

Self-Organized Operational Neural Network: Rethinking Atrial Fibrillation Detection Using ECG and Wrist-PPG

(This project report has been submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Electrical and Electronic Engineering)



Submitted By

Exam Roll: 63943

Exam Roll: 63937

Reg No: 2017214832

Reg No: 2017514857

Session: 2017-18

Session: 2017-18

DEPT. OF ELECTRICAL AND ELECTRONIC ENGINEERING
UNIVERSITY OF DHAKA

June 27, 2022

Abstract

Atrial fibrillation (AF) is a malfunctioning cardiovascular disease that can result in an arterial stroke. Early identification is essential for primary and secondary stroke prevention since it is the major contributor to cardiac stroke. A long-term electrocardiogram (ECG)-based Holter monitor device is utilized to diagnose AF. However, it is inconvenient and may already exhibit damage in a pre-symptomatic phase. To detect and diagnose early AF, we introduced a deep learning paradigm in this research utilizing ECG and a photoplethesographic (PPG)-based wrist-wearable optical sensor. We have introduced several advanced signal processing techniques to preprocess the disruptive PPG signal, including noise filtering, motion artifact reduction, and baseline wander removal. For classification, we have implemented the Self-Organized Operational Neural Network (Self-ONN) and its several architectures (i.e., Maclaurin approximation, $q = 3, 5, 7$) as well as the 1D Convolutional Neural Network (1D CNN, i.e., Maclaurin approximation, $q = 1$). Using the ECG signal, we have achieved a promising classification metric with $98.49 \pm 0.36\%$ accuracy, $98.32 \pm 0.93\%$ sensitivity, and $98.14 \pm 0.96\%$ specificity, as well as the PPG dataset with $92.96 \pm 0.36\%$ accuracy, $92.23 \pm 3.16\%$ sensitivity, and $91.50 \pm 3.16\%$ specificity. Self-ONN with the Maclaurin approximation, $q = 3$, outperforms on both ECG and PPG signals.

CONTENTS

Abstract	i
List of Figures	iv
List of Tables	v
1 INTRODUCTION	1
1.1 Problem definition:	1
1.2 Approaches to AF detection:	3
1.2.1 ECG Based Approach:	4
1.2.1.1 Single-lead ECG:	4
1.2.1.2 Multi-lead ECG:	4
1.2.2 Challenge in ECG Based Approach:	5
1.2.3 PPG Based Approaches:	5
1.2.4 Challenge in PPG Based Approach:	6
1.3 AF Studies Case:	7
1.3.1 Statistical Analysis Perspectives	8
1.3.2 Machine Learning Perspectives	9
1.3.3 Deep Learning Perspectives	9
1.4 Project Overview:	10
2 RELATED WORKS	12
2.1 AF Detection Methodology	12
2.1.1 Machine Learning Based AF Detection	13
2.1.2 Deep Learning Based AF Detection	15
3 THEORETICAL OVERVIEW	20
3.1 Sensor Description:	20
3.2 Dataset Description:	22
3.3 Deep learning and Classification Models	25
3.4 Convolutional Neural Networks (CNN)	26
3.5 Self-Organized Operational Neural Networks (Self-ONN):	28
4 METHODOLOGY	34

4.1	Data Pre-processing:	34
4.1.1	Resolving Drifting:	35
4.1.2	Noise Filtering:	36
4.1.3	Data Synchronization:	38
4.1.4	Baseline Wander:	40
4.2	Proposed Method	42
4.3	Models	43
4.3.1	Model Training	44
5	RESULT AND ANALYSIS	46
5.1	Evaluation Metrics	46
5.2	AUC - ROC Curve	48
5.2.1	Confusion Matrix	49
5.2.2	Training and Loss Curves	51
5.2.3	ROC-AUC Curve	57
5.2.4	Results for AF detection	59
6	CONCLUSION AND FUTURE Works	62
	Bibliography	64

LIST OF FIGURES

1.1	Different type of ECG and PPG data acquisition devices	2
1.2	Different techniques used in literature for AF detection	8
2.1	AF Detection using Pulse Rate Variability	13
2.2	Block Diagram of the Proposed Method	14
2.3	Complete Overview of AF Detection using PPG Data	15
2.4	Proposed model of AF Detetction using [1]	17
2.5	Complete DL Framework of AF Diagnosis	18
2.6	Deep Learning 1D-CNN model on ECG data[2]	19
3.1	(a) Bittium Faros™ 180 ECG device, reference ECG data acquisition protocol; (b) Wrist-worn device with embedded green LED for PPG data acquisition. Both of them with embedded acceleration sensor.	21
3.3	DIstribution of AF and Normal Samples	25
3.4	CNN Model Used in Classification Problem [3]	27
3.5	Depiction of the k^{th} neuron of a ONN (Right) and an CNN (left) along with the 3 successive ONN (right) and CNN (left) layers	30
3.6	Depiction of the 1D nodal operations with the 1D kernels of the i^{th} neuron of CNN (left), ONN (middle), and Self-ONN (right).	32
4.1	Change in PPG frequency (100 Hz) due to drifting	36
4.2	Responses of the Filter	37
4.3	Noise Removal of PPG and ECG Data	38
4.4	Alignment of PPG, VPG, APG and ECG Signals.	40
4.5	Baseline Shifting Function on PPG data.	42
4.6	Overall Block Diagram of our Proposed Method.	43
4.7	Models of Self-ONN with Two Different Structures.	44
5.1	ROC Curve obtained by a classifier model	49
5.2	Confusion Matrix for Self-ONN Model 1 on ECG Data	50
5.3	Confusion Matrix for Self-ONN Model 1 on PPG Data	51
5.4	Loss curve of Self-ONN Model 1 ($q=3$) on ECG Data	53
5.5	Accuracy curve of Self-ONN Model 1 ($q=3$) on ECG Data	54
5.6	Loss curve of Self-ONN Model 1 ($q=3$) on PPG Data	56
5.7	Accuracy curve of Self-ONN Model 1 ($q=3$) on PPG Data	57
5.8	Self-ONN model 1 Performances on ECG Data	58
5.9	Self-ONN model 1 Performances on PPG Data	58

LIST OF TABLES

3.1	Detailed demographic and clinical information for each AF patient .	23
4.1	Model Hyper-Parameters Information	45
5.1	Parameters used for evaluation.	46
5.2	Results for AF Detection on ECG Data using Model 1	59
5.3	Results for AF Detection on PPG Data using Model 1	60
5.4	Results for AF Detection on ECG Data using Model 2	60
5.5	Results for AF Detection on PPG Data using Model 2	61

CHAPTER 1

INTRODUCTION

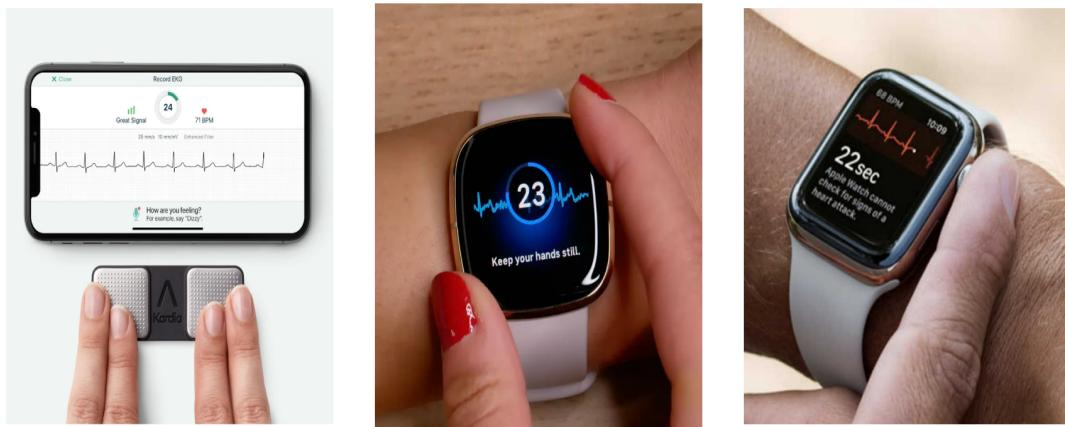
Atrial fibrillation (AF) is the most common cardiac arrhythmia worldwide. As per the 2017 Global Burden of Disease research [4], there were 37.57 million instances of AF globally, accounting for around 0.51% of the global population. This figure is constantly increasing. AF is distinguished by an irregular heartbeat caused by unregulated electrical impulses from the atria and throughout the heart. Long-term AF may cause inadequate blood perfusion, leading to fatigue, clots, strokes, and even death. The chance of developing AF is influenced by age and other genetically predisposed and lifestyle-related diseases and disorders, such as hypertension [5]. Additionally, several studies [3, 5, 6] indicate that exerting yourself can increase your risk of having an AF episode. Consequently, AF is classified as a disorder that worsens with time [7]. With time, AF episodes often become more frequent and more prolonged. Neither treatment nor a cure can entirely reverse AF, even though various drugs may assist minimize its symptoms. However, owing to the dynamic character of this field of research, rapid and precise AF diagnosis has received increased attention.

1.1 Problem definition:

Throughout the last several decades, various methods for diagnosing AF have been used and are still in use in medical settings. 12-lead electrocardiography (ECG), which was created based on the discoveries of Willem Einthoven, MD, PhD [8],

is one prominent example of this. The patient must be admitted to the hospital, which may be uncomfortable even though the 12-lead ECG produces accurate records of cardiac activity. Additionally, early AF, also known as paroxysmal AF, requires long-term surveillance because AF episodes are brief and infrequent at this stage of the illness. Ambulatory ECG equipment like the Holter monitor and event recorder has been created due to the immobility problem. The Holter monitor continuously records heart activity, while the event recorder only activates when the patient has symptoms. Even if such a device increases the patient's mobility, the patient may find it difficult to carry it.

Also used are electrodes, which provide more accurate recordings but may irritate the skin if worn continually. However, these devices are only advised if a patient has told medical professionals about any AF-specific symptoms they have had. It is more concerning, however, since AF could have started to show early signs of lasting damage. Naturally, even if no symptoms were present, the complexity and cost of such a device would discourage the general public from purchasing it. Due to this, AF cannot be detected early. As a result, a new strategy has to be created.



(a) AliveCor KardiaMobile. Courtesy of [8]. (b) Fitbit Sense. Courtesy of [11]. (c) Apple Watch Series 4. Courtesy of [12].

Figure 1.1: Different type of ECG and PPG data acquisition devices.

The focus has now shifted to developing practical, non-intrusive mobile devices to accomplish early AF detection. Artificial intelligence (AI) is also a part of this to help with automated detection. Wearable technology, or wearables, as these

devices are sometimes called, has developed fast over the last ten years to provide pro-active health monitoring. The AliveCor KardiaMobile is a portable device that can record single-lead ECGs and has received clinical validation [9]. The device may be paired with a bright phone to preserve the recordings, run an AI system, and email them to a doctor for a formal diagnosis. The Fitbit Sense [10], which may also be worn on the wrist, has comparable functions. Photoplethysmography (PPG) is used by the Apple Watch Series 4 [11] for AF detection and ECG recording. PPG is a light-based sensor for detecting heart rate and oxygen content in the blood. PPG is easier to include in a wearable than ECG, which is an advantage.

The issue with these devices is that it is not feasible to monitor heart activity continually. To record an ECG with these devices, the client must actively maintain physical contact with the electrodes, which is unreasonable over time. The Apple Watch only checks every 15 minutes, which is too long for early AF detection even if PPG is used to aid this problem.

1.2 Approaches to AF detection:

There are several methods for identifying and diagnosing AF. These methods may often be divided into two branches:

- By manually checking the pulse or by utilizing an AF automatic detection tool
- Integrated into case-finding initiatives or standard clinical practice

Several clinical tools, including Holter Monitoring, Electrocardiogram (ECG), and Photoplethysmogram, may record and track an AF episode by gathering clinical data (PPG). Several open-source datasets, ECG, and PPG are often used to identify AF. MIT-BIH DB [12], MIT-BIH AFDB [13], PhysioNet/CinC 2017 [71], MIT-BIH SRDB [14], MIT-BIH VFDB [12], and CU VTDB [15] are the most frequently used datasets.

1.2.1 ECG Based Approach:

The environment in which the ECG signal is being recorded or the equipment being utilized to record the data contributes to noise and artifacts in the signal. Several preprocessing techniques can denoise ECG signals, including Fourier cosine series operation to remove baseline wander and high-frequency components [16]. Noise reduction techniques include the use of a wavelet transform [4, 5, 9, 17] as well as an elliptical band-pass filter [18, 19]. For reducing motion artifacts, other filters including a notch filter [20], a band-pass Butterworth filter [21, 22, 23, 24], and a finite impulse response (FIR) filter [25] are also utilized. Besides, Z-score normalization [5, 17, 19] and a high-pass filter [26] are often employed for amplitude scaling and offset impact removal, respectively, to standardize the ECG signals for analysis.

The ECG data may be configured in various ways for use as input into DL models for AF detection, such as single or multiple lead(s), spectral analysis, or integrated features analysis.

1.2.1.1 Single-lead ECG:

Due to being computationally light and aiding model training, single-lead input is often utilized in published research [27, 28]. Lead II [29] accurately captures the P, QRS, and T waves, often seen in single-lead ECG input assessments. Another AF study has been attempted on the Physio Net/CinC 2017 dataset, consisting of modified Lead I ECG signals captured with medical-grade portable personal ECG monitoring equipment.

1.2.1.2 Multi-lead ECG:

Standard 12-lead ECG signal recordings are often available for studies that employ such input data [30]. Because the data dimensionality for 12-lead ECG signals is vast, some studies only utilized a portion of the 12-lead recordings to save processing costs. Due to removing four lead signals with minimal additional information, only eight leads (I, II, V1-6) were employed as inputs for the Attia et al. model.

Similarly, Baalman et al. [16] used Lead II of the 12-lead ECG dataset to train their model.

1.2.2 Challenge in ECG Based Approach:

Electrocardiography is insufficient for LVH screening on its own. Sensitivity ranges from 20% to 60% (usually less than 50%). LVH is only seen on the ECG in 3% of the general population and 5% of hypertension individuals. In individuals with hypertension, ECG criteria shouldn't be utilized to rule out LVH. As a result, given the limited sensitivity and variable specificity, the ECG for LVH identification in hypertension is not adequate.

Even when sex-specific criteria are utilized, ECG criteria perform worse in women than in men and have inadequate sensitivity for LVH at acceptable levels of specificity. In the general population, gender influences incidence rates and estimation results of ECG LVH criteria, with women having a higher predictive value than men. The combination of the Cornell voltage and the Sokolow-Lyon voltage seems to be a viable choice for clinical application. ECG alterations are a poor indicator of LVH (patients with clinically significant left ventricular hypertrophy seen on echocardiography may still have a relatively normal ECG). LVH is challenging to diagnose electrocardiographically in patients under the age of 40. Because young people often have large amplitude QRS complexes without left ventricular dysfunction, voltage criteria are insufficient for this population.

1.2.3 PPG Based Approaches:

Typically recorded at a peripheral location, a PPG waveform is produced throughout a heart cycle. As a result, it may be thought of as a pulse pressure waveform that begins with the movement of the heart and travels throughout the central vein. Numerous physiological factors might potentially be derived from the analysis of a PPG signal since brain, cardiac, and respiratory interactions regulate blood flow [31]. The PPG signal thus contains a wealth of information regarding physiological states [32].

AF appears as different pulse morphologies and pulse-to-pulse intervals in a PPG signal. On the other hand, a typical sinus rhythm can be identified by the regular spacing of PPG pulses and the comparable morphologies of successive pulses. In the presence of artifacts, it can occasionally be difficult to detect an irregularity in a PPG signal. It is possible to confuse artifacts with physiological irregularities. Data from accelerometry may be used to spot motion artifacts. Accelerometry sensors, which detect acceleration forces in several spatial directions, are a common feature of current wearable technology. PPGs contaminated with relics are often thrown away.

A typical AF detection method extracts features from the collected PPG signal (temporal, spectral, or architectural) and analyzes them to determine whether an AF rhythm is present. The temporal waveform has been represented as images in certain methodologies. After that, either traditional image processing techniques or ones based on artificial intelligence would be used to examine the resulting picture [33, 34, 35]. Since the tachogram (RR intervals) is a trustworthy indicator of heartbeats, major characteristics have historically been extracted from it [36, 37]. Beyond RR intervals, new characteristics were found. With promising outcomes, PPG time series and their picture representation (such as the entire signal plot, fast Fourier transform spectrum, or wavelet spectrogram) have been applied to detect physiological events [33, 38, 39].

1.2.4 Challenge in PPG Based Approach:

Poor signal quality and early beats are the two most frequent reasons why AF detection with PPG is inaccurate. PPG signals are disrupted by motion artifacts, particularly when users are moving their arms or walking or running.

The majority of PPG signal-based AF detection experiments take place in predetermined environments or with the patient in a predetermined postures [32, 40, 41].

As a result, we anticipated the relatively high percentage of false positives associated with ambulatory PPG signal-based AF detection. Individuals with paroxysmal AF had a decrease in specificity. The lower specificity in these people may be attributed to many premature beats and the limited number of participants

in the paroxysmal AF group. Longer assessment intervals and more participants, in our opinion, increase the algorithms' accuracy. The goal of AF detection with a smartwatch, on the other hand, is reducing strokes by early AF identification. The sensitivity is more crucial than the specificity when using a screening tool.

Some of the patients in this study had a history of atrial fibrillation (AF) or premature heartbeats, while others had symptoms of arrhythmia detection such as palpitations, vertigo, and hyper-tension. These individuals weren't in good health or subclinical AF sufferers. Participants with subclinical AF may have varied AF detection abilities. To assess the performance of participants with subclinical AF, more research must be done.

1.3 AF Studies Case:

Based on the methods used to construct an AF detector, studies were divided into three categories as follows:

- Statistical analysis perspectives
- Machine learning perspectives
- Deep learning perspectives

Using PPG signals to derive statistical parameters, classical statistical analysis may distinguish between AF and non-AF segments. For ML techniques, it is necessary to perform pre-selected feature extraction, a manual process that often gains from the inclusion of complicated physiological data. In order to develop an ML classifier, the attributes discovered from the training data samples are then applied. Deep learning (DL) approaches have attracted a lot of interest because to recent increases in processing power and the significant success of computer vision applications [77,78]. Unlike classical ML, which involves human feature engineering, DL incorporates an autonomous feature representation process with the incoming data. The AF detection analogy is shown in Figure 1.2

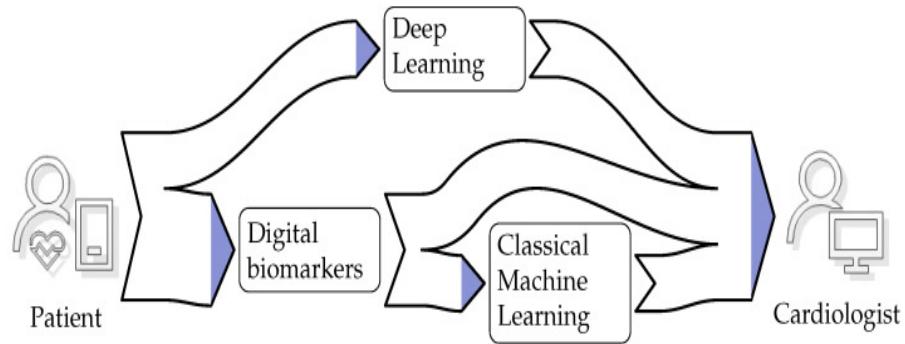


Figure 1.2: Different techniques used in literature for AF detection .

1.3.1 Statistical Analysis Perspectives

The RR-interval time series of well-annotated and openly available ECG datasets, such as the MIT-BIH atrial fibrillation, MIT-BIH normal sinus rhythm, or MIT-BIH arrhythmia database, are utilized to create statistical models for AF diagnosis [12, 13]. Most of the characteristics came from the RR-interval time series of annotated ECG waveforms. The histograms of each feature were looked at in both the presence and absence of AF and other cardiac rhythms to find the threshold that best separates the various rhythm classes. The same RR time series-based characteristics obtained from PPG signals were subjected to these criteria [12, 13]. Using a series of binary classifications, researchers may detect other arrhythmias such as premature atrial and ventricular contractions.

Other statistical approaches, such as logistic regression, may differentiate between AF and non-AF. A regularization strategy for regression and classification models [16] Logistic regression models use the logistic function rather than attempting to match the probability between 0 and 1. (corresponding to Non-AF and AF). Another statistical method for identifying AF is the Markov model. Using RR-interval time series features, this approach identifies the best distributions that match the data. The probability of certain rhythms may be predicted [13, 14, 15]. The elastic net performs variable selection and regularization to make the logistic regression model more predictable and understandable. Regularization techniques greatly enhanced the effectiveness of AF detection.

1.3.2 Machine Learning Perspectives

ML has been used for AF detection with promising results. A high degree of domain expertise is required to design features for a comprehensive representation of PPG waveforms and uncover class-differentiating patterns using ML techniques. The PPG time series may often obtain morphological descriptors, time-domain statistics, frequency domain statistics, nonlinear measures, wavelet-based measurements, and cross-correlation measures. A hyperplane separates the nearest data points from each class with a high margin that maximizes this distance. By projecting non-separable data into a higher-dimensional space, the method uses kernel functions like the Radial Basis Function (RBF) or polynomial to help SVMs address nonlinear classification issues [42].

Some studies integrated machine learning and threshold-based approaches. For instance, after using several feature criteria to first remove faulty pulses, an ML model was developed to recognize AF in the clean pulses[43, 44].

1.3.3 Deep Learning Perspectives

Thanks to DL, a trustworthy technique for locating anomalies in physiological data, applications for arrhythmia diagnosis employing ECG and PPG, including AF detection, have recently been encouraged. As opposed to machine learning, deep learning models automatically learn feature representations, which speeds up the laborious process of feature building. A neural network—a set of linked layers of computing nodes—is used in deep learning. Convolutional neural networks are the foundation of the most widely used DL methods for AF detection (CNN). CNN was used to solve automated feature extraction and classification issues. Only automated feature extraction was used in this study’s use of CNNs [33]. One study [33] suggested a hybrid model of two serially linked CNNs, the first detecting clean segments and the second identifying AF occurrences. Specific DL models were trained using hybrid input data to capture a wide variety of attributes covering more than one domain (i.e., time series and photos). Due to the lack of labeled data in biomedical applications, training a DL model from scratch requires a significant quantity of labeled training data. Transfer learning, which aims to

improve an advanced pre-trained DL, might be used to get around this issue [45]. The needed number of layers and the degree of fine-tuning complexity depend on the specific applications [18].

The top IT businesses have recently conducted extensive AF screening experiments to evaluate the efficacy of the most modern PPG-based smart gadgets.

1.4 Project Overview:

In Chapter 1 “Introduction”, a brief introduction on Atrial Fibrillation Detection has been provided along with the concepts and motivation of this project. A brief depiction of the various methods in the field of AF detection has also been described. Furthermore, different challenges that are faced while performing research in this field have been discussed. Lastly, the objectives of our research are reviewed in short.

In Chapter 2 “Related Works”, we have discussed other works related to this field. Various methods in AF detection of previous works along with their workflow diagram have been overviewed. Moreover, we have also made a comparison between methods of previous works along with their accuracy along with problems in respective works.

In Chapter 3 “Theoretical Overview”, the theoretical overview of this work was discussed. Sensors and datasets utilized in the study were described in detail. The theoretical features of the project’s models were then explained. Finally, the overview of different pre-processing pipeline and workflow are also described in this section.

In Chapter 4 “Methodology”, the research methodology of the project is described in detail. The pre-processing and Deep Learning techniques were explained further used in the project. Description of the models and their training parameters were added after that. In Chapter 5 “Result and Analysis”, Results obtained from training the models were shown and analyzed. Comparison between the performance of the models was given. Finally, insights obtained from comparison of the results obtained were discussed.

In Chapter 6 “Conclusion and Future Works”, conclusion has been drawn and scope of improvement in the future in this research has been discussed.

CHAPTER 2

RELATED WORKS

2.1 AF Detection Methodology

Atrial fibrillation (AF) is an abnormal cardiac rhythm characterized by a disorganized atrial activity [46, 47, 48]. AF is often undiagnosed, affects about 34 million people worldwide, and has an estimated prevalence of 3% in the adult population [49, 50]. Early detection of AF and timely initiated anticoagulation reduces AF related mortality and morbidity [46]. Unfortunately, the detection of AF is often challenging, up to 30% of AF patients are asymptomatic [3], and 25–60% of all AFs are paroxysmal [47]. Thus, despite a comprehensive clinical examination, almost 30% of all strokes remain cryptogenic and occult paroxysmal AF is one of the culprits [47]. Electrocardiograms (ECGs) are now the foundation for the majority of pertinent research on the automated diagnosis of AF. ECGs are used for ECG collection, pre-processing, feature extraction, and identification. The two most popular ways to diagnose AF are short-term ECG examinations and continuous dynamic ECG monitoring [51]. Burst AF typically exhibits no overt symptoms and is undetectable by short-term ECG [52]. Long-term ECG monitoring, however, is expensive and unfriendly to the patient since it uses large ECG sensors. Because of its simple setup and flexible sensor, AF monitoring with PPG has gained greater traction than ECG [53].

2.1.1 Machine Learning Based AF Detection

Machine learning (ML) and deep learning have been separated from current relevant studies based on PPG detection for AF (DL). The properties of the manual design are mostly used by the ML direction to identify AF. It has been shown in the past that photoplethysmography (PPG) may identify AF [32–38]. Research has concentrated on using PPG-based AF detection as an alternative to ECG-based solutions since it is a non-invasive technology that can record on most contemporary devices [54]. In 2017, Tang et al. [55] employed the logical regression model to detect AF using six factors, including the PPG’s RR time series and nonlinear analysis. Using the deep neural network shown in Figure 2.1. and the conventional signal processing method, the author of [56] developed a magnificent machine learning strategy to identify the AF episode.

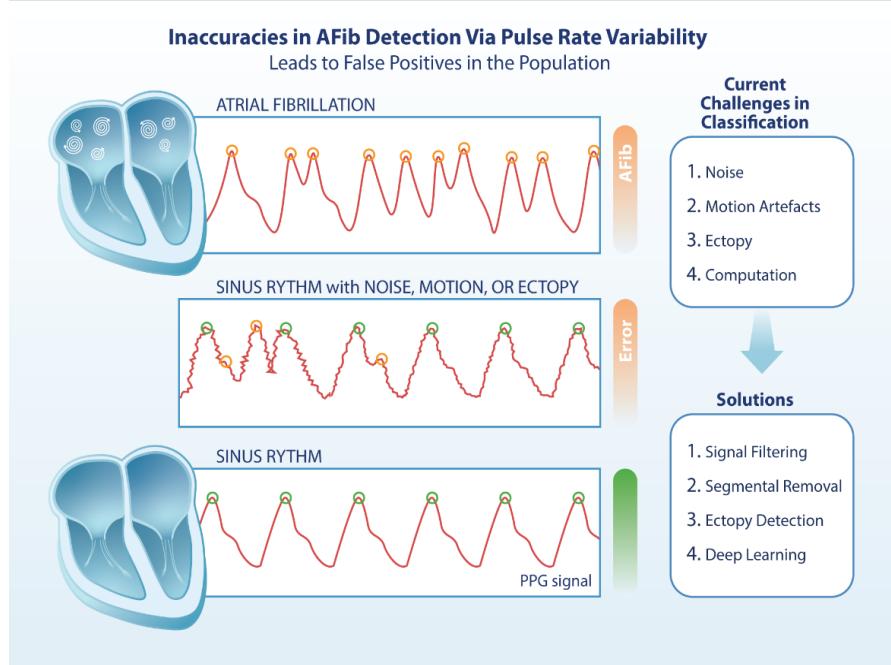


Figure 2.1: AF Detection using Pulse Rate Variability

In 2019, Fallet et al. [57] estimated the values of numerous parameters, such as the mean, standard deviation, and quartile in PPG. They identified these factors as inputs to the decision tree model. Both abnormalities of cardiac rhythm and the lack of p-wave detection approaches have been used in [58], in which the authors employ a Kardia band to capture the ECG and PPG data from a wristwatch

and use it to identify AF from normal sinus rhythm. In the same year, Linda et al. [59] incorporated PPG features, inter-pulse intervals (IPIs), and accelerometer data in the random forest (RF) model to categorize AF, atrial flutter (AFL), and other rhythms. In resting patients with continuous PPG monitoring, AF was diagnosed with a sensitivity of 81.0 %, specificity of 96.4%, and PPV of 86.6% [60]. In our investigation using night-time recordings, the hourly simulated AF alarm sensitivity for detecting AF was close to 78.2%, while the PPV was greater at 97.2%. KNN classification is a very basic clustering approach where a sample is categorized by a majority vote of its neighbours and allocated to the class based on the most frequent class among its k nearest neighbours [16]. SVM creates a hyperplane that divides two classes with a large margin that optimizes the distances between the closest data points from each class. SVMs appear to be effective in nonlinear classification problems by transferring non-separable features into a higher-dimensional space; a method called the kernel trick, which employs kernel functions such as Radial Basis Function (RBF) or polynomial [42]. Some investigations employed a mix of threshold-based and ML techniques. For example, the thresholds of several characteristics were initially employed to remove bad pulses, and then an ML model was created to identify AF in the clean pulses [18]. In the study [61], the author developed a technique consisting of first pre-processing the PPG signal followed by identifying motion and noise artifacts (MNA). Next, heart rate is estimated by peak detection from the segments of PPG signals found to be clean. The heart rate variations are categorized between AF and normal sinus rhythm. The block diagram of the suggested model is presented in fig 2.2.

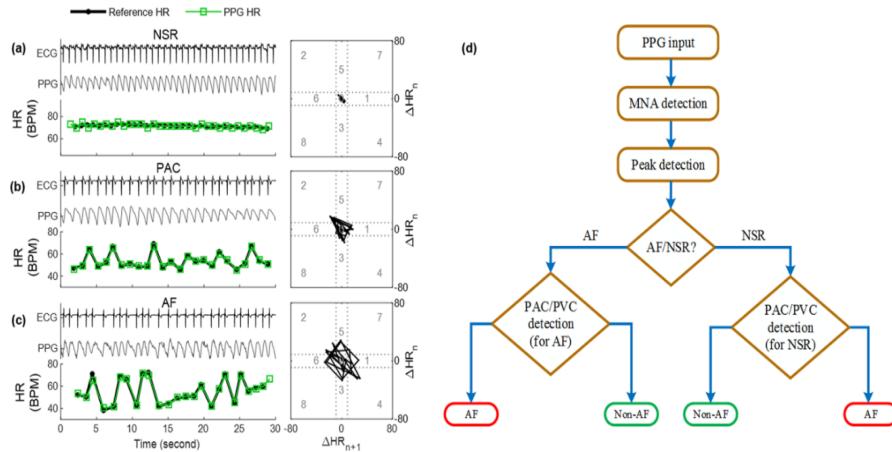


Figure 2.2: Block Diagram of the Proposed Method

A typical AF detection method extracts characteristics (temporal, spectral, or morphological) from the collected PPG signal and analyzes them to determine if an AF beat is observed. Image representation of the temporal waveform has been addressed in several techniques. The tachogram (RR intervals) has traditionally been used to extract significant characteristics since it is a dependable measure of heartbeats [36]. Realizing that PPG waveforms might include physiological information other than heart rate, additional properties other than RR intervals were developed [37]. The use of PPG time series and their picture representation (e.g., the basic plot of the signal, fast Fourier transform spectrum, or wavelet spectrogram-represented in Figure 2.3. (in PPG representation portion) were utilized with promising results in the identification of physiological events. Images for PPG representation in Figure 2.3 are a broad representation of the information format types employed by the various algorithms.

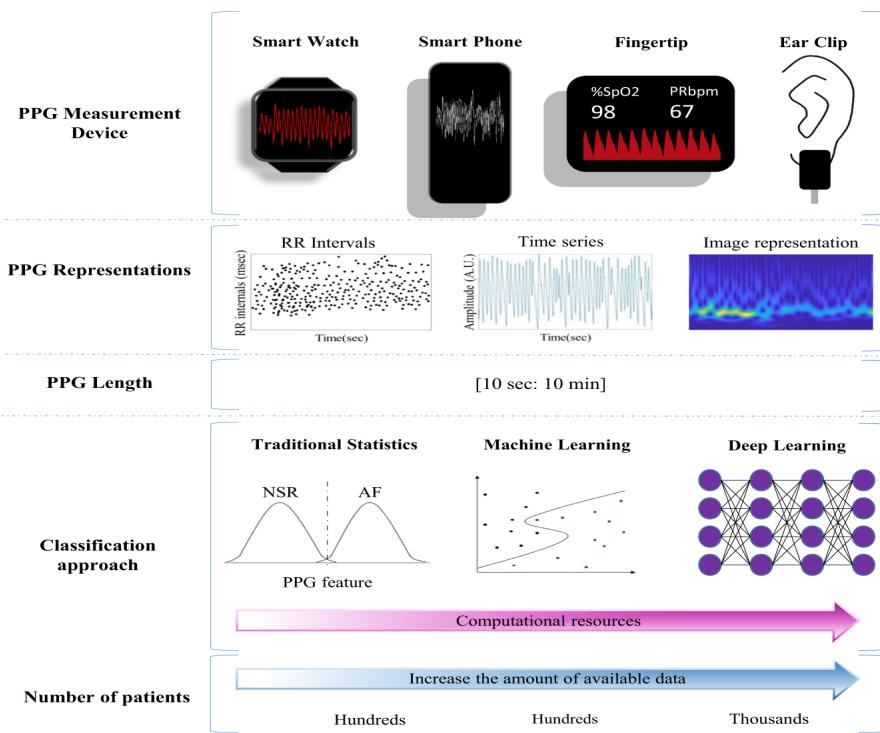


Figure 2.3: Complete Overview of AF Detection using PPG Data

2.1.2 Deep Learning Based AF Detection

Deep learning (DL) has recently emerged as a potent tool for identifying irregularities in physiological data, paving the way for arrhythmia diagnosis from ECG and

PPG, including AF detection. Deep learning models, unlike ML, automatically learn feature representations, saving the arduous effort of feature construction. DL employs a neural network, which is a collection of linked layers of computing nodes. Convolutional Neural Networks are the most often used DL techniques for AF detection (CNN). CNN was used to solve challenges with automated feature extraction and classification. CNNs were only employed in a few researches for automatic feature extraction. One research presented an aggregated model comprising two serially linked CNNs, where the former recognizes clean segments and the latter identifies occurrences of AF [62]. Some DL models were trained using hybrid input data (i.e., time series and pictures) to collect a diverse set of characteristics from many domains.

A substantial quantity of labelled training data is necessary to train a DL model from the start. The scarcity of labelled data is a key restriction in biological applications. Transfer learning, in which the aim is to fine-tune a complex pre-trained DL, might be used to overcome this constraint [45]. The number of layers necessary and the difficulty of fine-tuning is determined by the application [18]. In ref [1], a pre-trained ECG-based CNN model was fine-tuned using a limited set of labelled PPG segments to identify AF from PPG segments.

Deep learning has been employed in numerous disciplines such as image identification, computer vision, and pattern recognition due to the fast acceleration of GPU processing [1]. The author of [1] suggested an automated AF detection model that included 2D-CNN and LSTM models and compared the results to well-known deep learning models like VGGNet [26], GoogLeNet [63], and Microsoft ResNet [16]. The figure 2.4 shows the technique of [1]'s suggested model.

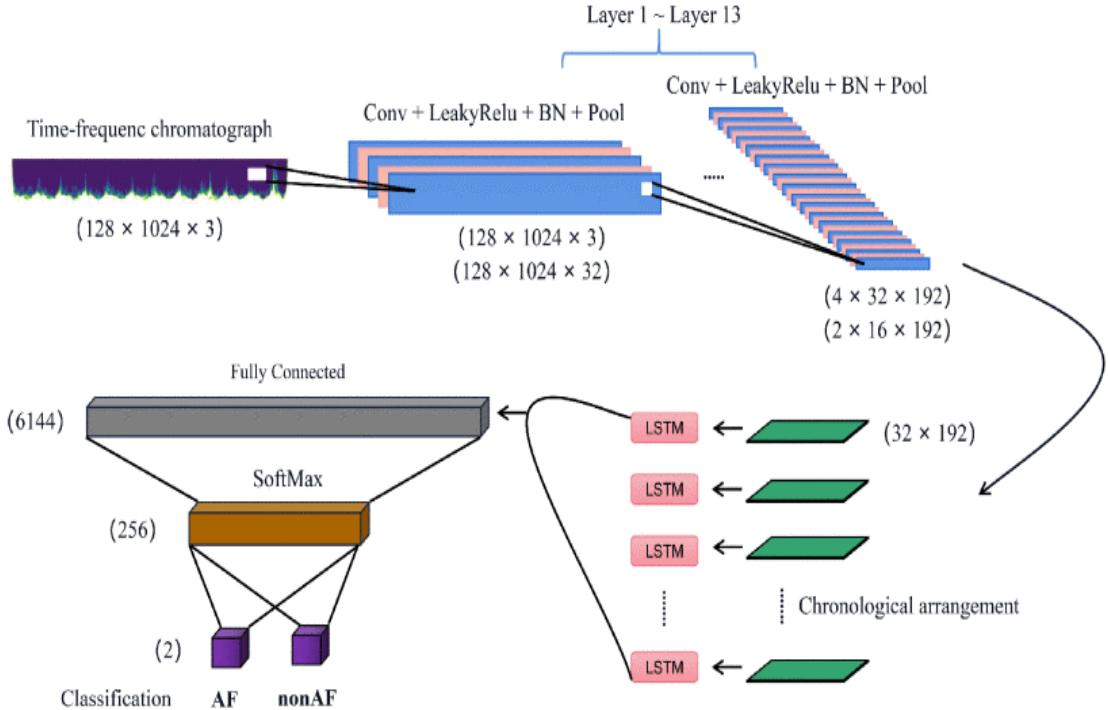


Figure 2.4: Proposed model of AF Detection using [1]

The accuracy, sensitivity, and specificity of these approaches for distinguishing AF from a clean sinus rhythm (SR; that is, SR without any premature atrial complexes or PACs) were promising [17, 64, 65]. The author of [66] created a DL-based algorithm that outperformed earlier algorithms in diagnosing AF using PPG data. They discovered that both 1D-CNN and RNN performed well in terms of diagnostic performance (AUC=0.998 and 0.996 for 1D-CNN and RNN, respectively; both DL-based algorithms performed better than previous well-known AF detection algorithms, even under a high PAC burden, and had the potential to improve as more samples were allowed to be trained; and most diagnoses by the DL classifiers were confident, and the respective calculated CLs provided an easily interpreted reliability. Fig 2.5 depicts the complete DL framework of AF diagnosis.

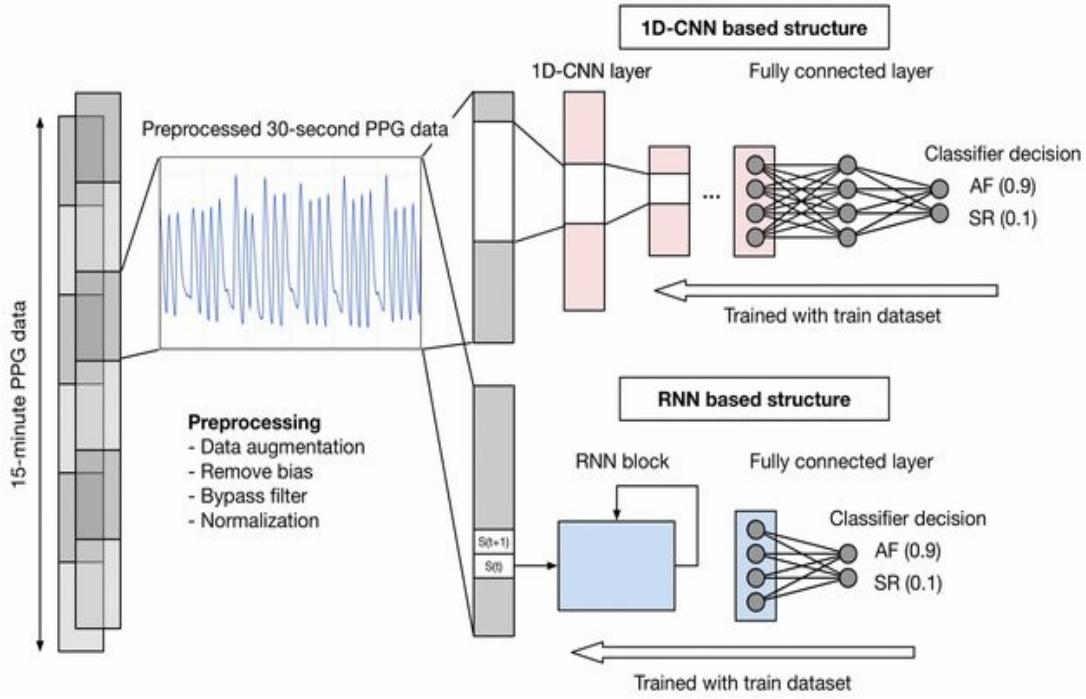


Figure 2.5: Complete DL Framework of AF Diagnosis

PACs, on the other hand, were often seen in individuals with paroxysmal AF or after successful cardioversion [23, 66, 67], making AF identification using PPG from filthy SRs less feasible. Previous techniques had a significant constraint in separating AF from SR using PACs during the feature extraction process [33, 34]. PPG monitoring should be made more practical by using more advanced AF detection techniques. The work shows that a deep learning system can identify AF with excellent accuracy using a raw PPG signal without feature engineering or considerable pre-processing [68]. Furthermore, this unique methodology outperformed typical approaches based on heart rate variability and traditional statistical methodologies, as well as a machine learning system fed just PPG-derived heart rate data. AF detection systems powered by artificial intelligence and wearables have the potential to reduce cardiovascular morbidity risks by identifying and treating millions of people who have undiagnosed AF. An explainable deep learning 1D-CNN model was presented by the author of [2] for application in intelligent healthcare systems using multipurpose devices like smartphones and smart wearables. Using the MIT-BIH ECG datasets and HRV characteristics as inputs,

the 1D-CNN model categorizes the NSR and AF from short-length ECG or PPG signals. Overall classification performance for the 1D-CNN model was 95.50 % accurate, 94.50% sensitive, 96.00% specific, 93.40% F1-score, and 95.30% AUC using a five-fold cross-validation strategy. The suggested model's flow diagram is shown in Fig 2.6

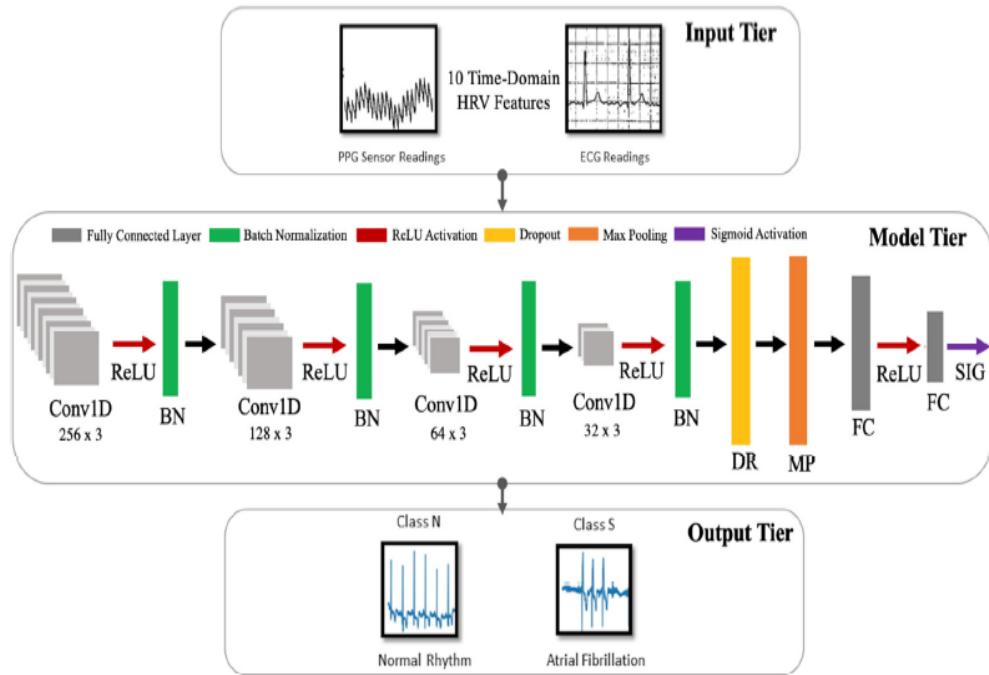


Figure 2.6: Deep Learning 1D-CNN model on ECG data[2]

CHAPTER 3

THEORETICAL OVERVIEW

In this project, we have suggested a deep learning paradigm to detect and classify Atrial Fibrillation (AF) from the ECG and PPG data. Due to the sensors' availability, many complex patterns are classified using machine learning and deep learning models. The wrist-based PPG and reference ECG with accelerometer sensor data are used to classify the AF episodes. This is especially appropriate for the creation of automated PPG-based AF detectors.

3.1 Sensor Description:

The dataset [69] we have used for the classification task consists of ECG and PPG signals with accelerometer sensor data. The acquisition of ECG and PPG was started for a week in the patients' homes. The Bittium Faros 180 ECG device and the Bittium OmegaSnap patch electrode were used to record the reference ECG data at a sampling frequency of 500 Hz. A green LED built within a wrist-worn gadget created by the Biomedical Engineering Institute was used to capture PPG data at a sampling rate of 100 Hz. Additionally, a tri-axial accelerometer sensor was used to acquire accelerometer data in both devices. The accelerometer data were obtained at 25 Hz and 50 Hz sampling frequencies for the reference ECG and wrist-worn devices shown in Figure 3.1.



Figure 3.1: (a) Bittium Faros™ 180 ECG device, reference ECG data acquisition protocol; (b) Wrist-worn device with embedded green LED for PPG data acquisition. Both of them with embedded acceleration sensor.

Bittium Faros™ 180 is an innovative kit for monitoring ECG for multi-purpose (i.e., AF episode detection) with the dimension of 49 x 29 x 12 mm of 16gm with smoothed design. It is an ideal choice for long-term ECG holtering and real-time online measurement. The sensor consists of 1 channel ECG holtering, capable of a wide range of sampling rates of 125 Hz, 250 Hz, 500 Hz, and 1000 Hz. A triaxial accelerometer sensor is embedded in the sensor, which has an adjustable sampling rate of up to 100 Hz. It also includes Bluetooth (API available), which ranges up to 100 meters. With a long durable, rechargeable battery and memory capability of up to 180 days of recording. Because of the lightweight and waterproof design, the device can be worn comfortably during everyday life and even during sleep. The wrist-worn wearable device provides a wide range of opportunities to monitor the different types of physiological conditions such as heart rate, AF episodes etc. It has become more popular due to its durability across longer time periods as well as with high resolution at a low price. An optical sensor (photoplethysmography; PPG) can collect information on the volumetric change in blood profusion, which can serve as a surrogate for clinical assistance. Wrist-band PPG sensors can have several distinct colours of light, such as green, ambient or red. Each of them is

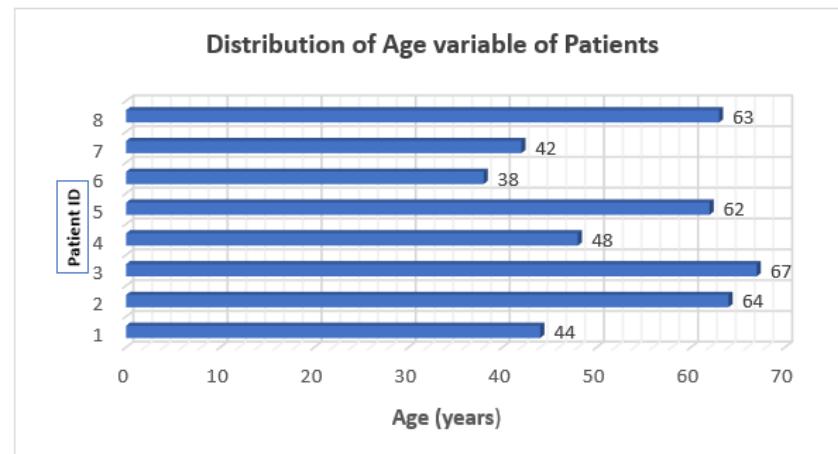
used for individual purposes. In our scenario, we have used green and ambient LED light. The green colour is suitable for different purposes, like It resists motion artifacts, has a shorter wavelength, and has an excellent signal-to-noise ratio. The light passes through and is received by the electrode situated at the opposite site. The absorption and transmission of the light depending upon the change in blood profusion. The sensor can acquire the data continuously at the sampling rate of 100 Hz with an embedding triaxial accelerometer sensor at a sampling frequency of 50 Hz. Thus, PPG allows for in vivo examinations of clinical processes in real-world settings.

3.2 Dataset Description:

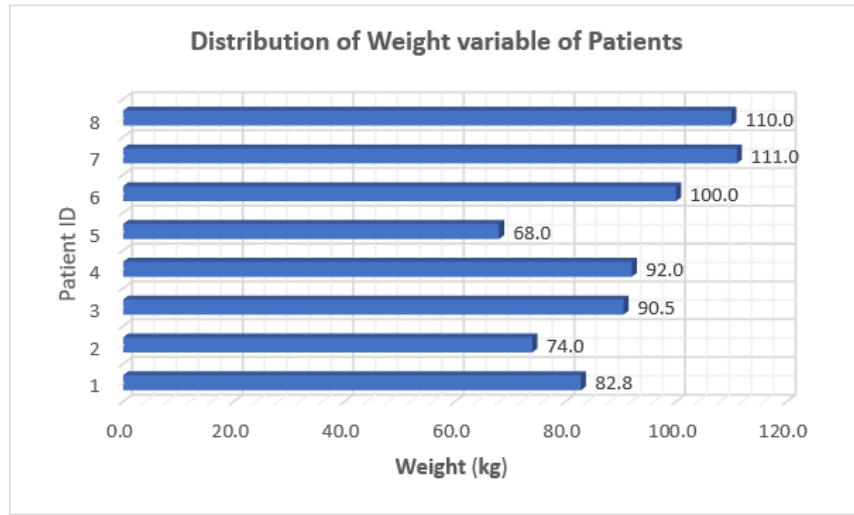
In this project, we have detected and classified the AF episode from the reference ECG and PPG data. We have used an open-source dataset [69], which is available in “Open Access Infrastructure for Research in Europe” (OpenAIRE). The dataset consists of long-term ECG and PPG monitoring signals from 8 individuals with suspected AF monitors for 5 to 8 days. Weight and Age information of individual patient is shown in fig 3.2

Table 3.1: Detailed demographic and clinical information for each AF patient

Patient ID	Alcohol Consumption	Familiar AF	Hyper Lipidaemia	Hyper-glycaemia	Arterial blood pressure >140/80 mmHg	Diagnose Hypertension
1	1	0	1	0	1	2
2	0	0	0	1	1	2
3	2	0	1	1	1	3
4	1	0	0	0	0	0
5	2	1	1	0	1	2
6	2	0	0	0	0	0
7	2	1	1	1	1	2
8	0	1	1	1	1	2



(a)



(b)

Figure 3.2: (a) Distribution of age, (b) Distribution of weight of each AF patient.

The dataset contains about 1306 hours of the acquisition of data (ECG and PPG with accelerometer data) for clinical purposes. The complementary detailed demographic and clinical information for each individual is listed in Table 3.1.

The data is acquired in the format of a .mat file with the format as XX_YYY_ZZ.mat, where XX is the patient ID, YYY is the name of the data (i.e., for ECG, it is “ECG”, for PPG, it is “PPG”) and the ZZ is the serial no. of the recordings. For individual patient, there is one continuous ECG recording file and several PPG files due to the technical issue of a long time to allow the battery to charge for wrist-worn devices.

Additionally, the raw ECG, PPG, and acceleration data include

1. a file header,
2. the number of data recorded within the recording process, and
3. the day when the recording process is started.

The QRS time indices, computed RR intervals, and beat-by-beat annotated AF are also included in the ECG data. Each PPG file also includes a timestamp vector that indicates the exact moment that each piece of data was captured by wrist-worn devices.

Physical inactivity is described as ≤ 5000 steps/day or ≤ 150 min/week of moderate-intensity exercise or ≤ 75 min/week of high-intensity exercise in the subject_info.xlsx file. Also excessive physical activity is defined as ≥ 750 min/week of moderate-intensity exercise. Hypertension is classified into stages based on systolic/diastolic blood pressure: stage I corresponding to 140/90–159/99 mmHg, stage II to 160/100–179/109 mmHg, and stage III to $\geq 180/100$ mmHg.

The PPG and ECG collection devices have not been synced, thus researchers utilizing the resource should be aware of this fact. As a result, the precise alignment of the signals should be accomplished based on certain physiological parameters, such as the heart rate deduced from the PPG and ECG data. The supplied time of day the PPG recording starts should be taken as an approximate estimate of the precise time because the wrist-worn device’s internal clock may somewhat drift throughout the course of the monitoring period. The devices’ sampling frequencies can vary somewhat, thus users should be aware of this. Utilizing the timestamp sec data record vector can help to reduce the impact of this problem on the PPG

signal, but an equivalent variable is not present for the ECG signal. The data distribution of both AF and NSR samples are shown below in fig 3.3

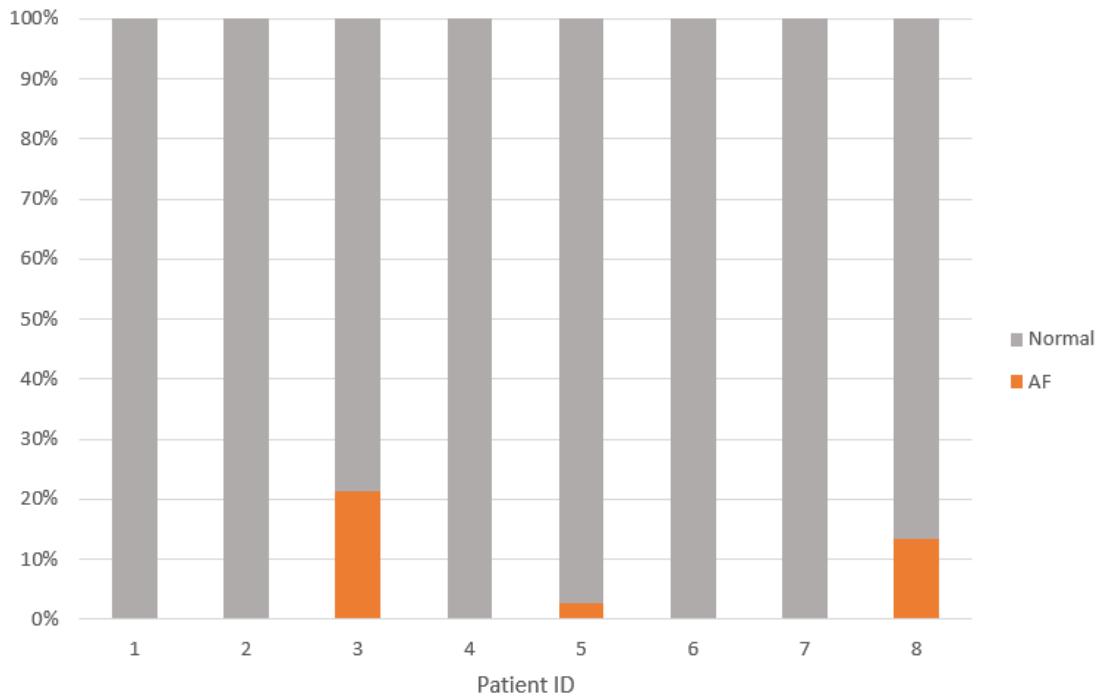


Figure 3.3: DIistribution of AF and Normal Samples

3.3 Deep learning and Classification Models

Atrial fibrillation (AF) is a serious atrial illness that affects many elderly persons who have cardiac problems [4]. AF is identified by a rapid and irregular heart beat that leads the atrium to lose its normal rhythm [5]. Machine learning (ML) and deep learning have been used to split recent relevant research for AF based on PPG detection (DL). To identify AF, the ML models depends heavily on the characteristics of the manual design. Deep learning is a subset of machine learning (ML) that uses data with a specific logical structure to find correlations and patterns similar to how the human brain works. Deep learning, also known as deep neural networks, employs several hidden layers in the neural network in contrast to traditional neural networks, which only have a few hidden layers. Deep learning models can create new features from a small set of features in the training dataset

without the need for further human interference. Deep learning can thus perform complex tasks that frequently necessitate large-scale feature engineering. One of the most appealing aspects of deep learning is its ability to deal with unstructured data. This is especially important in business because the vast majority of company data is unstructured. Text, images, and voice are the most common data forms used by organizations.

Deep learning models are sometimes referred to as deep neural networks because the majority of deep learning techniques use neural network topologies. The phrase “deep” is used to refer the number of hidden layers in the neural network models. Traditional neural networks have only 4-5 hidden layers, whereas deep networks can have up to 150. Deep learning models are taught utilizing massive amounts of labeled data and neural network topologies that learn features directly from the data, eliminating the requirement for manual feature extraction. The Most widely used methods are:

1. Multilayer Perceptron Networks (MLP)
2. Convolutional Neural Networks (CNN)
3. Long Short-Term Memory Recurrent Neural Networks

In this experiment, we will be using MLP, 1D CNN and Self-ONN model. Here, we propose Self-Organized Operational Neural Networks (Self-ONNs) with generative neurons to solve the disadvantages and limitations of CNN and MLP models. Self-ONNs have the potential to self-organize network operators during training, as the name indicates.

3.4 Convolutional Neural Networks (CNN)

Convolutional neural networks, sometimes referred to as CNNs or ConvNets, are a subclass of neural networks that are particularly adept at processing input with a grid-like architecture, like images. CNNs are also very good at identifying key prominent elements in a signal.

The main benefit of CNNs is that they offer automatic feature extraction. The given input data is first transferred to a feature extraction network, and the recovered features are then sent to a classifier network as shown in Figure 3.4. There are numerous convolutional and pooling layer pairs in the feature extraction network. To conduct the convolution operation on the input data, the convolutional layer is made up of a number of digital filters. The threshold is determined by the pooling layer, which also serves as a layer for dimensionality reduction. A number of parameters must be changed during back-propagation, which minimizes the connections inside the neural network architecture.

In the literature, there are several different CNN architecture variations. But they share a lot of fundamental elements. A CNN typically has three layers: a convolutional layer, a pooling layer, and a fully connected layer.

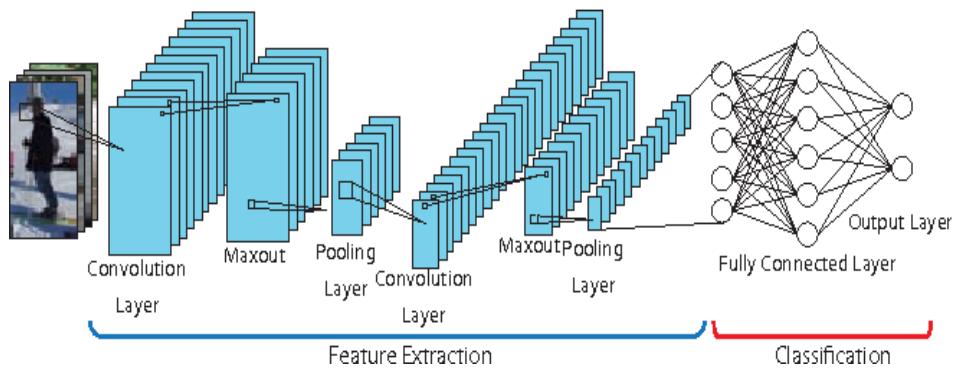


Figure 3.4: CNN Model Used in Classification Problem [3]

1. **Convolutional Layer:** The foundational component of the CNN is the convolution layer. It carries the majority of the computational load on the network. This layer creates a dot product between two matrices, one of which is the kernel—a collection of learnable parameters—and the other of which is the constrained area of the receptive field. The CNN convolves the entire image or signal with the intermediate feature maps in the convolutional layers, producing a variety of feature maps as a consequence. Suppose, input of size $W \times W \times D$ and Dout number of kernels F with stride S and amount of padding P , then the size of output layer can be calculated by:

$$W_{out} = \frac{W - F + 2P}{S} + 1 \quad (3.1)$$

This will produce the size of output volume : $W_{out} \times W_{out} \times D_{out}$.

2. Pooling Layers: The pooling layer change the output of the network at fixed positions by deriving a summary statistic of the nearby outputs. As a result, the representation's spatial size is reduced, which reduces the amount of computation and weights needed. Every slice of the representation is individually handled for the pooling operation.
3. Fully Connected Layer : As in a typical FCNN, all neurons in this layer are fully connected to all neurons in the layer before it and the layer after it. Because of this, it can be calculated using the standard method of matrix multiplication followed by a bias effect. The mapping of the representation between the input and output is facilitated by the FC layer.

Discussion of Activation Function: An activation function in a neural network describes how a node or nodes in a layer of the network translate the weighted sum of the input into an output. To convey complicated properties, CNNs can use a variety of activation functions. It is merely a function that you employ to obtain node's output. Transfer Function is another name for it. The activation function in this model is a unit that chooses which information should be sent to the following neuron, much like the neuron model in the real brain. The Sigmoid function, Relu function, Leaky Relu function, Tanh function, ELU function, and Soft-max activation function are some of the most popular activation functions.

3.5 Self-Orgaized Operational Neural Networks (Self-ONN):

In order to address the well-known limits and shortcomings of traditional CNNs, such as network homogeneity with the single linear neuron model, Operational Neural Networks (ONNs) [6] have recently been proposed as a superset of CNNs. In order to increase variety and learn extremely complicated and multi-modal functions or spaces with a minimum amount of network complexity and training data, ONNs are heterogeneous networks with a generalized neuron model. ONNs, like their prior network Generalized Operational Perceptrons (GOPs) [7, 11], Increase variety to learn very intricate and multimodal areas or functions with little

network complexity and practice data.

An "operator set" is a specific set of nodal, pool, and activation operators, and all potential operator sets are maintained in an operator set library. An ideal operator set per layer can be repeatedly found using the Greedy Iterative Search (GIS) method across numerous brief Back-Propagation (BP) training sessions. The best operator sets found can then be distributed among all neurons in the corresponding hidden layers to further personalize the final ONN. The results on numerous difficult learning problems demonstrate that 1) with the appropriate operator set, ONNs can perform the necessary linear or non-linear transformation in each layer/neuron to maximize learning performance, and 2) ONNs not only significantly outperform CNNs, but can also learn problems where CNNs completely fail.

By extending the nodal and pool operators' exclusive use of linear convolutions in the convolutional neurons, ONN expands the fundamental concept of GOP. These make up the operational layers and neurons, whereas traditional CNNs are directly responsible for the two major drawbacks of weight sharing and limited (kernel-wise) connection. This is also depicted in Figure 3.5 (right) where three operational layers and the k^{th} neuron with 3x3 kernels belong to an ONN. As illustrated, the input map of the k^{th} neuron at the current layer, x_k^l is obtained by pooling the final output maps, y_i^{l-1} of the previous layer neurons operated with its corresponding kernels, W_{ki}^l , as follows:

$$x_k^l = x_k^l + \sum_{i=1}^{N_{l-1}} Oper2D(W_{ki}^l, y_i^{l-1}, NoZeroPad') \quad (3.2)$$

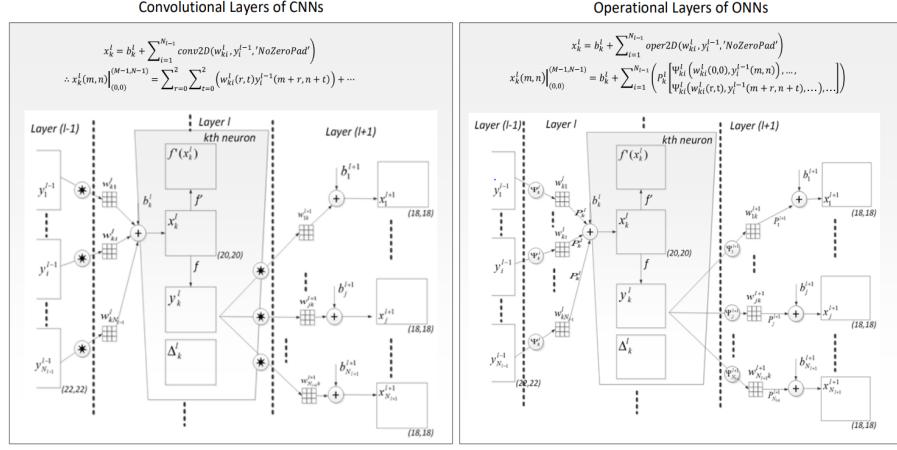


Figure 3.5: Depiction of the k^{th} neuron of a ONN (Right) and an CNN (left) along with the 3 successive ONN (right) and CNN (left) layers.

However, ONNs suggested in [6] also show some drawbacks. Low heterogeneity is produced when all of the neurons in a buried layer use the same set of operators. As a result, all kernel connections between each neuron and the neurons in the layer above must employ the same nodal operator. One notable limitation is that the operator set library's fixed-ahead operators, particularly the nodal operators, directly affect how well the ONN learns. In other words, learning performance will decrease if the appropriate operator set for proper learning is lacking. It is plainly impossible to cover every possible nodal operator because there are an infinite number of them. In addition, although they can be approximated, many operators cannot even be expressed using common nonlinear functions. Finally, the Greedy Iterative Search (GIS) method is a time-consuming local search technique that requires many BP runs. For deep networks trained on massive amounts of data, the best operator sets may not be suitable. GIS is a search technique used to locate the best operators in ONNs. Finding a single operator set each layer requires several training sessions. GIS thus causes a large computational complexity constraint.

In order to overcome these drawbacks and restrictions, Self-Organized Operational Neural Networks (Self-ONNs) with generative neurons have the capacity to self-organize the network operators during training. As a result, they don't need a

search strategy or previous operator set library to find the optimum nodal operator. Since each neuron can create any combination of nodal operators, which may or may not be a well-defined function such as linear, sinusoids, hyperbolic, exponential, or some other standard functions, "generative neurons" will actually address the limitation of using a single nodal operator for all kernel connections of each neuron.

It is true that the kernel (weights) parameters affect the nodal operator output; for example, the kernel parameters for the "Sinusoid" nodal operator of a certain neuron are separate frequencies. Pure sinusoids would not be the best option for this neuron; instead, a "composite operator" such as the linear combination of harmonics, hyperbolic and polynomial, or an arbitrary nodal operator function would be preferable. However, the final nodal operator function after training can only have a pure sine wave as a pattern or form. This is valid for biological neurons where synaptic connections may take any form or arrangement. To sum up, a generative neuron is a neuron that develops a composite nodal operator on its own during training. These generative neurons allow a Self-ONN to self-organize its nodal operators during training, and the training procedure "optimizes" the nodal operator functions to maximize learning performance.

Any nodal operator function can be created because to this tremendous flexibility. The training strategy that back-propagates the error across the operational layers of Self-ONNs is created in order to build the appropriate nodal functions for its neurons. We'll show that Self ONNs, which have superior computational efficiency than parameter-equivalent ONNs, can achieve comparable and frequently higher performance levels over the same set of difficult tasks in [6] with the same rigorous constraints. For instance, the generative neuron can have any arbitrary nodal function, whereas the CNN and ONN neurons in the sample illustration in Figure 3.6 have static nodal operators (linear and harmonic, respectively) for their 3x3 kernels. Ψ (including possibly standard types such as linear and harmonic functions) for each kernel element of each connection.

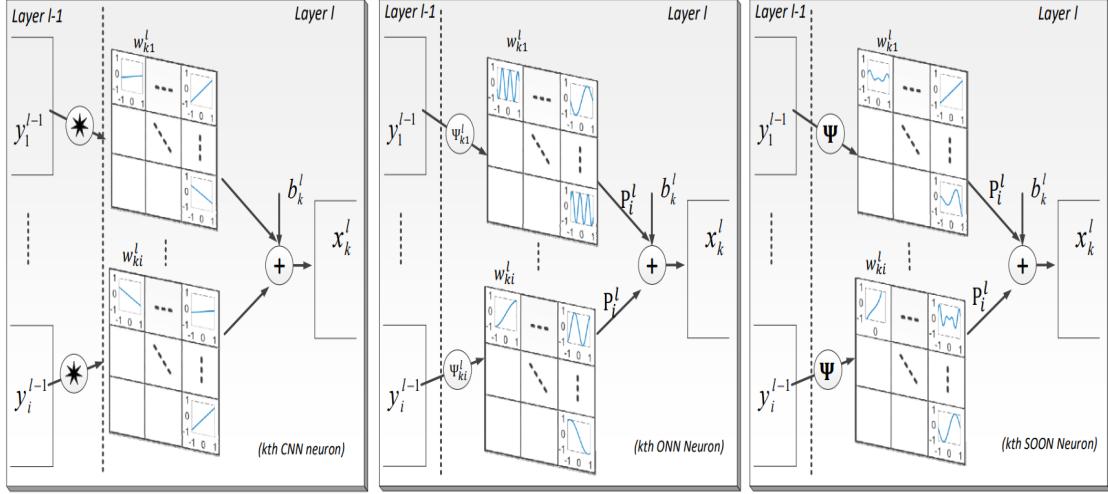


Figure 3.6: Depiction of the 1D nodal operations with the 1D kernels of the i -th neuron of CNN (left), ONN (middle), and Self-ONN (right).

Any nonlinear transformation, Ψ is carried out by the kernel components of each generating neuron of a Self-ONN. Ψ is a function that can be showed by the Taylor-series near the origin ($a = 0$),

$$\psi(x) = \sum_{n=1}^{\infty} \frac{\psi^{(n)}(0)}{n!} x^n \quad (3.3)$$

The Taylor polynomial, or Q^{th} order truncated approximation, is represented by the following finite summation:

$$\psi(x)^Q = \sum_{n=1}^Q \frac{\psi^{(n)}(0)}{n!} x^n \quad (3.4)$$

The above formulation can estimate any function $\psi(x)$ near 0. When the activation function boundaries, the input feature of the neuron maps close to 0 (e.g., tanh), the formulation in (3.4) can be employed to build a composite nodal operator where the power coefficients, $\frac{\psi^{(n)}(0)}{n!}$, can be the parameters of the network learned during training. It was shown in [9] and [10] that the nodal operator of the k^{th} generative

neuron in the l^{th} layer can take the following general form:

$$\widetilde{\psi}_k^l(W_{ik}^{l(Q)}(r), y_i^{l-1}(m+r)) = \sum_{q=1}^Q W_{ik}^{l(Q)}(r, q)(y_i^{l-1}(m+r))^q \quad (3.5)$$

Let $x_{ik}^l \in \mathbb{R}^M$ be the contribution of the i^{th} neuron's at the $(l-1)^{th}$ layer to the input map of l^{th} layer. Therefore, it can be expressed as,

$$\widetilde{x}_{ik}^l(m) = \sum_{r=0}^{k-1} \sum_{q=1}^Q W_{ik}^{l(Q)}(r, q)(y_i^{l-1}(m+r))^q \quad (3.6)$$

where $y_i^{l-1} \in \mathbb{R}^M$ is the output map of the i^{th} neuron's at the $(l-1)^{th}$ layer, $W_{ik}^{l(Q)}$ is a learnable Kernel of the network, which is a $K \times Q$ matrix, i.e., $W_{ik}^{l(Q)} \in \mathbb{R}^{K \times Q}$ formed as, $W_{ik}^{l(Q)} = [W_{ik}^{l(Q)}(r, 1), W_{ik}^{l(Q)}(r, 2), \dots, W_{ik}^{l(Q)}(Q)]$. By the Commutativity of the summation operation in (3.6), one can alternatively write:

$$\widetilde{x}_{ik}^l(m) = \sum_{q=q}^Q \sum_{r=0}^{K-1} W_{ik}^{l(Q)}(r, q-1)y_i^{l-1}(m+r)^q \quad (3.7)$$

One can simplify this as follows:

$$\widetilde{x}_{ik}^l(m) = \sum_{q=q}^Q Conv1D(W_{ik}^{l(Q)}, (y_i^{l-1})^q) \quad (3.8)$$

Hence, the formulation can be accomplished by applying Q 1D convolution operations. Finally, the output of this neuron can be formulated as follows:

$$x_k^l = b_k^l + \sum_{i=0}^{N_l-1} x_{ik}^l \quad (3.9)$$

where x_k^l is the bias associated with this neuron. The 0^{th} order term, $q=0$, The DC bias is disregarded since the neuron's learnable bias parameter can offset its additive effects. A generative neuron becomes a convolutional neuron with the $Q=1$ setting.

CHAPTER 4

METHODOLOGY

In this study, we have trained a variety of ECG and PPGs data to give the proposed model a higher generalization capability. The dataset used for our project is from a publicly accessible dataset [69] to detect and classify “Atrial Fibrillation” from raw data. The raw data collected from various sensors (i.e., A wrist-worn device has a green LED incorporated in it and Bittium Faros™ 180 ECG device with the Bittium OmegaSnap™ patch electrode) are very noisy and corrupted are not suitable to feed into deep learning models. Before training, several pre-processing pipelines are used to clean the data and remove different artifacts (i.e., baseline wander and motion artifact).

4.1 Data Pre-processing:

Due to the rapid exploration of machine learning and artificial intelligence, a generalized solution for any classification can be handled with the help of deep learning models. Deep learning models are very popular due to their diversity in handling problems. Deep learning models are hungry for large data. The more data we feed, the more the generalized performance of the models we get. But, the data should be in proper form to train the deep learning models; otherwise, the models’ performance will be down due to misclassifying. That is why a proper data pre-processing technique is required to train the large modes. All the necessary pre-processing kinds of stuff are listed below, which we have used in our project.

4.1.1 Resolving Drifting:

Multi-modal data has become a common fashion due to its popularity and performance over the mono-modal dataset. While dealing with the multi-modal dataset, there needs to be communication between the different sensors. In the dataset [69], the ECG and PPG acquisition devices mentioned above are not synchronous with each other. Each of the sensors has its own internal clock. In practice, they must be at the same frequency. But, while running a sensor for several hours/days, the performance of the sensors lags behind, resulting in the error in data acquisition. The authors of the dataset [69] mentioned that the internal clock wrist-worn device slightly drifted over the monitoring period. They have supplied the time of day the PPG recording begins, which may be used as a general indicator of the precise time, in order to address the drifting issue. Considering the ECG as the “Ground Truth”, they have provided us with the timestamp_sec_data_record vector for PPG signals. Due to the drifting issue, we have modified the timestamp vectors as follows:

$$PPG^i(n_k) \rightarrow Resample(PPG^i(n_k), f_s^i); \forall i = 1....N_{PPG}; \forall n_k \quad (4.1)$$

where n_k indicates the PPG time-onset while f_s^i indicates the resolved drifting frequency, which can be furthermore calculated as follow:

$$f_s^i = \frac{F_a^i}{F_s^i}; \forall i = 1....N_{PPG} \quad (4.2)$$

here, F_a^i and F_s^i are the data acquisition frequency and the sampling frequency of the PPG signal for ith subject respectively. The variable N_{PPG} indicates the number of subjects present in the PPG signals. The re-mapped PPG waveform, $PPG^i(n_k)$ are then utilized for others pre-processing stuffs. The resolved drifting frequencies, f_s^i along with the data acquisition frequency F_a^i are listed in Figure 4.1

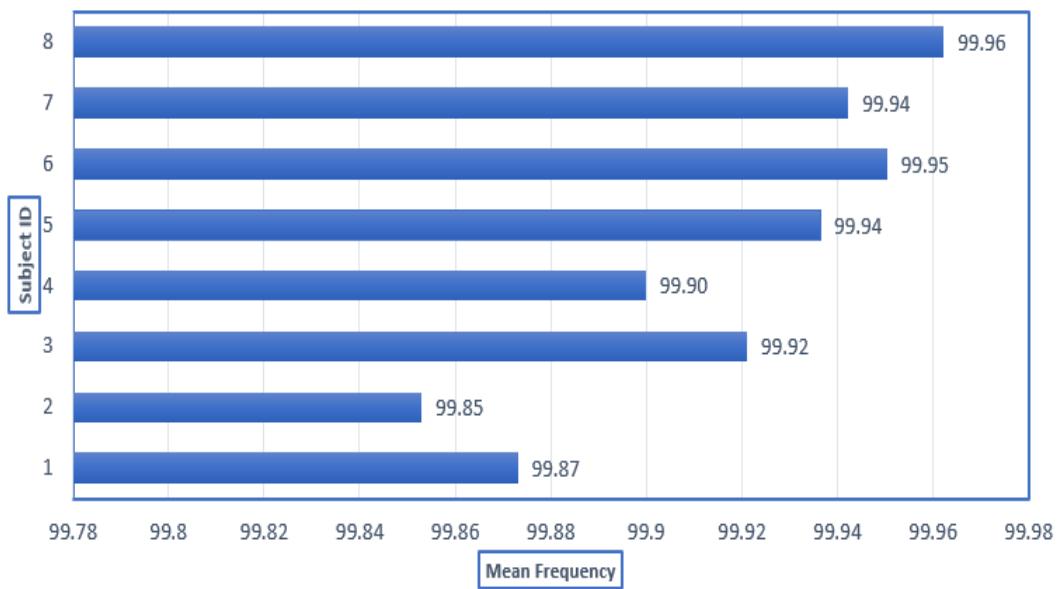


Figure 4.1: Change in PPG frequency (100 Hz) due to drifting

4.1.2 Noise Filtering:

Sensor data are very popular due to their simplicity and computation cost-efficient over image or video signals. But, the fatal disadvantage of the sensor data is exposed noise. The noise from the raw data must be filtered out before training the models. The proper noise reduction analogy can increase the models' performance quite a lot. The raw sensor data, both ECG and PPG, are very noisy.

Basically, noise is imposed due to its Stochastic nature. This signal's periodic frequency is equivalent to the subject's heart rate. The mechanical movement of the heart is about 1-20 Hz frequency range. For low-frequency range of both ECG and PPG signals, noise is added to the signal, which results in degrading the signal quality. However, it is quite challenging to separate the signal from the noise. To eliminate the noise influence from the signal, we have used a number of ways. This is accomplished by using low-pass and high-pass filters, which are intended to reduce both signal distortion and signal phase delay. Any gadget must have

a minimum delay (e.g., adjustment of the sensor at the measurement site). FIR (Finite Impulse Response) and IIR are the two basic kinds of available digital filters (Infinite Impulse Response). Despite having a linear phase and being consistently stable, symmetric FIR filters have significant delays when created with low cut-off frequencies. IIR filters may be quicker than FIR filters, however they often have a nonlinear phase response and hence generate frequency-related signal delays.

The PPG signal contains low-frequency information (between 1 and 20 Hz), but high-frequency noise often contaminates the signal. This is often the result of tests made in a setting with electrical noise or optical pick-up from outside illumination sources.

A low-pass filter with a cut-off frequency close to 20 Hz filters out the high frequencies. A moving average filter, a straightforward application of a FIR filter, is used to accomplish this. This filter features a low roll-off rate and a linear phase to prevent waveform distortion. Due of its very low (0.15 Hz) cut-off frequency, the high-pass filter requires extra design work. This filter's main objective is to suppress the dominating DC background, which the PPG signal is placed on. We have chosen a higher-order digital high-pass IIR filter with a cut-off of 0.05 Hz for processing and reaction speed. Although this filter has a nonlinear phase response and is an IIR filter, it is focused at very low frequencies below the filter cut-off frequency.

The filter responses are shown in Figure 4.2

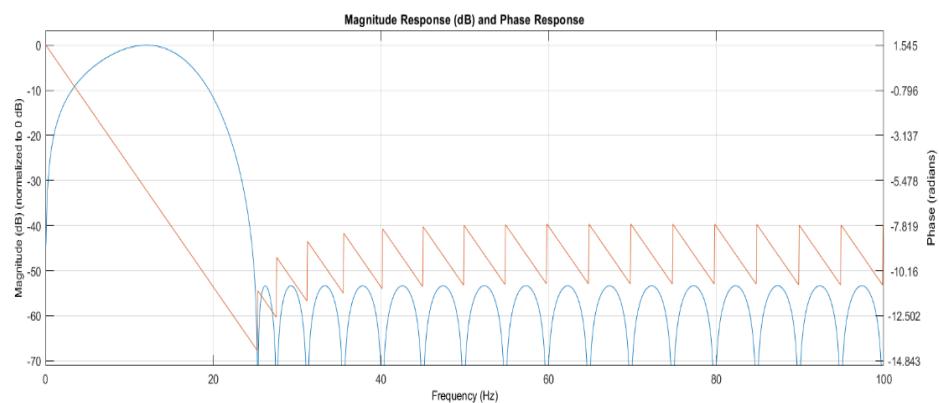


Figure 4.2: Responses of the Filter

The captured ECG signal is subjected to these filter settings. With the exception of the varied selection of the cut-off frequencies, this filtering is identical to that used on the PPG signal (i.e., a low-pass portion and high-pass part) (i.e., 0.5 Hz and 20 Hz, respectively, for high-pass and low-pass for the ECG signal). The variable cut-off frequencies help us to find out the optimal value for the PPG signal as it is quite noisier than that of the ECG signal. Besides this, several parameters are optimized for the filter (i.e., the filter order) to obtain a better-quality PPG signal. After that, the settings are set for the PPG signal to generate the noise-filtered ECG and PPG signal. In the process of parameter optimization, we have measured the similarity score between the ECG and PPG signal considering the ECG as the reference. The quality of the ECG and PPG signal after applying the noise filtering is shown in Figure 4.3

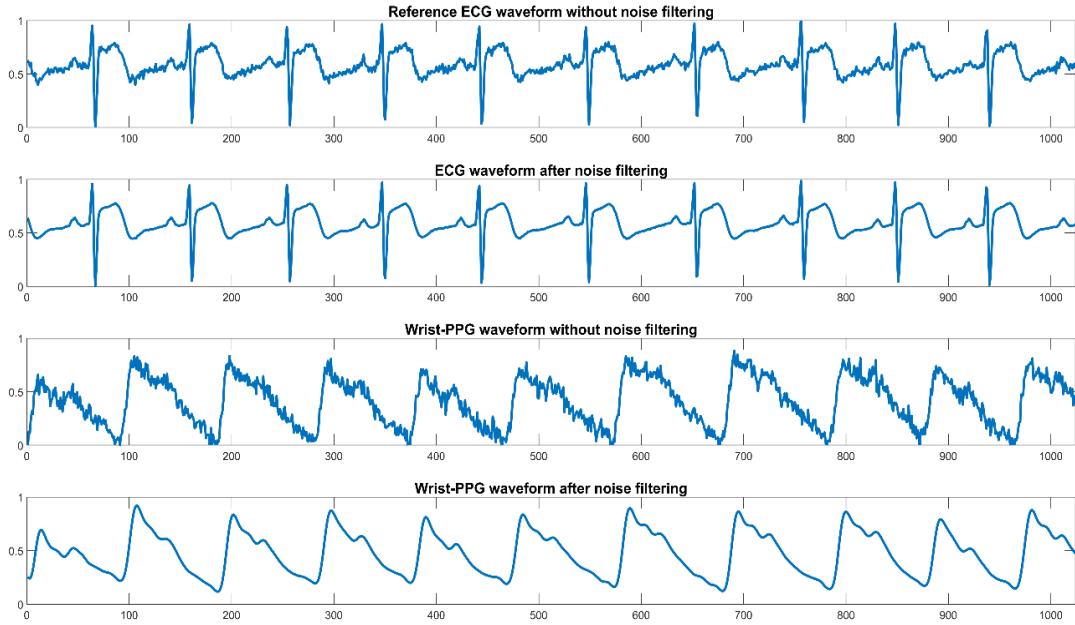


Figure 4.3: Noise Removal of PPG and ECG Data

4.1.3 Data Synchronization:

Due to the multi-modal sensor capability, it is quite issue for the synchronous type data. Both the ECG and PPG are acquired from the different devices mentioned above results in asynchronous between ECG and PPG data. In this section, we

will describe our noble methods to automatically data synchronization between the ECG and PPG signal.

Due to the multi-modal sensor capability, it is quite an issue for the synchronous type data. Both the ECG and PPG are acquired from the different devices mentioned above, resulting in asynchronous between ECG and PPG data. In this section, we will describe our noble methods to automatically data synchronization between the ECG and PPG signal.

In this pre-processing pipeline, we have extracted several from that the original ECG and PPG signals. The pipeline processes both ECG signals and PPG derivative signals so that the first-derivative PPG waveform (VPG) and the second-derivative PPG signal (APG) can be utilized to evaluate the corresponding ECG waveform by automatically segmenting the ECG from the same PPG time onset. This can be accomplished by normalizing the VPG and APG waveforms throughout the range [0, 1]. In order to time-align the peaks of the many signals, time-rescaling and shifting are carried out. The peaks of different signals included :

- The ECG waveform and the derivative PPG signal (VPG and APG)
- The ECG waveform and the ECG reference

By depending on the time alignment of the individual peaks, cross-correlation analysis of these signals may be made easier. The $ECG(n_k)$ to $VPG(n_k)$ and $APG(n_k)$ analysis are reported in the following equation:

$$ECG^i(n_k) \rightarrow ECG^i(n_k + \delta_k^i); \forall i = 1.....N_{ECG}; \forall n_k \quad (4.3)$$

where n_k refers to the PPG time-onset while δ_k^i refers to the offset required to align the $VPG(n_k)$ and $APG(n_k)$ peak with $ECG(n_k)$ signals. The variable N_{ECG} refers the total ECG segments. Alignment between $ECG_{ref}(n_k)$ and $ECG(n_k)$ is obtained throughout the Equation 4.3. After all, the degree to which $ECG(n_k)$, PPG waveform, and their derivatives (i.e., VPG, APG) are correlated is shown by the similarity score between them. The synchronization process is illustrated in Figure 4.4

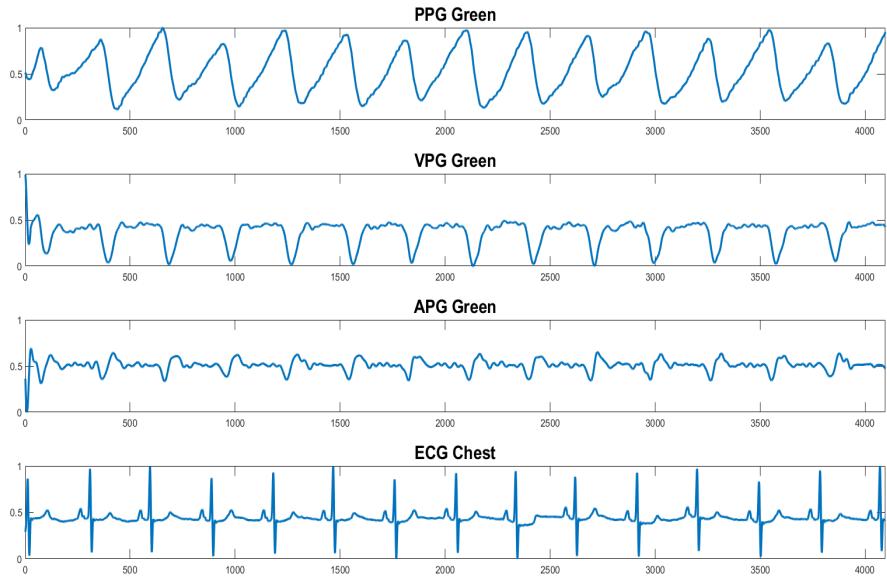


Figure 4.4: Alignment of PPG, VPG, APG and ECG Signals.

4.1.4 Baseline Wander:

Baseline wander is an unnecessary, low-frequency activity in the ECG and PPG that may obstruct the signal processing, resulting in erroneous and misleading clinical interpretation. For instance, because the isoelectric line is no longer clearly defined, it is impossible to compute ECG readings specified with reference to it. Baseline wander, which is frequently brought on by exercise, can be brought on by a number of noise-producing factors, including as sweat, breathing, body movements, and poor electrode contact. The size of the undesirable wander may be many times greater than the amplitude of the QRS complex. Although its spectral content is typically restricted to a range far below 1 Hz, intense activity may cause it to contain higher frequencies. In this part, signal processing methods for eliminating baseline wander are explained in depth.

High-pass filtering of ECG signals is a common technique for removing BW artifacts [70]. However, due to changes in the frequency spectrum of the ECG data,

high-pass filtering for BW removal is not favored as it can distort the shape of the ECG waveform. [70].

In our algorithmic pipeline, we have proposed several techniques to remove the baseline wander from the original signal. The process is segmented into three parts:

- Extraction of peaks from the signal.
- Follow the trend of peaks (i.e., polynomial fitting of the lower bound peaks)
- Remove the baseline and reconstruct the signal

The peaks of the signal are extracted using the standard peak finding algorithm. The algorithm helps to find the relevant peaks and ignore the redundant peaks which are very close to each other by constraining the minimum peak-peak interval threshold. In the dataset, they also provide us with the QRS index, which helps to improve and validate the detection accuracy. After that, the peak-peak intervals are extracted from the difference between the two adjacent peaks helps us to follow the trend of the peaks.

After extracting the peak-peak distances, we have fitted the trend lines of the lower peaks. There are various methods for removing the baseline, such as cubic splines, FIR and IIR filters, moving average filters and polynomial curve fitting. Among them, we have used the polynomial curve fitting, which performs best in our scenario. The degree of the polynomial function is also tuned to get better performance. We have used the median as the central tendency of the peak-peak distances, which are the better approximation since they are not affected by the outliers.

After that, to get the clean signal without the baseline, we have just subtracted the baseline trend from the raw signal. But, the subtraction of the baseline signal causes the amplitude change in the original signal, which may degrade the signal quality. To fix the slight change of amplitude due to the baseline shift, we have used the re-scaled the amplitude. The comparison between the raw signal with and without the baseline is illustrated in Figure 4.5

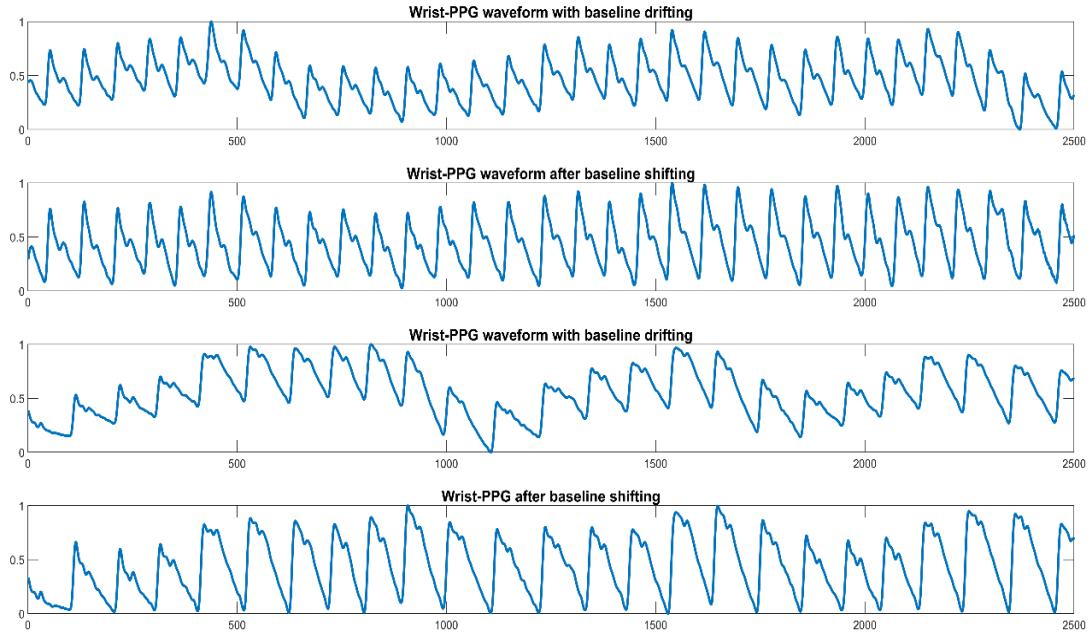


Figure 4.5: Baseline Shifting Function on PPG data.

4.2 Proposed Method

Our proposed method can be divided into 3 main Parts. They are: Data Pre-processing, Model Development and Training Evaluation. The overall block diagram of our proposed method is shown in the Figure 4.6 below:

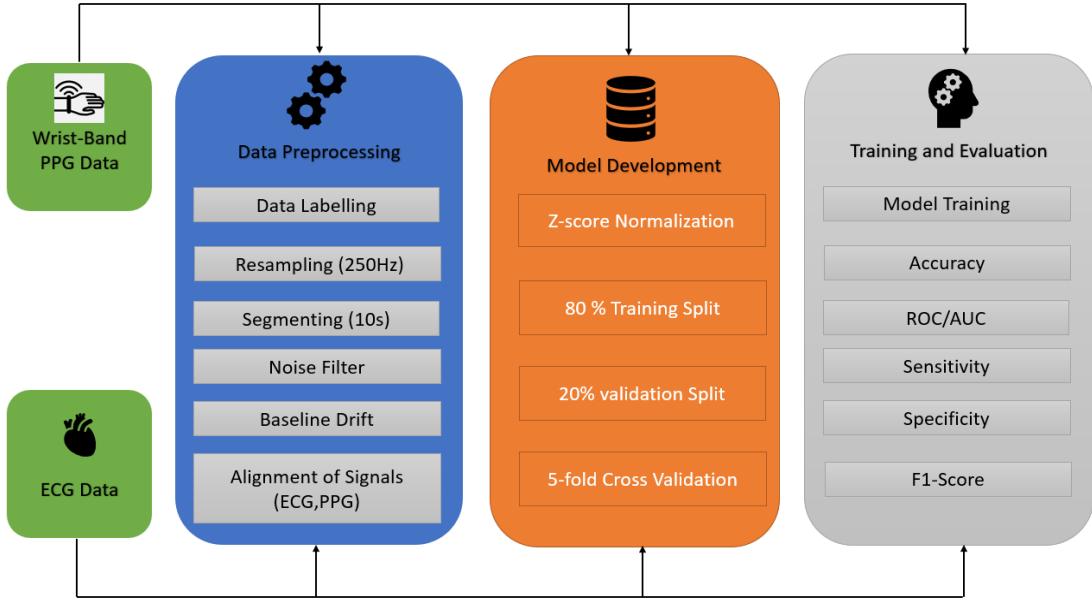


Figure 4.6: Overall Block Diagram of our Proposed Method.

For Model Development, we have divided our dataset into 3 parts. First, we have separated 20% data for testing purpose. The remaining 80% data are used for training and validation purpose. From which 80% data is used for training and 20% data is used for validation purpose. We have employed 5-fold Cross-Validation method to check the robustness and view the overall performance of our model.

4.3 Models

We have used 1D Self-ONN and 1D CNN model for the classification task. We have used several q values in Self-ONN to find the best model for the AF detection task. Here q is defined as qth order Maclaurin approximation, common values are : {1,3,5,7,9}. When q=1, then Self-ONN is equivalent to conventional CNN model. We have used 2 Self-ONN models where each of the model has four Self-ONN layers. After 4th Self-ONN layer, Model1 has 1 fully connected layer and model2 has 2 fully connected layers. The illustration of the 2 models are shown in Fig 4.7

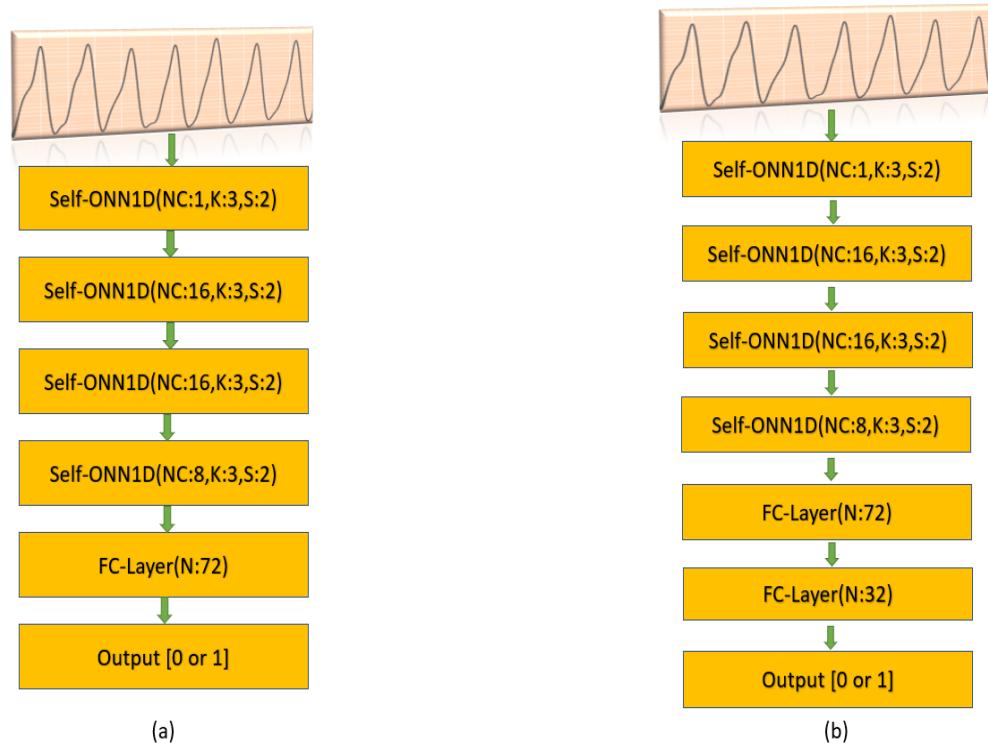


Figure 4.7: Models of Self-ONN with Two Different Structures.

At the first Self-ONN layer, number of input channel is 1 and the output channel is 16, kernel size is of 3 and stride is set as 2. kernal size and number of stride is same for all 4 operational layers. For the 2nd, 3rd and 4th layer, input channel is 16, 16 and 8 respectively. In Model1, there is a single fully connected layer of 72 neurons before the output layer shown in Fig 4.7 (a). In Model 2, there is 2 fully connected layer before the output layer which have 72 and 32 neurons sequentially. In both models, output layer has one neuron to predict whether the data sample is from AF patient or healthy person.

4.3.1 Model Training

All of the models were trained using the following model hyper-parameters shown in Table 3.1

Table 4.1: Model Hyper-Parameters Information

Model Parameter	Value
Batch_Size	16
Optimizer	SGD
Learning Rate	0.001
Number of Epoch	20
Epoch Patience	7

CHAPTER 5

RESULT AND ANALYSIS

In this chapter, we discuss and evaluate the results of our models and analyze its performance in various scenarios and test setups using a number of parameters. The testing and training phases have shown very good results both theoretically and practically. The further implications of our work and in-depth analysis of the success and failures of the result is discussed below.

5.1 Evaluation Metrics

Since a lot of the results are prediction based, we can classify them into 4 parameters. These parameters are used for the Confusion Matrix, precision, sensitivity and F1 score.

Table 5.1: Parameters used for evaluation.

Actual Class	Predicted Class	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

There must be some established factors or criteria in order to compare and examine model performance. Each prediction can be one of the four outcomes, based on how it matches up to the actual value:

1. True Positive (TP): Predicted True and True in reality.
2. True Negative (TN): Predicted False and False in reality.
3. False Positive (FP): Predicted True and False in reality.
4. False Negative (FN): Predicted False and True in reality.

Some factors or metrics that can give information about the performance of a model are specified based on the occurrences of these outcomes.

Accuracy: It is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (5.1)$$

Precision: It is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5.2)$$

Recall: It is the ratio of correctly predicted positive observations to the all observations in actual class.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5.3)$$

Specificity: It gives the capability of the system to correctly identify the proportion of true negatives predicted by the model.

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (5.4)$$

F1 Score: It simply measures the percentage of correct predictions that a machine learning model has made. F1 Score is the weighted average of Precision and Recall.

$$\text{F1Score} = \frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (5.5)$$

5.2 AUC - ROC Curve

An assessment metric for binary classification issues is the Receiver Operator Characteristic (ROC) curve. In essence, it separates the "signal" from the "noise" by plotting the TPR against the FPR at different threshold values. The capacity of a classifier to differentiate between classes is measured by the Area Under the Curve (AUC), which is used as a summary of the ROC curve. The higher the AUC, the better the performance of the model at distinguishing between the positive and negative classes.

Performance evaluation is a crucial role in machine learning. We can therefore rely on an AUC-ROC Curve when it comes to a classification task. An indicator of performance for classification issues at different threshold levels is the AUC-ROC curve. It reveals how well the model can differentiate across classes. Higher the AUC, the better the model is at predicting 0 classes as 0 and 1 classes as 1. For example, the Higher the AUC, the better the model is at distinguishing between patients with the disease and no disease.

The True Positive Rate (TPR) and False Positive Rate (FPR) are represented on the y-axis and x-axis, respectively, of the ROC curve. When AUC is between 0.5 and 1, there is a high chance that the classifier will be able to distinguish the positive class values from the negative class values. This is so because the classifier is able to detect more numbers of True positives and True negatives than False negatives and False positives. $AUC = 1$ indicates that the classifier can accurately distinguish between all Positive and Negative class points. The classifier would be predicting all Negatives as Positives and all Positives as Negatives, however, if the AUC had been 0.

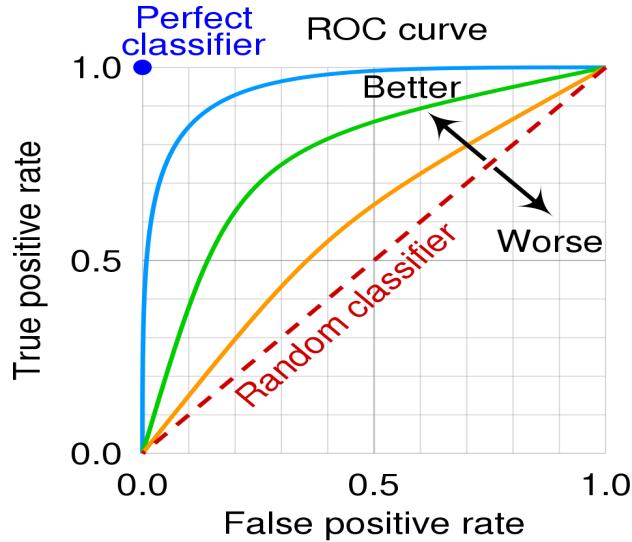


Figure 5.1: ROC Curve obtained by a classifier model

5.2.1 Confusion Matrix

A confusion matrix is a table that is used to define the performance of a classification algorithm. The effectiveness of a classification algorithm is shown and summarized via a confusion matrix. We have achieved slightly less accuracy in Self-ONN Model 2 compared to Self-ONN Model 1. Performance of both models are illustrated in Table 5.2 to 5.5. Confusion matrix for both ECG and PPG Dataset results obtained from the CNN and Self-ONN model 1 are shown below:

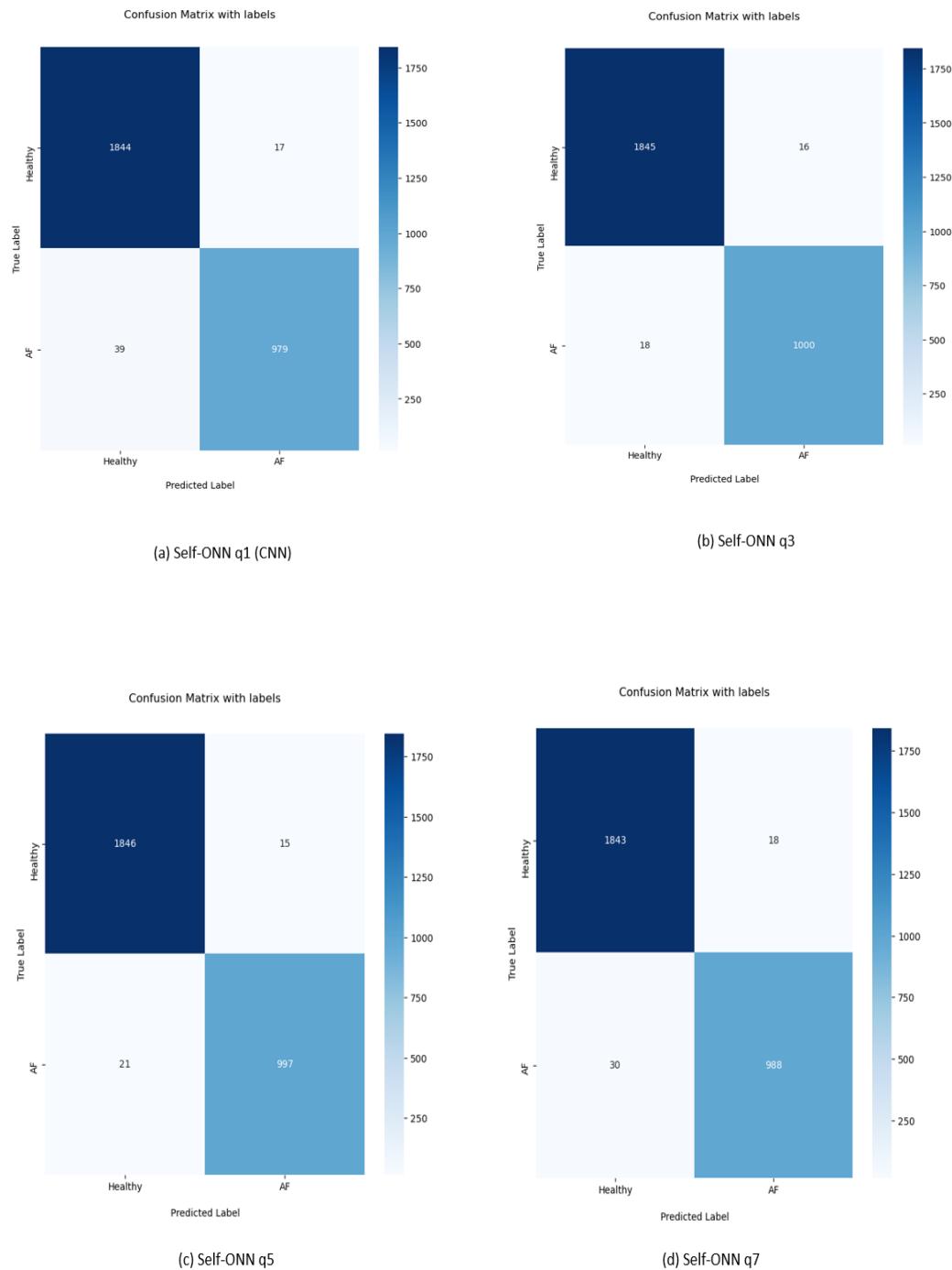


Figure 5.2: Confusion Matrix for Self-ONN Model 1 on ECG Data

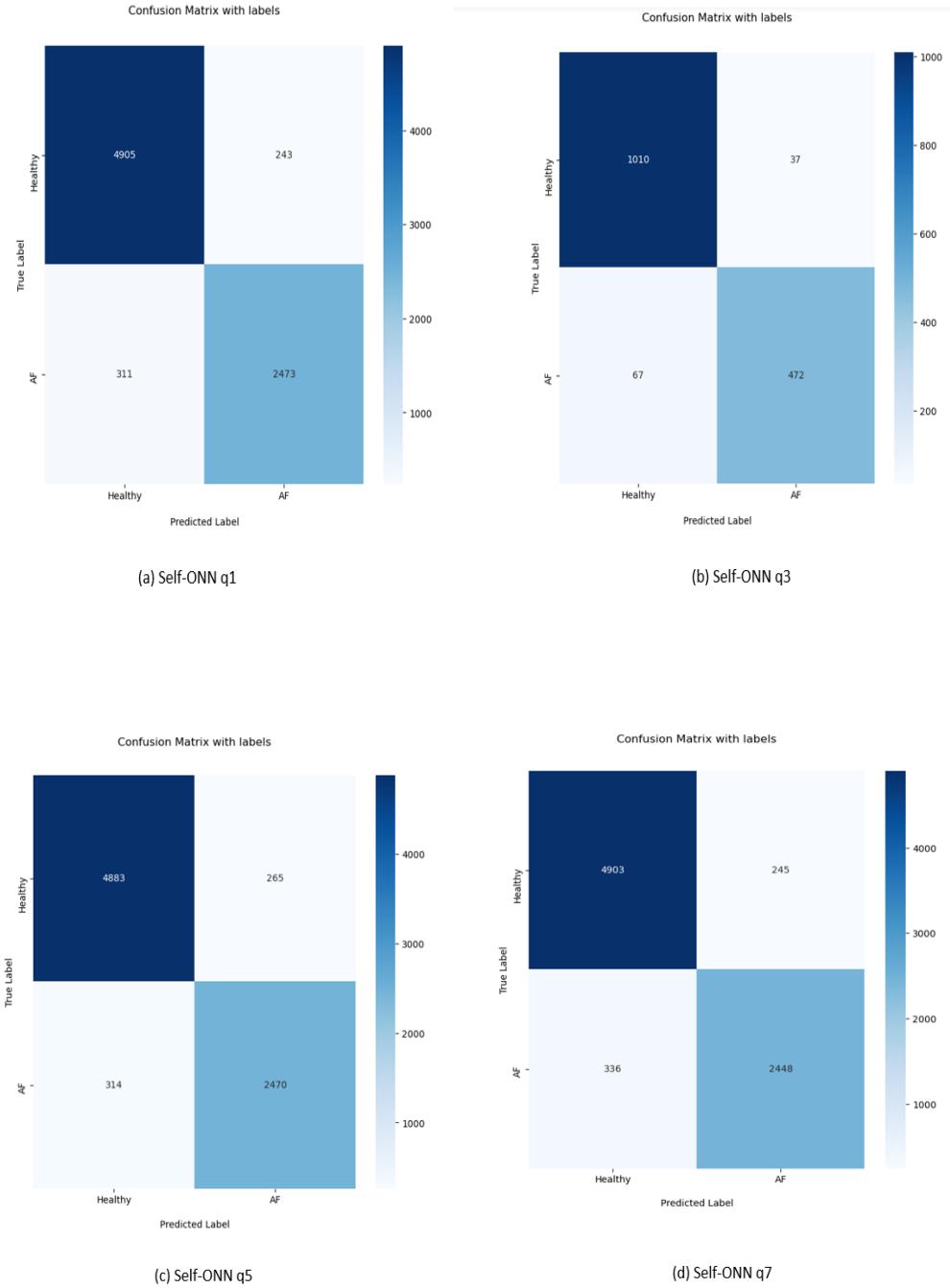
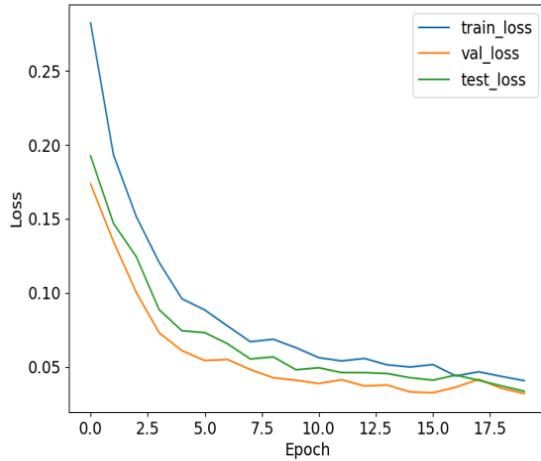


Figure 5.3: Confusion Matrix for Self-ONN Model 1 on PPG Data

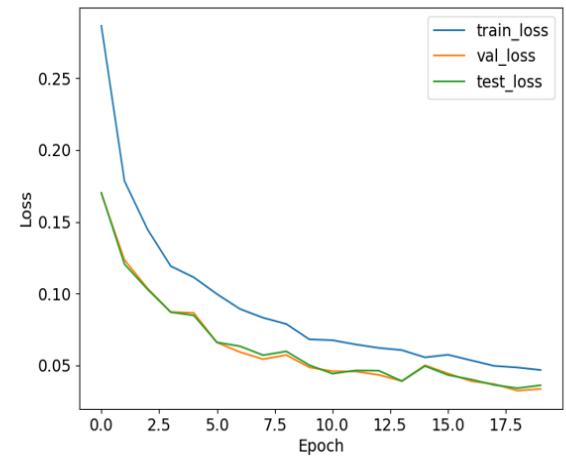
5.2.2 Training and Loss Curves

We have applied 5-fold cross validation method to test the performances of our proposed method. We have achieved highest accuracy using $q=3$ in Self-ONN

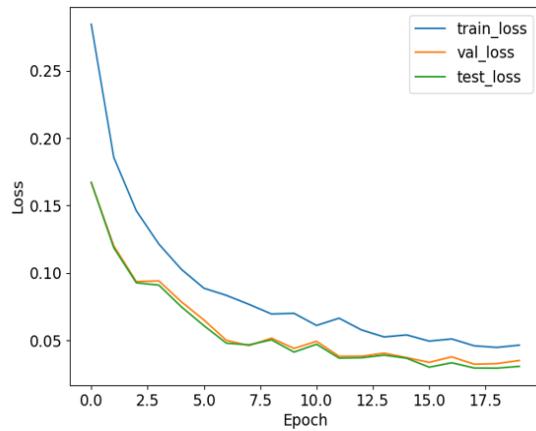
model 1. The loss and accuracy curves for both ECG and PPG data for this model are shown below



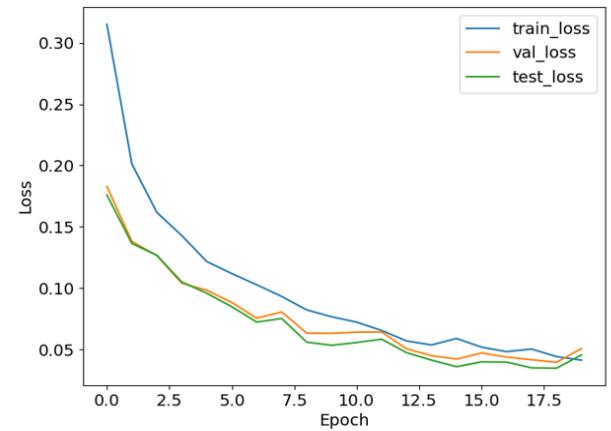
(a) Fold-1



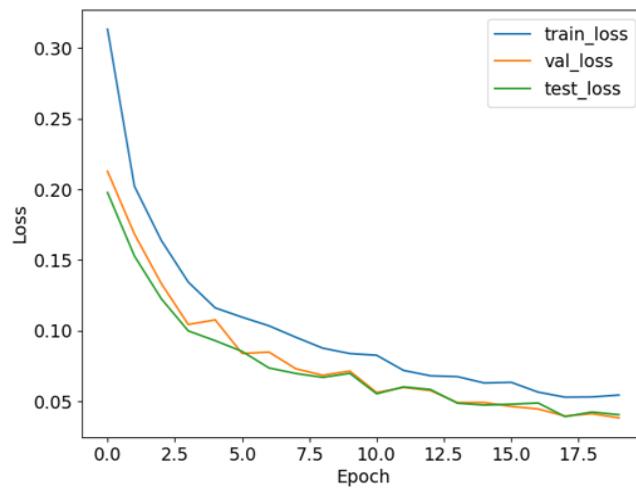
(b) Fold-2



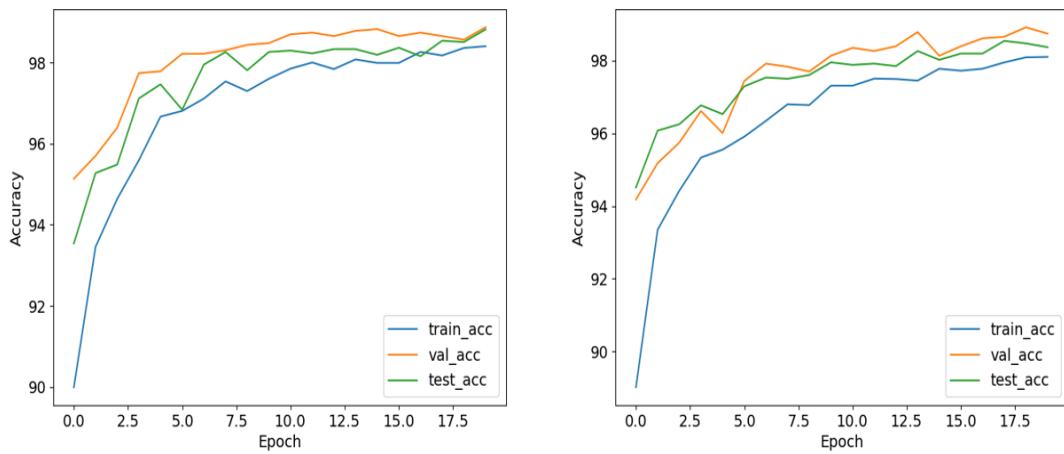
(c) Fold-3



(d) Fold-4

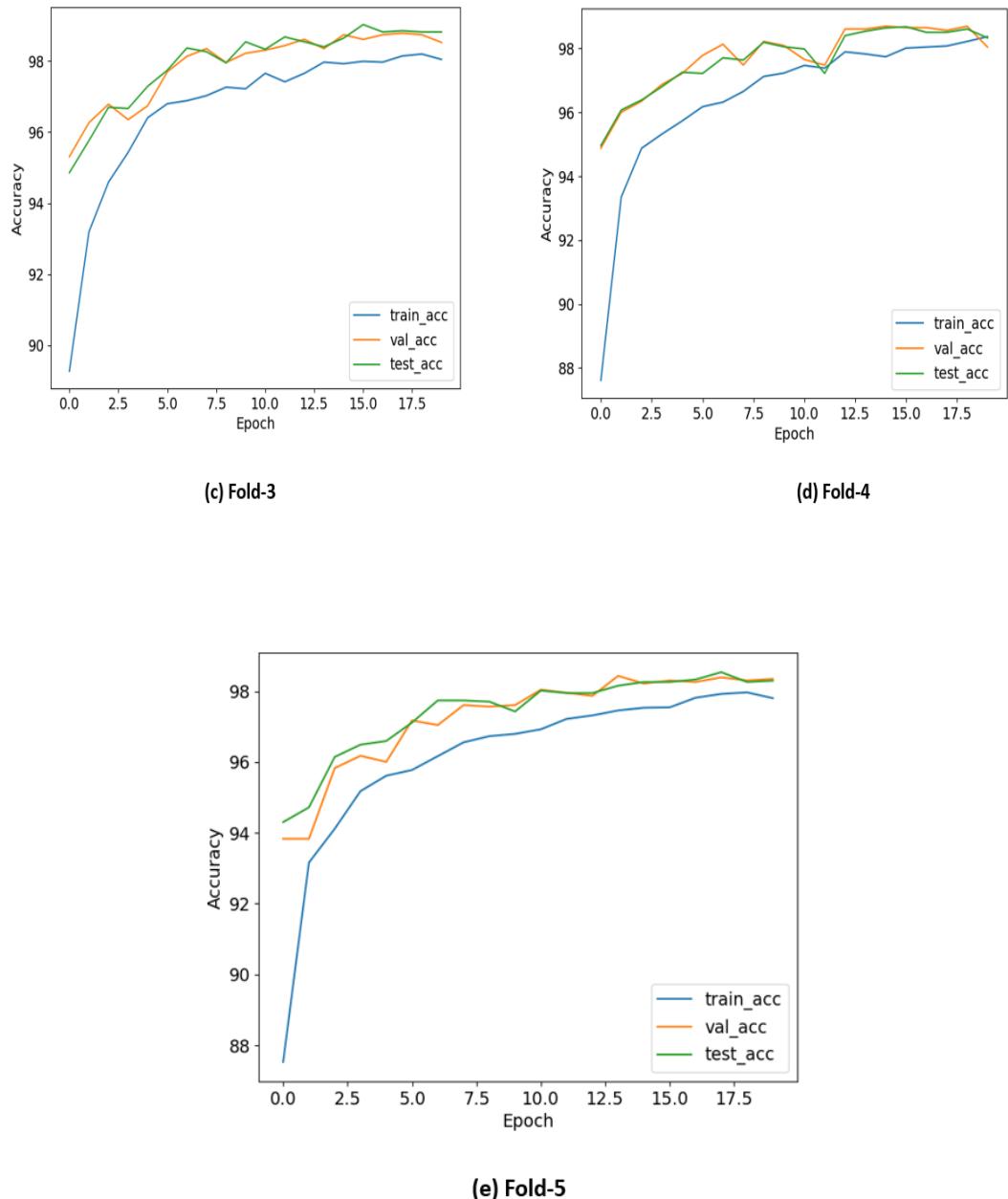


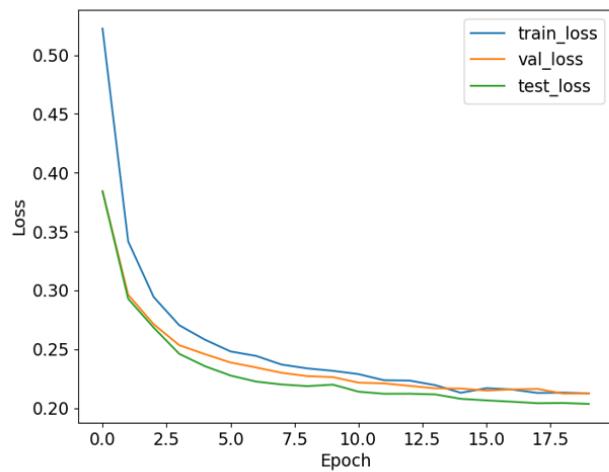
(e) Fold-5

Figure 5.4: Loss curve of Self-ONN Model 1 ($q=3$) on ECG Data

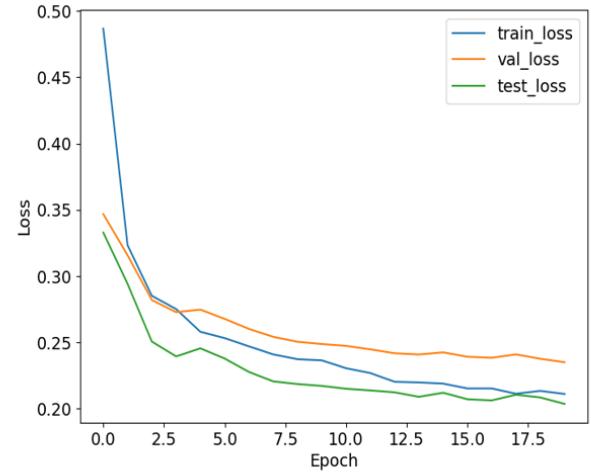
(a) Fold-1

(b) Fold-2

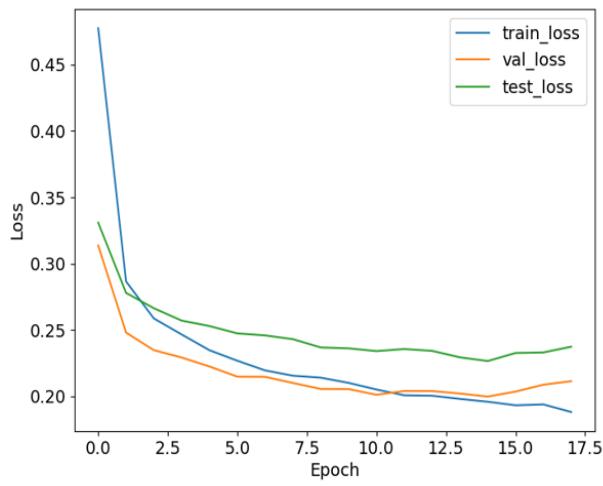
Figure 5.5: Accuracy curve of Self-ONN Model 1 ($q=3$) on ECG Data



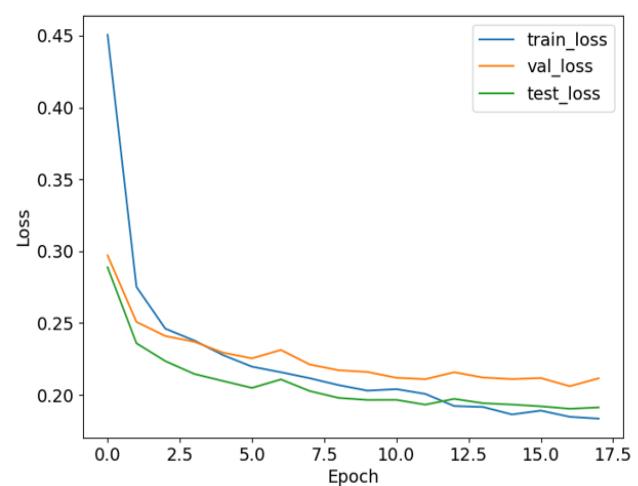
(a) Fold-1



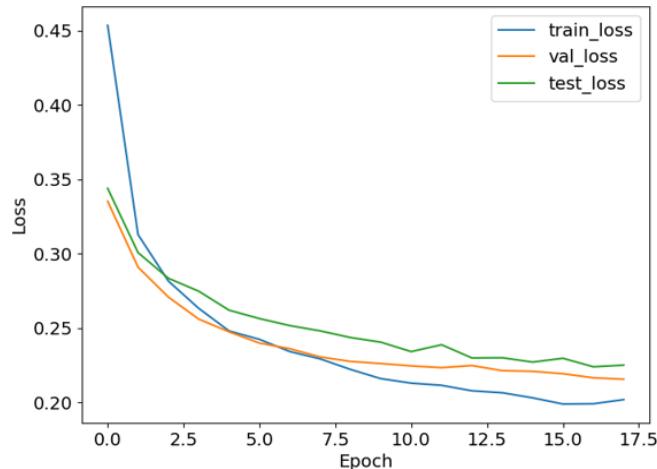
(b) Fold-2



(c) Fold-3



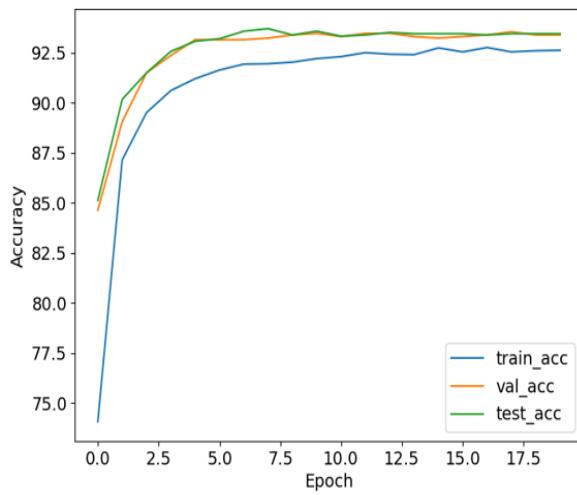
(d) Fold-4



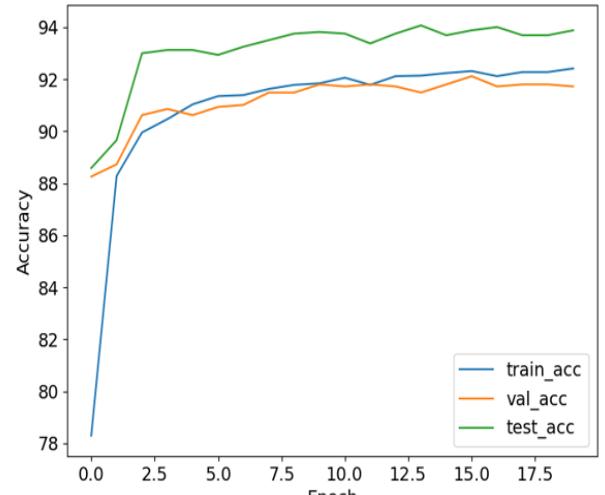
(e) Fold-5

Figure 5.6: Loss curve of Self-ONN Model 1 ($q=3$) on PPG Data

On ECG dataset, the average test accuracy of Self-ONN model 1 ($q=3$) is 98.82%. For PPG dataset, this model achieves an average test accuracy of 93.44%. In Table 5.2 and 5.3, we can evaluate the overall performance in different metrics of Self-ONN Model 1 on both ECG and PPG data for various q values.



(a) Fold-1



(b) Fold-2

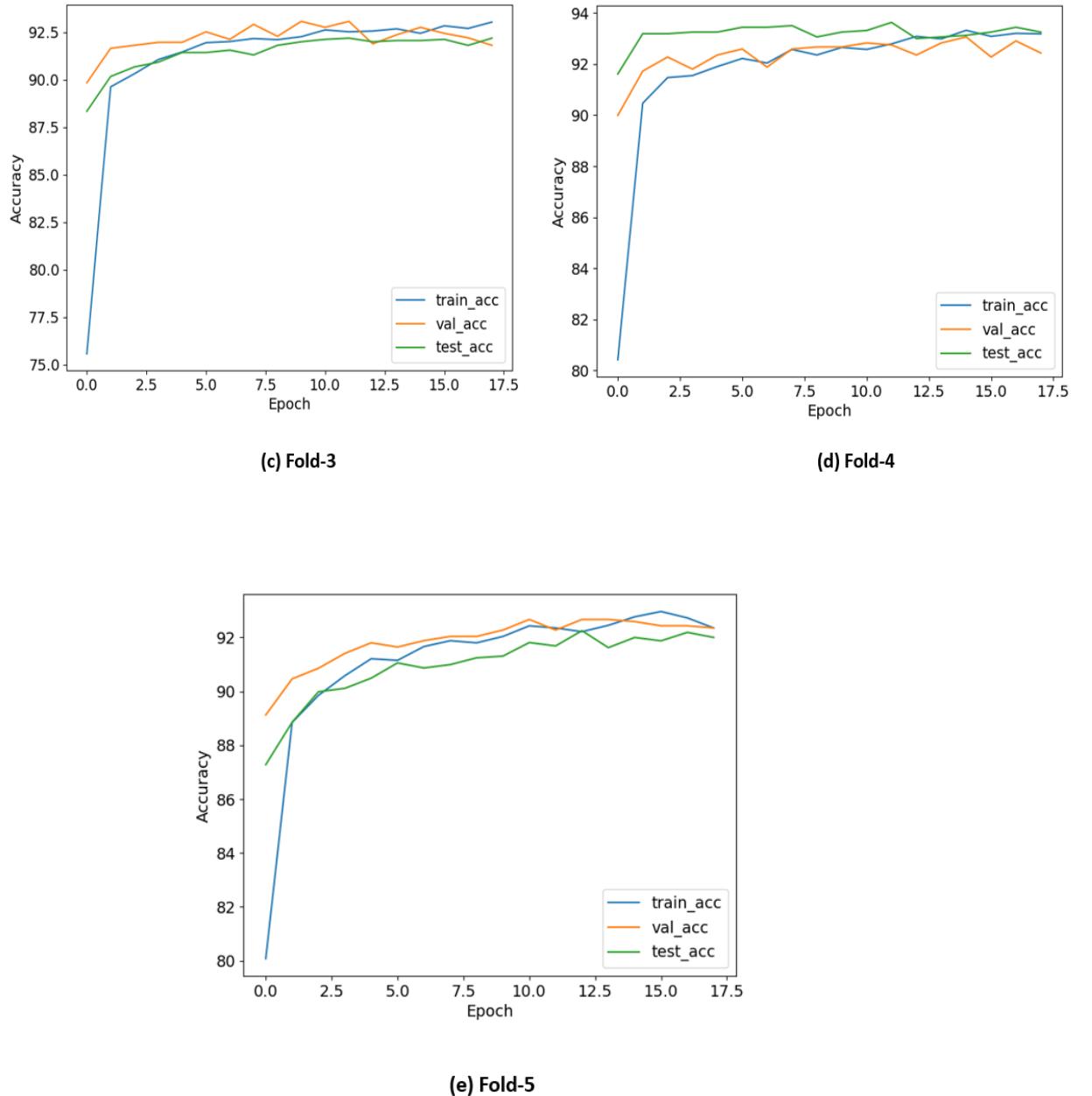


Figure 5.7: Accuracy curve of Self-ONN Model 1 ($q=3$) on PPG Data

5.2.3 ROC-AUC Curve

We have achieved an excellent AUC value of 0.999 using $q=5$ in Self-ONN on ECG Data. Our proposed method may successfully differentiate between data from healthy and AF, according to the ROC curve of ECG data. We have also got a very good AUC vaalue of 0.945 using $q=5$ on PPG Data. We have got less

accuracy on PPG Data because there are lots of noise, motion artifacts cut-off signals over a certain period.

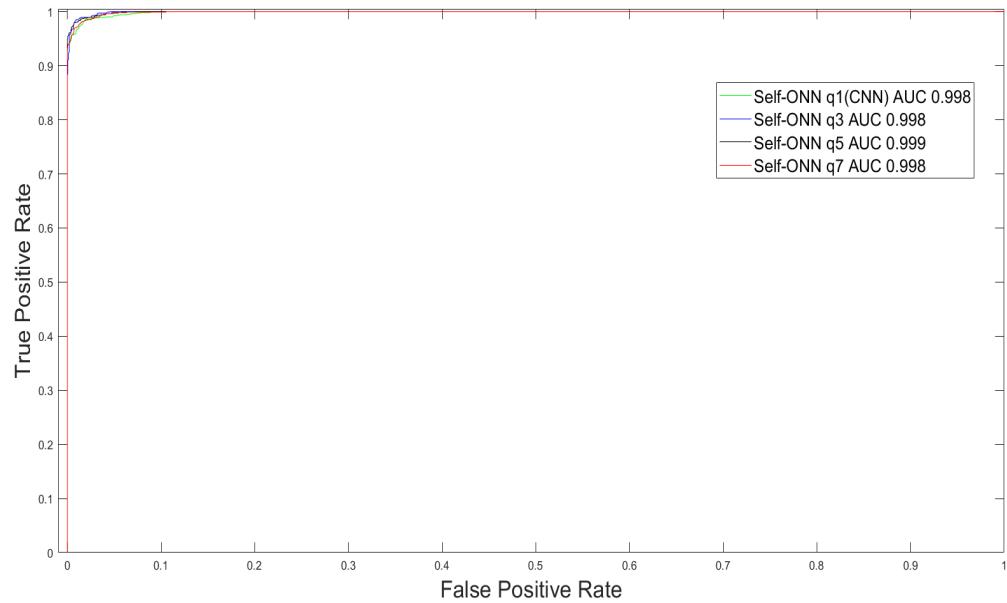


Figure 5.8: Self-ONN model 1 Performances on ECG Data

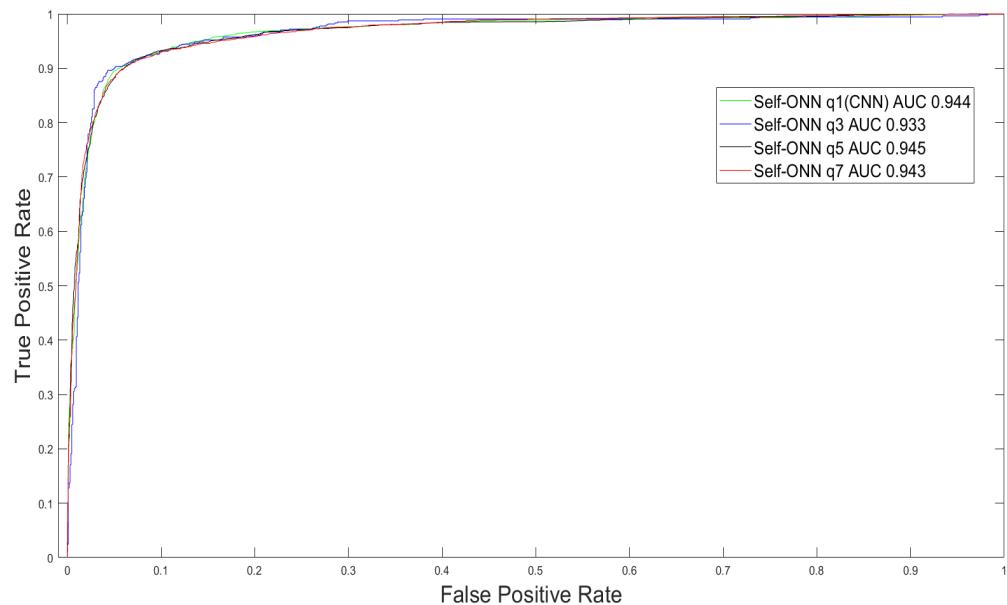


Figure 5.9: Self-ONN model 1 Performances on PPG Data

5.2.4 Results for AF detection

We have found the highest accuracy using Model 1 of Self-ONN. Table 5.2 sums up about the performance of the of the AF detection with Self-ONN Model 1 on ECG Data. We obtain highest accuracy of 98.82% specificity of 99.14% and F1 score of 98.81% using q=3. Setting q value to 3 and 5 in Self-ONN, we have obtained satisfactory results. When q=1, the Self-ONN model is equivalent to 1D-CNN model. we have also achieved promising results using 1D-CNN model.

Table 5.2: Results for AF Detection on ECG Data using Model 1

		Accuracy	Precision	Sensitivity	F1_score	Specificity
Self-ONN q=1(CNN)	Healthy	98.05	97.93	99.09	98.51	96.17
	AF		98.29	96.17	97.22	99.09
	Weighted Average		98.06	98.06	98.05	97.2
Self-ONN q=3	Healthy	98.82	99.03	99.14	99.08	98.23
	AF		98.43	98.23	98.33	99.14
	Weighted Average		98.82	98.82	98.81	98.55
Self-ONN q=5	Healthy	98.75	98.88	99.19	99.03	97.94
	AF		98.52	97.94	98.23	99.19
	Weighted Average		98.75	98.75	98.75	98.38
Self-ONN q=7	Healthy	98.33	98.4	99.03	98.71	97.05
	AF		98.21	97.05	97.63	99.03
	Weighted Average		98.33	98.33	98.33	97.75

The results of the AF detection using the Self-ONN Model-1 on PPG data are summarized in Table 5.3. With q=3, we achieve the greatest F1 score of 93.4%, specificity of 90.59% percent , and accuracy of 93.44%. We have obtained less accuracy on PPG data compared to ECG Data. It is because there is a lot of noise and motion artifacts, which can occasionally cause the signal's shape to alter. Due to high level of noise, some there are signals cut-off for a certain period of time.

The Performance of AF detection using Self-ONN Model 2 on ECG Data is summarized in Table 5.4. Utilizing q=7, we achieve the greatest accuracy of 98.47 %, specificity of 97.78 %, and F1 score of 98.47% . We have achieved good results in Self-ONN with q values set to 3 and 5. The Self-ONN model and 1D-CNN model are interchangeable when q=1. Using the 1D-CNN model, we also obtained satisfactory results.

Table 5.5 provides a summary of the performance of AF detection using the Self-ONN Model 2 on PPG Data.

Table 5.3: Results for AF Detection on PPG Data using Model 1

		Accuracy	Precision	Sensitivity	F1_score	Specificity
Self-ONN q=1(CNN)	Healthy	93.02	94.04	95.28	94.66	88.83
	AF		91.05	88.83	89.93	95.28
	Weighted Average		92.99	93.02	93	91.09
Self-ONN q=3	Healthy	93.44	93.78	96.47	95.11	87.57
	AF		92.73	87.57	90.08	96.47
	Weighted Average		93.42	93.45	93.4	90.59
Self-ONN q=5	Healthy	92.7	93.96	94.85	94.4	88.72
	AF		90.31	88.72	89.51	94.85
	Weighted Average		92.68	92.7	92.68	90.87
Self-ONN q=7	Healthy	92.68	93.59	95.24	94.41	87.93
	AF		90.9	87.93	89.39	95.24
	Weighted Average		92.65	92.67	92.65	90.5

Table 5.4: Results for AF Detection on ECG Data using Model 2

		Accuracy	Precision	Sensitivity	F1_score	Specificity
Self-ONN q=1(CNN)	Healthy	98.26	98.45	98.87	98.66	97.15
	AF		97.92	97.15	97.53	98.87
	Weighted Average		98.26	98.26	98.26	97.76
Self-ONN q=3	Healthy	98.4	98.25	99.3	98.77	96.76
	AF		98.7	96.76	97.72	99.3
	Weighted Average		98.41	98.4	98.4	97.66
Self-ONN q=5	Healthy	98.44	98.2	99.41	98.8	96.66
	AF		98.89	96.66	97.76	99.41
	Weighted Average		98.44	98.44	98.43	97.63
Self-ONN q=7	Healthy	98.47	98.35	99.3	98.82	96.95
	AF		98.7	96.95	97.82	99.3
	Weighted Average		98.47	98.47	98.47	97.78

The highest accuracy, specificity, and F1 score are obtained when q=3 is used. They are 93.04 %, 90.91% percent, and 93.01 %, respectively. With q values set to 3 and 7, we had the best results in Self-ONN. We have found overall satisfactory classification performance on both ECG and PPG Data using Self-ONN Model. We have achieved overall satisfactory classification performance on both ECG and PPG Data using Self-ONN Model.

Table 5.5: Results for AF Detection on PPG Data using Model 2

		Accuracy	Precision	Sensitivity	F1_score	Specificity
Self-ONN q=1(CNN)	Healthy	92.85	93.9	95.16	94.53	88.58
	AF		90.83	88.58	89.69	95.16
	Weighted Average		92.82	92.85	92.83	90.89
Self-ONN q=3	Healthy	93.04	93.84	95.55	94.69	88.4
	AF		91.49	88.4	89.92	95.55
	Weighted Average		93.02	93.04	93.02	90.91
Self-ONN q=5	Healthy	92.81	93.43	95.65	94.53	87.57
	AF		91.59	87.57	89.53	95.65
	Weighted Average		92.78	92.81	92.78	90.41
Self-ONN q=7	Healthy	92.91	93.49	95.75	94.61	87.68
	AF		91.77	87.68	89.68	95.75
	Weighted Average		92.89	92.92	92.88	90.51

CHAPTER 6

CONCLUSION AND FUTURE WORKS

This study suggested the design and implementation of an explainable deep learning Self-ONN and 1D-CNN model for application in smart healthcare systems utilizing general-purpose devices such as smart wearables and smartphones. Self-ONN with the Maclaurin approximation, $q = 3$, performed the best on both ECG and PPG signals using a 5-fold cross validation approach. Furthermore, the suggested methodology reduces noise and motion artifacts from commercial PPG-sensors within a framework for health monitoring. Photoplethysmography (PPG) is a simple, low-cost, non-invasive approach for measuring blood vessel pulse waves using an electro-optic technology. Though the current study is focused on AF monitoring, PPG signals may also be used to measure blood oxygen levels, systolic blood pressure (BP), cardiac output, and respiratory rate. For feature extraction, analysis, and parameter estimation, all of these applications require a clean and improved PPG signal. It is a very challenging task to acquire a PPG signal free of motion artifacts (MA) caused by the subject's deliberate or involuntary movement. Long-term wrist-band PPG recording does not allow for the acquisition of PPG signals devoid of many types of noise and motion artifacts, particularly the cut-off signal condition. Despite the fact that we used many signal processing techniques to eliminate MA and signal cut-off conditions, the results were not always satisfactory in some circumstances. Using Cycle-GAN, we can eliminate this type of difficulty. The Cycle Generative Adversarial Network, or Cycle-GAN, is a method for training a deep convolutional neural network for specially image-to-image translation problems. Using an unpaired dataset, the

Network learns the mapping between input and output pictures. Authors in [71] has already converted noisy and MAs ECG signals to cleaned and good quality ECG signals. Since this model does not require paired dataset, we can transform bad quality and noisy signals to clean and good quality PPG signals in future.

BIBLIOGRAPHY

- [1] P. Cheng, Z. Chen, Q. Li, Q. Gong, J. Zhu, and Y. Liang, “Atrial fibrillation identification with ppg signals using a combination of time-frequency analysis and deep learning,” *IEEE Access*, vol. 8, pp. 172 692–172 706, 2020.
- [2] J. Ramesh, Z. Solatidehkordi, R. Aburukba, and A. Sagahyoon, “Atrial fibrillation classification with smart wearables using short-term heart rate variability and deep convolutional neural networks,” *Sensors*, vol. 21, no. 21, p. 7233, 2021.
- [3] H. Fukui, T. Yamashita, Y. Yamauchi, H. Fujiyoshi, and H. Murase, “Pedestrian detection based on deep convolutional neural network with ensemble inference network,” in *2015 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2015, pp. 223–228.
- [4] S. Björck, B. Palaszewski, L. Friberg, and L. Bergfeldt, “Atrial fibrillation, stroke risk, and warfarin therapy revisited: a population-based study,” *Stroke*, vol. 44, no. 11, pp. 3103–3108, 2013.
- [5] H. Lieder, G. Breithardt, and G. Heusch, “Fatal attraction—a brief pathophysiology of the interaction between atrial fibrillation and myocardial ischemia,” *International Journal of Cardiology*, vol. 254, pp. 132–135, 2018.
- [6] S. Kiranyaz, T. Ince, A. Iosifidis, and M. Gabbouj, “Operational neural networks,” *Neural Computing and Applications*, vol. 32, no. 11, pp. 6645–6668, 2020.
- [7] S. Kiranyaz, T. Malik, Ince, A. Iosifidis, and M. Gabbouj, “Generalized model of biological neural networks: progressive operational perceptrons.”

- [8] S. S. Barold, “Willem einthoven and the birth of clinical electrocardiography a hundred years ago,” *Cardiac electrophysiology review*, vol. 7, no. 1, pp. 99–104, 2003.
- [9] S. Kiranyaz, J. Malik, H. B. Abdallah, T. Ince, A. Iosifidis, and M. Gabbouj, “Self-organized operational neural networks with generative neurons,” *Neural Networks*, vol. 140, pp. 294–308, 2021.
- [10] J. Malik, S. Kiranyaz, and M. Gabbouj, “Self-organized operational neural networks for severe image restoration problems,” *Neural Networks*, vol. 135, pp. 201–211, 2021.
- [11] S. Kiranyaz, H. B. Malik, T. Ince, A. Iosifidis, and M. Gabbouj, “Progressive operational perceptrons,” *Neurocomputing*, vol. 224, pp. 142–154, 2017.
- [12] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals,” *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [13] C. A. Millán, N. A. Girón, and D. M. Lopez, “Analysis of relevant features from photoplethysmographic signals for atrial fibrillation classification,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 2, p. 498, 2020.
- [14] G. D. Clifford, C. Liu, B. Moody, H. L. Li-wei, I. Silva, Q. Li, A. Johnson, and R. G. Mark, “Af classification from a short single lead ecg recording: The physionet/computing in cardiology challenge 2017,” in *2017 Computing in Cardiology (CinC)*. IEEE, 2017, pp. 1–4.
- [15]
- [16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [17] D. D. McManus, J. W. Chong, A. Soni, J. S. Saczynski, N. Esa, C. Napolitano, C. E. Darling, E. Boyer, R. K. Rosen, K. C. Floyd et al., “Pulse-smart:

- pulse-based arrhythmia discrimination using a novel smartphone application,” Journal of cardiovascular electrophysiology, vol. 27, no. 1, pp. 51–57, 2016.
- [18] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1299–1312, 2016.
- [19] Q. Wang, J. Wan, and X. Li, “Robust hierarchical deep learning for vehicular management,” IEEE Transactions on Vehicular Technology, vol. 68, no. 5, pp. 4148–4156, 2018.
- [20] R. Assabumrungrat, S. Sangnark, T. Charoenpattarawut, W. Polpakdee, T. Sudhawiyangkul, E. Boonchieng, and T. Wilairasitporn, “Ubiquitous affective computing: A review,” IEEE Sensors Journal, 2021.
- [21] H. He, C.-K. Wen, S. Jin, and G. Y. Li, “Model-driven deep learning for mimo detection,” IEEE Transactions on Signal Processing, vol. 68, pp. 1702–1715, 2020.
- [22] S. Nemati, M. M. Ghassemi, V. Ambai, N. Isakadze, O. Levantsevych, A. Shah, and G. D. Clifford, “Monitoring and detecting atrial fibrillation using wearable technology,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 3394–3397.
- [23] D. Wallmann, D. Tüller, N. Kucher, J. Fuhrer, M. Arnold, and E. Delacretaz, “Frequent atrial premature contractions as a surrogate marker for paroxysmal atrial fibrillation in patients with acute ischaemic stroke,” Heart, vol. 89, no. 10, pp. 1247–1248, 2003.
- [24] M. Ksiāczyk, A. Debska-Kozłowska, I. Warchoł, A. Lubiński et al., “Enhancing healthcare access—smartphone apps in arrhythmia screening,” JMIR mHealth and uHealth, vol. 9, no. 8, p. e23425, 2021.
- [25] S. Purushotham, C. Meng, Z. Che, and Y. Liu, “Benchmarking deep learning models on large healthcare datasets,” Journal of biomedical informatics, vol. 83, pp. 112–134, 2018.

- [26] Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, “Multi-class arrhythmia detection from 12-lead varied-length ecg using attention-based time-incremental convolutional neural network,” *Information Fusion*, vol. 53, pp. 174–182, 2020.
- [27] B. A. Schoonderwoerd, M. D. Smit, L. Pen, and I. C. Van Gelder, “New risk factors for atrial fibrillation: causes of ‘not-so-lone atrial fibrillation’,” *Europace*, vol. 10, no. 6, pp. 668–673, 2008.
- [28] A. Aizer, J. M. Gaziano, N. R. Cook, J. E. Manson, J. E. Buring, and C. M. Albert, “Relation of vigorous exercise to risk of atrial fibrillation,” *The American journal of cardiology*, vol. 103, no. 11, pp. 1572–1577, 2009.
- [29] D. R. Seshadri, B. Bittel, D. Browsky, P. Houghtaling, C. K. Drummond, M. Y. Desai, and A. M. Gillinov, “Accuracy of apple watch for detection of atrial fibrillation,” *Circulation*, vol. 141, no. 8, pp. 702–703, 2020.
- [30] Z. I. Attia, P. A. Noseworthy, F. Lopez-Jimenez, S. J. Asirvatham, A. J. Deshmukh, B. J. Gersh, R. E. Carter, X. Yao, A. A. Rabinstein, B. J. Erickson et al., “An artificial intelligence-enabled ecg algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction,” *The Lancet*, vol. 394, no. 10201, pp. 861–867, 2019.
- [31] H. Njoum and P. Kyriacou, “Investigation of finger reflectance photoplethysmography in volunteers undergoing a local sympathetic stimulation,” in *Journal of Physics: Conference Series*, vol. 450, no. 1. IOP Publishing, 2013, p. 012012.
- [32] T. Pereira, I. Santos, T. Oliveira, P. Vaz, T. Pereira, H. Santos, H. Pereira, C. Correia, and J. Cardoso, “Pulse pressure waveform estimation using distension profiling with contactless optical probe,” *Medical Engineering & Physics*, vol. 36, no. 11, pp. 1515–1520, 2014.
- [33] S. P. Shashikumar, A. J. Shah, G. D. Clifford, and S. Nemati, “Detection of paroxysmal atrial fibrillation using attention-based bidirectional recurrent neural networks,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 715–723.

- [34] Y. Liang, Z. Chen, R. Ward, and M. Elgendi, “Photoplethysmography and deep learning: enhancing hypertension risk stratification,” *Biosensors*, vol. 8, no. 4, p. 101, 2018.
- [35] A. Ciccone and H.-T. Wu, “How nonlinear-type time-frequency analysis can help in sensing instantaneous heart rate and instantaneous respiratory rate from photoplethysmography in a reliable way,” *Frontiers in physiology*, vol. 8, p. 701, 2017.
- [36] J. W. Chong, D. D. McManus, and K. H. Chon, “Arrhythmia discrimination using a smart phone,” in *2013 IEEE International Conference on Body Sensor Networks*. IEEE, 2013, pp. 1–4.
- [37] J. W. Chong, N. Esa, D. D. McManus, and K. H. Chon, “Arrhythmia discrimination using a smart phone,” *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 815–824, 2015.
- [38] A. Reiss, P. Schmidt, I. Indlekofer, and K. Van Laerhoven, “Ppg-based heart rate estimation with time-frequency spectra: A deep learning approach,” in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018, pp. 1283–1292.
- [39] R. Raj, C. Kattankulathur, J. Selvakumar, and M. Anburajan, “Evaluation of hypotension using wavelet and time frequency analysis of photoplethysmography (ppg) signal,” in *International Conference on Advances in Computational Intelligence in Communication*, vol. 14, 2016, pp. 57–61.
- [40] L. Krivoshei, S. Weber, T. Burkard, A. Maseli, N. Brasier, M. Kühne, D. Connen, T. Huebner, A. Seeck, and J. Eckstein, “Smart detection of atrial fibrillation,” *EP Europace*, vol. 19, no. 5, pp. 753–757, 2017.
- [41] S. K. Bashar, D. Han, S. Hajeb-Mohammadipour, E. Ding, C. Whitcomb, D. D. McManus, and K. H. Chon, “Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches,” *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

- [42] J. Nalepa and M. Kawulok, “Selecting training sets for support vector machines: a review,” *Artificial Intelligence Review*, vol. 52, no. 2, pp. 857–900, 2019.
- [43] S. K. Bashar, D. Han, A. Soni, D. D. McManus, and K. H. Chon, “Developing a novel noise artifact detection algorithm for smartphone ppg signals: Preliminary results,” in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE, 2018, pp. 79–82.
- [44] J. W. Chong, C. H. Cho, F. Tabei, D. Le-Anh, N. Esa, D. D. McManus, and K. H. Chon, “Motion and noise artifact-resilient atrial fibrillation detection using a smartphone,” *IEEE journal on emerging and selected topics in circuits and systems*, vol. 8, no. 2, pp. 230–239, 2018.
- [45] R. Avram, M. Ramsis, A. D. Cristal, V. Nathan, L. Zhu, J. Kim, J. Kuang, A. Gao, E. Vittinghoff, L. Rohdin-Bibby et al., “Validation of an algorithm for continuous monitoring of atrial fibrillation using a consumer smartwatch,” *Heart Rhythm*, vol. 18, no. 9, pp. 1482–1490, 2021.
- [46] E. Hariri, A. Hachem, G. Sarkis, and S. Nasr, “Optimal duration of monitoring for atrial fibrillation in cryptogenic stroke: a nonsystematic review,” *BioMed research international*, vol. 2016, 2016.
- [47] C. A. Morillo, A. Banerjee, P. Perel, D. Wood, and X. Jouven, “Atrial fibrillation: the current epidemic,” *Journal of geriatric cardiology: JGC*, vol. 14, no. 3, p. 195, 2017.
- [48] G. Hindricks, T. Potpara, N. Dagres, E. Arbelo, J. J. Bax, C. Blomström-Lundqvist, G. Boriani, M. Castella, G.-A. Dan, P. E. Dilaveris et al., “2020 esc guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the european association for cardio-thoracic surgery (eacts) the task force for the diagnosis and management of atrial fibrillation of the european society of cardiology (esc) developed with the special contribution of the european heart rhythm association (ehra) of the esc,” *European heart journal*, vol. 42, no. 5, pp. 373–498, 2021.
- [49] S. S. Chugh, R. Havmoeller, K. Narayanan, D. Singh, M. Rienstra, E. J. Benjamin, R. F. Gillum, Y.-H. Kim, J. H. McAnulty Jr, Z.-J. Zheng et al.,

- “Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study,” *Circulation*, vol. 129, no. 8, pp. 837–847, 2014.
- [50] L. Friberg and L. Bergfeldt, “Atrial fibrillation prevalence revisited,” *Journal of internal medicine*, vol. 274, no. 5, pp. 461–468, 2013.
- [51] J. Chovančík, V. Bulková, M. Fiala, J. Gandalovičová, S. Královec, R. Neuwirth, H. Tolaszová, O. Jiravský, J. Brada, and J. Januška, “A comparison of two methods of long-term external ecg telemonitoring in patients after ablation for atrial fibrillation,” *Vnitřní lekarství*, vol. 58, no. 9, pp. 633–639, 2012.
- [52] K. Lee, H. O. Choi, S. D. Min, J. Lee, B. B. Gupta, and Y. Nam, “A comparative evaluation of atrial fibrillation detection methods in koreans based on optical recordings using a smartphone,” *IEEE Access*, vol. 5, pp. 11 437–11 443, 2017.
- [53] G. Borianì and L. Padeletti, “Management of atrial fibrillation in bradycardias,” *Nature Reviews Cardiology*, vol. 12, no. 6, pp. 337–349, 2015.
- [54] S. Nemati, M. M. Ghassemi, V. Ambai, N. Isakadze, O. Levantsevych, A. Shah, and G. D. Clifford, “Monitoring and detecting atrial fibrillation using wearable technology,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 3394–3397.
- [55] S.-C. Tang, P.-W. Huang, C.-S. Hung, S.-M. Shan, Y.-H. Lin, J.-S. Shieh, D.-M. Lai, A.-Y. Wu, and J.-S. Jeng, “Identification of atrial fibrillation by quantitative analyses of fingertip photoplethysmogram,” *Scientific reports*, vol. 7, no. 1, pp. 1–7, 2017.
- [56] M. J. C. Büschges, “New carotid stiffness population centiles in the young and association with measures of general and abdominal obesity background: Data on carotid stiffness (cs) in unselected adolescents and young adults are scarce. to validly investigate associations with established risk factors, centiles are needed by age and growth. although evidence has accumulated linking obesity in the young to elevated cimt, studies rarely included cs or compare obesity parameters.”

- [57] S. Fallet, M. Lemay, P. Renevey, C. Leupi, E. Pruvot, and J.-M. Vesin, “Can one detect atrial fibrillation using a wrist-type photoplethysmographic device?” *Medical & biological engineering & computing*, vol. 57, no. 2, pp. 477–487, 2019.
- [58] J. M. Bumgarner, C. T. Lambert, A. A. Hussein, D. J. Cantillon, B. Baranowski, K. Wolski, B. D. Lindsay, O. M. Wazni, and K. G. Tarakji, “Smart-watch algorithm for automated detection of atrial fibrillation,” *Journal of the American College of Cardiology*, vol. 71, no. 21, pp. 2381–2388, 2018.
- [59] L. M. Eerikäinen, A. G. Bonomi, F. Schipper, L. R. Dekker, H. M. de Morree, R. Vullings, and R. M. Aarts, “Detecting atrial fibrillation and atrial flutter in daily life using photoplethysmography data,” *IEEE journal of biomedical and health informatics*, vol. 24, no. 6, pp. 1610–1618, 2019.
- [60] A. Kashiwa, F. Koyama, K. Miyamoto, T. Kamakura, M. Wada, K. Yamagata, K. Ishibashi, Y. Inoue, H. Okamura, S. Nagase et al., “Performance of an atrial fibrillation detection algorithm using continuous pulse wave monitoring,” *Annals of Noninvasive Electrocardiology*, vol. 24, no. 2, p. e12615, 2019.
- [61] S. K. Bashar, D. Han, S. Hajeb-Mohammadalipour, E. Ding, C. Whitcomb, D. D. McManus, and K. H. Chon, “Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches,” *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.
- [62] S. Nemati, M. M. Ghassemi, V. Ambai, N. Isakadze, O. Levantsevych, A. Shah, and G. D. Clifford, “Monitoring and detecting atrial fibrillation using wearable technology,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 3394–3397.
- [63] X. Jin, L. Wu, X. Li, X. Zhang, J. Chi, S. Peng, S. Ge, G. Zhao, and S. Li, “Ilgnnet: inception modules with connected local and global features for efficient image aesthetic quality classification using domain adaptation,” *IET Computer Vision*, vol. 13, no. 2, pp. 206–212, 2019.

- [64] J. Rajeswari and M. Jagannath, "Advances in biomedical signal and image processing—a systematic review," *Informatics in Medicine Unlocked*, vol. 8, pp. 13–19, 2017.
- [65] K. Lee, H. O. Choi, S. D. Min, J. Lee, B. B. Gupta, and Y. Nam, "A comparative evaluation of atrial fibrillation detection methods in koreans based on optical recordings using a smartphone," *IEEE Access*, vol. 5, pp. 11 437–11 443, 2017.
- [66] C. Timmermans, L.-M. RODRIGUEZ, J. L. Smeets, and H. J. Wellens, "Immediate reinitiation of atrial fibrillation following internal atrial defibrillation," *Journal of Cardiovascular electrophysiology*, vol. 9, no. 2, pp. 122–128, 1998.
- [67] D. Wallmann, D. Tuller, K. Wustmann, P. Meier, J. Isenegger, M. Arnold, H. P. Mattle, and E. Delacrétaz, "Frequent atrial premature beats predict paroxysmal atrial fibrillation in stroke patients: an opportunity for a new diagnostic strategy," *Stroke*, vol. 38, no. 8, pp. 2292–2294, 2007.
- [68] K. Aschbacher, D. Yilmaz, Y. Kerem, S. Crawford, D. Benaron, J. Liu, M. Eaton, G. H. Tison, J. E. Ogin, Y. Li et al., "Atrial fibrillation detection from raw photoplethysmography waveforms: A deep learning application," *Heart rhythm O2*, vol. 1, no. 1, pp. 3–9, 2020.
- [69] J. Bacevičius, Abramikas, and I. Badaras, "Long-term electrocardiogram and wrist-based photoplethysmogram recordings with annotated atrial fibrillation episodes," Feb. 2022, The collection of data was supported by the European Regional Development Fund (01.2.2-LMT-K-718-03-0027) under grant agreement with the Research Council of Lithuania (LMLT). [Online]. Available: <https://doi.org/10.5281/zenodo.5815074>
- [70] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, "Proceedings of the royal society of london a: mathematical, physical and engineering sciences," 1998.
- [71] S. Kiranyaz, O. C. Devecioglu, T. Ince, J. Malik, M. Chowdhury, T. Hamid, R. Mazhar, A. Khandakar, A. Tahir, T. Rahman et al., "Blind ecg restoration by operational cycle-gans," *arXiv preprint arXiv:2202.00589*, 2022.