MDPI

Article

Detection of Atrial Fibrillation Using a Machine Learning Approach

Sidrah Liaqat [1], Kia Dashtipour [2], Adnan Zahid [2,3], Khaled Assaleh [4], Kamran Arshad [4] and Naeem Ramzan [1,*]

[1] School of Engineering and Computing, University of the West of Scotland, Glasgow G72 0LH, UK; sidra193@gmail.com
[2] James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK; kia.dashtipour@glasgow.ac.uk (K.D.); a.zahid.1@research.gla.ac.uk (A.Z.)
[3] School of Engineering and Physical Science, Heriot-Watt University, Edinburgh EH14 4AS, UK
[4] Faculty of Engineering and IT, Ajman University, Ajman 346, UAE; k.assaleh@ajman.ac.ae (K.A.); k.arshad@ajman.ac.ae (K.A.)
* Correspondence: Naeem.Ramzan@uws.ac.uk

check for updates

Abstract: The atrial fibrillation (AF) is one of the most well-known cardiac arrhythmias in clinical practice, with a prevalence of 1–2% in the community, which can increase the risk of stroke and myocardial infarction. The detection of AF electrocardiogram (ECG) can improve the early detection of diagnosis. In this paper, we have further developed a framework for processing the ECG signal in order to determine the AF episodes. We have implemented machine learning and deep learning algorithms to detect AF. Moreover, the experimental results show that better performance can be achieved with long short-term memory (LSTM) as compared to other algorithms. The initial experimental results illustrate that the deep learning algorithms, such as LSTM and convolutional neural network (CNN), achieved better performance (10%) as compared to machine learning classifiers, such as support vectors, logistic regression, etc. This preliminary work can help clinicians in AF detection with high accuracy and less probability of errors, which can ultimately result in reduction in fatality rate.

Keywords: atrial fibrillation; machine learning; cardiovascular; deep learning; healthcare

1. Introduction

Atrial fibrillation (AF) is the major concern with irregular heart rhythms, especially if a person approaches the age of 65. The heart is a pump used for efficient operation and it is regulated by an internal pacemaker that controls the heartbeat, the electrical impulse usually starts from the sinoatrial node [1–3] and moves from atria to ventricles, which can cause regular rhythmic contractions of the chambers. AF risk remains unacknowledged in some patients, because of its unawareness, and other patients may be mindful of the erratic behaviour of heartbeat and may be uncomfortable with sensation. The irregular heartbeat indicates the symptoms of any accidental strokes that may result in further prolonged illness and, as a result, it leads to ultimate heart failure [4–8]. The AF detection methods are mainly focused on the RR intervals, short term study of heart rate variability, and sequential examination to check the existence of P-wave [9]. The current studies are mostly focused on the feature extraction techniques; therefore, the features have significant fictional effect on the final outcome of the models. The major improvements in electrocardiogram (ECG) data and the advancement of Dynamic Neural Network (DNNs) and, in particular, the implementation of algorithms such as long short term memory (LSTM) and convolutional neural network (CNN). These algorithms can be

directly trained on the large-scaled dataset that RR Intervals of the ECG signal requirements and the performance of the ECG models is dramatically improved [10]. It is worth mentioning that, in the past decade, deep learning (DL) classifiers have been successfully applied in different fields, such as speech recognition, image classification [11], and many other domains such as natural language processing [12]. However, DL algorithms have not been widely applied for AF detection. However, there are few studies applied aforementioned classifier for AF prediction, but the comparison results show that the performance was unsatisfactory when compared to the image classification and image recognition [13]. It is to be noted that there is not any study available to detect the AF using DL classifiers.

In order to overcome the current issues, we build six models that are based on the feature-based approaches and DL approaches including support vector machine (SVM), multilayer perceptron (MLP), CNN and LSTM. The feature-based model is trained based on the manually extracted features, and DL methods are trained on raw data without any feature engineering. It is envisioned that this is the only exploiting automated feature engineering that focuses on DL for the automatic AF detection and classification. Additionally, the DL classifiers are compared with shallow learning classifiers.

In summary, the paper reports three major contributions that are outlined below:

- We developed a novel deep learning architecture for convolutional neural network (CNN) and long short-term memory (LSTM) to automatically detect AF. In addition, in depth comparison has been done with state-of-the-art approaches as well as baseline models, such as ResNet and Convolutional LSTM.
- Comparative analysis of the proposed approach with two widely online benchmark datasets.
- It is to be noted that, unlike the traditional machine learning algorithms, the deep learning methods have integrated feature extraction into the model, thus the handcrafted features are not needed. In addition, these methods can mine well different types of data sources and have good generalization ability, allowing for the computer to automatically learn and extract related features for any given issues. We developed an end-to-end approach that is based on deep learning approaches, which does not require feature selection and feature extraction technique.
- Additionally, we developed novel framework that can detect AF based on raw ECG signals than instead of other ECG features.

This paper structured, as follows: in Section 2, related work on the machine learning (ML) and DL classifiers for AF detection are presented. In Section 3, our novel approach for automated AF detection using ML and DL are described. In Section 4, the experimental results are discussed. Section 5 provides the discussion of the experimental results. Lastly, a conclusion is drawn out in Section 6 with future recommendations.

2. Related Work

In this section, the current state-of-the-art approaches for Atrial Fibrillation automatic detection are discussed. There are different ML and DL methods that are used to detect AF; however, the data should be transformed into an acceptable representation that enables classifiers to recognize the most suitable classes. Therefore, the DL methods break these rules and outperform the state-of-art efficiency method in many fields, such as text classification, image classification, etc. [14–16]. The DL methods that basically extract and classify the abstract features of the raw data [17–19].

2.1. ML Methods

Bruser et al. [20] proposed an approach for employing bed-mounted sensors for the detection of AF from the cardiac vibration signals. The approach remotely monitors patients, as well as various ML classifiers are used in order to evaluate the performance of the approach. Moreover, Xiong et al. [21] presented an approach based on k-means to identify the AF. The PubMed dataset is used in order to evaluate the performance of the approach, the empirical results show that the maximum entropy outperforms other algorithms. Hurnanen et al. [22] introduced a method based on

the linear classification of the spectral entropy and SCG signals. The performance of the approach is tested on the thirteen patients. The technique does not require the identification of heartbreak peaks from SCG data, as it works well on low-quality SCG signals.

Asgari [23] presented a method that employs stationary wavelet transform and ML algorithms, such as SVM, to detect AF. The method does not require P and R peak detection. The comparison results with different algorithms demonstrated that the presented method obtains considerable performance. In addition, Andersen et al. [24] presented a method that is based on ML algorithms in order to predict AF by using long-term ECG recordings. The method used extracted features from the data. The experimental results reveal that the model achieved better performance when compared to state-of-the-art approaches on the same benchmark dataset and the method is more computationally efficient for processing all recording in less than one minute. Additionally, the procedure has been tested on healthy subjects with SNR in order to detect any false positive and negative predictions to clarify the robustness of the technique. The study of misclassified data shows that the classifiers produce plenty of false positives that impact the approach's output. The potential solution to this problem is to locate and manually delete the noise level in the ECG signal to maximise the method's efficiency, but this is computationally costly. Wu et al. [25] used different DL algorithms, such as CNN and LSTM, to improve the performance and, as neural networks, obtained better results without the feature extraction process. The hybrid approach for CNN and LSTM has been used in order to improve the performance of AF detection.

Nemati et al. [26] developed an approach that is based on ML algorithms in order to detect multiple diseases. The approach can detect AF based on the noise that is recorded from the wrist. In addition, the time of the movement and also ambulatory pulsatile is recorded. The experimental results were based on the 10-fold cross-validation and the proposed approach obtained a high performance. The presented approach was accurate in monitoring AF and this approach is the first of its kind algorithm to receive high performance in general population AF detection.

Recently, Aschbacher et al. [27] introduced a system that can employ a PPG signal in order to classify the AF without extensive feature engineering; therefore, the system required lots of pre-processing and used heart rate variability and even statistical approaches for achieving high accuracy. The AF detection tools that are based on artificial intelligence have great potential for minimising the risk of AF and it can deliver the early treatment to patients. However, the main limitation of the system is a lack of comparison with current state-of-the-art approaches. Wang et al. presented an [4] automated AF detection for the correlation among wavelet coefficient series in ECG signals. However, the method had performed well in detecting AF in clinical diagnosis and, in order to evaluate the performance of the feature construction strategy MIT-BIH, the AF database is used in order to explore the performance of the approach. There are a variety of wireless sensors are used to detect the AF; however, the short noisy ECG signal is still a big challenge. The DL classifiers are correctly identified AF from wireless ECG records without feature extraction [28]. Three different models are using artificial neural network (ANN), binary decision tree, and SVM on 10 ECG signals are used for comparison. The best classifier for Atrial Fibrillation was a binary decision tree, which split signal equal to 100 and the worst case is SVM while using one feature [29].

2.2. Feature-Based Methods

Sadr et al. [30] developed a method that was based on different features, including RR inter-beat intervals, time domain, frequency domain, and also distribution features, there are three different classifiers, such as linear classifier, SVM, and quadratic neural networks, are used to assess the efficiency of the method. Lim et al. [31] implemented a feature selection technique for AF of ECG morphological features and heart rate variability and several ML classifiers are used to understand the performance of the of the feature selection technique. However, there is lack of comparison results with current feature selection technique, such as information gain and mutual information. Lahdenoja et al. [32] introduced a system to collect data whlie using a mobile phone that was equipped with Google Android OS to

detect AF. The system can pre-process, extract features and classify the data. The applications obtained high performance; however, the application is not available online.

Xia et al. [33] proposed an approach for converting one dimensional ECG into two-dimensional form by short-time Fourier transform and wavelet transform to detect AF while using SVM. The experimental results on MIT-BIH, Massachusetts Institute of Technology (MIT)—Beth Israel Hospital (BIH) AFIB dataset displays the superior performance on the detection of AF. Cao et al. [34] proposed a data augmentation technique for combining of the ECG episodes to improve the diversity of samples. The two layered benchmark LSTM is used to train the method. The method was successfully detect AF in small and imbalanced dataset. However, the key issue is that they have used a large number of histogram data and threshold values that are difficult to detect AF.

2.3. Wearable Devices for AF Detection

There are different approaches proposed to detect AF using wearable devices, such as smart watches, For example, Dorr et al. [35] introduced an approach for detecting AF based on the smart watches with high accuracy. However, it is challenging to obtain high performance from the wrist. The device can improve the detection of AF in patients with real-time monitoring. However, the proposed device is required to test on the wider population to monitor AF. ML algorithms improve outcomes, especially when the diagnose are provided from large datasets or complicated patterns of data, such as in AF. However, in order to reduce the data, the wavelet transformation is used for the compression of an image and then compressed image is processed by SVM classifier [2]. From ECG signals, heart rates are calculated and R peaks are extracted while using R-R intervals. Three classifiers K-nearest neighbour and random forest classifiers were used [36]. Zhang et al. [37] developed a mobile app for monitoring AF with active measurement. The sensitivity and accuracy used for monitoring AF were used in order to evaluate the performance of the approach.

3. Methodology

This section describes our proposed methodology for detecing AF. Figure 1 depicts the proposed conceptual framework and its details are presented in subsequent sections.
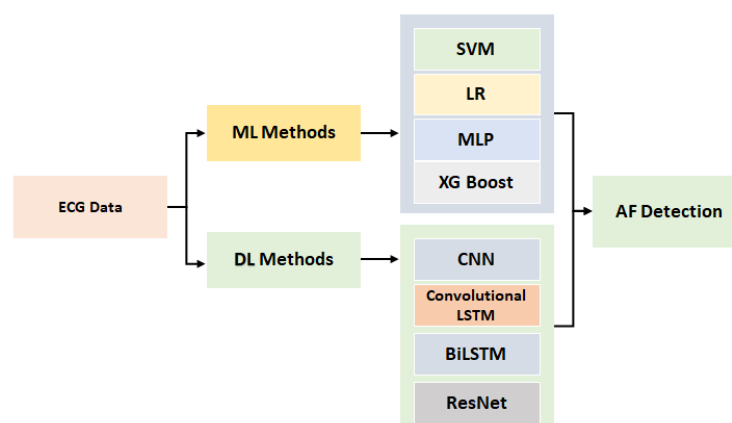

Figure 1. An overview of the proposed framework for atrial fibrillation (AF) Detection.

Pre-processing: there are two different pre-processing methods that were applied before feeding the data into the models, the normalization consists of removing noises from ECG signals and then zero padding were used:

$$X = \frac{X - \mu}{\sigma} \tag{1}$$

where x denotes the matrix that is composed by ECG recording with the shape of 8528, 18,000, μ and σ are the mean and standard deviation of the matrix. please confirm.

In order to train ML classifiers, it is necessary to use fast fourier transform (FFT) in order to convert signal from time domain to the recorded frequency domain and then proceed to perform the feature extraction, each of the heart rhythm can be converted into components rhythms. In order to extract ECG features, first, each ECG is divided into 10 s segments and four QRS detectors have been applied, gqrs, Pan-Tompkins, maxima search, and match filtering. The voting system is applied to mix the results of QRS detector, based on the ECG records and QRS labels. There are 83 various features extracted, such as R-R intervals, single-wave features, and full-wave features. The features, such as age, gender, and information on commodities, has been used. In order to train the ML models the principal component analysis (PCA) is used to reduce the dimensionally of the data. However, in order to train the DL methods, the raw data to train the model.

3.1. ML Models

After feature extraction, there are different machine learning classifiers, including SVM, logistic regression, MLP, XGBoost, CNN, and LSTM, has been applied in order to evaluate the performance of the approach.

Support Vector Machine (SVM): the SVM is one of the most well-known machine learning algorithm that is used to determine the hyperplane for the n-dimensional space of data, the SVM can be applied to non-linear classifier problem by using different kernel functions [38]. The main idea behind the SVM is how to divide the space with a decision boundary between data points. The w presents the vector per perpendicular to a median of the decision boundary, u is unknown vectors, and b is a constraint. Equation (2) and (3) show the positive and negative equations, respectively:

$$w \cdot u + b > 0 \tag{2}$$

$$w \cdot u + b < 0 \tag{3}$$

Multilayer Perceptron (MLP): the MLP consists of interconnected processing elements called neurons, the neurons in MLP is connected to inputs with different weights. The output of a neuron is the summation of all the connected inputs, and then the non-linear processing unit is used which is called a transfer function. The aim of MLP is to transform the inputs into an understandable output by learning the relation between input and the output and offer a solution to unseen problems [39–41]. Figure 2 displays the architecture for MLP.
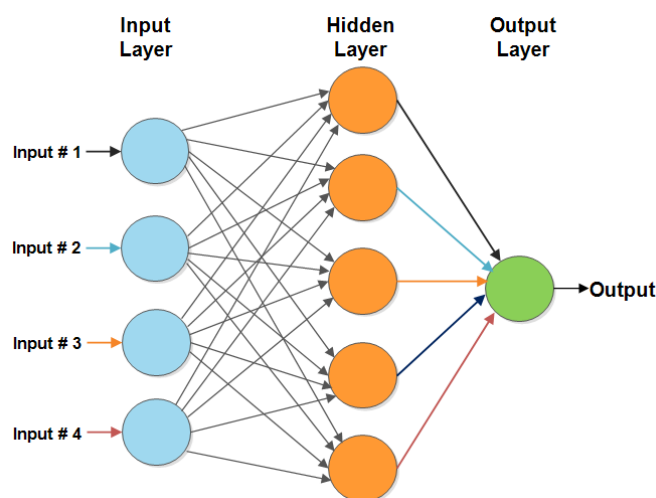

Figure 2. Multilayer Perceptron.

In our experiment, the features are extracted from the fully connected layer in order to train MLP classifier to identify the AF automatically. The experimental results show that the MLP classifier contains one hidden layer. The best performance of the MLP classifier is with a learning rate of 0.09. The number of epochs that is used to train MLP classifier is 100.

Logistic Regression: the logistic regression is also called logit regression which is a mathematical method used for a statistic to predict the probability of the event that is being processed in the previous data. The logistic regression can work on binary and multi-class data where the event can be happening, or the event is not happening [42–44].

XGBoost: XGBoost has been widely used and it is well-known classifier for the detection of AF, as it has shown superior performance on a large dataset. It is highly flexible and most of the regression, classification and ranking problems. In order to use XGBoost, we used skicit-learn package to train the classifier.

3.2. DL Models

Convolutional Neural Networks (CNN) architecture: CNN is a type of classifier that can be applied for different tasks, such as analysing images, text, and many other fields. However, CNN is also called shift invariant that can share the weights and also other characteristics. The CNN can be applied to sentiment analysis, recommendation system, image classification, healthcare, etc. [43,45–47]. CNN is modern version of multi-layer perceptron that means fully connected layers that each neuron is connected to all other neurons. The regularization technique is added to some characteristics to lose function. The CNN usually used a novel method towards regularization, the main benefits of these types of methods are more complex methods to use a small and simple pattern in order to identify the prediction. CNN has been inspired by biological processes with a connectivity pattern between the neuron and visual cortex. However, in deep learning, there is not any requirement for extraction of the features, sometimes in machine learning algorithms, the extracted features can reduce the performance of the approach, because the appropriate features are not selected by the framework [48–51].

We present the novel architectural details of the proposed of the CNN model used for the approach. Recently, the CNN classifier is one of the most well-known supervised learning techniques in the field of computer vision due to parameter sharing and sparse connectivity, which make the model computationally efficient [52,53]. The main benefit of using CNN classifier for this problem is to automatically detect important features while using any human intervention. For instance, given lots of data into CNN, it can learn distinctive features for each class itself. In addition, the CNN is computationally efficient. It is to be noted that most of the DL classifiers do not require any feature engineering or feature selection, which can cause saving time during the pre-processing and feature engineering stage of developing a framework. Table 1 displays the structure of the proposed CNN model. The Co-presented convolutional layer, the Max shows Maxpooling layer, Global shows Global Max Pooling layer, F shows Fully connected layer, and ReLU is Rectified Linear Unit Activation

Table 1. CNN Architecture, Co [a], the Max [b], Global [c], F [d], ReLU [e].

Layer	1	2	3	4	5	6	7	8	9	10
Type	Co	Max	Co		Co	Max	Co	Global	F	F
Filters	16		32		64		128			
Kernal Size	3	2	3	2	3	2	3			
Neurons									128	2
Activation	ReLU		ReLU		ReLU		ReLU		ReLU	SoftMax

[a] Convolutional layer; [b] Maxpooling layer; [c] GlobalMax Pooling layer; [d] Fully connected layer; [e] Rectified Linear Unit Activation.

Long Short-Term Memory (LSTM): the LSTM architecture is made of input and two different LSTM layers and, in addition to that, it has one fully connected layer. The LSTM that is made up of two layers consists of 128 and 64 cells followed by dropout of layers with 0.2 probability and one dense layer with two different neurons. The main proposed architecture was trained based on Adam optimizer. The LSTM function is defined, as follows [54]:

$$i_t = \sigma \left(W_i \cdot [h_{t-1}, x_t] + b_i \right) \tag{4}$$

$$f_t = \sigma \left(W_f \cdot [h_{t-1}, x_t] + b_f \right) \tag{5}$$

$$l_t = tanh \left(W_l \cdot [h_{t-1}, x_t] + b_l \right) \tag{6}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot l_t \tag{8}$$

$$h_t = o_t \odot tanh(c_t) \tag{9}$$

σ is denoted the sigmoid function. *tanh* is a tangent function that provides an output of $[-1, 1]$, $\odot$ is component-wise multiplication. The old memory is controlled and discarded by f_t and i_t and they are used to control information that is stored in the new memory.

ResNet: The Residual Networks is classic neural network used as classification for many computer vision tasks. The model was winner of ImageNet challenge in 2015. The fundamental with ResNet was allowed to train the deep neural networks with two-layer ResNet block, two generalized residual blocks (ResNet Init), two-layer ResNet block from two generalized residual blocks (grayed out connections are 0), and two-layer RiR block.

4. Experimental Results

In this section, we explain the experimental setup, followed by the results and discussions. In order to evaluate the performance of the proposed approach, there are different evaluation metrics, including accuracy, precision, recall, and f-measure are used [54].

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F\_measure = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{12}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

where *TP* denotes true positive, *TN* presents true negative, *FP* is false positive, and *FN* represents false negative respectively.

PhysioNet Dataset: In order to understand the performance of our proposed approach, we used the open labeled dataset of 2017 Physionet challenge dataset which consists of 8528 single channel ECG waveforms that were donated by the AliveCor. The waveform was recorded on the average of 30 s with a short waveform of 9 s and a long waveform of 61 s. The records were manually labeled into four different classes, including normal rhythm, AF rhythm, other rhythms, and noisy recording. The dataset consists of 8528 ECG recordings. The dataset consists of duration for each recording from 9 s to 60 s and it has four different class, such as normal, AF, noisy, and other. The sampling frequency of the record of 300 Hz.

MIT-BIH Atrial Fibrillation Dataset: the dataset consists of 23 ECG recording of human subjects with AF arrhythmia. Each of the ECG recording signals are sampled with 250 Hz with 12-bit resolution

over range of ±10 millivolts. In this study, we used four segments and labelled each of them based on the threshold parameter.

In order to evaluate the performance of the approach, the dataset is divided into 70% of training, 20% of testing and 10% of validation. The generalisation model has been evaluated while using the performance of the model on 48 short-term recording. The dataset consists of four different subjects that are more likely to learn the feature. Table 2 shows the summary of parameters to train ML and DL models.

Table 2. Summary of Parameters.

Algorithms	Parameters
MLP	Max iteration = 300
SVM	Kernel Linear
CNN, ResNet	Adam Optimizer, 10 Layer
LSTM, Convolutional LSTM	2-LSTM Layer, 0.2 probability
XGBoost	Kernel Linear
Logistic Regression	Random-state = 0

Table 3 shows the summary of ML and DL classifiers. The DL algorithms, such as CNN and LSTM, achieved better performance as compared to ML classifiers using PhysioNet, as shown in the Table. Table 4 show the summary of ML and DL classifiers. As shown in the Table the DL algorithms such as CNN and LSTM achieved better performance as compared to ML classifiers while using MIT-BIH Atrial Fibrillation. Table 5 shows the summary of results for ML and DL classifiers to identify the AF. The comparative experimental results show that the deep learning algorithms, such as LSTM and CNN, achieved better performance as compared to traditional ML classifiers. The experimental results show that CNN and LSTM achieved better performance when compared to traditional ML classifier. The proposed CNN architecture does not require any feature engineering as compared to machine learning classifier and it can generate the best performance as compared to machine learning algorithms such as MLP and logistic regression. We need to point out that the proposed CNN is more accurate to use the feature learning as compared to MLP and logistic regression, because the proposed approach can improve the AF detection by using DL classifiers. However, the CNN results show that the capability of learning features of identify of AF which can outperform in another ML algorithm. In order to evaluate the current results the Convolutional LSTM and ResNet have been added. However, the baseline ResNet and Convolutional LSTM achieved lower performance as compared to the proposed architecture for CNN and LSTM. However, the training time for Convolutional LSTM is lower than the training time for CNN and LSTM. Tables 3–6 has been updated. In addition, the ResNet took a long time to train the model. It is to be noted that the major drawback for ResNet is that is takes a long time to train the model. In addition, the convolutional LSTM required lots of resources and trained data to be trained on the AF dataset; therefore, it obtained lower performance as compared to proposed architecture for CNN and LSTM.

Table 3. Summary of machine learning (ML) and deep learning (DL) classifiers—PhysioNet.

Models	F1-Score						
	Training Accuracy	Testing Accuracy	N	A	O	~	Average
SVM	0.9652	0.737	0.836	0.700	0.527	0.333	0.722
MLP	0.9596	0.664	0.764	0.643	0.518	0.373	0.673
Logistic Regression	0.9686	0.712	0.756	0.691	0.521	0.356	0.701
XGBoost	0989	0.764	0.852	0.688	0.608	0.585	0.765
CNN	0.9905	0.865	0.90	0.865	0.809	0.675	0.860
LSTM	0.9949	0.875	0.921	0.869	0.812	0.681	0.863
Convolutional LSTM	0.8652	0.811	0.78	0.75	0.71	0.70	0.81
ResNet	0.8352	0.792	0.77	0.76	0.75	0.72	0.75

Table 4. Summary of ML and DL classifiers MIT-BIH Atrial Fibrillation.

Models	F1-Score						
	Training Accuracy	Testing Accuracy	N	B	O	~	Average
SVM	0.923	0.712	0.786	0.697	0.56	0.486	0.78
MLP	0.915	0.657	0.709	0.65	0.534	0.431	0.699
Logistic Regression	0.92	0.708	0.741	0.684	0.54	0.472	0.70
XGBoost	09525	0.682	0.711	0.677	0.56	0.481	0.74
CNN	0.982	0.812	0.835	0.781	0.733	0.718	0.826
LSTM	0.984	0.829	0.84	0.793	0.762	0.751	0.788
Convolutional LSTM	0.972	0.801	0.80	0.79	0.79	0.772	0.78
ResNet	0.953	0.784	0.77	0.76	0.74	0.72	0.74

Table 5 presents the summary of results for ML and deep learning classifiers in order to identify the AF using MIT-BIH Atrial Fibrillation. The initial experimental results demonstrate that the deep learning algorithms, such as LSTM and CNN, achieved better performance when compared to traditional machine learning classifiers.

Table 5. Summary of Results—PhysioNet.

	Accuracy	Precision	Recall	F1-Score	Time
SVM	0.737	0.72	0.71	0.722	2 min 24 s
MLP	0.664	0.67	0.66	0.673	1 min 40 s
Logistic Regression	0.712	0.70	0.69	0.701	1 min 38 s
XGBoost	0.764	0.76	0.75	0.765	2 min 3 s
CNN	0.865	0.86	0.85	0.860	5 min 32 s
LSTM	0.875	0.86	0.85	0.86	6 min 28 s
Convolutional LSTM	0.811	0.81	0.80	0.81	5 min 02 s
ResNet	0.792	0.78	0.78	0.79	13 min 21 s

The proposed DL approaches, including CNN and LSTM, achieved better performance as compared with traditional machine learning classifiers, such as SVM, MLP, Logistic regression, and XGBoost, as shown in Table 6. The overall accuracy of LSTM is 87.5%. The main reason why deep learning model achieved better performance as compared to a machine learning model is the feature extraction ability of DL classifiers. However, the main problems of deep learning classifiers is the computation speed. The TensorFlow Python package is used in order to train the model. The models trained used two hundred epochs with back propagation, the Adam optimizer is used for minimising the categorical cross-entropy loss function. In order to find the best accuracy for CNN, the different layers of CNN have been used in order to find out the best accuracy for CNN. The 2017 Physionet challenge dataset has been evaluated with a different number of convolutional layers, as shown in

Table 7. The initial experimental results show that, as the number of convolutional layers are increased, the performance is improved and, also, once the number of convolutional layers increased to five, the performance is decreased gradually. Therefore, we have chosen four convolutional layers, as the performance is improved. It is worth to mention that, as number of convolutional layers increased, the model takes more time to train. According to the results shown in Tables 7 and 8, the precision, recall and f-measure for layer 4 is better as compared to other layers. In addition, the cohen's kappa statistical values were above 0.8, where it shows that our proposed system is the most perfect approach in detecting AF.

Table 6. Summary of Results—MIT-BIH Atrial Fibrillation.

	Accuracy	Precision	Recall	F1-Score	Time
SVM	0.712	0.78	0.77	0.78	2 min 2 s
MLP	0.657	0.69	0.68	0.699	1 min 37 s
Logistic Regression	0.708	0.70	0.70	0.70	1 min 12 s
XGBoost	0.682	0.74	0.73	0.74	2 min 28 s
CNN	0.812	0.82	0.81	0.826	5 min 2 s
LSTM	0.829	0.78	0.77	0.788	5 min 31 s
Convolutional LSTM	0.8012	0.80	0.79	0.80	4 min 32 s
ResNet	0.784	0.78	0.77	0.78	14 min 18 s

Table 7. CNN-Results-Physionet.

Layer	Accuracy	Precision	Recall	F1-Score
Layer 1	79.29	0.79	0.78	0.79
Layer 2	81.54	0.81	0.80	080
Layer 3	82.67	0.82	0.81	0.82
Layer 4	86.5	0.86	0.85	0.86
Layer 5	83.91	0.83	0.82	0.83

Table 8. CNN-Results- MIT-BIH Atrial Fibrillation.

Layer	Accuracy	Precision	Recall	F1-Score
Layer 1	76.18	0.76	0.75	0.76
Layer 2	78.21	0.78	0.77	0.78
Layer 3	80.04	0.80	0.78	0.78
Layer 4	81.2	0.82	0.81	0.82
Layer 5	79.31	0.79	0.78	0.79

There are different layers of LSTM that have been used to find the best model architecture to detect the AF in order to find the best accuracy of LSTM. As shown in Table 9, the 2017 Physionet challenge dataset has been evaluated with a different number of LSTM layers. The experimental results reveal that an increased number of LSTM layer increases the time to train the model and, as the number of layers are decreased, the model is faster to train. According to results, as the number of layer increase from one to two, the performance of the model is rapidly improved. However, as the number of the layers increased from two to three layer the accuracy decreased gradually. The two-layer LSTM achieved better performance in terms of other evaluation metrics, such as precision, recall, and f-measure. In addition, Table 10 evaluated a different number of LSTM layers with MIT-BIH Atrial Fibrillation dataset. As shown, the LSTM algorithm with two-layer achieved better performance when compared to other layers.

Table 9. Long Short-Term Memory (LSTM) Results—-Physionet.

Layer	Accuracy	Precision	Recall	F1-Score
Layer 1	75.89	0.75	0.74	0.75
Layer 2	87.5	0.86	0.85	0.86
Layer 3	83.96	0.83	0.82	0.83

Table 10. LSTM Results—MIT-BIH Atrial Fibrillation.

Layer	Accuracy	Precision	Recall	F1-Score
Layer 1	75.89	0.75	0.74	0.75
Layer 2	87.5	0.86	0.85	0.86
Layer 3	83.96	0.83	0.82	0.83

Comparison with State-of-the-Art Approaches

Table 11 demonstrates the comparison of our proposed approach with other state-of-the-art approaches. As demonstrated, our proposed approach achieved better performance in terms of accuracy, precision, recall, and f-measure.

Table 11. Comparison with state-of-the-art Approaches.

Ref.	Accuracy	Precision	Recall	F1-Score
Rubin et al. [55]	80	0.80	0.79	0.80
Warrick et al. [56]	77.16	0.77	0.76	0.76
Our Proposed Approach	86.5	0.86	0.85	0.86

5. Discussion

The performance achieved on PhysioNet and MIT-BIH Atrial Fibrillation datasets on deep learning classifiers, such as CNN and LSTM, beats the traditional machine learning classifiers, such as MLP and logistic regression. The performance achieved on these datasets indicates that our proposed unique data-driven DL models outperform traditional feature extraction techniques. It is to be noted that the proposed model has been trained on the NVIDIA GeForce 940M GPU with 384 Cuda cores and 2 GB DDR3.

The major strength of deep learning algorithms that their ability to generalise well on the learning good representation of data. It is to be noted that the datasets used in our study consist od lots of noise and invariance to factors that may contribute to irregularity of heart rhythms, which can cause low performance in the proposed deep learning classifiers. However, in our method, the current invariance can be obtained while using combination of time frequency and attention to mechanism that can detect the AF with better performance. For example, change the position the high heart rate in the time-frequency. One of the main limitations with machine learning methods is lack of gold standard labels datasets. Therefore, in this study, there are two widely online benchmark datasets used, which are clinically annotated. It is to be noted that there are still lots of noise available in dataset. In summary, this paper provides motivation for researcher to utilize the deep learning models in order to detect the AF, because the preliminary experimental results demonstrate that the deep learning models achieved better performance when compared to traditional machine learning classifiers. To the best of our knowledge, this is the first study to consider the AF prediction. Our future work will be more focused on the ensemble classification of machine learning and deep learning classifiers to achieve better performance in terms of detecting AF.

6. Conclusions

The abnormal activity of the heart can cause cardiac arrhythmias, which can cause atrial fibrillation (AF). The AF is one of the most common types of heart failure, which can lead to a high mortality rate. More than 70% of AF occurred without the notice of patients and around eight million people in the world are suffering from this disease. Therefore, it is required to develop a novel technique for accurately and proactively detecting AF. In this study, we presented a novel framework to detect AF while using both machine and deep learning techniques. Most of the current approaches required feature selection; however, in this work, we proposed a deep learning algorithm, which does not require any feature engineering. The experimental results demonstrated that the deep learning approaches achieved better performance as compared to traditional shallow learning classifiers, such as SVM. As future work, we are planning to develop a real-time approach in order to detect AF without the requirement of any labelled data.

Author Contributions: Methodology, S.L.; Software, S.L. and N.R.; Formal Analysis, S.L.; Resources, A.Z., K.A. (Kamran Arshad) and N.R.; Data Curation, K.D.; Supervision, N.R.; Funding Acquisition, K.A. (Khaled Assaleh). All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shen, M.; Zhang, L.; Luo, X.; Xu, J. Atrial Fibrillation Detection Algorithm Based on Manual Extraction Features and Automatic Extraction Features. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Hulun Buir, China, 28–30 August 2020.
2. Lown, M.; Brown, M.; Brown, C.; Yue, A.M.; Shah, B.N.; Corbett, S.J.; Lewith, G.; Stuart, B.; Moore, M.; Little, P. Machine learning detection of Atrial Fibrillation using wearable technology. *PLoS ONE* **2020**, *15*, e0227401. [CrossRef] [PubMed]
3. Zolotarev, A.M.; Hansen, B.J.; Ivanova, E.A.; Helfrich, K.M.; Li, N.; Janssen, P.M.; Mohler, P.J.; Mokadam, N.A.; Whitson, B.A.; Fedorov, M.V.; et al. Optical Mapping-Validated Machine Learning Improves Atrial Fibrillation Driver Detection by Multi-Electrode Mapping. *Circ. Arrhythmia Electrophysiol.* **2020**, *13*, e008249. [CrossRef] [PubMed]
4. Wang, J.; Wang, P.; Wang, S. Automated detection of atrial fibrillation in ECG signals based on wavelet packet transform and correlation function of random process. *Biomed. Signal Process. Control.* **2020**, *55*, 101662. [CrossRef]
5. Shankar, M.G.; Babu, C.G. An exploration of ECG signal feature selection and classification using machine learning techniques. *Int. J. Innov. Technol. Explor. Eng. Regul* **2020**, *9*, 797–804.
6. Shi, H.; Wang, H.; Qin, C.; Zhao, L.; Liu, C. An incremental learning system for atrial fibrillation detection based on transfer learning and active learning. *Comput. Methods Programs Biomed.* **2020**, *187*, 105219. [CrossRef]
7. Parvaneh, S.; Rubin, J.; Babaeizadeh, S.; Xu-Wilson, M. Cardiac arrhythmia detection using deep learning: A review. *J. Electrocardiol.* **2019**, *57*, S70–S74. [CrossRef]
8. Jiang, F.; Dashtipour, K.; Hussain, A. A Survey on Deep Learning for the Routing Layer of Computer Network. In Proceedings of the 2019 UK/China Emerging Technologies (UCET), Glasgow, UK, 21–22 August 2019; pp. 1–4.
9. Gotlibovych, I.; Crawford, S.; Goyal, D.; Liu, J.; Kerem, Y.; Benaron, D.; Yilmaz, D.; Marcus, G.; Li, Y. End-to-end deep learning from raw sensor data: Atrial fibrillation detection using wearables. *arXiv* **2018**, arXiv:1807.10707.
10. Wu, J.; Roy, J.; Stewart, W.F. Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches. *Med. Care* **2010**, *48*, S106–S113. [CrossRef]
11. Chan, T.H.; Jia, K.; Gao, S.; Lu, J.; Zeng, Z.; Ma, Y. PCANet: A simple deep learning baseline for image classification? *IEEE Trans. Image Process.* **2015**, *24*, 5017–5032. [CrossRef]

12. Dashtipour, K.; Gogate, M.; Li, J.; Jiang, F.; Kong, B.; Hussain, A. A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks. *Neurocomputing* **2020**, *380*, 1–10. [CrossRef]

13. Liu, S.H.; Liu, H.C.; Chen, W.; Tan, T.H. Evaluating Quality of Photoplethymographic Signal on Wearable Forehead Pulse Oximeter With Supervised Classification Approaches. *IEEE Access* **2020**, *8*, 185121–185135, [CrossRef]

14. Kwon, S.; Hong, J.; Choi, E.K.; Lee, B.; Baik, C.; Lee, E.; Jeong, E.R.; Koo, B.K.; Oh, S.; Yi, Y. Detection of Atrial Fibrillation Using a Ring-Type Wearable Device (CardioTracker) and Deep Learning Analysis of Photoplethysmography Signals: Prospective Observational Proof-of-Concept Study. *J. Med Internet Res.* **2020**, *22*, e16443. [CrossRef]

15. Gogate, M.; Dashtipour, K.; Bell, P.; Hussain, A. Deep Neural Network Driven Binaural Audio Visual Speech Separation. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–7.

16. Dashtipour, K.; Ieracitano, C.; Morabito, F.C.; Raza, A.; Hussain, A. An Ensemble Based Classification Approach for Persian Sentiment Analysis. In *Progresses in Artificial Intelligence and Neural Systems*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 207–215.

17. Grégoire, J.; Gilon, C.; Subramanian, N.; Bersini, H. Forecasting episodes of atrial fibrillation using a new machine learning algorithm. *Arch. Cardiovasc. Dis. Suppl.* **2020**, *12*, 103. [CrossRef]

18. Bundy, J.D.; Heckbert, S.R.; Chen, L.Y.; Lloyd-Jones, D.M.; Greenland, P. Evaluation of Risk Prediction Models of Atrial Fibrillation (from the Multi-Ethnic Study of Atherosclerosis [MESA]). *Am. J. Cardiol.* **2020**, *125*, 55–62. [CrossRef] [PubMed]

19. Wang, Q.C.; Wang, Z.Y. Big Data and Atrial Fibrillation: Current Understanding and New Opportunities. *J. Cardiovasc. Transl. Res.* **2020**, 1–9. [CrossRef] [PubMed]

20. Bruser, C.; Diesel, J.; Zink, M.D.; Winter, S.; Schauerte, P.; Leonhardt, S. Automatic detection of atrial fibrillation in cardiac vibration signals. *IEEE J. Biomed. Health Inf.* **2012**, *17*, 162–171. [CrossRef]

21. Xiong, Z.; Liu, T.; Tse, G.; Gong, M.; Gladding, P.A.; Smaill, B.H.; Stiles, M.K.; Gillis, A.M.; Zhao, J. A machine learning aided systematic review and meta-analysis of the relative risk of atrial fibrillation in patients with diabetes mellitus. *Front. Physiol.* **2018**, *9*, 835. [CrossRef]

22. Hurnanen, T.; Lehtonen, E.; Tadi, M.J.; Kuusela, T.; Kiviniemi, T.; Saraste, A.; Vasankari, T.; Airaksinen, J.; Koivisto, T.; Pänkäälä, M. Automated detection of atrial fibrillation based on time–frequency analysis of seismocardiograms. *IEEE J. Biomed. Health Inf.* **2016**, *21*, 1233–1241. [CrossRef]

23. Asgari, S.; Mehrnia, A.; Moussavi, M. Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. *Comput. Biol. Med.* **2015**, *60*, 132–142. [CrossRef]

24. Andersen, R.S.; Peimankar, A.; Puthusserypady, S. A deep learning approach for real-time detection of atrial fibrillation. *Expert Syst. Appl.* **2019**, *115*, 465–473. [CrossRef]

25. Wu, X.; Sui, Z.; Chu, C.H.; Huang, G. Detection of Atrial Fibrillation from Short ECG Signals Using a Hybrid Deep Learning Model. In *International Conference on Smart Health*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 269–282.

26. Nemati, S.; Ghassemi, M.M.; Ambai, V.; Isakadze, N.; Levantsevych, O.; Shah, A.; Clifford, G.D. Monitoring and detecting atrial fibrillation using wearable technology. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3394–3397.

27. Aschbacher, K.; Yilmaz, D.; Kerem, Y.; Crawford, S.; Benaron, D.; Liu, J.; Eaton, M.; Tison, G.H.; Olgin, J.E.; Li, Y.; et al. Atrial fibrillation detection from raw photoplethysmography waveforms: A deep learning application. *Heart Rhythm. O2* **2020**, *1*, 3–9. [CrossRef]

28. Gadaleta, M.; Rossi, M.; Steinhubl, S.R.; Quer, G. Deep learning to detect atrial fibrillation from short noisy ecg segments measured with wireless sensors. *Circulation* **2018**, *138*, A16177.

29. Sieciński, S.; Kostka, P.S.; Tkacz, E.J. Comparison of Atrial Fibrillation Detection Performance Using Decision Trees, SVM and Artificial Neural Network. In *International Conference on Information Technology & Systems*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 693–701.

30. Sadr, N.; Jayawardhana, M.; Pham, T.T.; Tang, R.; Balaei, A.T.; de Chazal, P. A low-complexity algorithm for detection of atrial fibrillation using an ECG. *Physiol. Meas.* **2018**, *39*, 064003. [CrossRef] [PubMed]

31.   Lim, H.W.; Hau, Y.W.; Lim, C.W.; Othman, M.A. Artificial intelligence classification methods of atrial fibrillation with implementation technology. *Comput. Assist. Surg.* **2016**, *21*, 154–161. [CrossRef]

32.   Lahdenoja, O.; Hurnanen, T.; Iftikhar, Z.; Nieminen, S.; Knuutila, T.; Saraste, A.; Kiviniemi, T.; Vasankari, T.; Airaksinen, J.; Pänkäälä, M.; et al.   Atrial fibrillation detection via accelerometer and gyroscope of a smartphone. *IEEE J. Biomed. Health Inf.* **2017**, *22*, 108–118. [CrossRef]

33.   Xia, Y.; Wulan, N.; Wang, K.; Zhang, H. Detecting atrial fibrillation by deep convolutional neural networks. *Comput. Biol. Med.* **2018**, *93*, 84–92. [CrossRef]

34.   Cao, P.; Li, X.; Mao, K.; Lu, F.; Ning, G.; Fang, L.; Pan, Q. A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation. *Biomed. Signal Process. Control* **2020**, *56*, 101675. [CrossRef]

35.   Dörr, M.; Nohturfft, V.; Brasier, N.; Bosshard, E.; Djurdjevic, A.; Gross, S.; Raichle, C.J.; Rhinisperger, M.; Stöckli, R.; Eckstein, J.   The WATCH AF trial: SmartWATCHes for detection of atrial fibrillation. *JACC Clin. Electrophysiol.* **2019**, *5*, 199–208. [CrossRef]

36.   Choudhury, S.; Bhowal, A. Comparative analysis of machine learning algorithms along with classifiers for network intrusion detection. In Proceedings of the 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Chennai, India, 6–8 May 2015; pp. 89–95.

37.   Zhang, H.; Zhang, J.; Li, H.B.; Chen, Y.X.; Yang, B.; Guo, Y.T.; Chen, Y.D. Validation of single centre pre-mobile atrial fibrillation apps for continuous monitoring of atrial fibrillation in a real-world setting: Pilot cohort study. *J. Med. Internet Res.* **2019**, *21*, e14909. [CrossRef]

38.   Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]

39.   Naraei, P.; Abhari, A.; Sadeghian, A. Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data.   In Proceedings of the 2016 Future Technologies Conference (FTC), San Francisco, CA, USA, 6–7 December 2016; pp. 848–852.

40.   Yang, X.; Shah, S.A.; Ren, A.; Zhao, N.; Zhang, Z.; Fan, D.; Zhao, J.; Wang, W.; Ur-Rehman, M. Freezing of gait detection considering leaky wave cable. *IEEE Trans. Antennas Propag.* **2018**, *67*, 554–561. [CrossRef]

41.   Cui, T.J.; Zoha, A.; Li, L.; Shah, S.A.; Alomainy, A.; Imran, M.A.; Abbasi, Q.H. Revolutionizing Future Healthcare using Wireless on the Walls (WoW). *arXiv* **2020**, arXiv:2006.06479.

42.   Dashtipour, K.; Raza, A.; Gelbukh, A.; Zhang, R.; Cambria, E.; Hussain, A. PerSent 2.0: Persian Sentiment Lexicon Enriched with Domain-Specific Words. In *International Conference on Brain Inspired Cognitive Systems*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 497–509.

43.   Jiang, F.; Kong, B.; Li, J.; Dashtipour, K.; Gogate, M. Robust Visual Saliency Optimization Based on Bidirectional Markov Chains. *Cogn. Comput.* **2020**. [CrossRef]

44.   Hussien, I.O.; Dashtipour, K.; Hussain, A. Comparison of sentiment analysis approaches using modern Arabic and Sudanese Dialect.   In *International Conference on Brain Inspired Cognitive Systems*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 615–624.

45.   Adeel, A.; Gogate, M.; Farooq, S.; Ieracitano, C.; Dashtipour, K.; Larijani, H.; Hussain, A. A survey on the role of wireless sensor networks and IoT in disaster management. In *Geological Disaster Monitoring Based on Sensor Networks*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 57–66.

46.   Dashtipour, K.; Gogate, M.; Adeel, A.; Hussain, A.; Alqarafi, A.; Durrani, T. A comparative study of persian sentiment analysis based on different feature combinations. In *International Conference in Communications, Signal Processing, and Systems*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 2288–2294.

47.   Liu, S.H.; Li, R.X.; Wang, J.J.; Chen, W.; Su, C.H. Classification of photoplethysmographic signal quality with deep convolution neural networks for accurate measurement of cardiac stroke volume. *Appl. Sci.* **2020**, *10*, 4612. [CrossRef]

48.   Gogate, M.; Dashtipour, K.; Adeel, A.; Hussain, A. Cochleanet: A robust language-independent audio-visual model for speech enhancement. *arXiv* **2020**, arXiv:1909.10407.

49.   Dashtipour, K.; Gogate, M.; Adeel, A.; Algarafi, A.; Howard, N.; Hussain, A.   Persian named entity recognition. In Proceedings of the 2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC), Oxford, UK, 26–28 July 2017; pp. 79–83.

50. Ozturk, M.; Gogate, M.; Onireti, O.; Adeel, A.; Hussain, A.; Imran, M.A. A novel deep learning driven, low-cost mobility prediction approach for 5G cellular networks: The case of the Control/Data Separation Architecture (CDSA). *Neurocomputing* **2019**, *358*, 479–489. [CrossRef]

51. Mammone, N.; Ieracitano, C.; Morabito, F.C. A deep CNN approach to decode motor preparation of upper limbs from time–frequency maps of EEG signals at source level. *Neural Netw.* **2020**, *124*, 357–372. [CrossRef]

52. Gogate, M.; Adeel, A.; Hussain, A. Deep learning driven multimodal fusion for automated deception detection. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017; pp. 1–6.

53. Gogate, M.; Adeel, A.; Marxer, R.; Barker, J.; Hussain, A. DNN driven speaker independent audio-visual mask estimation for speech separation. *arXiv* **2018**, arXiv:1808.00060.

54. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]

55. Rubin, J.; Parvaneh, S.; Rahman, A.; Conroy, B.; Babaeizadeh, S. Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings. *J. Electrocardiol.* **2018**, *51*, S18–S21. [CrossRef]

56. Warrick, P.A.; Homsi, M.N. Ensembling convolutional and long short-term memory networks for electrocardiogram arrhythmia detection. *Physiol. Meas.* **2018**, *39*, 114002. [CrossRef] [PubMed]