Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats

Oh, Shu Lih; Ng, Eddie Yin Kwee; Tan, Ru San; Acharya, U. Rajendra

2018

Oh, S. L., Ng, E. Y. K., Tan, R. S., & Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. Computers in biology and medicine, 102, 278-287. doi:10.1016/j.compbiomed.2018.06.002

https://hdl.handle.net/10356/136847

https://doi.org/10.1016/j.compbiomed.2018.06.002

Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length Heart beats

Shu Lih Oh [a], Eddie Y K Ng [b], Ru San Tan [c], U. Rajendra Acharya [a,d,e,*]

[a] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[b] School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore.
[c] National Heart Centre Singapore, Singapore
[d] Department of Biomedical Engineering, School of Science and Technology, Singapore University of Social Sciences, Singapore
[e] Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia
[*]Postal Address: Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489.
Telephone: +65 6460 6135; Email Address: aru@np.edu.sg

Abstract

Arrhythmia is a cardiac conduction disorder characterized by irregular heartbeats. Abnormalities in the conduction system can manifest in the electrocardiograhic (ECG) signal. However, it can be challenging and time-consuming to visually assess the ECG signals due to the very low amplitudes. Implementing an automated system in the clinical setting can potentially help expedite diagnosis of arrhythmia, and improve the accuracies. In this paper, we propose an automated system using a combination of convolutional neural network (CNN) and long short-term memory (LSTM) for diagnosis of normal sinus rhythm, left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature beats (APB) and premature ventricular contraction (PVC) on ECG signals. The novelty of this work is that we used ECG segments of variable length from the MIT-BIT arrhythmia physio bank database. The proposed system demonstrated high classification performance in the handling of variable-length data, achieving an accuracy of 98.10%, sensitivity of 97.50% and specificity of 98.70% using ten- fold cross validation strategy. Our proposed model can aid clinicians to detect common arrhythmias accurately on routine screening ECG.

Keywords: arrhythmia, ambulatory electrocardiogram, computer aided detection, deep learning, convolutional neural network, long short-term memory, automation, diagnosis

1. Introduction

According to a report by United Nation on global population in 2017, 962 million persons are 60 years of age or older. The number of elderly is expected to double by 2050, rising to become nearly 2.1 billion. Furthermore, the number of people more than 80 years of age is projected to triple globally: rising from 137 million to 425 million between 2017 and 2050 [1]. As one ages, the physiological reserve starts to decline and the cardiovascular system becomes weak [2]. As a result, old people are more susceptible to diseases [3]. The cardiovascular system in old age undergoes a multitude of changes, resulting in blood vessels losing their elasticity, thickening of left ventricle muscle wall and accompanying diastolic dysfunction. These changes are caused by the enlargement of myocytes and apoptosis of neighboring tissues followed by the accumulation of fibrofatty plaque on the myocardium [4]. Infiltration of these fatty tissue disrupts the cardiac conduction system and overall heart function, which eventually leads to arrhythmias. Arrhythmias are most often responsible for sudden deaths, and frequently accompany other high-risk cardiac diseases [5-10]. Left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature beats (APB) and premature ventricular contraction (PVC) are the most common types of conduction abnormalities and arrhythmias manifested in older adults [2].

The screening of arrhythmias requires careful study of the ECG records by experienced cardiologists and this process is tedious and time consuming. Moreover, there may be minute changes in the ECG that are overlooked by the naked eye. Hence, several computer-aided diagnosis (CAD) algorithms have been employed over the last decades to automate the identification of arrhythmias.

Many computer-aided diagnosis (CAD) algorithms presented in Table 4 are conventional techniques which require signal preprocessing, waveform detection, features extraction and the use of hand crafted features for classification **[11-22]**. In these systems, hard coded features are often engineered and selected by trial-and-error or experience. Furthermore, such systems also tend to generate more false positives which may lead to misdiagnosis and inappropriate therapy [23]. Therefore, deep learning technique is employed in this work to overcome these challenges faced by the traditional systems and achieve better detection accuracy without the use of any hard coded features.

Recently, many healthcare researchers have incorporated deep learning algorithms to solve various complex tasks ranging from brain scan segmentation [24] to provide useful intervention opinions upon diagnosis [25]. The convolutional neural network (CNN) has achieved great success in the field of computer vision research and is currently widely used in image processing tasks [26-29]. Due to its unique ability of capturing position and translation invariant patterns, CNN is applied to physiological signals [30-33] for morphological analysis. The CNN is relatively less sensitive to noise capable of extracting useful information even when the signals are noisy [34]. This attribute is built within the hierarchical deep structure. As the layers in the network progresses, features are learnt and represented in a more abstract manner. Acharya et al. [30] has investigated the effect of detecting myocardial infraction with and without noise presence in ECG signals using the CNN. The study showed promising results on both sets of data with minor drop in the accuracy during the noisy ECG signals.

In addition to CNN, long short-term memory (LSTM) is another type of deep learning algorithm widely used to analyze the time series. It has been used in many applications like natural language processing [35], speech recognition [36-39] as well as handwriting recognition [28, 39, 40]. The connections between LSTM units allows information to cycle through a loop across the adjacent time-steps. This creates an internal state of feedbacks which allows the network to understand the concept of time and learn about the temporal dynamics within the presented data. LSTMs units also possess the ability to remember or forget information selectively by maintaining a memory state. Information which are of high importance are retained and back propagated while the irrelevant ones are forgotten and thrown away. Recently Tan et al. [31] implemented the LSTM network on ECG segments to diagnose coronary artery disease. The method involved slicing the 5-second ECG signals into short sequences and performing convolution on them to reduce the number of data points. After that, the LSTM is applied to extract the temporal features from the convolved sequences. Their model attained a diagnosis accuracy of 99.85%.

Both the CNN and LSTM have performed relatively well on ECG signals. Furthermore deep learning models do not require any extraction of hand crafted features and they are relatively easy to implement. Hence we have employed the combination of these two algorithms in this work for the diagnosis of arrhythmias.

2. Data Description

In this work, the dataset and the annotations used were downloaded from the PhysioNet public database [41]. A total of forty eight ECG recordings were obtained from forty seven subjects. Each MIT-BIH recording was digitized at 360 hertz and accompanied by a set of beat labels rendered at the R peak. The labels were annotated independently by multiple cardiologists and the discrepancies in their diagnosis were resolved through mutual consensus. In this study, we have used only the signals obtained from modified limb lead II. This is a bipolar lead, similar to the standard limb lead II except that, it is acquired though the attachment of electrodes on the torso which is commonly used in Holter recording.

2.1 Preprocessing

The segmentation process is accomplished by assigning the uninterrupted sequences of arrhythmia beats into the corresponding arrhythmia groups. Each segment consists of only one ECG beat type with 99 samples to the left of the first R peak and 160 samples to the right of the last identified uninterrupted R peak. Table 1 shows an overview of the total number of ECG segments obtained from the MIT-BIH arrhythmia database with the corresponding sample length range.

Table 1: Extracted ECG segments and the corresponding sample length ranges.

	No. of Segments	No. of Segments < 1000 samples	No. of Segments > 1000 samples	Sample Length (range)	Average sample length ± SD
Normal	8245	5866	2378	260-512780	7551.74±15126.83
LBBB	344	97	247	260-364825	6764.85±26725.71
RBBB	660	492	168	260- 103072	3400.35±10515.51
APB	1004	847	157	260-18639	617.77±980.78
PVC	6248	6233	13	260-14012	285.71±276.4
Total	16499	13535	2963		

*SD: standard deviation.

It can be seen from Table 1 that, the variations in the sample length between the segments are sometimes very large. In order to standardize the input length of the model and to reduce any unnecessary training time, the longer segments were arbitrarily truncated down from the end to

1000 samples each. The segments of less than 1000 samples were padded with zeros. This enables the LSTM cells of the subsequent layers to be trained at a faster rate without having to iterate through the long segments of padded zeros for the shorter length signals. A total of 2963 segments were truncated and 13535 segments were post padded with zeros.

To further accelerate the training process, Z-score normalization is applied to the input data to eliminate the offset effect and standardize the ECG signal amplitude. [42]. Range of values in the raw data varies. The normalization technique helps to standardize the values by squashing the original data values to a smaller range. This improves the gradient flow in the network during gradient decent and thus helps to speed up the convergence rate. The typical ECG segments of normal, LBBB, RBBB, APB and PVC signals are shown in Figure 1.

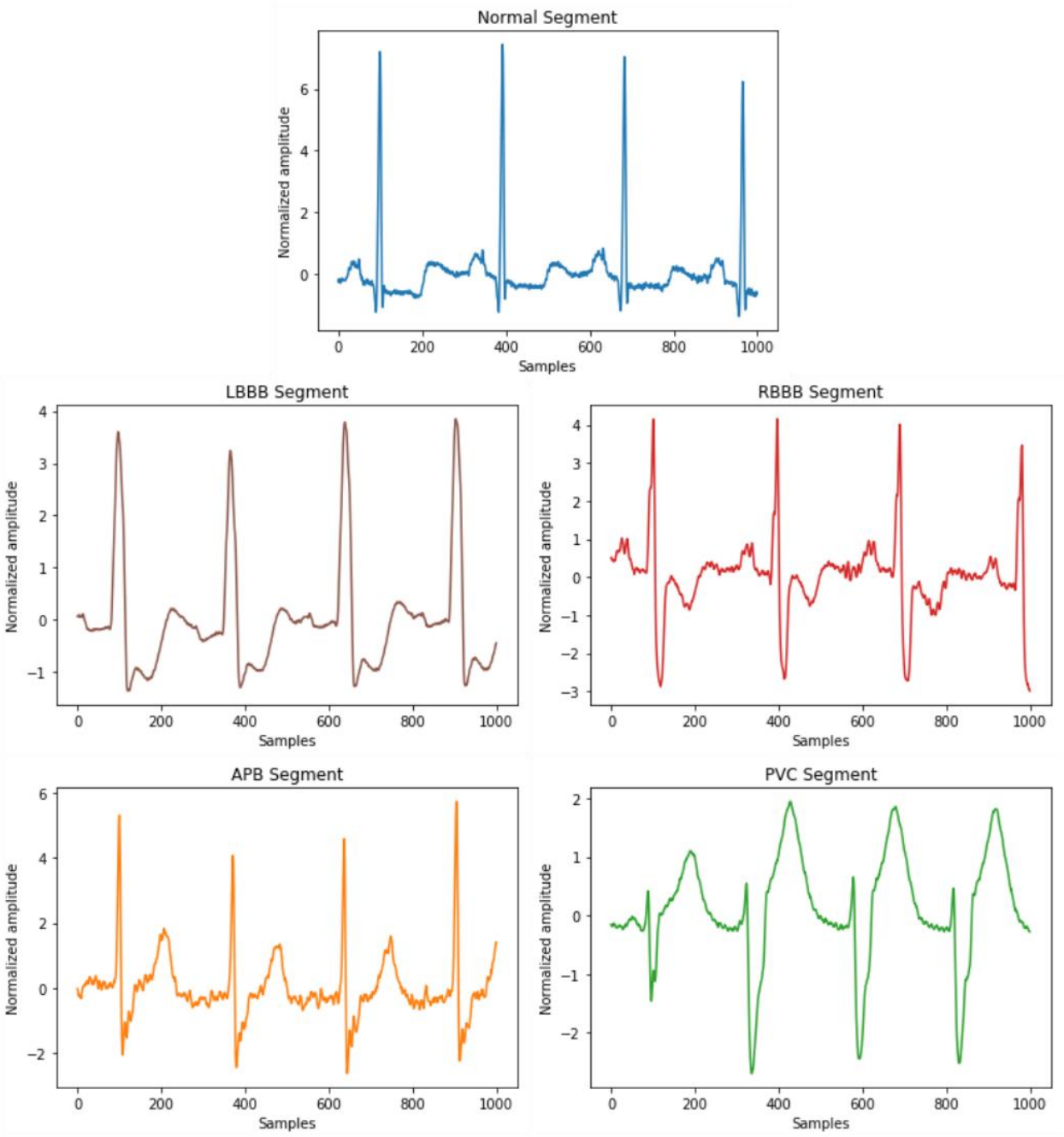

Figure 1: Typical sketch of normal and ECG arrhythmia segments.

3. Proposed CNN-LSTM Model

In this work, a novel technique to detect five different classes of arrhythmias using the ECG signals automatically is proposed. Arrhythmias are erratic and irregular, may come in single or multiple beats. Therefore the proposed network should be capable of handling such signals of variable in lengths. The proposed network is a hybrid deep learning model which combines CNN with LSTM to meet this requirement. Figure 2 illustrates the proposed architecture and Table 2 gives a detail overview of the structure.

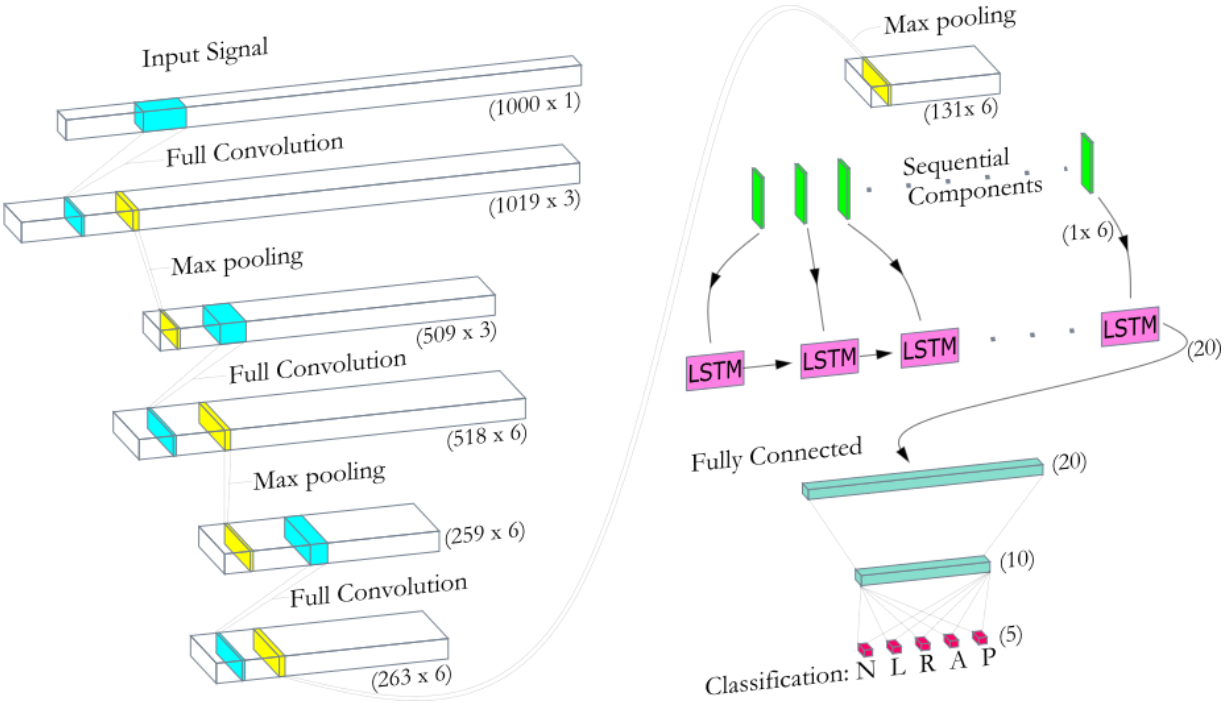

Figure 2: An illustration of the proposed CNN-LSTM architecture.

Layers 1 to 6 of the structure are convolution layers coupled with max pooling layers whereas layer 7 is a LSTM layer. In the final stage of the network is a series of fully-connected layers used to predict the output. The convolution layers are good in extracting the spatial features maps and the succeeding LSTM layer helps the model to capture the temporal dynamics present within these features maps [43]. The ECG segments are classified depending on the output of the final LSTM cell via the fully connected layers.

Table 2: Detailed overview of the proposed CNN-LSTM model.

Layers	Types	Activation function	Output Shapes	Kernel Size	No. of Filters	Stride	No. of trainable parameters	Scheme A	Scheme B
0	Input	-	1000 x 1	-	-	-	0	-	-
1	1D Full convolution without bias	ReLU	1019 x 3	20 x 1	3	1	60	-	-
2	1D Max-pooling	-	509 x 3	2 x 1	3	2	0	-	-
3	1D Full convolution without bias	ReLU	518 x 6	10 x 1	6	1	180	-	-
4	1D Max-pooling	-	259 x 6	2 x 1	6	2	0	-	-
5	1D Full convolution without bias	ReLU	263 x 6	5 x 1	6	1	180	-	-
6	1D Max-pooling	-	131 x 6	2 x 1	6	2	0	-	-
7	LSTM		20	-	-	-	2160	Recurrent dropout (20%) Dropout (20%)	Recurrent dropout (20%)
8	Fully-connected	ReLU	20	-	-	-	420	-	Dropout (20%)
9	Fully-connected	ReLU	10	-	-	-	210	-	Dropout (20%)
10	Fully-connected	Softmax	5	-	-	-	55	-	-
						Total	**3265**		

3.1 System Architecture

The proposed model consists of *three* convolution layers with stride 1. Each convolution operation is performed by shifting the kernel across the input vector one sample at a time, where by the superimposed matrices are multiplied and summed. During the training process, the weights of kernel *k* are constantly adjusted by the network to pick up meaningful spatial information present in the data. Instead of valid convolution, full convolution is performed in this work. This is because the shorter length segments are already padded with zeros. Additionally, to retain the integrity of the zero padding, no bias is added during the

convolution operation. The output of the zero padded sequence from the convolution layer will therefore still be treated as zero. To reduce the size of the input representation by half, max pooling filter of size 2 with non-overlapping stride is applied to the feature maps after each convolutional layer.

The LSTM layer is used subsequently to extract the temporal information from these feature maps. The features extracted from the convolution and pooling process were broken down into sequential components and fed to the recurring LSTM units for temporal analysis. Only the output from the last step of the LSTM is then fed in the fully connected layer for arrhythmia prediction.

3.2 Dropout Regularization

Overfitting of model during training is a considerable obstacle especially when the number of ECG segments is few. In this work, two schemes of dropouts were deployed and tested at different layers of the network to prevent the overfitting. The idea of randomly dropping out part of the network during the training phase is to prevent the neurons from adapting exceedingly well to the training data. As the neurons were dropped out, the connecting weights will be excluded from updating, this forces the network to learn from the imperfect patterns and thus improve the generalization of the model [44].

3.3 Training and Evaluation

The ten-fold cross-validation strategy is employed in this work to evaluate the robustness of the proposed model. Stratified random sampling was performed on the ECG segments by dividing the data into ten equal portions. Each fold 90 % of the ECG segments are used for training the CNN-LSTM network, while the remaining 10 % is used to test the performance of our proposed system. The process is iterated 10 times and each time a newly initialized model will be trained with a new set of training and test data. In order to monitor the training process and prevent the model from overfitting, 20% of the k-fold training data is used to validate the performance at the end of each training epoch. The details of data distribution for each training fold is shown in Figure 3.

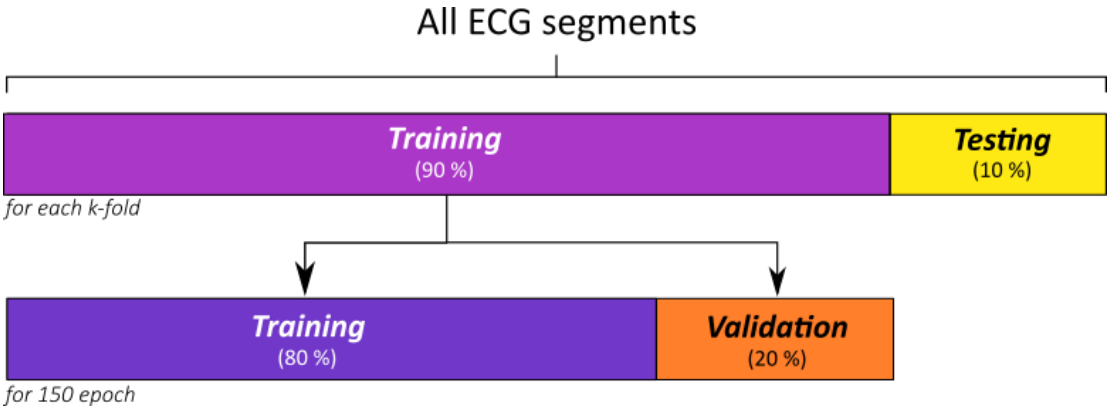

Figure 3: Details of ECG segments used for training and testing the proposed network.

In each fold, the weights from the network were reinitialized using the Xavier (also known as Glorot) algorithm [45] and then the model was then trained end-to-end for 150 epochs using backpropagation algorithm with a batch size of 10. The learning the rate (η) of the model is set to 0.001 and is used in conjunction with Adam optimizer to accelerate the learning process of the network. Weighted loss is also applied during the updates to counter the class imbalance problem.

The performance of each fold is evaluated based on accuracy (Acc), sensitivity (Sens), specificity (Spec) and positive predictive value (PPV).

4. Results

The deep network in this work is developed and evaluated in python language using Keras [46] with Tensorflow as backend [47]. The workstation used consists of two Intel Xeon 2.40 GHz (E5620) processors and a 24GB RAM to train the network. Each training epoch took approximately 138.12 seconds to complete. In total, three different schemes are tested. The first scheme is a trained network without any dropout and is introduced as reference to examine the effect between a regular network and dropout network. The other two are dropout schemes. For Scheme A, 20% of the recurrent and input connections of the LSTM layer are dropped out. In Scheme B, 20% of the LSTM recurrent connections followed by 20% of the two fully connected layers are dropped out. The training and validation accuracy curves for each of these schemes are presented in Figure 4.

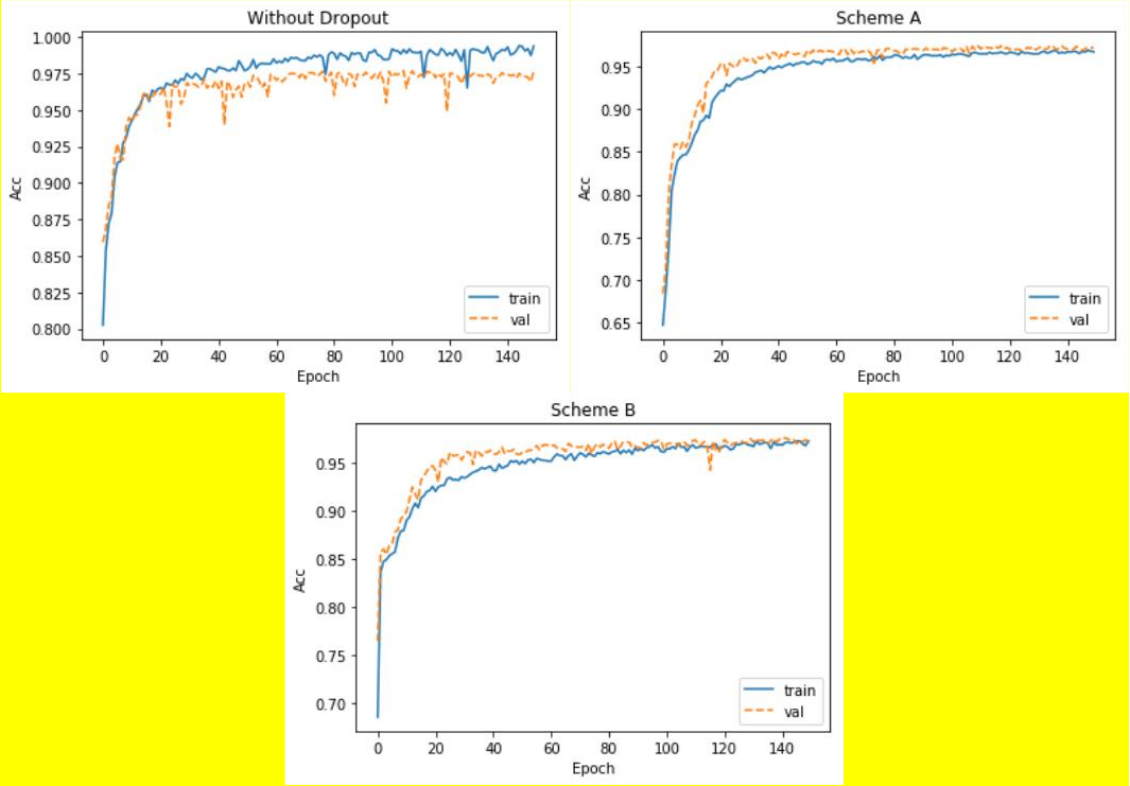

Figure 4: Accuracy plots for the various schemes during training.

It can be observed that the scheme without drop out is clearly overfitting beyond epoch 25 as the validation accuracy starts to plateau while the training accuracy continues to rise. The accuracy of the dropout networks on the other hand are fairly stable. Both the validation and training curves increased at a steady state and eventually meet at the end of 150 epochs.

The overall cross validation performances for all three schemes are summarized in Table 3. The highest accuracy, sensitivity, specificity and PPV of 98.42%, 98.07%, 98.76% and 98.76% respectively are achieved by the over fit model and Scheme A is the lowest performing scheme among the three training schemes.

Table 3: Average classification performance for different schemes.

Scheme	Acc % ± SD		Sens % ± SD		Spec % ± SD		PPV % ± SD	
Without Dropout ***(test reference)***	98.42	±0.31	98.07	±0.62	98.76	±0.35	98.76	±0.34
A	97.88	±0.31	97.26	±0.59	98.50	±0.4	98.48	±0.4
B	98.10	±0.2	97.50	±0.5	98.70	±0.3	98.69	±0.3

*Acc=accuracy, Sens=sensitivity, Spec=specificity

The confusion matrix for Scheme A and B are presented in Figure 5 and Figure 6 respectively. It is observed that Scheme B model is generally better in predicting the normal, LBBB, RBBB and APB classes, whereas Scheme A is a better predictor of PVC class. Both models have performed badly in the classification of APB segments. In Scheme A, 1.8% of the segments are misclassified, while in Scheme B, 1.3 % of the segments are misclassified. Both the models have consistently confused APB with normal segments and vice versa. This may be due to the subtle morphological difference between an ectopic P wave and normal P wave. The amplitude of such waves is very small and hence makes it more difficult detect during the presence of noise.

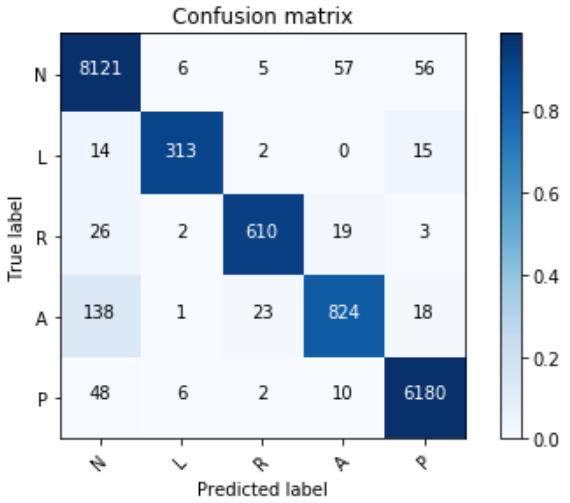

Figure 5: Scheme A - Confusion matrix of the classified ECG segments (N = Normal, L = LBBB, R = RBBB, A = APB, P = PVC).

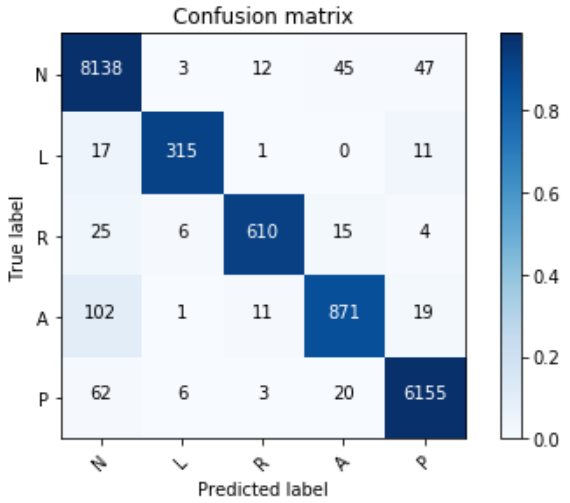

Figure 6: Scheme B - Confusion matrix of the classified ECG segments (N = Normal, L = LBBB, R = RBBB, A = APB, P = PVC).

5. Discussion

Over the years many studies on ECGs have been tested using the MIT-BIH arrhythmia database. Table 4 summarizes the studies on the automated detection of the arrhythmias. The classification of arrhythmia beats used in these studies are not the same. Few studies have used linear [11, 12, 20], nonlinear [13-16, 18, 19, 22], and discrete wavelet transform (DWT) [15] algorithms for feature extraction and classification. Yeh et al.[20] extracted morphological features from ECG components and used them in linear discriminant classification analysis. The

detection of ECG is achieved through the difference operation method (DOM), whereby a series of filters and thresholds are used to identify the PQRST waves. Their study managed to obtain an accuracy of 96.23% using four selected morphological features. In another study [18], higher order cumulants of the ECG beats are extracted using the Hermitian basis functions. The method is shown to be effective in suppressing the intraclass morphological variations and reduced the effects of Gaussian noise. Martis et al. [17] have implemented principal component analysis (PCA) on the segmented ECG beats and reported 98% using only twelve PCA derived components. In the subsequent year, Martis et al. [16] applied the PCA on Higher Order Statistics (HOS) coefficients to capture the subtle morphological variations in the ECG beats for arrhythmias detection. The similar reduction technique is also used by Li et al. [14] to reduce the discrete wavelet transform (DWT) coefficients before extracting the multi-domain features. The system is able to obtain complex characteristics of ECG via a series of multiple reduction techniques with an accuracy of 97.3%. All of these studies confirm that the conduction disorder of heart directly reflects upon the morphology changes of the ECG signals. Hence, is evident that the efficient use of CAD system can help in detecting the subtle changes in ECG signals and improve the detection accuracy of arrhythmia.

Recently, deep learning techniques for arrhythmia detection have been used. Many have investigated the use of convolutional neural network (CNN) for arrhythmia detection and have achieved promising results [48-51]. Yildirim et al. [52] have explored the use of DWT layer coupled with bidirectional long short-term memory (LSTM) for arrhythmia classification and obtained an accuracy of 99.39%.

Table 4: Studies on the automated detection of the arrhythmias.

Author, Year	Notable features	Database	ECG beats	Classifier	Performance
Conventional Methods					
Sahoo et al., 2017 [11]	Morphological and heart rate based features	MITDB	Normal LBBB RBBB Paced **Total (109494)**	Square-Support Vector Machine	Accuracy: 98.39% Sensitivity: 99.87%, PPV: 99.69%
Li et al., 2016 [12]	RR-intervals and Shannon entropy (SE), log energy entropy (LEE), Renyi Entropy (RE) and Tsallis entropy (TE)		Normal (90082) Supra-ventricular ectopic	Random forest	Accuracy: 94.61%

	extracted from Wavelet Packet Decomposition (WPD)		(2779) Ventricular ectopic (7009) Fusion (803) Unknown (15) **Total (100688)**		
Elhaj et al., 2016 [13]	Discrete wavelet coefficients reduced by Principal Components Analysis (PCA) and Higher-Order Statistics (HOS) cumulants reduced by Independent component analysis (ICA)	MITDB	Normal (90580) Supra-ventricular ectopic (2973) Ventricular ectopic (7707) Fusion (1784) Unknown (7050) **Total (110094)**	Square-Support Vector Machine (Radial Basis Function)	Accuracy: 98.91% Sensitivity: 98.91%, Specificity: 97.85%
Li et al., 2016 **[**14**]**	Discrete wavelet combined with kernel-independent component analysis (kernel ICA) and Principal Components Analysis (PCA)	MITDB	Normal (400) LBBB (400) RBBB (400) PVC (400) APB (200) **Total (1800)**	Support Vector Machine	Accuracy: 98.8% Sensitivity: 98.50%, Specificity: 99.69% PPV: 98.91%
Martis et al ., 2013 [15]	Independent Component Analysis (ICA) applied on Discrete wavelet transform (DWT) sub bands	MITDB	Normal (90580) Supra-ventricular ectopic (2973) Ventricular ectopic (7707) Fusion (1784) Unknown	Probabilistic Neural Network (PNN)	Accuracy: 99.28% Sensitivity: 97.97%, Specificity: 99.83% PPV: 99.21%

			(7050) **Total (110094)**		
Martis et al ., 2013 **[**16**]**	Bispectrum and Principal Components Analysis (PCA)	MITDB	Normal (10000) LBBB (8069) RBBB (7250) PVC (7126) APB (2544) **Total (34989)**	Least Square-Support Vector Machine (Radial Basis Function)	Accuracy: 93.48%, Sensitivity: 99.27% Specificity: 98.31%
Martis et al ., 2012 **[**17**]**	Principal Components Analysis (PCA) of ECG beat segments	MITDB	Normal (10000) LBBB (8069) RBBB (7250) PVC (7126) APB (2544) **Total (34989)**	Least Square-Support Vector Machine (Radial Basis Function)	Accuracy: 98.11%, Sensitivity: 99.90% Specificity: 99.10% PPV: 99.61%
Karimifard et al., 2011 **[**18**]**	Hermite model of the Higher-Order Statistics (HOS)	MITDB	Normal (2000) LBBB (1456) RBBB (2251) PVC (2938) APB (722) **Total (9367)**	1-Nearest Neighborhood	Specificity: 99.67% Sensitivity: 98.66%
Martis et al., 2011 [19]	Higher order spectra (HOS) cumulants of Wavelet packet decomposition (WPD)	MITDB	Normal (641) Ventricular premature contractions and atrial premature	Support Vector Machine (Radial Basis Function)	Accuracy: 98.48%, Sensitivity: 98.90% Specificity: 98.04%

			contractions (606) Total (1247)		PPV: 98.13%
Yeh et al., 2009 [20]	Morphological and heart rate based features	MITDB	Normal (75054) LBBB (8074) RBBB (9259) PVC (7129) APB (2544) **Total (102060)**	Linear Discriminant Analysis	Accuracy: 96.23% Sensitivity: Normal 98.97% LBBB 91.07%, RBBB 95.09% PVC 92.63% APB 84.68%
Yu et al., 2008 [21]	Independent component analysis (ICA) and RR interval		Normal (800) LBBB (200) RBBB (200) PVC (1060) APB (364) Paced (200) Ventricular flutter wave (472) Ventricular escape (104) **Total (3400)**	Probabilistic Neural Network (PNN)	Accuracy: 98.71%
Osowski et al., 2008 [22]	Higher-Order Statistics (HOS) cumulants and Hermite coefficients of QRS complex	MITDB	Normal and 12 types of arrhythmias **Total (12785)**	Support Vector Machine	Accuracy: 98.71%
Deep Learning Methods					
Yildirim et al., 2018 [52]	Discrete wavelet transform (DWT) sub bands	MITDB	Normal (2190) LBBB (1870) RBBB (1356)	Deep bidirectional LSTM networks	Accuracy: 99.39%

			PVC (510) Paced (1450) **Total (7376)**		
Acharya et al., 2017 **[**48**]**	Deep learning	MITDB AFDB CUDB	Normal, atrial fibrillation, atrial flutter, and ventricular fibrillation **Two seconds (21709)** **Five seconds (8683)**	11 layers convolutional neural networks (CNNs)	Two seconds Accuracy: 92.50%, Sensitivity: 98.09% Specificity: 93.13% Five seconds Accuracy: 94.90%, Sensitivity: 99.13% Specificity: 81.44%
Acharya et al., 2017 **[**49**]**	Deep learning	MITDB	Normal (90592) Supra-ventricular ectopic (2781) Ventricular ectopic (7235) Fusion (802) Unknown (8039) **Total (109449)**	9 layers convolutional neural networks (CNNs)	Accuracy: 94.03%, Sensitivity: 96.71% Specificity: 91.54%
Zubair et al., 2016 **[**50**]**	Deep learning	MITDB	Normal, Supra-ventricular ectopic, Ventricular ectopic, Fusion and Unknown **Total (100389)**	Convolutional neural networks (CNNs)	Accuracy: 92.70%
Kiranyaz et al., 2016 **[**51**]**	Deep learning	MITDB	Normal, Supra-ventricular ectopic, Ventricular	Convolutional neural networks (CNNs)	Accuracy: 99.00%, Sensitivity: 93.90%

			ectopic, Fusion and Unknown **Total (83648)**		Specificity: 98.90%
This study	Deep learning	MITDB	Normal (8245) LBBB (344) RBBB (660) APB (1004) PVC (6348) **Total (16499)**	LSTM and CNN	Accuracy: 98.1%, Sensitivity: 97.5% Specificity: 98.7%

In this study, end-to-end deep learning structure is employed without any hand-crafted feature. In most of the machine learning the common problem is overfitting. The introduction of dropout regularization helps to prevent the CNN-LSTM model from overfitting during training. By dropping out 20% of the recurrent connections and subsequent dense layers of the model (Scheme B) has led to the better generalization effect as compared to the 20% drop out in LSTM recurrent and input connections (Scheme A). It can also be observed in Table 3 that the high performance of Scheme B is relatively much stable. It has yielded the smallest standard deviation values for all the calculated performance measures. Although the over fitted model in this case has achieved better performance than the proposed schemes, it is only being used as a test reference to prove the efficacy of dropout regularization. This is because such models tend to underperform with a new unknown dataset [53].

The proposed hybrid system of CNN and LSTM learning model has delivered promising results as compared to the other conventional studies **[11-22]**. The high classification accuracy of 98.10% is achieved by dropping out the recurrent connections and dense connections of the network. The CNN is good at picking up spatial features while the LSTM is better at learning temporal features. Hence, by merging the two modalities has not only allows the net to yield better diagnostic accuracy, but has also enhanced the functionality of the model in classifying cardiac signals of varying sequence lengths. Since the proposed deep learning network is

trained end to end using noisy ECG signals, preprocessing techniques like noise filtering and features extraction are therefore not required.

The highlights of the proposed model are as follows:

1. Proposed system is fully automated.
2. Predictions made by the system are reproducible with no inter- and intra-observer biases.
3. Noise filtering, feature extraction and selection techniques are not required.
4. The ECG signals of varying length are used.
5. Ten -fold cross-validation is employed to make the system repeatable.

The pitfalls of the proposed algorithm are as follows:

1. Requires the first R peak of the segment to be detected.
2. It is assumed that each ECG segment will have one type of arrhythmia.
3. Algorithm is not robust in detecting the APB from normal ECG segment.
4. Model is computationally intensive and learning is slow.
5. Limited availability of variable length data.
6. The system is developed using imbalance dataset.

Our proposed system is able to classify the ECG signal of varying lengths assuming that the test segment consists of only a single type of arrhythmia. This is not always true, as the ECG signal in the real-world may contain more than one type of arrhythmia. In future, we can use auto encoder network on these ECG data for element wise analysis, by associating each pixel with a class label. This will enable us to parallelize the process of beat detection and also the classification process. High-end graphics cards can also be used to accelerate the training process of the model. Also, instead of using weighted loss for training, data augmentation can be utilized as an up sampling technique to boost the variability of training data and counter the class imbalance problem [54]. Other cardiac diseases like coronary artery disease and ischemic heart disease can be explored in the future.

6. Conclusion

Early diagnosis of cardiac abnormalities is important to reduce the risk of cardio vascular events and morbidity rate. An automated accurate diagnosis system can help the clinicians in early diagnosis and also patients can get early treatment. Our developed automated system is effective detecting the heart conditions, the proposed system used a combination of state-of art techniques (CNN and LSTM) in the detection of arrhythmias using variable length of ECG beats. To the best of our knowledge this is the first paper to use the variable length heart beats for the arrhythmia detection. We have obtained the high classification accuracy of 98.10% without noise elimination. In future, we intend to use this model with long and short duration heart beat signals.

7. References

1. Nations, U., *World population ageing 2017 Highlights.* New York: Department of Economic and Social Affairs, 2017.
2. Chow, G.V., J.E. Marine, and J.L. Fleg, *Epidemiology of Arrhythmias and Conduction Disorders in Older Adults.* Clinics in geriatric medicine, 2012. **28**(4): p. 539-553.
3. Mak, K., *The Normal Physiology of Aging*, in *Colorectal Cancer in the Elderly*, K.-Y. Tan, Editor. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 1-8.
4. Anversa, P., et al., *Myocyte cell loss and myocyte cellular hyperplasia in the hypertrophied aging rat heart.* Circulation Research, 1990. **67**(4): p. 871-885.
5. Schneider, J.F., et al., *Newly acquired left bundle-branch block: The framingham study.* Annals of Internal Medicine, 1979. **90**(3): p. 303-310.
6. Fahy, G.J., et al., *Natural history of isolated bundle branch block.* The American Journal of Cardiology, 1996. **77**(14): p. 1185-1190.
7. Thrainsdottir, I.S., et al., *The epidemiology of right bundle branch block and its association with cardiovascular morbidity — The Reykjavik Study.* European Heart Journal, 1993. **14**(12): p. 1590-1596.
8. Binici, Z., et al., *Excessive Supraventricular Ectopic Activity and Increased Risk of Atrial Fibrillation and Stroke.* Circulation, 2010. **121**(17): p. 1904.
9. Engström, G., et al., *Cardiac Arrhythmias and Stroke.* Stroke, 2000. **31**(12): p. 2925.
10. Fleg, J.L. and H.L. Kennedy, *Cardiac Arrhythmias in a Healthy Elderly Population.* CHEST. **81**(3): p. 302-307.
11. Sahoo, S., et al., *Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities.* Measurement, 2017. **108**: p. 55-66.
12. Li, T. and M. Zhou, *ECG Classification Using Wavelet Packet Entropy and Random Forests.* Entropy, 2016. **18**(8).
13. Elhaj, F.A., et al., *Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals.* Computer Methods and Programs in Biomedicine, 2016. **127**: p. 52-63.
14. Li, H., et al., *Arrhythmia Classification Based on Multi-Domain Feature Extraction for an ECG Recognition System.* Sensors (Basel, Switzerland), 2016. **16**(10): p. 1744.
15. Martis, R.J., U.R. Acharya, and L.C. Min, *ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform.* Biomedical Signal Processing and Control, 2013. **8**(5): p. 437-448.
16. Martis, R.J., et al., *Cardiac decision making using higher order spectra.* Biomedical Signal Processing and Control, 2013. **8**(2): p. 193-203.
17. Martis, R.J., et al., *Application of principal component analysis to ECG signals for automated diagnosis of cardiac health.* Expert Systems with Applications, 2012. **39**(14): p. 11792-11800.
18. Karimifard, S. and A. Ahmadian, *A robust method for diagnosis of morphological arrhythmias based on Hermitian model of higher-order statistics.* BioMedical Engineering OnLine, 2011. **10**: p. 22-22.

19. Martis, R.J., et al. *Application of higher order cumulants to ECG signals for the cardiac health diagnosis*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011.
20. Yeh, Y.-C., W.-J. Wang, and C.W. Chiou, *Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals.* Measurement, 2009. **42**(5): p. 778-789.
21. Yu, S.-N. and K.-T. Chou, *Integration of independent component analysis and neural networks for ECG beat classification.* Expert Systems with Applications, 2008. **34**(4): p. 2841-2846.
22. Osowski, S., L.T. Hoai, and T. Markiewicz, *Support vector machine-based expert system for reliable heartbeat recognition.* IEEE Transactions on Biomedical Engineering, 2004. **51**(4): p. 582-589.
23. Lee, J.-G., et al., *Deep Learning in Medical Imaging: General Overview.* Korean Journal of Radiology, 2017. **18**(4): p. 570-584.
24. Akkus, Z., et al., *Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions.* Journal of Digital Imaging, 2017. **30**(4): p. 449-459.
25. Pham, T., et al., *Predicting healthcare trajectories from medical records: A deep learning approach.* Journal of Biomedical Informatics, 2017. **69**: p. 218-229.
26. Kahali, S., S.K. Adhikari, and J.K. Sing, *Convolution of 3D Gaussian surfaces for volumetric intensity inhomogeneity estimation and correction in 3D brain MR image data.* IET Computer Vision, 2018. **12**(3): p. 288-297.
27. Li, Q., Q. Peng, and C. Yan, *Multiple VLAD Encoding of CNNs for Image Classification.* Computing in Science & Engineering, 2018. **20**(2): p. 52-63.
28. Yang, W., et al. *Improved deep convolutional neural network for online handwritten Chinese character recognition using domain-specific knowledge*. in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. 2015.
29. Zhang, F., et al., *Image denoising method based on a deep convolution neural network.* IET Image Processing, 2018. **12**(4): p. 485-493.
30. Acharya, U.R., et al., *Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals.* Information Sciences, 2017. **415-416**: p. 190-198.
31. Tan, J.H., et al., *Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals.* Computers in Biology and Medicine, 2018. **94**: p. 19-26.
32. Zhang, J. and Y. Wu, *A New Method for Automatic Sleep Stage Classification.* IEEE Transactions on Biomedical Circuits and Systems, 2017. **11**(5): p. 1097-1110.
33. Zhang, Q., D. Zhou, and X. Zeng, *HeartID: A Multiresolution Convolutional Neural Network for ECG-Based Biometric Human Identification in Smart Health Applications.* IEEE Access, 2017. **5**: p. 11805-11816.
34. Qian, Y., et al., *Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2016. **24**(12): p. 2263-2276.
35. Wu, Y., et al., *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. 2016.

36. Kim, M., et al., *Speaker-Independent Silent Speech Recognition From Flesh-Point Articulatory Movements Using an LSTM Neural Network.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017. **25**(12): p. 2323-2336.
37. Song, E., F.K. Soong, and H.G. Kang, *Effective Spectral and Excitation Modeling Techniques for LSTM-RNN-Based Speech Synthesis Systems.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017. **25**(11): p. 2152-2161.
38. Sundermeyer, M., H. Ney, and R. Schlüter, *From Feedforward to Recurrent LSTM Neural Networks for Language Modeling.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015. **23**(3): p. 517-529.
39. Greff, K., et al., *LSTM: A Search Space Odyssey.* IEEE Transactions on Neural Networks and Learning Systems, 2017. **28**(10): p. 2222-2232.
40. Zhang, X.Y., et al., *End-to-End Online Writer Identification With Recurrent Neural Network.* IEEE Transactions on Human-Machine Systems, 2017. **47**(2): p. 285-292.
41. Goldberger, A.L., et al., *PhysioBank, PhysioToolkit, and PhysioNet.* Components of a New Research Resource for Complex Physiologic Signals, 2000. **101**(23): p. e215-e220.
42. LeCun, Y., et al., *Efficient BackProp*, in *Neural Networks: Tricks of the Trade, this book is an outgrowth of a 1996 NIPS workshop*. 1998, Springer-Verlag. p. 9-50.
43. Hochreiter, S., #252, and r. Schmidhuber, *Long Short-Term Memory.* Neural Comput., 1997. **9**(8): p. 1735-1780.
44. Srivastava, N., et al., *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. Vol. 15. 2014. 1929-1958.
45. Bengio, Y. and X. Glorot, *Understanding the difficulty of training deep feed forward neural networks*. 2010. 249-256.
46. Chollet, F., *Keras*. 2015.
47. Abadi, M., et al. *TensorFlow: A System for Large-Scale Machine Learning*. in *OSDI*. 2016.
48. Acharya, U.R., et al., *Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network.* Information Sciences, 2017. **405**: p. 81-90.
49. Acharya, U.R., et al., *A deep convolutional neural network model to classify heartbeats.* Computers in Biology and Medicine, 2017. **89**: p. 389-396.
50. Zubair, M., J. Kim, and C. Yoon. *An Automated ECG Beat Classification System Using Convolutional Neural Networks*. in *2016 6th International Conference on IT Convergence and Security (ICITCS)*. 2016.
51. Kiranyaz, S., T. Ince, and M. Gabbouj, *Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks.* IEEE Transactions on Biomedical Engineering, 2016. **63**(3): p. 664-675.
52. Yildirim, Ö., *A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification.* Computers in Biology and Medicine, 2018. **96**: p. 189-202.
53. Hawkins, D.M., *The Problem of Overfitting.* Journal of Chemical Information and Computer Sciences, 2004. **44**(1): p. 1-12.
54. C. Wong, S., et al., *Understanding Data Augmentation for Classification: When to Warp?* 2016. 1-6.