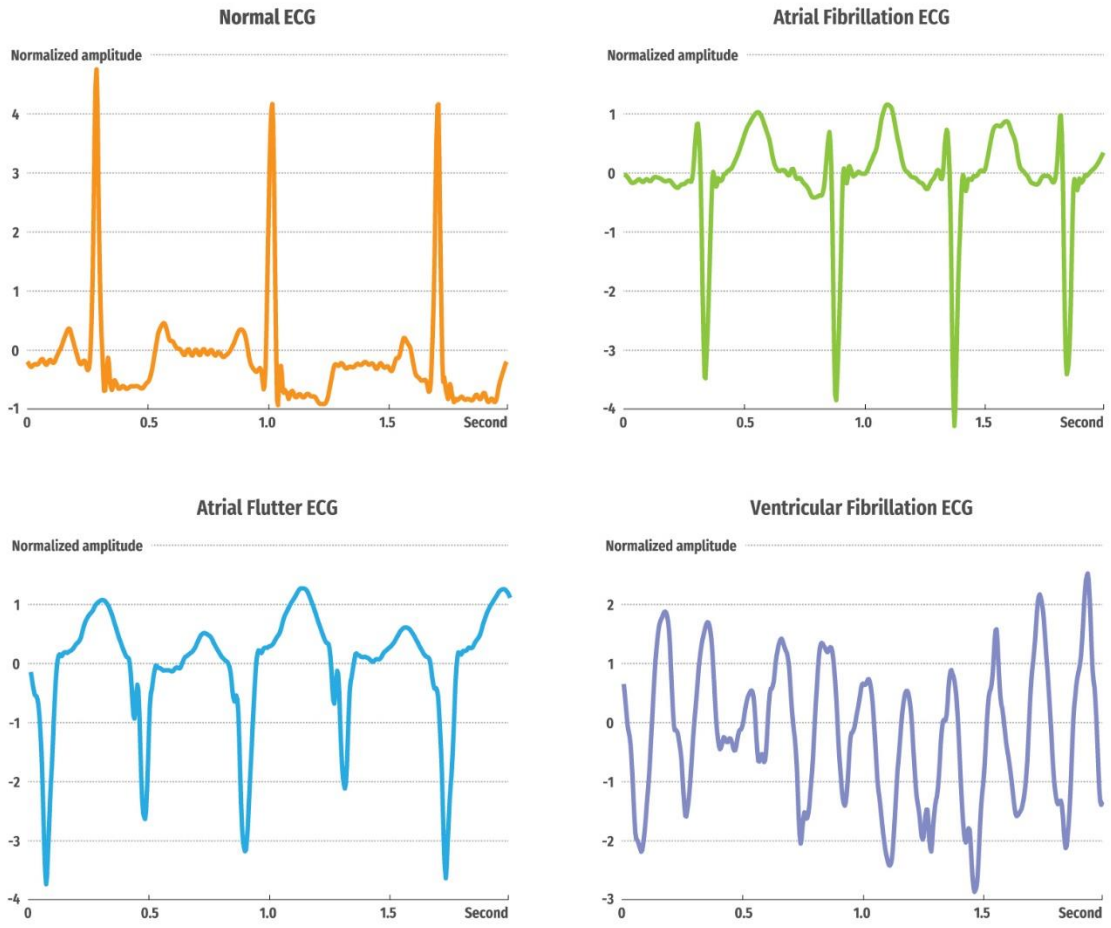

Figure 1. An illustration of ECG segments for net A.

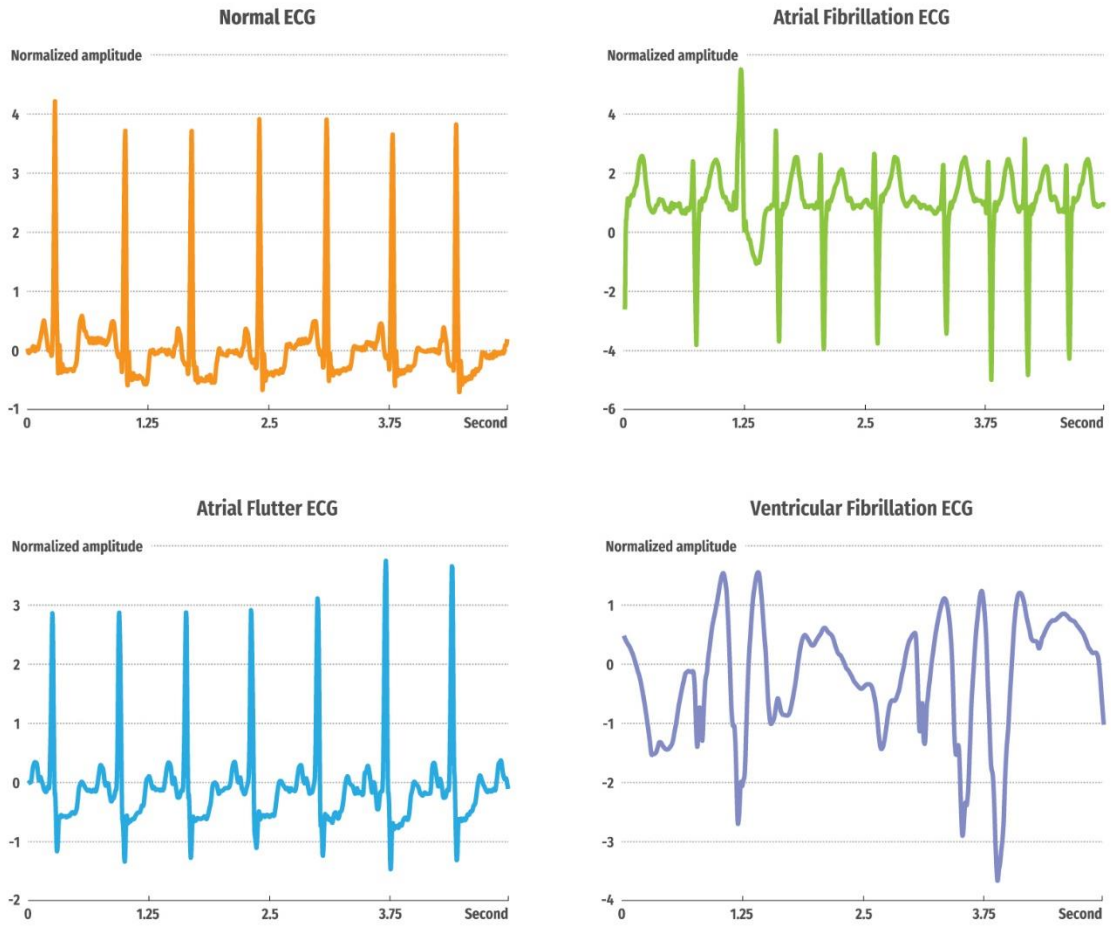

Figure 2. An illustration of ECG segments for net B.

3.2 Convolutional Neural Network (CNN)

Convolutional neural network (CNN) is first introduced by Fukushima in 1980 [10] and later improved by LeCun et al [34]. It is a form of deep learning where the structure is made up of many hidden layers and parameters [34]. Further, the CNN can self-learn and self-organize which does not require supervision [10]. CNN has been applied in diverse applications such as object recognition [16], image classification [1], and handwriting classification [6]. It is also employed in the medical field as an automated diagnostic tool to aid clinicians [13, 15, 18, 32, 37].

It is noted that CNN eliminates the need for pre-processing and separate feature extraction technique [35]. Therefore, it can help to reduce the burden during training and selecting the best feature extraction technique for the automated detection of arrhythmias. Also, there is a possibility of attaining better performance if we can achieve a fitting learning based on the trained hidden layers by learning the structure of the data. Thus, we used CNN in this study for these reasons.

3.3 The Architecture

The primary operations involved in CNN are convolution, non-linearity, pooling, and classification [8, 18]. Two architectures of CNN (net A and net B) are proposed in this work. Figure 3 illustrates the working architecture of net A with 500 input samples. The architecture for net B is illustrated in Figure 4 with 1,250 input samples.

Stride refers to the number of samples the filter matrix slides over the input matrix. Therefore, in this work, we have used 1 and 2 strides (see Table 3 and Table 4). When the stride is 1, the filter is moved from one sample to another at a time and when the stride is 2, the filter moves 2 samples at a time. A bigger stride will result in smaller feature maps and vice versa.

For both net A and net B, the input layer (layer 0) is convolved with a kernel size of 27 to produce layer 1. A max-pooling of size 2 is applied onto every feature map (layer 2). Then, the feature maps from layer 2 are convolved with a kernel size of 14 (net A) and 15 (net B) respectively to obtain layer 3. A max-pooling of size 2 is again applied to every feature map (layer 4). The feature maps from layer 4 are then convolved with a kernel size of 3 (net A) and 4 (net B) to produce layer 5 in net A and net B respectively. A max-pooling of size 2 is applied onto every feature map (layer 6). Then, the feature maps from layer 6 are once again, convolved with a kernel size of 4 (net A) and 3 (net B) to obtain layer 7 for both net A and net B accordingly. A max-pooling of size 2 is again applied to every feature map (layer 8). Finally, the neurons of every feature maps in layer 8 are fully connected to 30 neurons in layer 9, which is also fully connected to 10 and 4 outputs in layers 10 and 11 respectively.

The leaky rectifier linear unit [11] is used as an activation function for layer 1, 3, 5, 7, 9, and 10. We have used the softmax function for the last layer (layer 11) and Xavier initialization [33] for the weights of layers 1, 3, 5, 7, 9, and 10.

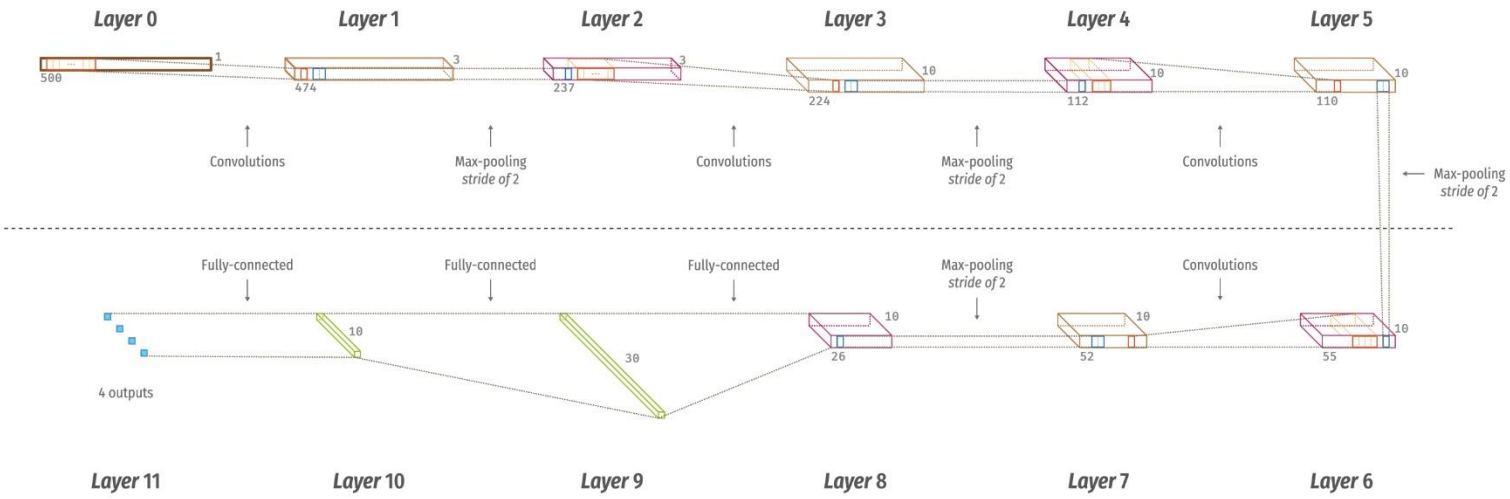

Figure 3. The architecture of the proposed CNN for net A.

Table 3
The details of CNN structure for net A.

Layers	Type	Number of Neurons (Output Layer)	Kernel Size for Each Output Feature Map	Stride
0-1	Convolution	474 x 3	27	1
1-2	Max-pooling	237 x 3	2	2
2-3	Convolution	224 x 10	14	1
3-4	Max-pooling	112 x 10	2	2
4-5	Convolution	110 x 10	3	1
5-6	Max-pooling	55 x 10	2	2
6-7	Convolution	52 x 10	4	1
7-8	Max-pooling	26 x 10	2	2
8-9	Fully-connected	30	-	-
9-10	Fully-connected	10	-	-
10-11	Fully-connected	4	-	-

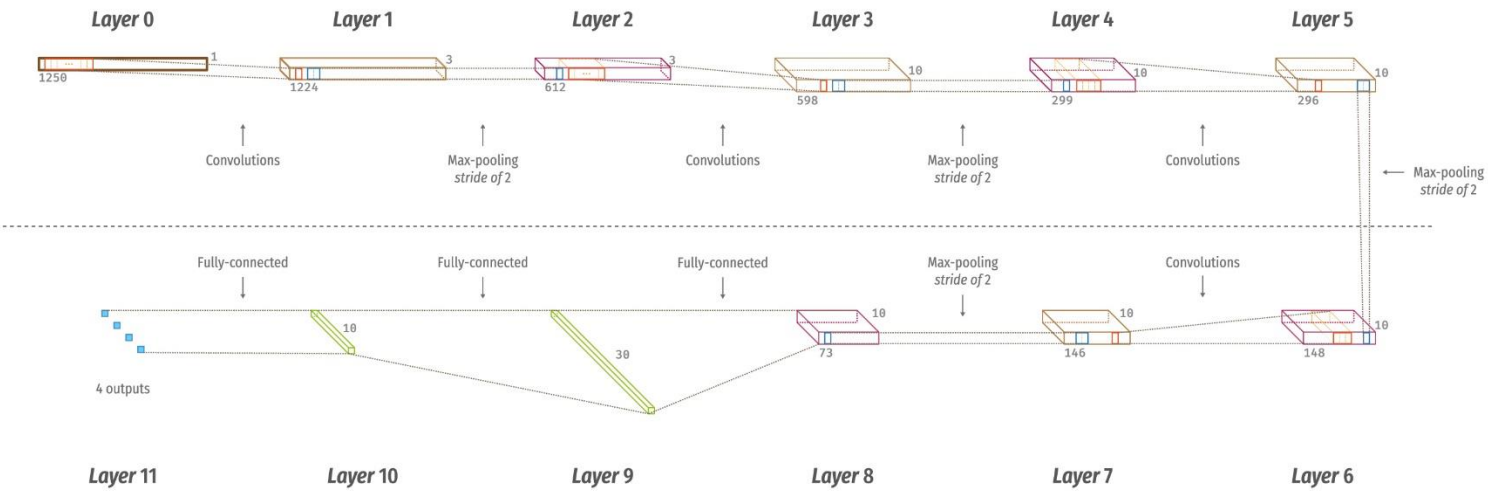

Figure 4. The architecture of the proposed CNN for net B.

Table 4
The details of CNN structure for net B.

Layers	Type	Number of Neurons (Output Layer)	Kernel Size for Each Output Feature Map	Stride
0-1	Convolution	1224 x 3	27	1
1-2	Max-pooling	612 x 3	2	2
2-3	Convolution	598 x 10	15	1
3-4	Max-pooling	299 x 10	2	2
4-5	Convolution	296 x 10	4	1
5-6	Max-pooling	148 x 10	2	2
6-7	Convolution	146 x 10	3	1
7-8	Max-pooling	73 x 10	2	2
8-9	Fully-connected	30	-	-
9-10	Fully-connected	10	-	-
10-11	Fully-connected	4	-	-

3.4 Training

Standard backpropagation [9] with a batch size of 10 is implemented for stochastic learning. The weights are updated according to *equation (1)*.

$$w_l = \left(1 - \frac{n_\lambda}{ts}\right) w_{l-1} - \frac{n}{x}\frac{\partial c}{\partial w} \tag{1}$$

where w, l, n, $\boldsymbol{\lambda}$, ts, x, and c denotes the weight, layer number, learning rate, regularization parameter, total number of training samples, batch size, and cost function respectively. In addition, the biases are updated through *equation (2)*.

$$b_l = b_{l-1} - \frac{n}{x}\frac{\partial c}{\partial w} \tag{2}$$

In this work, we have used learning rate, regularization, and momentum parameters. The parameters are set at 0.002, 0.2, and 0.7 respectively.

3.5 Testing

After every round of training epoch is completed, our algorithm performs a test on the CNN model. We used *30*% of the training set (*90*%) for validation of the algorithm after every epoch. A total of twenty epochs of training and testing iterations was run. An illustration of the distribution of ECG segments used for training and testing procedures can be seen in Figure 5.

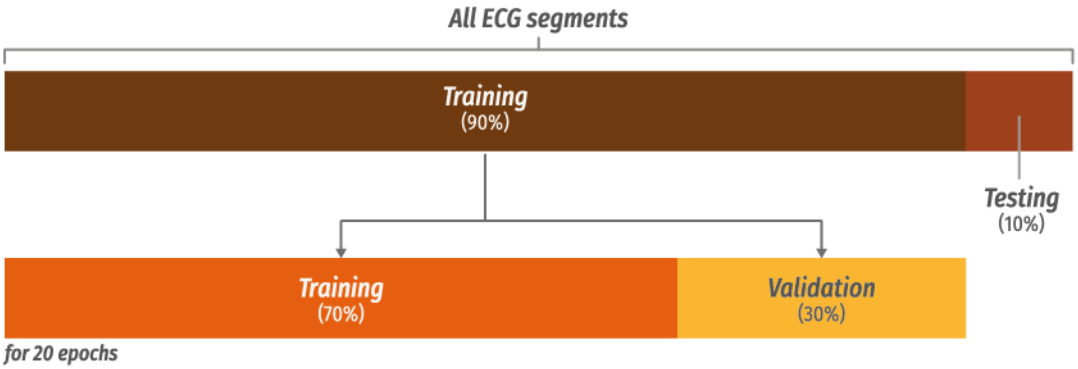

Figure 5. The distribution of ECG segments used for training and testing.

3.6 k-fold Cross-validation

We have used ten-fold cross-validation [24] in this work. Therefore, the total ECG segments for net A (21,709 segments) and net B (8,683 segments) are divided into *ten* equal portions. *Nine* out of *ten* portions are used for training and the rest (one-tenth) of the ECG segments are used for testing. This procedure is repeated *ten* times by shifting the testing data portion. In each fold, the performance (sensitivity, specificity, and accuracy) is evaluated. The average of all *ten*-folds gives the total performance of the system.

4. Result

We have trained our algorithm on a workstation with two Intel Xeon 2.40 GHz (E5620) processor and a 24GB RAM. It took an average of 557.812 seconds to complete an epoch of training for net A and 256.332 seconds for net B.

Table 5 and Table 6 show the confusion matrix for two-second and five-second segment respectively. It can be seen from Table 5 that; 93.13% ECG segments are correctly classified as NSR class. 92.89% of ECG segments are correctly classified as A-Fib. A total of 8.64% AFL ECG segments is wrongly classified as NSR, A-Fib, and V-Fib. Furthermore, more than a third of V-Fib is wrongly classified as A-Fib.

Also in Table 6, 18.56% of NSR ECG segments are wrongly classified as A-Fib and AFL. 7.11% A-Fib segments are incorrectly classified as NSR, AFL, and V-Fib ECG segments. Out of 736 AFL ECG segments, 86.96% are accurately classified as AFL. Again, more than a third of the V-Fib segments are wrongly classified as A-Fib.

The overall classification results for net A and net B is tabulated in Table 7. An accuracy of 92.50% and a sensitivity and specificity of 98.09% and 93.13% respectively is achieved using net A. Also, an average accuracy, sensitivity, and specificity of 94.90%, 99.13%, and 81.44% respectively are obtained for net B.

Table 5
Confusion matrix for net A.

Original/ Predicted	NSR	A-Fib	AFL	V-Fib	Acc (%)	PPV (%)	Sen (%)	Spec (%)
NSR	840	45	17	0	97.88	67.85	93.13	98.09
A-Fib	363	17,467	597	377	92.82	98.75	92.89	92.39
AFL	32	115	1,681	12	96.41	73.02	91.36	96.87
V-Fib	3	61	7	92	97.88	19.13	56.44	98.19

* *Acc* = Accuracy, *PPV* = Positive Predictive Value, *Sen* = Sensitivity, *Spec* = Specificity

Table 6
Confusion matrix for net B.

Original/ Predicted	NSR	A-Fib	AFL	V-Fib	Acc (%)	PPV (%)	Sen (%)	Spec (%)
NSR	294	55	12	0	98.40	80.33	81.44	99.13
A-Fib	57	7,289	116	59	95.32	97.67	96.92	85.03
AFL	15	75	640	6	97.37	82.90	86.96	98.34
V-Fib	0	44	4	17	98.70	20.73	26.15	99.25

Table 7
The overall classification results for the classification of NSR, A-Fib, AFL and V-Fib classes.

Segment Length	*TP*	*TN*	*FP*	*FN*	*Acc* (%)	*PPV* (%)	*Sen* (%)	*Spec* (%)
Two seconds	20,409	840	62	398	92.50	99.70	98.09	93.13
Five seconds	8,250	294	67	72	94.90	99.19	99.13	81.44

**TP* = True Positive, *TN* = True Negative, *FP* = False Positive, *FP* = False Negative

5. Discussion

The number of V-Fib segments (Table 2) used in this work is too few (163 and 65 ECG segments in net A and net B respectively) and hence resulted in low sensitivity rate.

In this work, net B (five seconds long ECG signal) performed slightly better than net A (two seconds long ECG signal) as there are additional three seconds of additional information on ECG morphology. However, the results of two (two and five second) time durations are comparable.

Also, CNN is invariant to translation. Therefore, in this work, the ECG segments are not affected by time shifting and scaling thus there is no need to perform QRS detection in the pre-processing stage. Normally, the primary steps involved in analyzing ECG signals are (i) filtering of noise, (ii) detection of QRS complex, (iii) extraction of R-peak, and (iv) formulation of feature set [30]. Nonetheless, we did not implement step (ii) and (iii) in this work. Most of the works reported in Table 8 have detected QRS wave in their study. Our results for net A and net B are comparable to the previous works reported (in Table 8) which proves that the detection of QRS wave is not necessary for the classification of arrhythmia.

In addition, the sensitivity rate achieved for net A (98.09%) and net B (99.13%) is comparable to those studies summarized in Table 8. Our group [29] obtained a sensitivity of 99.30% using a total of 614,526 ECG beats (75,815 NSR beats, 520,292 A-Fib beats, 14,257 AFL beats, and 4,162 V-Fib beats). In this present work, we obtained a sensitivity of 98.09% and 99.13% for two and five seconds' durations with a total of 21,709 and 8,683 ECG segments for net A and net B respectively.

Additionally, in contrast to the authors [21, 23, 28, 29] in Table 8, we analyzed the ECG signals in short-term duration (two-second and five-second segments) instead of analyzing one beat of ECG signal. Normally, doctors analyze a short-duration of ECG signals, not just an ECG beat for diagnosis. Therefore, it is more realistic to feed two and five seconds of ECG signals to the CNN structure for the automated detection of arrhythmias. Hence, in this study, we segmented our ECG signals into two-second and five-second ECG segments.

Table 8
Summary of selected studies conducted for the detection and diagnosis of arrhythmias using the same database.

Author, Year	Database	Special Characteristics	ECG Rhythms	Classifier	Performance
			Three-Class		
Wang et al., 2001 [36]	mitdb	• No QRS detection performed • Analysis of 1.2 second ECG segment • Analysis of 1.8 second ECG segment • Analysis of 2.4 second ECG segment • Two-layer fuzzy Kohenen network	A-Fib, V-Fib, VT	Fuzzy Kohonen network	A-Fib: *Acc* = 99.40% *Sen* = 98.30% *Spec* = 100.00% V-Fib: *Acc* = 97.20% *Sen* = 98.30% *Spec* = 96.700% VT: *Acc* = 97.80% *Sen* = 95.00% *Spec* = 99.20%
Martis et al., 2013 [21]	afdb, mitdb	• QRS detection performed • Analysis of one ECG beat (2,383 beats)	A-Fib, AFL, NSR	K-nearest neighbor	*Acc* = 99.50% *Sen* = 100.00% *Spec* = 99.22%
Martis et al., 2014 [23]	afdb, mitdb	• QRS detection performed • Analysis of one ECG beat (2,942 beats)	A-Fib, AFL, NSR	K-nearest neighbor	*Acc* = 99.45% *Sen* = 99.61% *Spec* = 100.00%
			Four-Class		
Fahim et al., 2011 [26]	MIT-BIH physiobank	• QRS detection performed • Analysis of ten-second ECG segment (800 segments)	A-Fib, Atrial premature beat, Premature ventricular contraction, V-Fib or VFL	Rule-based	*Acc* = 97.00% (average)
Acharya et al., 2016 [29]	afdb, cudb, mitdb	• QRS detection performed • Analysis of one ECG beat (614,526 beats)	A-Fib, AFL, V-Fib, NSR	Decision tree	*Acc* = 96.30% *Sen* = 99.30% *Spec* = 84.10%

Desai et al., 2016 [28]	afdb, cudb, mitdb	• QRS detection performed • Analysis of one ECG beat (3,858 beats)	A-Fib, AFL, V-Fib, NSR	Rotation forest	*Acc* = 98.37%
Current study	**afdb, cudb, mitdb**	• **No QRS detection performed** • **Analysis of two-second ECG segment (21,709 segments)** • **Analysis of five-second ECG segment (8,683 segments)** • **No feature extraction or feature selection involved** • **Eleven-layer deep CNN**	**A-Fib, AFL, V-Fib, NSR**	**Convolutional neural network**	**Net A:** ***Acc* = 92.50%** ***Sen* = 98.09%** ***Spec* = 93.13%** **Net B:** ***Acc* = 94.90%** ***Sen* = 99.13%** ***Spec* = 81.44%**

*_Acc_ = Accuracy, _Sen_ = Sensitivity, _Spec_ = Specificity

*A-Fib = Atrial fibrillation, AFL = Atrial flutter, V-Fib = Ventricular flutter, VFL = Ventricular flutter, VT = Ventricular tachycardia, NSR = Normal sinus rhythm

*afdb = MIT-BIH atrial fibrillation, cudb = Creighton university ventricular tachyarrhythmia, mitdb = MIT-BIH arrhythmia

It is evident that our proposed algorithm is more robust as compared to the rest of the works mentioned in Table 8. Overall, our proposed system does not require any QRS detection. Also in this work, the feature extraction and selection and classification are merged into one single model. Furthermore, we have validated the performance of our deep learning model in this work using net A and net B ECG segments.

To the best of our knowledge, this is the first study to implement an eleven-layer CNN for the automated detection system of A-Fib, AFL, NSR, and V-Fib ECG signals without the detection of QRS complex.

The main highlights of our proposed algorithm are as follows:

i. CNN is invariant to translation, therefore, no pre-processing of QRS detection is needed in this work.
ii. No QRS detection is required in this work.

iii. Feature extraction, feature selection, and classification steps are merged in the CNN algorithm.
iv. *Ten*-fold cross-validation is used for the evaluation of CNN performance in this work. Hence, the reported performance is robust.

The drawbacks of our proposed algorithm are as follows:

i. Requires a lot of data (big data) for training.
ii. Takes more time to train the data.

6. Conclusion

Generally, the presence of arrhythmia is reflected in the ECG morphology. Essentially, with many elderly affected by serious arrhythmias, there is a need to design an efficient and robust CAD system to accurately and automatically detect various types of arrhythmias. In this work, we have developed a CNN to automatically classify the four classes (NSR, A-Fib, AFL, and V-Fib) using 21,709 ECG segments of net A and 8,683 ECG segments of net B. Our proposed algorithm achieved an accuracy, sensitivity, and specificity of 92.50%, 98.09%, and 93.13% respectively for net A. Also, we obtained an average accuracy, sensitivity, and specificity of 94.90%, 99.13%, and 81.44% respectively for net B. Hence, it is evident that our developed system has potential to be implemented in clinical settings. Our proposed toolkit can serve as an adjunct tool to assist the clinicians to cross-check their findings. Moreover, clinicians can recommend appropriate treatments promptly and avoid further deterioration of cardiac condition. Further, the robustness of the proposed system can be further improved by using large arrhythmia database with more number of V-Fib, A-Fib, AFL, and NSR ECG segments. In future, we intend to use huge database and employ the Keras models [14] for the validation of the CNN instead of k-fold cross-validation strategy. Also, we propose to automatically classify the ECG signals using CNN ***without performing any noise filtering*** in our future work.

7. References

1. A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems: 12: 1097-1105, 2012.
2. A. L. Goldberger. Clinical Electrocardiography: a simplified approach. Mosby, St. Louis, MO, USA, 2012.
3. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. H. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley. PhysioBank, PhysioToolkit, and PhyisoNet: Components of a new research resource for complex physiologic signals: Circulation 101(23): e215-e220, 2000.
4. B. N. Singh, A. Tiwari. Optimal selection of wavelet basis function applied to ECG signal denoising. Digital Signal Processing 16(3): 275-287, 2006.
5. C. Berry, A. C. Rankin, A. J. B. Brady. Bradycardia and tachycardia occurring in older people: An introduction. British Journal of Cardiology 11(1), 2004.
6. D. C. Ciresan, U. Meier, L. M. Gambardella, J. Schmidhuber. Convolutional neural network committees for handwritten character classification. IEEE International Conference on Document Analysis and Recognition, 1135-1139, 2011.
7. G. V. Chow, J. E. Marine, J. L. Fleg. Epidemiology of arrhythmias and conduction disorders in older adults. Clinics in Geriatric Medicine 28(4): 539-553, 2012.
8. I. Goodfellow, Y. Bengio, A. Courville. Deep learning. MIT Press, http://www.deeplearningbook.org, 2016.
9. J. Bouvrie. Notes on convolutional neural network, 2007.
10. K. Fukushima. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics 36: 193-202, 1980.
11. K. He, X. Zhang, S. Ren, J. Sun. Delving deep into rectifiers: Surpassing human-level performance on image net classification, 1026-1034, 2015.
12. K. Najarian, R. Splinter. Biomedical signal and image processing, Second edition. CRC Press, Taylor and Francis Group, Boca Raton, 2012.
13. K. Sirinukunwattana, S. E. A. Raza, Y. W. Tsang, D. R. J. Snead, I. A. Cree, N. M. Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Transactions on Medical Imaging 35(5): 1196-1206, 2016.
14. Keras Documentation. About Keras models. https://keras.io/models/about-keras-models/. (Last accessed: 09 March 2017).

15. M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, C. I. Sánchez. Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. IEEE Transactions on Medical Imaging 35(5): 1273-1284, 2016.
16. M. Oquab, L. Bottou, I. Laptev, J. Sivic. Is object localization for free? - weakly-supervised learning with convolutional neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 685-694, 2015.
17. M. Zubair, J. Kim, C. W. Yoon. An automated ECG beat classification system using convolutional neural networks. IEEE 6th International Conference on IT Convergence and Security, 2016.
18. N. Hatipoglu, G. Bilgin. Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships. Medical and Biology in Engineering and Computing, 1-20, 2017.
19. National Institute on Aging – turning discovery into health. Aging hearts and arteries: A scientific quest, Chapter 4: Blood vessels and aging: The rest of the journey. https://www.nia.nih.gov/health/publication/aging-hearts-and-arteries/chapter-4-blood-vessels-and-aging-rest-journey (Last accessed: 24 February 2017).
20. National Institute on Aging – turning discovery into health. Global health and aging. Assessing the costs of aging and health care. https://www.nia.nih.gov/research/publication/global-health-and-aging/assessing-costs-aging-and-health-care. (Last accessed: 24 February 2017).
21. R. J. Martis, U. R. Acharya, H. Prasad, K. C. Chua, C. M. Lim, J. S. Suri. Application of higher order statistics for atrial arrhythmia classification. Biomedical Signal Processing and Control, 8: 888-900, 2013.
22. R. J. Martis, U. R. Acharya, H. Adeli. Current methods in electrocardiogram characterization. Computers in Biology and Medicine 48: 133-149, 2014.
23. R. J. Martis, U. R. Acharya, H. Adeli, H. Prasad, J. H. Tan, K. C. Chua, C. L. Too, S. W. J. Yeo, L. Tong. Computer-aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation. Biomedical Signal Processing and Control 13: 295-305, 2014.
24. R. O. Duda, P. E. Hart, D. G. Stork. Pattern Classification 2nd Edition. New York, John Wiley and Sons, 2001.

25. S. Bawany. This is the economic impact of an aging Singaporean workforce. Singapore Business Review, 2013. (Last accessed: 24 February 2017).
26. S. Fahim, I. Khalil. Diagnosis of cardiovascular abnormalities from compressed ECG: A datamining-based approach. IEEE Transactions on Information Technology in Biomedicine, 15: 33-39, 2011.
27. S. Kiranyaz, T. Ince, M. Gabbouj. Real-time patient-specific ECG classification by 1-D convolutional neural network. IEEE Transactions on Biomedical Engineering 63(3): 664-675, 2016.
28. U. Desai, R. J. Martis, U. R. Acharya, C. G. Nayak, G. Seshikala, S. K. Ranjan. Diagnosis of multiclass tachycardia beats using recurrence quantification analysis and ensemble classifiers. Journal of Mechanics in Medicine and Biology, 16(1), 2016.
29. U. R. Acharya, H. Fujita, M. Adam, S. L. Oh, J. H. Tan, V. K. Sudarshan, J. E. W. Koh. Automated characterization of Arrhythmias using nonlinear features from tachycardia ECG beats. IEEE International Conference on Systems, Man, and Cybernetics, 2016.
30. U. R. Acharya, J. S. Suri, J. A. E. Spaan, S. M. Krishnan. Advances in Cardiac Signal Processing. New York, Springer-Verlag Berlin Heidelberg, 2007.
31. United Nations. Department of Economic and Social Affairs Population Division. World population aging 2015. New York, 2015.
32. V. Golkov, A. Dosovitskiy, J. I. Sperl, M. I. Menzel, M. Czisch, P. Sämann, T. Brox, D. Cremers. q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Transactions on Medical Imaging 35(5): 1344-1351, 2016.
33. X. Glorot, Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. Aistats, 2010.
34. Y. LeCun, L. Bottou, Y. Bengio, P Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86 (11): 2278-2324, 1998.
35. Y. LeCun, Y. Bengio. Convolutional networks for images, speech, and time-series. In: The handbook of brain theory and neural networks, MIT Press Cambridge, MA, USA, 1998.
36. Y. Wang, Y. S. Zhu, N. V. Thakor, Y. H. Xu. A short-time multifractal approach for arrhythmia detection based on fuzzy neural network. IEEE Transactions on Biomedical Engineering 48(9): 989-995, 2001.
37. Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, S. Zhang, D. N. Metaxas, X. S. Zhou. Multi-instance deep learning: discover discriminative local anatomies for bodypart recognition. IEEE Transactions on Medical Imaging 35(5): 1332-1343, 2016.