

WPPG Cycle-GAN: Robust Atrial Fibrillation Detection Using Restored Wrist PPG and ECG Signals Integrating 1D Cycle-GAN and Self-MobileNet

(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: 100543

Exam Roll: 100544

Reg No: 2017214832

Reg No: 2017514857

Session: 2017-18

Session: 2017-18

**DEPT. OF ELECTRICAL AND ELECTRONIC
ENGINEERING UNIVERSITY OF DHAKA**

January 10, 2023

Abstract

Atrial fibrillation (AF) increases the risk of stroke and causes damage to the heart's ability to pump blood. The most accurate way to identify AF is through an electrocardiogram (ECG). However, ECGs have short boards, shorter monitoring cycles, challenges with data collection, issues with screening, and sex-specific criteria. For these drawbacks, photoplethysmography (PPG), a non-invasive optical device that is simple to use and appropriate for long-term monitoring, is becoming more popular. We have developed a novel approach to detect AF from smart watch PPG signals that incorporates a Cycle-GAN and 1D Self-MobileNet model with several Self-ONN layers for classification. To the best of our knowledge, this is the first study to employ Cycle-GAN for reconstruction of PPG signals since the quality of wrist PPG signals is very poor due to motion artifacts and mismatch of devices. We have used 21,278 10-second wrist PPG segments to verify our proposed approach. Our classification pipeline with Cycle-GAN achieves an accuracy of 96.41% and 97.09% for PPG test segments 1 and 2, respectively. Using reconstructed PPG signals, we have achieved a significant improvement in accuracy of 2.94% and 5.08% as compared to raw PPG segments for test segments 1 and 2, respectively. We have also performed AF detection task using ECG signals that matched to PPG segments to assess the validity of employing reconstructed PPG based AF classification. We have achieved an accuracy of 98.07% and 98.97% using ECG segments, which is extremely comparable to the performance of AF detection utilizing reconstructed wrist PPG segments.

CONTENTS

Abstract	i
List of Figures	iv
List of Tables	vi
1 INTRODUCTION	1
1.1 AF Classification Studies	2
1.2 Reconstruction of Wrist PPG	3
1.3 Project Overview:	4
2 RELATED WORKS	6
2.1 AF Detection Methodology	6
2.1.1 Machine Learning Based AF Detection	7
2.1.2 Deep Learning-Based AF Detection	9
3 THEORETICAL OVERVIEW	13
3.1 Sensor Description:	13
3.2 Dataset Description	15
3.3 Data Processing Steps	17
3.3.1 Resolving Drifting	18
3.3.2 Noise Filtering	19
3.3.3 Data Synchronization	21
3.3.4 Baseline Wander	23
4 METHODOLOGY	25
4.1 Data Pre-processing	25
4.2 Self-ONN	26
4.3 Self-MobileNet	28
4.4 Cycle-GAN	30
5 RESULT AND ANALYSIS	34
5.1 Evaluation Metrics	34
5.2 AUC - ROC Curve	36
5.3 Experimental Results and Evaluation	37

5.3.1	Dataset	37
5.3.2	Invalid Signal Detection	38
5.3.3	Evaluation	38
5.3.4	ROC-AUC Curve	43
5.3.5	Quantitative Evaluation	43
6	CONCLUSION AND FUTURE WORKS	45
	Bibliography	47

LIST OF FIGURES

2.1	AF Detection using Pulse Rate Variability.	7
2.2	Block Diagram of the Proposed Method.	8
2.3	Complete overview of AF detection using ECG Data.	9
2.4	Proposed model of AF detection by DL method.	10
2.5	Complete DL framework of AF diagnosis.	11
2.6	Deep learning 1D-CNN model on ECG data [1].	12
3.1	(a) Bittium Faros™ 180 ECG device, reference ECG data acquisition protocol; (b) A wrist-worn device with an embedded green LED for PPG data acquisition. Both of them have embedded acceleration sensors.	14
3.3	Change in PPG frequency (100 Hz) due to drifting.	19
3.4	Responses of the filter.	20
3.5	Noise removal of PPG and ECG Data.	21
3.6	Alignment of PPG, VPG, APG, and ECG Signals.	23
4.1	ECG signals after pre-processing.	26
4.2	PPG signals after pre-processing.	26
4.3	Self-ONN operation on an input vector with pooling operator P and nodal operator Ψ	28
4.4	Self-MobileNet block designed with Self-ONN layers.	29
4.5	Self-MobileNet architecture for classifying AF and Non-AF.	29
4.6	Block diagram of our proposed CycleGAN architecture.	30
4.7	Generator network of our proposed Cycle-GAN architecture.	31
4.8	Discriminator network of our proposed Cycle-GAN architecture.	32
5.1	ROC Curve obtained by a classifier model.	36
5.2	Noisy PPG and their corresponding reconstructed PPG signal with Non-AF patient.	40
5.3	Noisy PPG and their corresponding reconstructed PPG signal with AF patient.	40
5.4	Noisy PPG and their corresponding reconstructed PPG signal with Non-AF patient.	41
5.5	Noisy PPG and their corresponding reconstructed PPG signal with AF patient.	41
5.6	ROC curves using test segment 2 ECG signals (left) and test segment 1 ECG signals (right).	43

- 5.7 ROC curves using test segment 2 PPG signals (left) and test segment 1 PPG signals (right). 44

LIST OF TABLES

5.1	Parameters used for evaluation.	34
5.2	Partition of dataset with the number of recorded files and segments for each class.	37
5.3	Performance measure of detecting invalid wrist PPG segments. . . .	38
5.4	Confusion matrix of detecting invalid wrist PPG segments.	38
5.5	Performance measure of both Raw PPG Test segments.	39
5.6	Performance measure of both Reconstructed PPG Test segments. . . .	39
5.7	Representation of confusion matrix of both Test PPG segments. . . .	42
5.8	Performance comparison of the Self-MobileNet model with different q orders using reconstructed PPG segments.	42
5.9	Performance measure of both ECG Test segments.	43
5.10	Evaluation of Cycle-GAN for different passes with entropy functions.	44

CHAPTER 1

INTRODUCTION

Atrial fibrillation (AF) is the most common cardiac arrhythmia worldwide. As per the 2017 Global Burden of Disease research [2], there were 37.57 million instances of AF globally, accounting for around 0.51% of the global population. 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 [3]. Additionally, several studies [4, 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. 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 help minimize its symptoms. However, owing to the dynamic character of this field of research, rapid and precise AF diagnosis has received increased attention.

Throughout the last several decades, various methods for diagnosing AF have been used and are still in use in medical settings based on electrocardiography (ECG) [7], ballistocardiograms [8], and photoplethysmography (PPG) signals [9, 10]. Among them, ECG and PPG signals are more frequently utilized to efficiently detect AF, including single or multi-lead ECGs [11]. With these sensors, long-term patient recording techniques can gather more knowledge and enable better detection of arrhythmias, thereby increasing the overall classification accuracy of AF.

There are several methods and clinical tools, including Holter monitoring, ECGs, and PPG signals, that may record and track an AF episode by gathering clinical data. Single-lead and multi-lead [12] ECG signals are studied in the detection of AF with promising accuracy despite being uncomfortable for the patients. ECG data can be denoised using a variety of pre-processing methods, such as Fourier cosine series operation to remove baseline drift and high-frequency components [13]. Noise reduction techniques include the use of a wavelet transform as well as an elliptical band-pass filter [14, 15]. For reducing motion artifacts, other filters, including a notch filter [16], a band-pass Butterworth filter, and a finite impulse response (FIR) filter [17] are also utilized. Besides, Z-score normalization and a high-pass filter [17] are often employed for amplitude re-scaling and impact removal to standardize the ECG signals for analysis.

1.1 AF Classification Studies

Various robust approaches have been used to diagnose arrhythmia from ECG signals. In [18] authors have proposed a multi-scale grouped CNN for atrial fibrillation detection and found 97.07% accuracy in 5-fold cross validation method. Zahid et al. [19] proposed a 1D Self-ONN model and achieved 99.21% accuracy in ECG classification using the MIT-BIH arrhythmia database. In [20], authors used a 1D CNN model and transfer learning on both ECG and wrist PPG datasets. They have achieved 95.5% accuracy in the 5-fold cross validation method on ECG data and 95.1% accuracy on unseen PPG data. For PPG classification, they have used a very limited number of samples. Numerous physiological factors might potentially be derived from the analysis of a PPG signal since brain, cardiac, and respiratory interactions regulate blood flow [21] resulting in the PPG signal's wealth of information regarding physiological states [22]. 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. With promising outcomes, PPG time series and their picture representation (such as the entire signal plot, fast Fourier transform spectrum, or wavelet spectrum) have been applied to detect physiological events [23, 24]. Cheng et al. [25] used

a combination of time-frequency chromatography and deep learning networks (2-CLSTM) on both wrist and finger PPG data and achieved 98.21% accuracy for the AF classification task. Bashar et al. [26] proposed a motion and noise artifact (MNA) detection algorithm to extract only clean wrist PPG signals from the "Umass" and "Chonlab" datasets and perform a classification task. They have used only 289 clean PPG segments (detected by MNA) for AF classification. In [27] authors have used optimized signal quality assessment (SQA) and obtained promising results on a classification task. They have used both wrist PPG and finger PPG from different databases.

1.2 Reconstruction of Wrist PPG

Several studies show that the detection accuracy of AF using ECG signals may be degraded for many reasons, such as screening, sex-specific criteria, and different age ranges. Despite the aforementioned disadvantages, PPG signals are most commonly used to diagnose and detect AF because of their promising performance. Poor signal quality due to motion artifacts when users are moving their arms or walking or running and early beats are the two most frequent reasons why AF detection with wrist PPG is inaccurate. As a result, we anticipated the relatively high percentage of false positives associated with ambulatory PPG signal-based AF detection as well as individuals with paroxysmal AF having a decrease in specificity. Several statistical, machine learning, and deep learning techniques are studied to resolve these problems. Generative Adversarial Network (GAN) are very outperforming in image doming for image synthesis since 2014 [28]. Cycle-consistent Generative Adversarial Networks (Cycle-GANs) [29] are used to translate uncorrelated datasets. To meet our objective, we implemented a 1D version of Cycle-GANs to transform corrupted PPG signals into clean PPG signals. Since Cycle-GANs are consistent in preserving the major patterns of the translated signals, all the characteristics of PPG signals are preserved, and we have obtained the improved and denoised quality of PPG signals. Self-Organized Operational Neural Networks (Self-ONN) [29, 30, 31, 32, 33] are non-linear neural models that can achieve diverse and increased learning capabilities and are also used for PPG restoration. There is not any robust study that uses a large amount

of wrist PPG data to train a classification model for AF detection using wrist PPG signals, and the low signal quality of those data forces us to use only clean data. Otherwise, the performance of the classification model will significantly degrade. The novel and significant contributions of this study are as follows:

- This is the first study attempting to remove motion artifacts and baseline drifts and restore wrist PPG signals with the aid of **1D-Cycle-GAN** for AF classification.
- Instead of using conventional CNN layers, we have proposed the recent heterogeneous and nonlinear network model, **Self-ONNs** to build 1D-CycleGAN and **Self-MobileNet**.
- We have proposed a robust **1D-Self-MobileNet** model for an efficient atrial fibrillation classification task using wrist PPG signals and corresponding ECG signals to compare the performance on both signals.
- We have used the same **1D-Self-MobileNet** model to extract a large amount of clean and motion-corrupted wrist PPG signals with robust AF classification and remove invalid PPG signals.

1.3 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. Next, approaches to improving the signal quality of motion-corrupted wrist PPG signals are discussed. Lastly, the novel and most significant contributions of our study are briefly reviewed.

In Chapter 2, "Related Works," we have discussed other works related to this field. Various methods for AF detection of previous works, along with their workflow diagrams, have been reviewed. Moreover, we have also made a comparison between

the methods of previous works along with their accuracy and the problems in their 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. We briefly explain traditional CNN layers, and finally, an overview of different pre-processing pipelines and workflows is 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 and used in the project. After that, descriptions of the models and details of their training parameters were added. In Chapter 5, "Result and Analysis," results obtained from training the models were shown and analyzed. A comparison between the performances of the two models was given. Finally, insights obtained from a comparison of the results obtained were discussed.

In Chapter 6, "Conclusion and Future Works," a conclusion was reached and the scope of future research improvements was discussed.

CHAPTER 2

RELATED WORKS

2.1 AF Detection Methodology

Atrial fibrillation (AF) is an irregular cardiac rhythm distinguished by abnormal atrial activity. [34, 35, 36]. About 34 million people worldwide suffer from AF, which is frequently misdiagnosed and is thought to impact 3% of adult population. [37, 38]. Early identification of AF and timely initiating AF-related mortality and morbidity [34]. Unfortunately, it might be difficult to recognize AF because up to 30% of individuals have no symptoms. [34], and 25–60% of all AFs are paroxysmal [35]. Thus, hidden paroxysmal AF is one of the causes of roughly 30% of all strokes that remain cryptogenic despite a thorough clinical assessment [35]. 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 [39]. Burst AF typically exhibits no overt symptoms and is undetectable by short-term ECG [40]. 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 [41].

2.1.1 Machine Learning Based AF Detection

Machine learning (ML) and deep learning (DL) have been separated from current relevant studies based on PPG detection for ML. 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 [42, 43]. Research has concentrated on using wrist PPG-based AF identification as an replacement to ECG-based solutions since it is a non-invasive technology that can record on most contemporary devices [44]. In 2017, Tang et al. [45] 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 [46] developed a magnificent machine learning strategy to identify the AF episode using pulse rate variability characteristics.

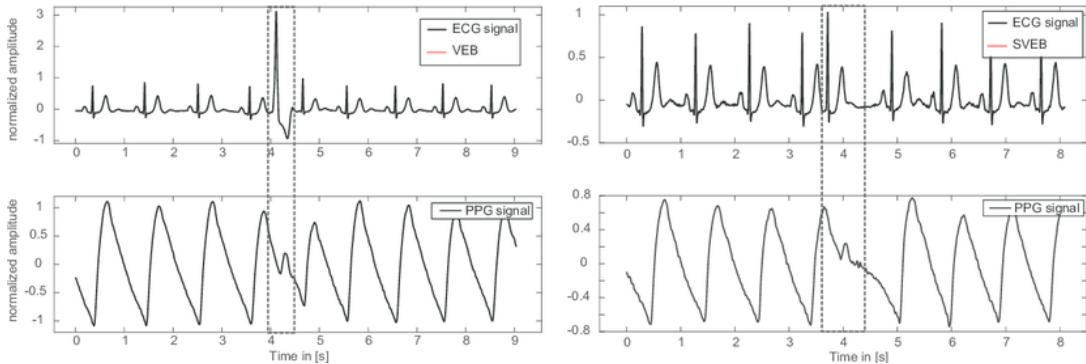


FIGURE 2.1: AF Detection using Pulse Rate Variability.

In 2019, Fallet et al. [47] estimated the values of several parameters (mean, standard deviation, and quartile) in wrist 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 [48], in which the authors employ a Kardia band to capture the PPG and ECG data from a wrist-watch and utilize it to identify atrial fibrillation from normal sinus rhythms. Linda et al. [49] incorporated PPG features, inter-pulse intervals, and accelerometer signals in random forest (RF) model to categorize AF and other rhythms. With continuous PPG

monitoring in resting individuals, AF was identified with a sensitivity of 81.0%, specificity of 96.4%, and positive predictive value (PPV) of 86.6% [50]. In this investigation, using nighttime recordings, the hourly AF 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 neighbors and allocated to the class based on the most frequent class among its k nearest neighbors [17]. SVMs appear to be effective in non-linear detection problems by transferring features into a higher-dimensional state space, a method called the "kernel trick" that employs kernel functions as the basis function (RBF) or polynomial [51]. Some investigations employed a mix of threshold-based and ML techniques. For instance, the thresholds of several characteristics are initially employed to remove bad signals, and then an ML model is created to identify AF in the clean pulses [18]. In the study [52], the author developed a technique consisting of first pre-processing the PPG signal and identifying motion and noise artifacts (MNA). Next, heart rate is estimated by peak detection from the examples of PPG signals found to be clean. The heart-rate variations are categorized into AF and non-AF rhythms. The block diagram of the suggested model is presented in Figure 2.2.

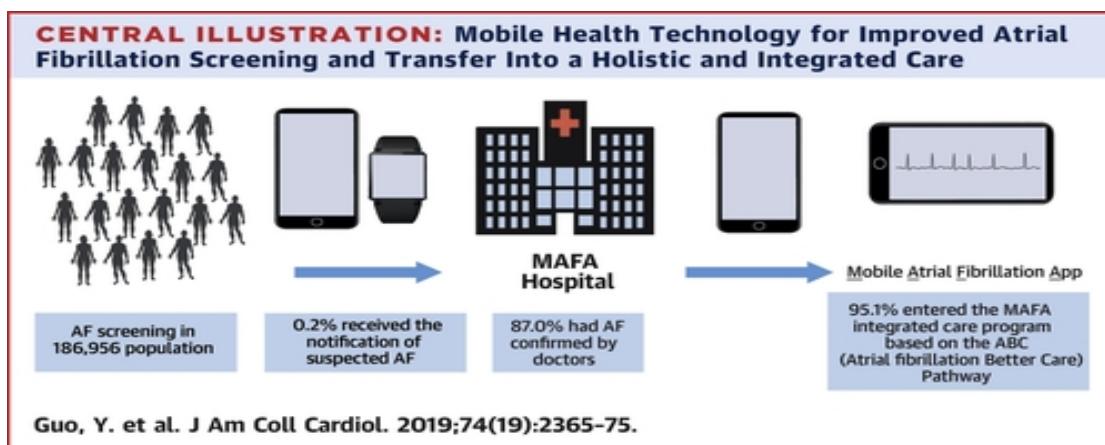


FIGURE 2.2: Block Diagram of the Proposed Method.

A classical AF detection method extracts characteristics from the collected PPG signals and analyzes them to determine if an AF beat is observed. Image information of the temporal signals has been addressed in several techniques. The

tachogram (RR intervals) has traditionally been used to extract significant characteristics as it is a dependable measure of heartbeats [53]. Considering that PPG signals might include physiological information other than heart rate, additional properties other than R-R intervals were developed [54]. The utilization of ECG signals and their visual representation yielded promising results in the identification of physiological events. Images for ECG representation in Figure 2.3 are a broad representation of the information format types employed by the various algorithms.

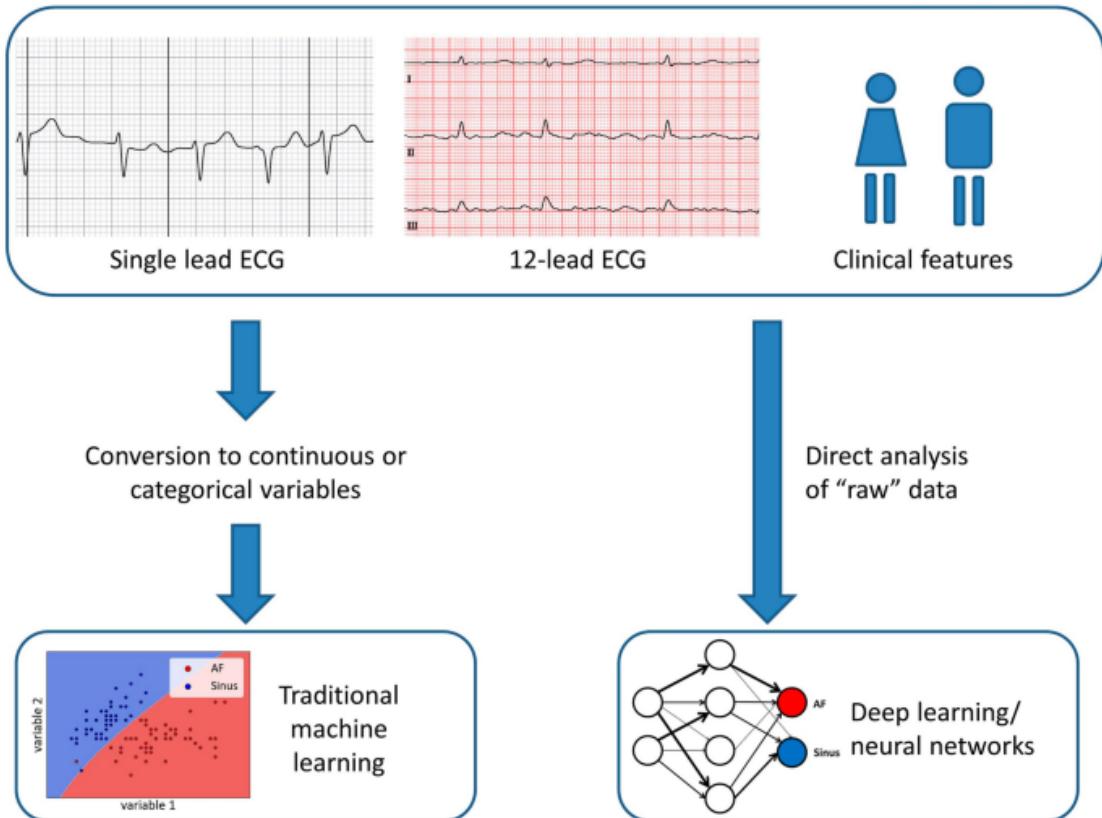


FIGURE 2.3: Complete overview of AF detection using ECG Data.

2.1.2 Deep Learning-Based AF Detection

Deep learning has emerged as a potent tool for identifying irregularities in physiological data, paving the way for arrhythmia diagnosis from ECG and PPG. Deep learning models, unlike ML, automatically learn significant feature characteristics, saving the arduous effort of feature construction. Convolutional neural networks

(CNN) are the most often used DL techniques for AF detection. CNN was used to solve challenges with automated feature extraction and classification. CNNs were only employed in a few research projects for automatic feature extraction. One study presented an integrated model comprising two linked CNNs, where the former recognizes clean PPG segments and the latter detects occurrences of AF [55]. 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 labeled training data is necessary to train a DL model from the start. The scarcity of labeled 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 [56]. The number of layers necessary and the difficulty of fine-tuning are determined by the application [57]. In [58], a pre-trained ECG-based CNN network was fine-tuned using a limited number of labelled PPG signals to identify AF from PPG segments.

Due to the fast acceleration of GPU processing, deep learning has been employed in numerous studies such as image identification, computer vision, and pattern recognition. The author of [58] suggested an automated AF detection model that included 2D-CNN and LSTM models and compared the results to DL models like VGGNet [59], GoogLeNet [60], and Microsoft ResNet [61]. The Figure 2.4 shows the technique of [58]'s suggested model.

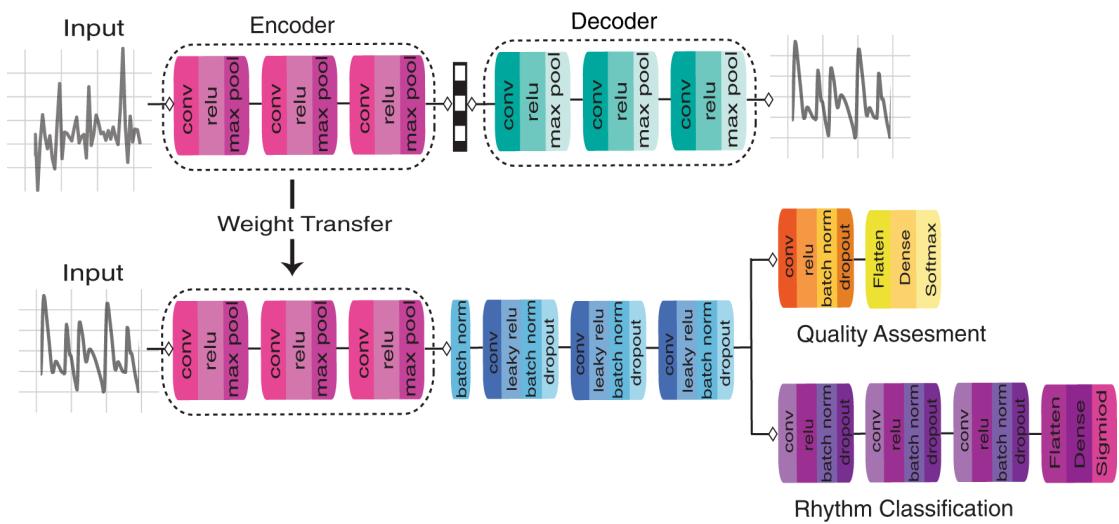


FIGURE 2.4: Proposed model of AF detection by DL method.

The accuracy, recall, and specificity of these approaches for distinguishing AF signals from a clean sinus rhythm (SR; that is, SR without any premature atrial complexes or PACs) were satisfactory [62, 63, 64]. The author of [65] 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 CNN and RNN models, respectively), and that both deep learning-based approaches performed better than previous AF classification algorithms. The majority of diagnoses made by the DL classifiers were confident, and the corresponding estimated CLs provided an easily perceived reliability even under a high PAC burden. They also had the potential to get better if more samples were permitted to be trained. Figure 2.5 depicts the complete DL framework of AF diagnosis.

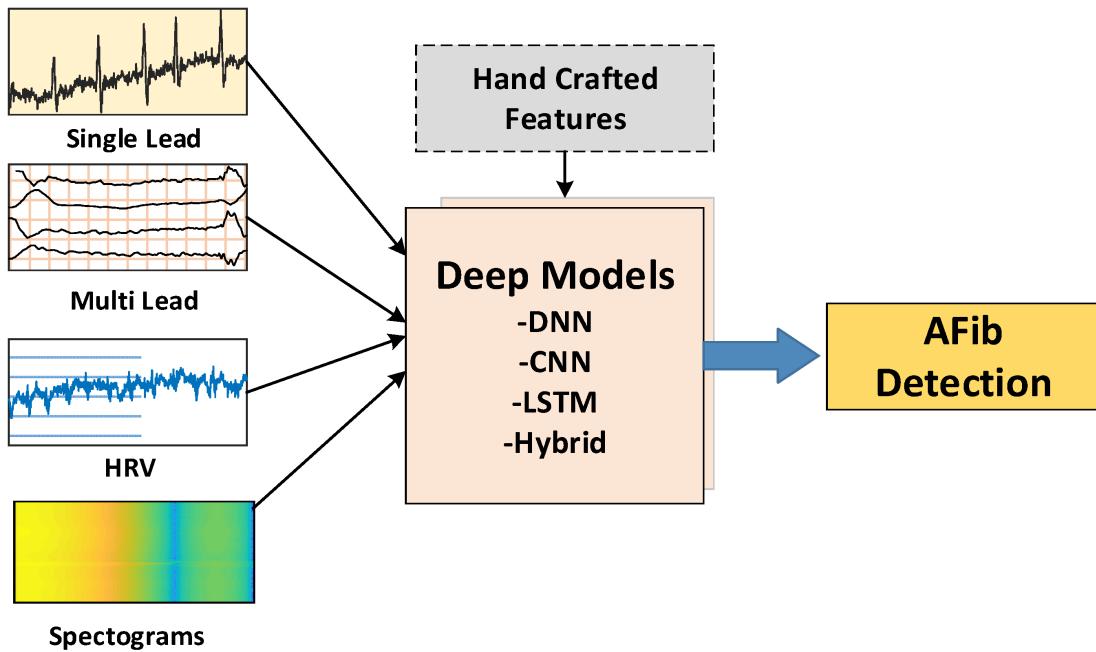


FIGURE 2.5: Complete DL framework of AF diagnosis.

PACs, on the other hand, were often seen in individuals with paroxysmal AF or after successful cardioversion [65, 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 [68, 69].

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 [70]. Furthermore, this unique methodology outperformed typical approaches based on heart rate variability and traditional statistical methodologies, as well as a machine learning system fed only 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 [1] 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 Figure 2.6

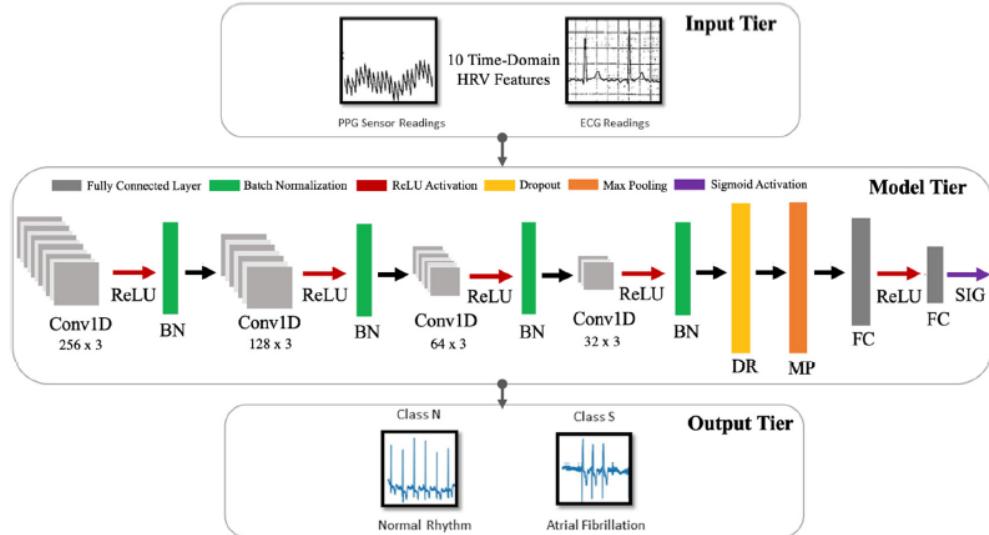


FIGURE 2.6: Deep learning 1D-CNN model on ECG data [1].

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. We have also employed Cycle-GAN to transform motion-corrupted wrist PPG signals into clean PPG signals. There is a significant improvement in AF classification using reconstructed PPG signals compared to the raw original PPG signals. Due to the sensors' availability, many complex patterns are classified using ML and DL 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 [71] 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 and continued for a week in the patients' homes. At a 500 Hz sampling frequency, the Bittium OmegaSnap and the Bittium Faros 180 are used to acquire data. A green LED embedded in a wrist-worn device 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) A wrist-worn device with an embedded green LED for PPG data acquisition. Both of them have embedded acceleration sensors.

Bittium Faros™ 180 is an instrument for recording ECG for multi-purpose (i.e., AF episode detection) with the dimensions of 49 x 29 x 12 mm and a smooth design. The sensor consists of a single channel for 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, a long, durable, rechargeable battery, and memory capability of up to 180 days of recording. The wrist-worn wearable device provides a wide range of opportunities to monitor different types of physiological conditions, such as heart rate, AF episodes, etc. It has become more popular due to its durability over longer time periods as well as its high resolution at a low price. An optical sensor (photoplethysmography; PPG) can collect information on the volumetric change in blood perfusion, which can serve as a surrogate for clinical assistance. Wrist-band PPG sensors can detect several distinct colors 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 color 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 depend

on the change in blood profusion. The sensor can acquire the data continuously at a sampling rate of 100 Hz with an embedded triaxial accelerometer sensor at a sampling frequency of 50 Hz. As a result, PPG enables *in vivo* examinations of clinical processes in real-world settings.

3.2 Dataset Description

We have used an open-source dataset [33], 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 monitored for 5 to 8 days. The dataset contains about 1306 hours of data acquisition (ECG and PPG with accelerometer data) for clinical purposes. We have used this publicly accessible dataset to detect and classify AF from raw data collected from various sensors. The ECG and wrist-band PPG data are not suitable for direct use with deep learning models because of their high levels of noise and severe corruption. To clean the data and get rid of various artifacts (such as baseline drift and motion artifacts), several pre-processing pipelines are employed before training. Figure 3.2 depicts weight and age information for individual patients.

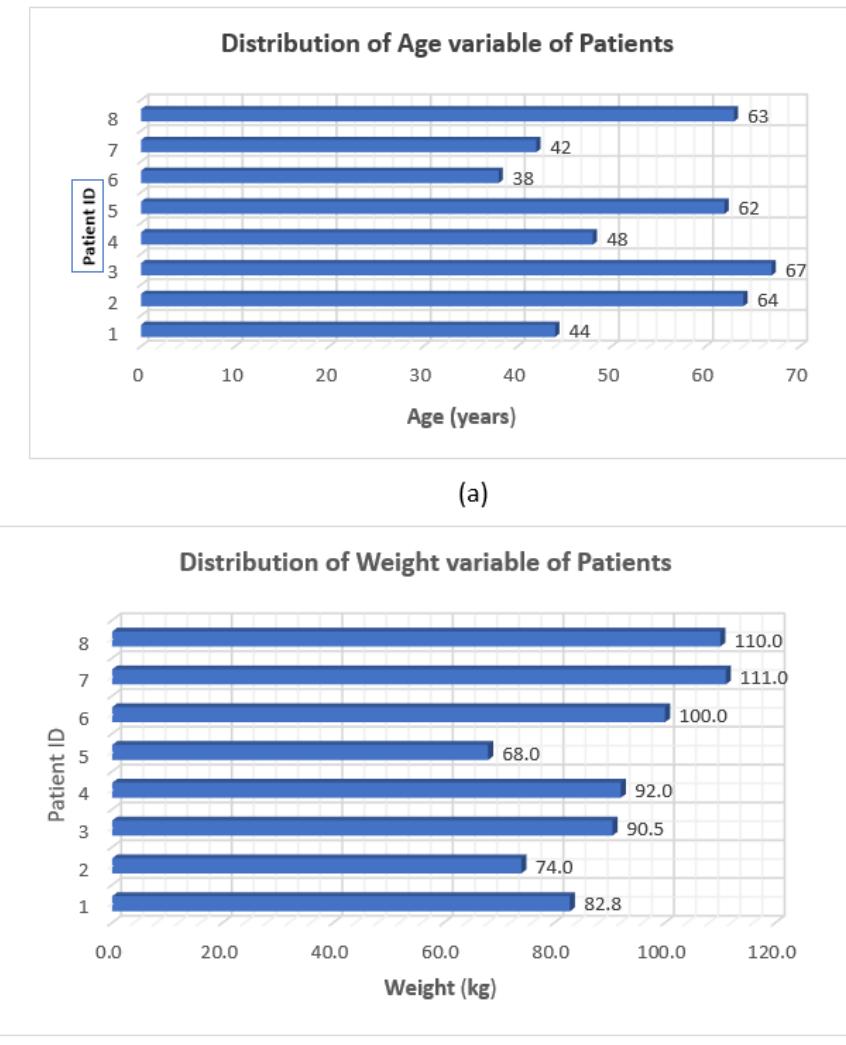


FIGURE 3.2: (a) Distribution of age, (b) distribution of weight of each AF patient.

The data is acquired in the format of a .mat file with the format AA_BBB_CC.mat, where AA is the patient ID, BBB is the name of the data (i.e., for ECG, it is "ECG", for PPG, it is "PPG") and the CC is the serial no. of the recordings. The batteries of wrist-worn devices are charged after long periods of collecting data.

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.

The PPG and ECG collection devices are not synchronized, and researchers utilizing the resource should be aware of this fact. As a result, the alignment should be done based on certain physiological parameters. The supplied time of day when 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, so users should be aware of this. Utilizing the timestamp-sec data record vector can help reduce the impact of this problem on the PPG signal, but an equivalent variable is not present for the ECG signal.

3.3 Data Processing Steps

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 versatility in problem solving. Deep learning models are hungry for large data sets. The more data we feed the models, the more generalized their performance becomes. But the data should be in proper form to train the deep learning models; otherwise, the models' performance will be lower due to misclassification. 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.

3.3.1 Resolving Drifting

Multi-modal data has become common due to its popularity and superior 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 [71], 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 sensor lags behind, resulting in errors in data acquisition. The authors of the dataset [71] mentioned that the internal clock of the 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:

$$PPG^i(n_k) \rightarrow Resample(PPG^i(n_k), f_s^i); \forall i = 1.....N_{PPG}; \forall n_k \quad (3.1)$$

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

$$f_s^i = \frac{F_a^i}{F_s^i}; \forall i = 1.....N_{PPG} \quad (3.2)$$

here, F_a^i and F_s^i are the data acquisition frequency and the sampling frequency of the PPG signal for i_{th} 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 3.3

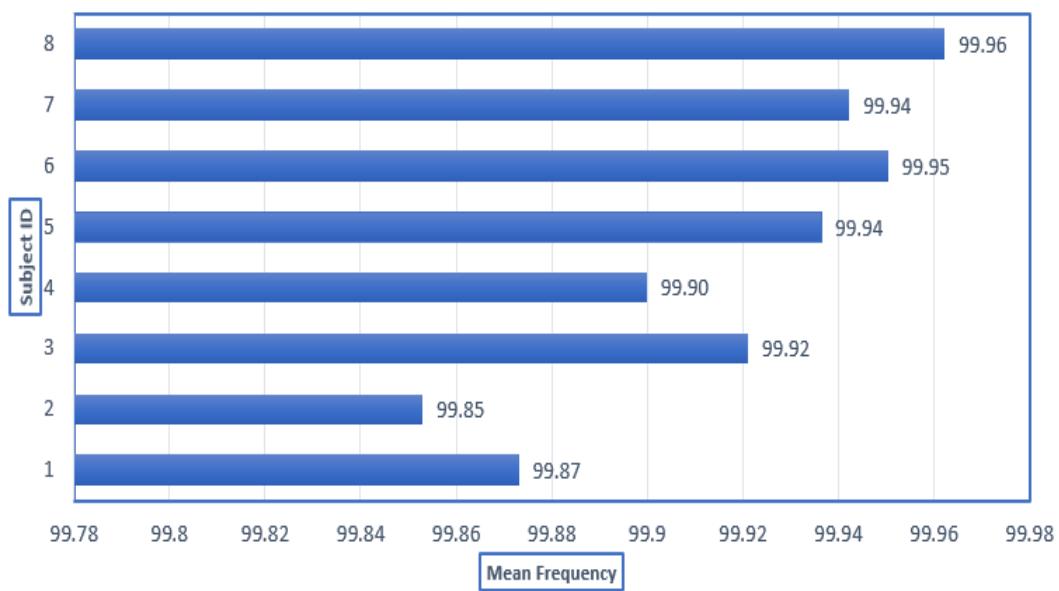


FIGURE 3.3: Change in PPG frequency (100 Hz) due to drifting.

3.3.2 Noise Filtering

Sensor data are very popular due to their simplicity and computation efficiency over image or video signals. However, the sensor data has a fatal flaw: 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 movement of the heart is in the 1–20 Hz frequency range. In the low-frequency range of both ECG and PPG signals, noise is added to the signal, which results in a degradation of the signal quality. However, it is quite challenging to separate the signal from the noise. We used a variety of techniques to remove the noise influence from the signal. 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. FIR 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 non-linear 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 where cut-off frequency is 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 to 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, on which the PPG signal is placed. We have chosen a higher-order digital high-pass IIR filter with a cut-off of 0.05 Hz for processing and reaction speed. The filter responses are shown in Figure 3.4.

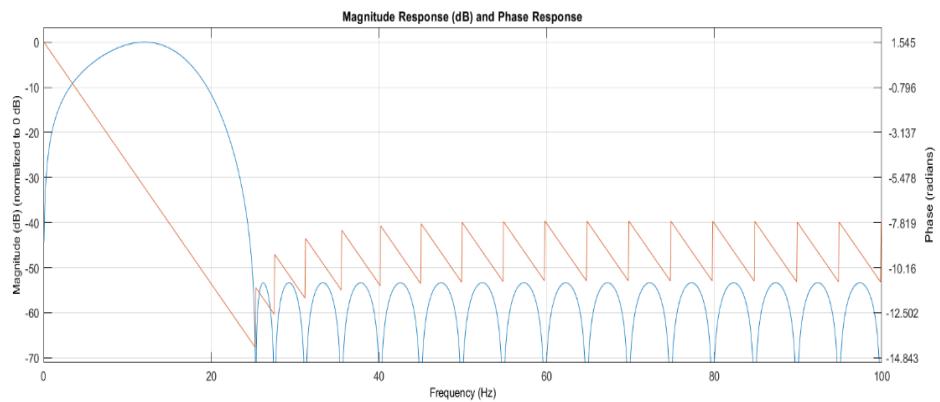


FIGURE 3.4: Responses of the filter.

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., 0.05 Hz and 100 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 the ECG signal. Besides this, several parameters are optimized for the filter (i.e., the filter order) to obtain a better-quality PPG signal. The noise-filtered ECG and PPG signals are then generated using the PPG signal settings. In the process of parameter optimization, we have measured the similarity score between the ECG and PPG signals, considering the ECG as the reference. Figure 3.5 depicts the quality of the ECG and PPG signals after noise filtering.

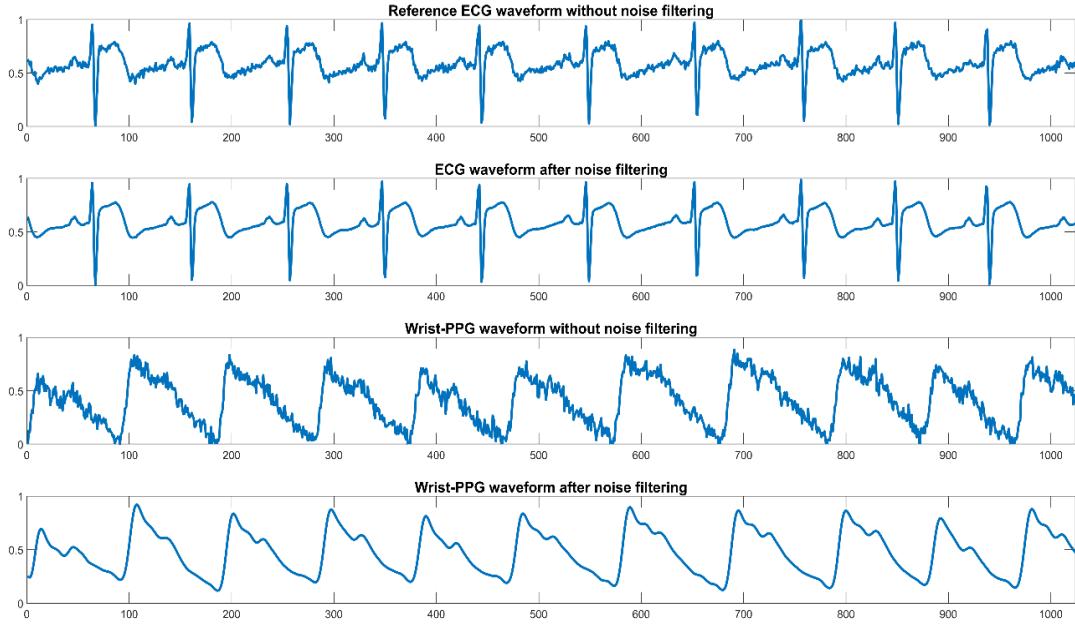


FIGURE 3.5: Noise removal of PPG and ECG Data.

3.3.3 Data Synchronization

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

Due to the multi-modal sensor capability, it is quite an issue for the synchronous type of data. Both the ECG and PPG are acquired from the different devices

mentioned above, resulting in asynchronous ECG and PPG data. In this section, we will describe our noble methods for automatically synchronizing the ECG and PPG signals.

In this pre-processing pipeline, we have extracted several components from 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 waveform throughout the range [0, 1]. In order to time-align the peaks of the many signals, time re-scaling and shifting are carried out. The peaks of different signals include the ECG waveform and the derivative PPG signal (VPG and APG) and also 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:

$$ECG^i(n_k) \rightarrow ECG^i(n_k + \delta_k^i); \forall i = 1 \dots N_{ECG}; \forall n_k \quad (3.3)$$

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 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 3.6

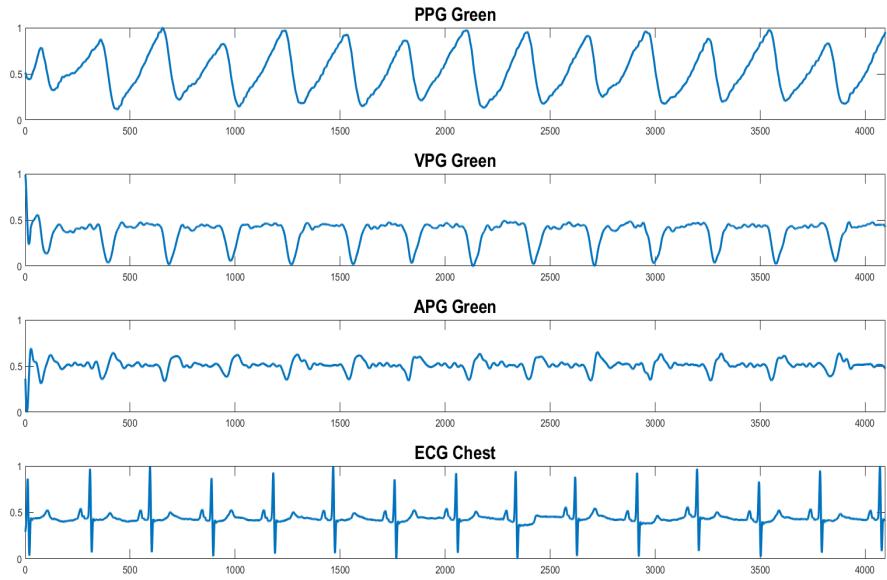


FIGURE 3.6: Alignment of PPG, VPG, APG, and ECG Signals.

3.3.4 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 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 [72]. 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 [72].

In our algorithmic pipeline, we have proposed several techniques to eliminate 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 that 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-to-peak intervals are extracted from the difference between the two adjacent peaks, which helps us follow the trend of the peaks.

After extracting the peak-to-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 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-to-peak distances, which are a better approximation since they are not affected by the outliers.

After that, to get the clean signal without the baseline, we just subtracted the baseline trend from the raw signal. But the subtraction of the baseline signal causes an amplitude change in the original signal, which may degrade the signal quality. To fix the slight change in amplitude due to the baseline shift, we have re-scaled the amplitude.

CHAPTER 4

METHODOLOGY

4.1 Data Pre-processing

ECG and PPG data were first resampled at 250 Hz from 500 Hz and 125 Hz, respectively. For the classification and restoration tasks, we employed 10-second PPGs and matched ECG records. As a result, the signal length for each example is 2500. The classification model performs poorly if baseline drift is not properly removed. Baseline drift is a low-frequency signal variation problem that results in unintentional amplitude shifts in the signal. The built-in MATLAB functions "movmin," "polyfit," and "polyval" were used to perform baseline drift correction. Initially, an array of approximate minimum points serving as a baseline approximation for the waveform was found using the "movmin" or moving minimum function. The higher-order polynomial was then fitted using the "polyfit" function together with the estimated points, and the "polyval" function was used to create the polynomial based on the "polyfit" result, which is the estimated baseline. The baseline drift adjusted signal was then created by subtracting the baseline from the raw signal. After baseline correction, we performed z-score normalization, followed by a range of normalized between 0 and 1 per segment. Band-pass filtering of the ECG signal is very important to ensure the band-limited ECG signals are used for further analysis. To remove any potential unwanted noises, we used a band-pass infinite impulse response (IIR) filter with cutoff frequencies of 0.05 Hz and 100 Hz. For PPG data, we have selected a cut-off frequency of 0.5 Hz and 25 Hz for the

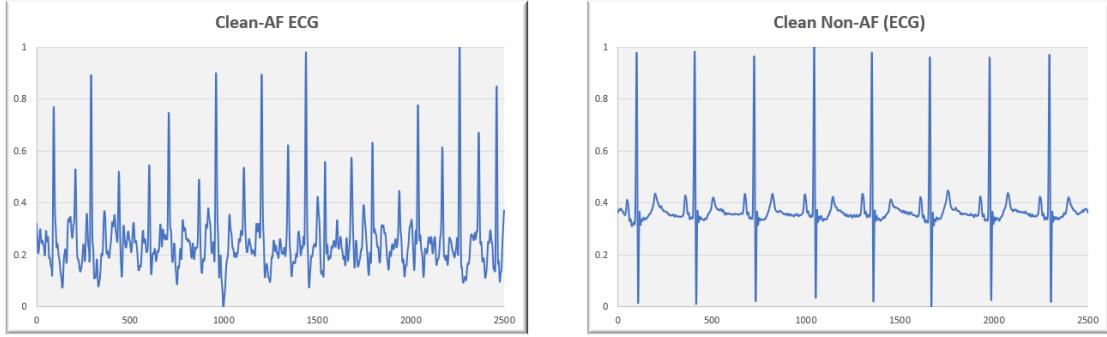


FIGURE 4.1: ECG signals after pre-processing.

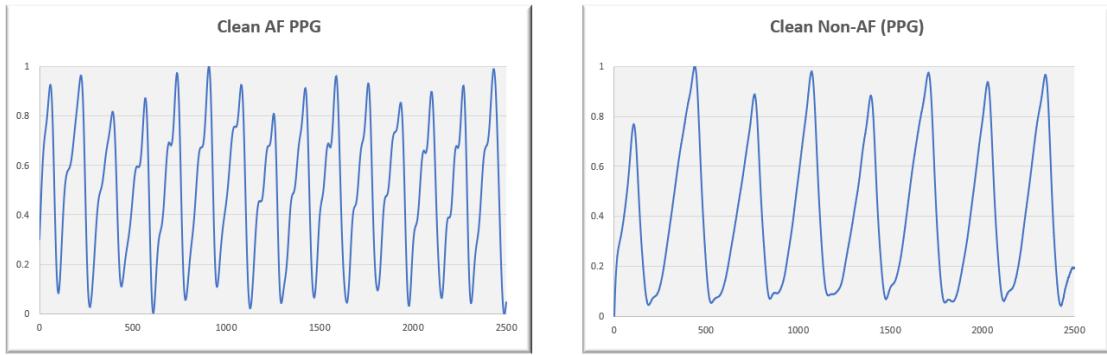


FIGURE 4.2: PPG signals after pre-processing.

band-pass filter. After applying the data pre-processing technique, we obtained clean, noiseless ECG and PPG signals, which are illustrated in Figure 4.1 and 4.2.

4.2 Self-ONN

Homogeneous network structure along with a linear neuron model is a common drawback of Convolutional Neural Networks (CNNs) [73]. This drawback is addressed in the concept of Generative Operational Perceptions or, in short, "GOPs" [73, 74, 75, 76]. Operational Neural networks or ONNs incorporated the concept of GOPs [73] and later, Self-ONN, a variant of ONNs, was proposed to introduce non-homogeneity in neural networks [30, 32]. Self-ONN mimics the biological neuron's activity during activation or signal propagation [31, 74]. In biological neuron, lots of neuro-chemical activities take place, such as non-linear synaptic connections and integration of signals in the Soma [73, 74]. Similarly,

Self-ONN mimics this non-linear process in the neural network. "Nodal operator" or Ψ in "Nodal operation" of Self-ONN is analogous to synaptic connections of biological neurons, and "Pool operator" or "P" is analogous to the integration in the Soma. Self-ONN does not need any pre-determined nodal operator set as ONN, while Self-ONN can generate any composite nodal operator, which may not be a pre-defined function such as hyperbolic, sinusoidal, exponential, etc. If x_m^n is the output in m^{th} neuron of n^{th} layer due to the input y^{n-1} , the calculation can be presented as follow:

$$x_m^n = b_m^n + P \left(\sum_{i=1}^{N_{n-1}} \Psi_{m_i}^n \cdot (\omega_{m_i}^n, y_i^{n-1}) \right) \quad (4.1)$$

Here in Equation (1), $\omega_{m_i}^n$ is the weights and b_m^n is the bias term. The ω is an array of parameters in q dimensions made up of internal parameters and weights for each unique function. The combination of these arrays and inputs is combined using the nodal operator. The Taylor series approximation is used to estimate the nodal operator. So, the approximation is as follows:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^q(x_0)}{q!}(x - x_0)^q \quad (4.2)$$

$$f(x) = f(0) + \frac{f'(x_0)}{1!}(x) + \frac{f''(x_0)}{2!}(x)^2 + \dots + \frac{f^q(x_0)}{q!}(x)^q \quad (4.3)$$

$$f(x) = \omega_0 + \omega_1(x) + \omega_2(x)^2 + \dots + \omega_q(x)^q \quad (4.4)$$

In Equation 4.2 and 4.3, first, second, and q^{th} order derivative are defined as f' , f'' , and f^q respectively. The x_0 in Equation 4.2 can be approximated to 0 by bounding the input between [-1,1] using Tanh activation in the previous layer. So, the weight ω can be expanded to q^{th} order and ω_0 is the bias coefficient which is compensated by each neuron's bias term available in Equation 4.1. Figure 4.3 illustrates the Self-ONN operation on an input vector with nodal operation and pooling operation with q order of 3.

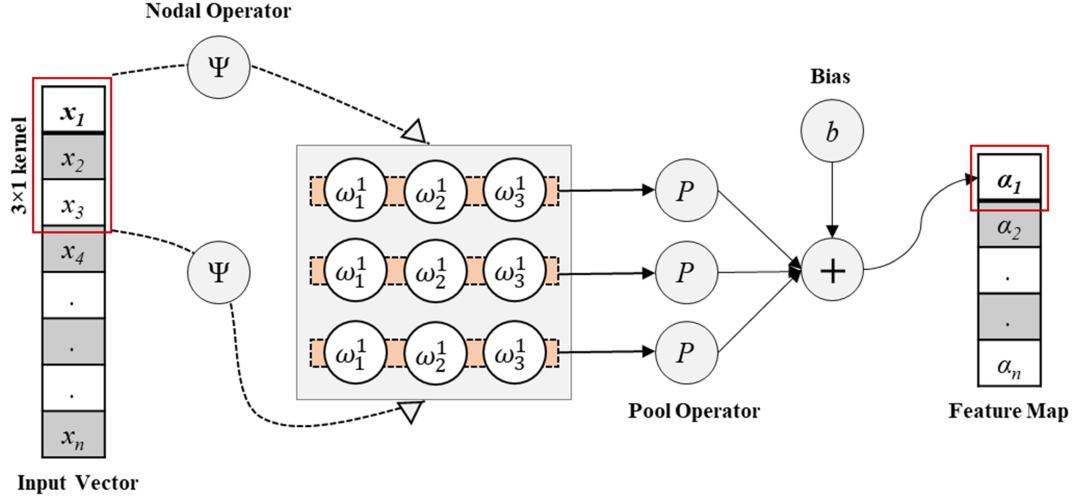


FIGURE 4.3: Self-ONN operation on an input vector with pooling operator P and nodal operator Ψ .

4.3 Self-MobileNet

In the literature, the architecture of cutting-edge CNN models, such as ResNet, was redesigned with Self-ONN layers [77] for image super-resolution. In this research, a 1D Self-ONN version of MobileNet [39] was adopted. Figure 2 illustrates a Self-MobileNet block used in this research, which was completely designed using Self-ONN layers. From Figure 2, it can be seen that point-wise and depth-wise 1D operations are employed in this Self-MobileNet block. First, a point-wise operation was designed to increase the number of channel and second point-wise operation was designed by decreasing the number of channels to create a bottleneck using kernel size of 1. So, the depths-wise operation between two point-wise operations was done using 1D Self-ONN layer. A Tanh activation was applied to the features formulated from the residual connection between the input feature and bottleneck feature to obtain the output feature from a self-mobile network. The Self-MobileNet block using Self-ONN layers is shown in Figure 4.4

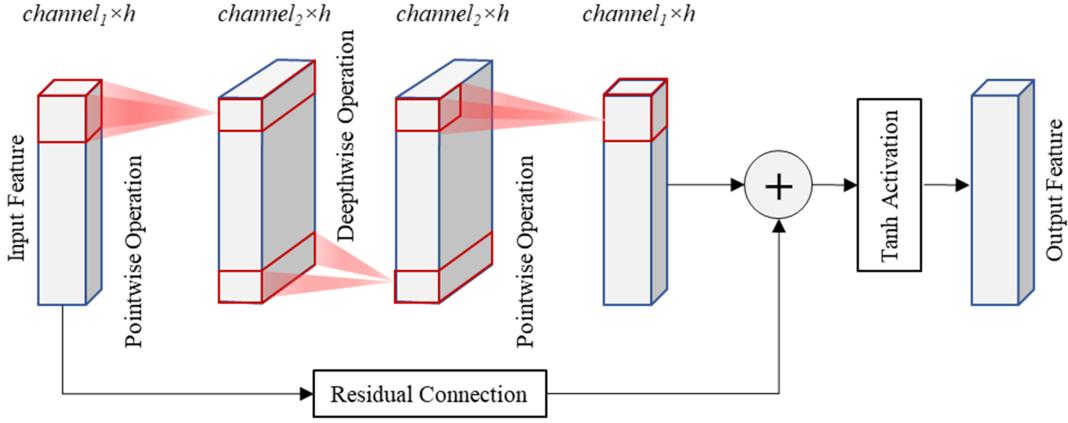


FIGURE 4.4: Self-MobileNet block designed with Self-ONN layers.

Three Self-MobileNet blocks were used to build the Self-MobileNet architecture and an illustration of Self-MobileNet architecture is available in Figure 3. From Figure 3, every Self-ONN layer is followed by a Batch-Norm and Tanh activation layer. The output feature after three Self-MobileNet blocks are pooled to a feature size of 12 in each channel using Adaptive Average pooling. Instead of conventional MLP, Self-MLP, a Self-ONN based multi-layer perceptron was used in this study for classifier part. Overall, a lightweight and shallow 1D Self-ONN based Self-MobileNet is proposed in this study to classify AF and Healthy signals from Wrist PPG signal are shown in Figure 4.5.

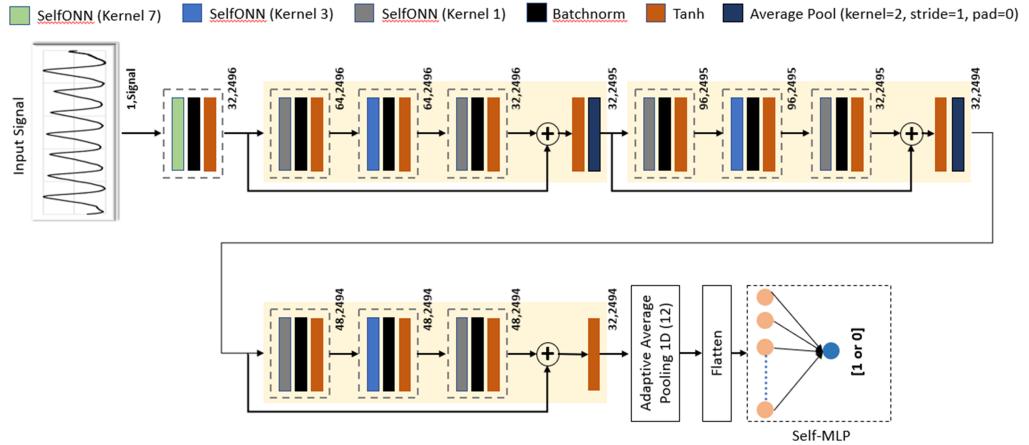


FIGURE 4.5: Self-MobileNet architecture for classifying AF and Non-AF.

4.4 Cycle-GAN

Our model employs a Self-ONN based model as the discriminator and a ResNet12 block model as the generator. Here, we used 12 residual blocks in the generator model since it reconstructs the motion-corrupted PPG signals better than ResNet9 block model. Traditional GANs only employ adversarial loss. Cycle-GAN also works to reduce identity and cycle-consistency loss. In addition to adversarial loss, there are two different kinds of generators. The generator $G_{X \rightarrow C}$ learns how to reconstruct clean wrist PPG signals from motion-corrupted PPG signals, whereas the generator $G_{C \rightarrow X}$ learns to produce motion-artifacts wrist PPG signals from clean PPG signals. The block diagram of Cycle-GAN is illustrated in Figure 4.6.

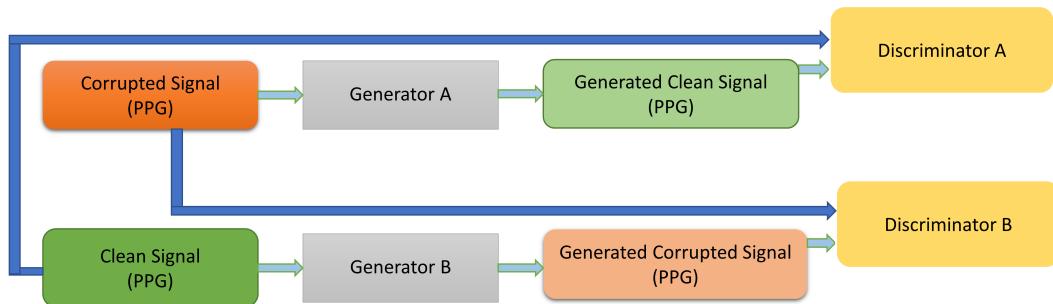


FIGURE 4.6: Block diagram of our proposed CycleGAN architecture.

In contrast, DC and DX , which are the analogous discriminators for $GX2C$ and $GC2X$, work to improve the adversarial loss function and direct the generators to create more accurate transformations. The adversarial loss functions are formulated in Equation 4.5 and Equation 4.6

$$Loss_{adv1}(GX2C, DC, WPPG_X) = \frac{1}{m} \cdot \sum_{i=1}^m \left(1 - DC(GX2C(WPPG_X(i))) \right)^2 \quad (4.5)$$

$$Loss_{adv2}(GC2X, DX, WPPG_C) = \frac{1}{m} \cdot \sum_{i=1}^m \left(1 - DX(GC2X(WPPG_C(i))) \right)^2 \quad (4.6)$$

Here, $WPPG_X$ and $WPPG_C$ are motion-corrupted and clean wrist PPG data respectively, and Here, m is the number of segments used in the training set. The discriminator DC tries to discriminate between real clean and fake (generated) clean PPG segments and guides the generator $GX2C$ to generate more realistic clean PPG wave-forms while DX tries to do just the opposite. The cycle consistency loss measures the difference between the PPG segment input to Generator 1 and the PPG segment produced by Generator 2 and the generator models are updated accordingly to reduce the difference in the PPG segments. The cycle-consistency loss combines the consistency losses from the two pathways in an effort to get more realistic end results. The generator (ResNet12 block) and the discriminator model (Self-ONN) of our proposed methodology are illustrated in Figure 4.7 and 4.8.

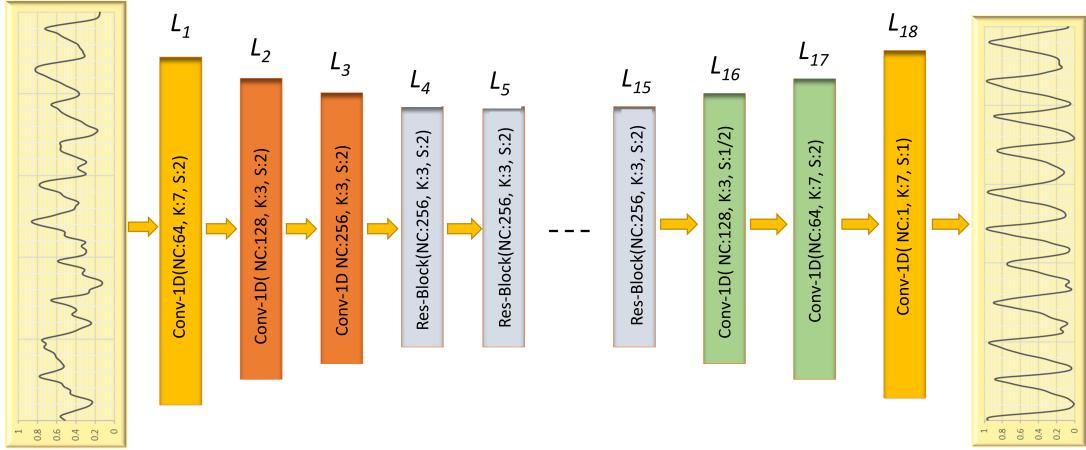


FIGURE 4.7: Generator network of our proposed Cycle-GAN architecture.

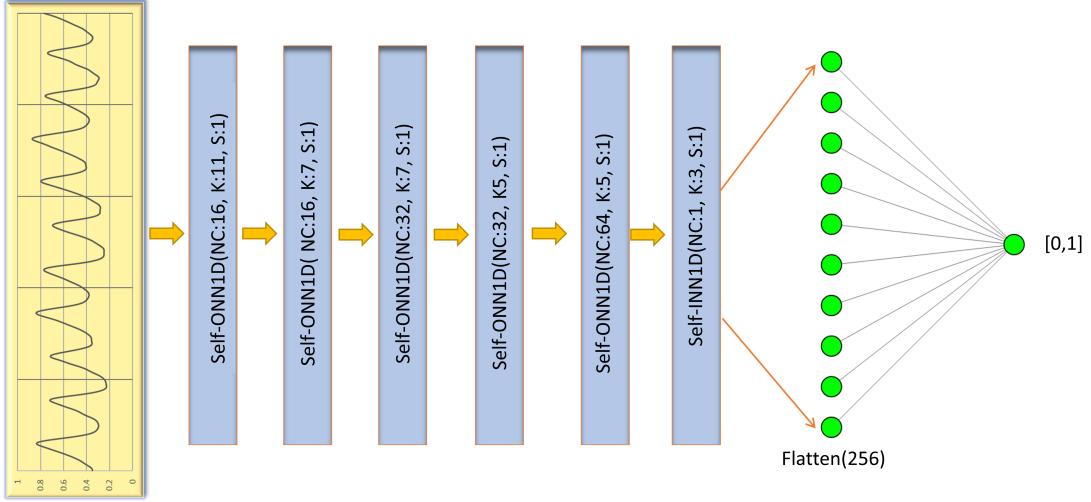


FIGURE 4.8: Discriminator network of our proposed Cycle-GAN architecture.

$$\begin{aligned}
 & Loss_{cyc}(GX2C, GC2X, WPPG_X, WPPG_C) \\
 &= \frac{1}{m} \sum_{i=1}^m \left(GC2X(GX2C(WPPG_X(i))) - WPPG_X(i) \right) \\
 &\quad + \frac{1}{m} \sum_{i=1}^m \left(GX2C(GC2X(WPPG_C(i))) - WPPG_C(i) \right)
 \end{aligned} \tag{4.7}$$

Identity loss retains the shape of the signal and prevents morphological changes in the input PPG segments when the input PPG segment is clean.

$$\begin{aligned}
 & Loss_{ide}(GX2C, GC2X, WPPG_X, WPPG_C) \\
 &= \frac{1}{m} \sum_{i=1}^m \left(\left(GX2C(WPPG_C(i)) \right) - WPPG_C(i) \right) \\
 &\quad + \frac{1}{m} \sum_{i=1}^m \left(\left(GC2X(WPPG_X(i)) \right) - WPPG_X(i) \right)
 \end{aligned} \tag{4.8}$$

The objective of training Cycle-GAN is to reduce the total loss formulated in Equation 4.9

$$Loss_{total} = Loss_{adv1} + Loss_{adv2} + \lambda \cdot Loss_{cyc} + \beta \cdot L_{ide} \quad (4.9)$$

Here, λ and β are weights that are tuned before training.

CHAPTER 5

RESULT AND ANALYSIS

In this chapter, we discuss and evaluate the results of our models and analyze their 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 successes and failures of the result are discussed below.

5.1 Evaluation Metrics

We can categorize the results into four categories because many of them are prediction-based. 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	TP	FN
Negative	FP	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, depending on how well it matches up to the actual value.

- TP: True Positive.
- TN: True Negative.
- FP: False Positive.
- FN: False Negative.

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

Accuracy: It is simply a ratio of correctly classified observations over total observations.

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

Precision: It is the ratio of correctly classified positive samples over the total predicted positive samples.

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

Recall: It is the ratio of correctly classified positive samples over all samples in the actual class.

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

Specificity: It gives the system the capability 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. The 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 measurement technique for binary classification issues is the Receiver Operator Characteristic (ROC) curve. In essence, it distinguishes the "signal" from the "noise" by introducing the TPR against the FPR at several thresholds. The performance is measured by the Area Under the Curve (AUC). The higher the AUC, the better the quality of the model at distinguishing between the two classes. 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. The higher the AUC, the better it can detect AF.

The TPR and FPR are on the y-axis and x-axis of the ROC curve. AUC between 0.5 and 1 indicates that positive and negative samples are detected very well. AUC = 1 indicates that the classifier is perfect. The classifier would be predicting all negatives as positives and all positives as negatives, however, if the AUC had been 0.

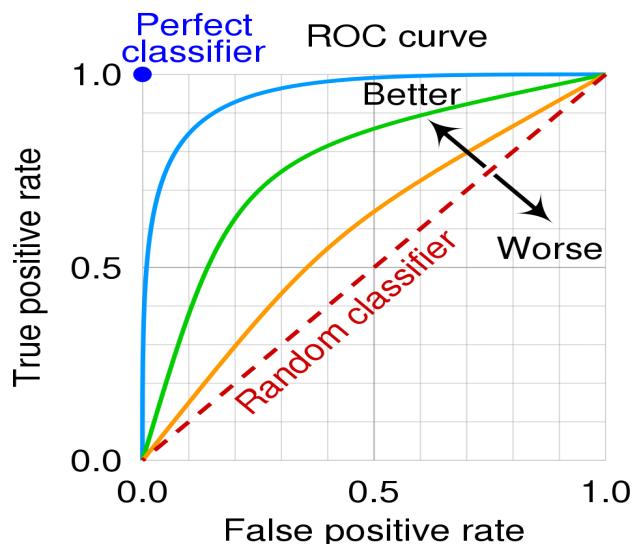


FIGURE 5.1: ROC Curve obtained by a classifier model.

5.3 Experimental Results and Evaluation

Because long-term continuous labelled public wrist PPG datasets for AF identification tasks are extremely rare, we first studied the dataset [78] and used it for training and evaluating our proposed architecture. We will present the partition of the dataset for a robust AF detection task. Next, we propose the same 1D Self-MobileNet model to extract the clean and motion-corrupted PPG signals and remove the invalid signals. Then, we will present a comprehensive set of experiments and evaluations on both wrist PPG and corresponding ECG signals over the used open source dataset.

5.3.1 Dataset

In this section, we first demonstrate the experimental design of the dataset that was used for the training and evaluation of our proposed integrated model. We have selected five subjects out of eight. As the clinical recording of these subjects contains more than one hundred hours of data, we have taken some recording files for model training and evaluation. We have taken Subject 3 for both training and evaluation because it contains more than 80 percent of the AF segments. In Table 5.2, we present the partition of PPG recordings and the total number of segments for each class.

TABLE 5.2: Partition of dataset with the number of recorded files and segments for each class.

Segment	Segment_ID	Recording_ID	Healthy	AF	Total Segments
Test 1	3	8	5846	3958	9804
	4	8			
	5	10			
	8	5			
Test 2	2	5	7993	3481	11474
	3	2			
	8	1			
	8	4			

5.3.2 Invalid Signal Detection

There are many invalid wrist PPG signals with a random shape due to some mismatch of the wrist-worn device as each PPG record was collected for some hours with clean and motion-corrupted PPG signals. It takes a long time to manually remove those invalid PPG segments and is not practical for real-world applications to. Therefore, we first trained a 1D Self-MobileNet model using limited number of PPG segments to detect the invalid PPG signals. We obtained an overall accuracy of 92.09% using 13,067 PPG segments and a 5-fold cross-validation method, as shown in Table 5.3. Then, we applied this model in the whole PPG dataset to remove the invalid PPG signals. We have found 21,278 PPG valid segments from total 40,115 PPG segments.

TABLE 5.3: Performance measure of detecting invalid wrist PPG segments.

Classifier	Type	Accuracy	Precision	Sensitivity	F1_Score	Specificity
Self-MobileNet (q=3)	Invalid	92.09	93.86	89.09	91.41	94.78
	Valid	92.09	90.64	94.78	92.66	89.09
	Weighted Average	92.09	92.16	92.09	92.07	91.78

The number of invalid detection of wrist PPG segments is shown below as a confusion matrix.

TABLE 5.4: Confusion matrix of detecting invalid wrist PPG segments.

		Predicted Class	
		Invalid	Valid
Ground Truth	Invalid	5503	674
	Valid	360	6530

5.3.3 Evaluation

We carried out two experiments to compare the 1D Self-MobileNet model's performance in real-world circumstances and show the robustness of the model. The PPG segments in Test Segments 2 in Table 1 were first considered unseen subject segments, and those records served as the test set while Test Segments 1 was used for model training. After that, Test Segment 2 will be used for training the model, whereas Test Segment 1 will serve as an evaluation set. We have compared the performance of the Self-MobileNet model with reconstructed PPG data using

Cycle-GAN to the original preprocessed PPG segments. We have found significant improvements in AF classification using reconstructed PPG signals from Cycle-GAN. We demonstrate the results of these two experiments with both raw original PPG signals and reconstructed PPG signals by Cycle-GAN in Table 5.5 and Table 5.6.

TABLE 5.5: Performance measure of both Raw PPG Test segments.

Classifier	Evaluation	Types	Accuracy	Precision	Sensitivity	F1_Score	Specificity
Self-MobileNet (q=3)	Test 2	Healthy	92.01	97.06	92.29	94.61	91.13
		AF		78.83	91.13	84.53	92.29
		Weighted Average		92.69	92.01	92.19	91.41
	Test 1	Healthy	93.47	94.03	97.41	95.69	82.02
		AF		91.59	82.02	86.54	97.41
		Weighted Average		93.41	93.47	93.35	85.96

We have found 97.09% accuracy with 5.08% improvement when utilizing reconstructed PPG data compared to original preprocessed signals when using Test Data 2 as unseen data. When we have used Test data 1 as unseen data. We have found that AF classification performance has also increased. When comparing reconstructed AF signals to raw original PPG signals, we can see a significant improvement in sensitivity (10.09%). Considering both experiments using Test Segments 1 and 2, we can see that overall classification performance of AF detection significantly improves on reconstructed PPG signals using Cycle-GAN. The noisy PPG and their corresponding reconstructed signals for both AF and non-AF are illustrated below.

TABLE 5.6: Performance measure of both Reconstructed PPG Test segments.

Classifier	Evaluation	Types	Accuracy	Precision	Sensitivity	F1_Score	Specificity
Self-MobileNet (q=3)	Test 2	Healthy	97.07	97.24	98.98	98.1	91.09
		AF		96.57	91.09	93.75	98.98
		Weighted Average		97.08	97.09	97.06	92.98
	Test 1	Healthy	96.41	97.3	97.89	97.59	92.11
		AF		93.75	92.11	92.92	97.89
		Weighted Average		96.39	96.41	96.4	93.59

We have conducted a number of experiments using the Self-MobileNet model for classification tasks with various q orderings for the Self-ONN layers. The Self-MobileNet model's performance against that of the Self-ONNs and CNN layers is compared in Tables 5.8. It is clear that Self-ONN layers with q=3, 5 and 7 in the Self-MobileNet model perform better than traditional CNN layers, where q = 1 in Self-ONN layers.

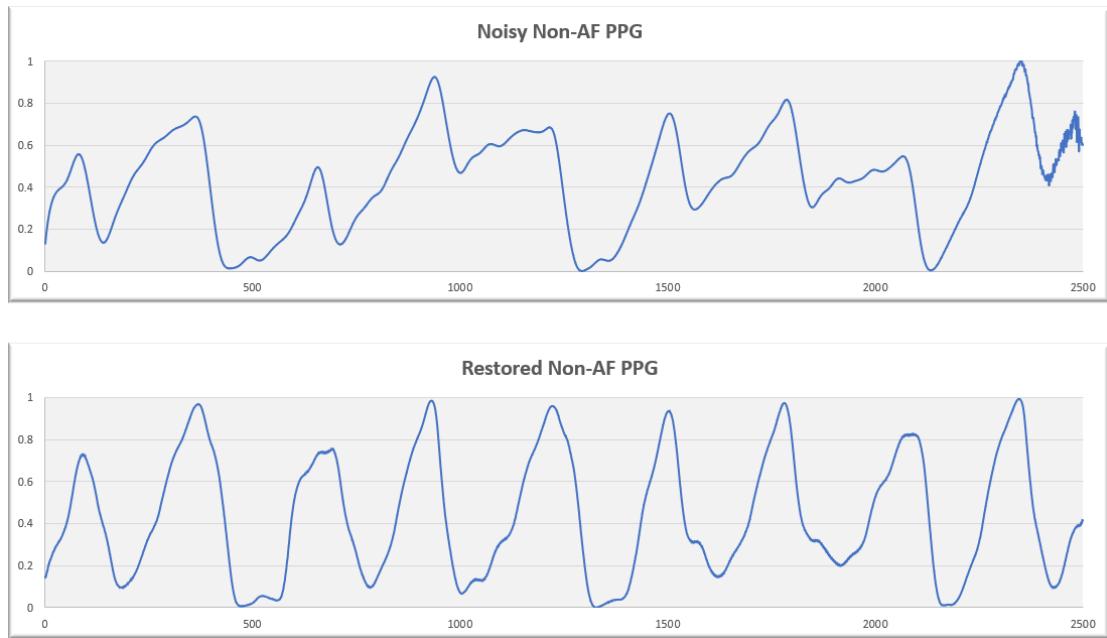


FIGURE 5.2: Noisy PPG and their corresponding reconstructed PPG signal with Non-AF patient.

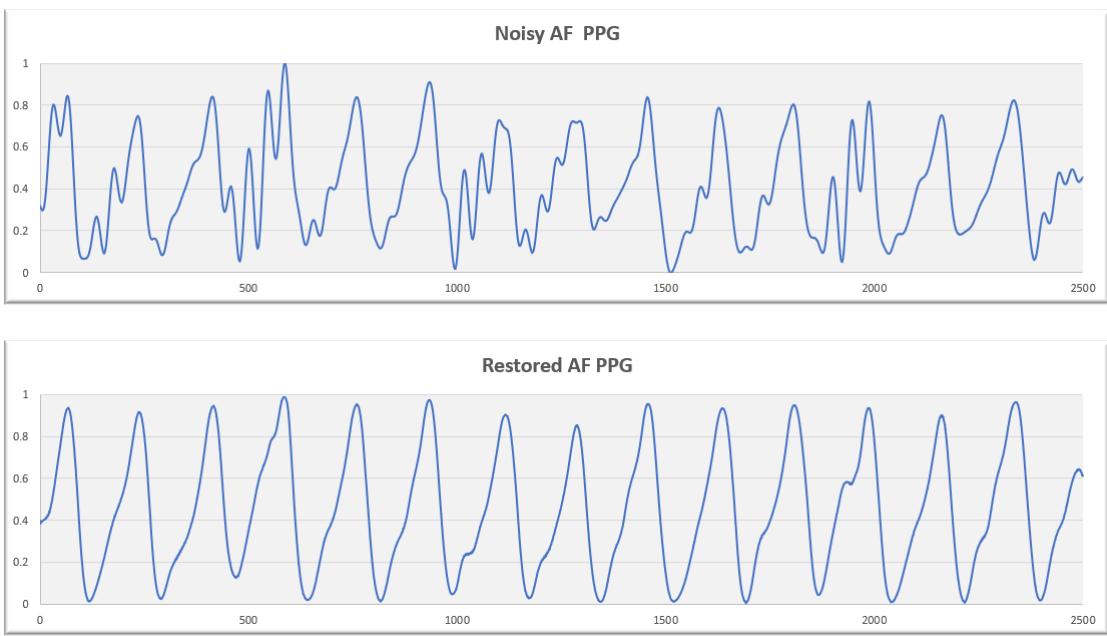


FIGURE 5.3: Noisy PPG and their corresponding reconstructed PPG signal with AF patient.

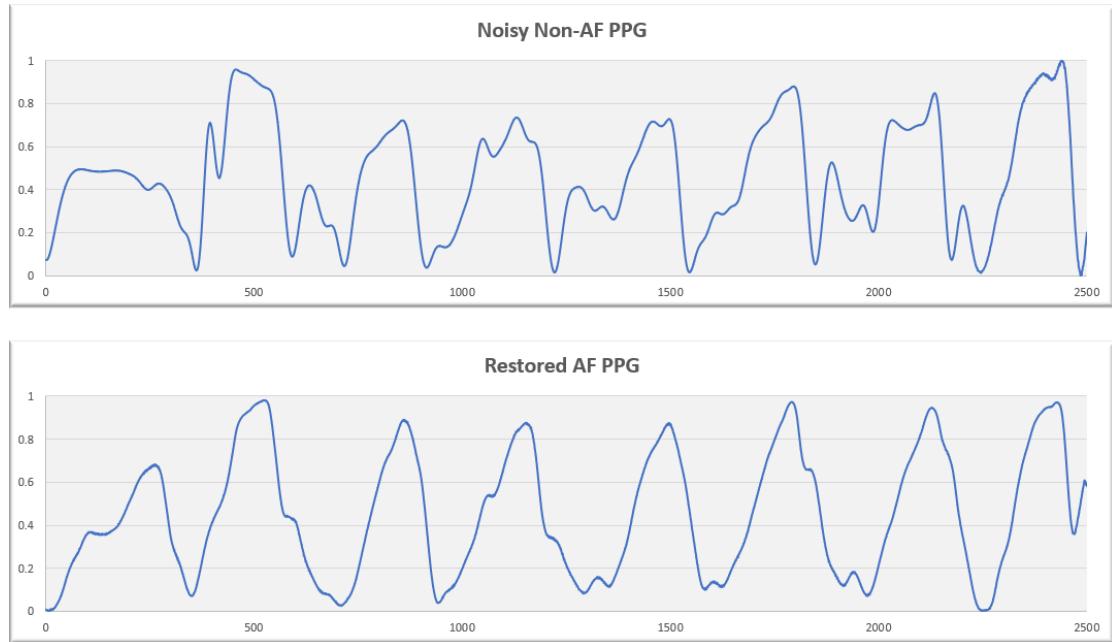


FIGURE 5.4: Noisy PPG and their corresponding reconstructed PPG signal with Non-AF patient.

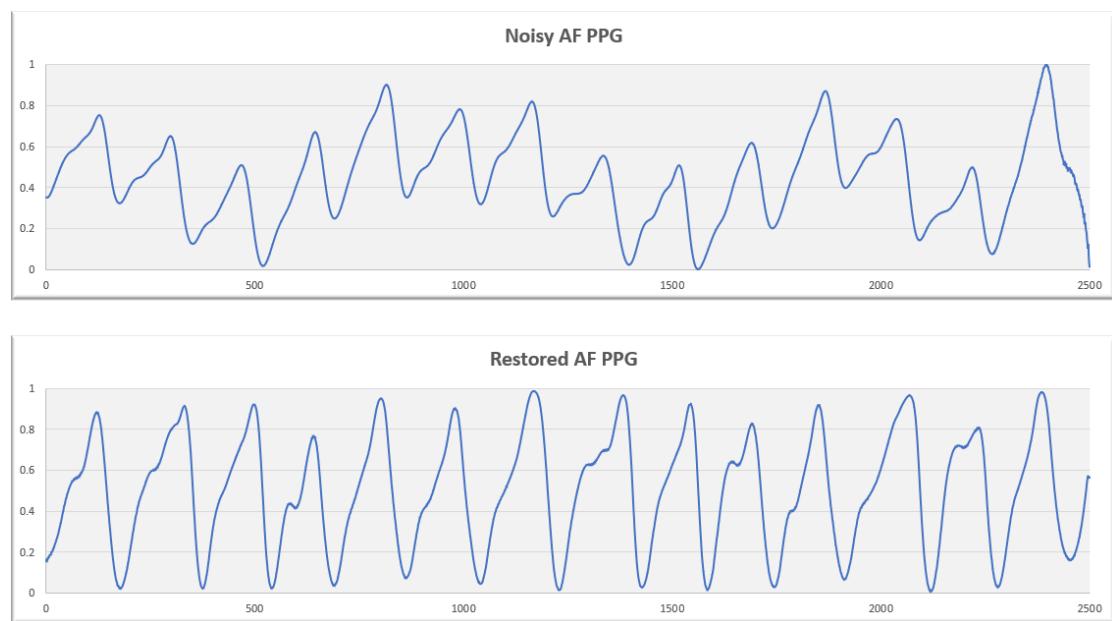


FIGURE 5.5: Noisy PPG and their corresponding reconstructed PPG signal with AF patient.

TABLE 5.7: Representation of confusion matrix of both Test PPG segments.

		Predicted Class			
Ground Truth	Raw PPG Test 2		Reconstructed PPG Test 2		
		Healthy	AF		Healthy
	Healthy	8051	637	Healthy	8635
	AF	244	2506	AF	245
	Raw PPG Test 1		Reconstructed PPG Test 1		
		Healthy	AF		Healthy
	Healthy	7107	189	Healthy	7142
	AF	451	2057	AF	198
					2310

TABLE 5.8: Performance comparison of the Self-MobileNet model with different q orders using reconstructed PPG segments.

Classifier	Evaluation	Accuracy	Precision	Sensitivity	F1_Score	Specificity
Self-MobileNet (q=1) (CNN)	Test 2	95.69	95.84	95.69	95.74	94.61
Self-MobileNet (q=3)		97.09	97.08	97.09	97.06	92.98
Self-MobileNet (q=5)		96.82	96.81	96.82	96.81	94.14
Self-MobileNet (q=7)		96.94	96.93	96.94	96.91	92.63
Self-MobileNet (q=1) (CNN)	Test 1	94.58	94.55	94.58	94.49	88.12
Self-MobileNet (q=3)		96.41	96.39	96.41	96.4	93.59
Self-MobileNet (q=5)		96.34	96.33	96.34	96.34	93.92
Self-MobileNet (q=7)		96.31	96.33	96.31	96.31	94.44

For the AF identification task, we have also employed ECG data that corresponded to PPG signals. We have partitioned the ECG signals for evaluation into 2 parts likewise the wrist PPG segments. Similar to wrist PPG signals, we have divided the ECG signals into two portions for evaluation. ECG signals are more reliable for any task involving the detection of abnormal cardiac rhythms because they are relatively clear and less motion-corrupted than wrist PPG signals. We have carried out an ECG-based AF classification task to test the reliability of our PPG-based AF classification pipeline, which incorporates Cycle-GAN to restore the motion-corrupted PPG signals. We present the performance outcomes of AF classification using ECG segments in Table 5.10. We have achieved the best performance using Self-ONN layers of $q=3$ in the Self-MobileNet model. It is clear that the results of restored PPG-based AF classification performance using Cycle-GAN closely match the results of ECG-based AF classification performance.

TABLE 5.9: Performance measure of both ECG Test segments.

Classifier	Evaluation	Types	Accuracy	Precision	Sensitivity	F1_Score	Specificity
Self-MobileNet (q=3)	Test 2	Healthy	98.97	98.97	99.55	98.74	99.14
		AF		98.97	98.1	99.31	98.7
		Weighted Average		98.97	98.98	98.97	98.97
	Test 1	Healthy	98.07	97.9	99.55	98.72	93.78
		AF		98.62	93.78	96.14	99.55
		Weighted Average		98.08	98.07	98.06	95.26

5.3.4 ROC-AUC Curve

We have achieved an excellent AUC value of 0.999 using q=3 in Self-MobileNet on ECG Data. According to the ROC curve of ECG data shown in Figure 5.6, our proposed method can successfully differentiate (nearly perfect classifier) between data from healthy and AF. On the contrary, we have obtained very handsome AUC value for the reconstructed PPG data of test 1 and test 2 (0.987 and 0.984) which is the good indication of predicting the AF and Non-AF samples from only the reconstructed data generated by Cycle-GAN is shown in Figure 5.7.

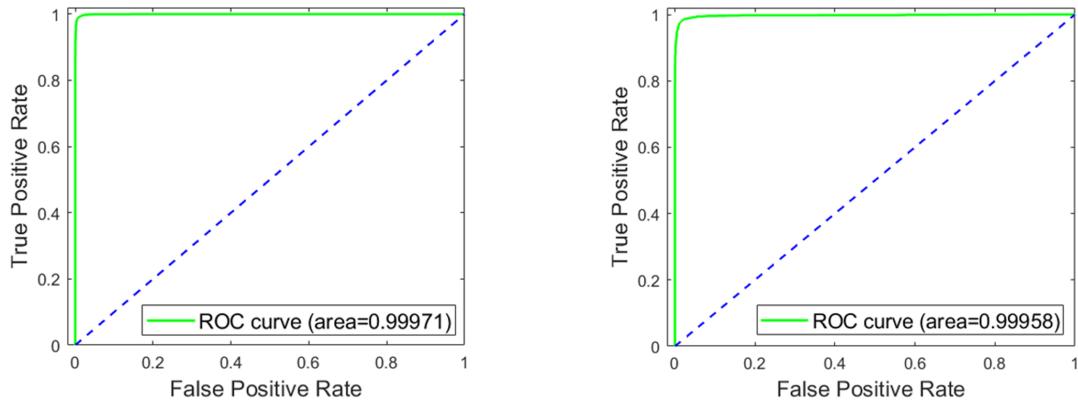


FIGURE 5.6: ROC curves using test segment 2 ECG signals (left) and test segment 1 ECG signals (right).

5.3.5 Quantitative Evaluation

For wrist PPG reconstruction using Cycle-GAN, we do not have any exact clean PPG signals for corresponding motion-corrupted signals. Therefore, we cannot use any proper metrics like RMSE, MAE since it has no proper ground truth. To measure the amount of irregularity and unpredictability of wrist PPG signals, we

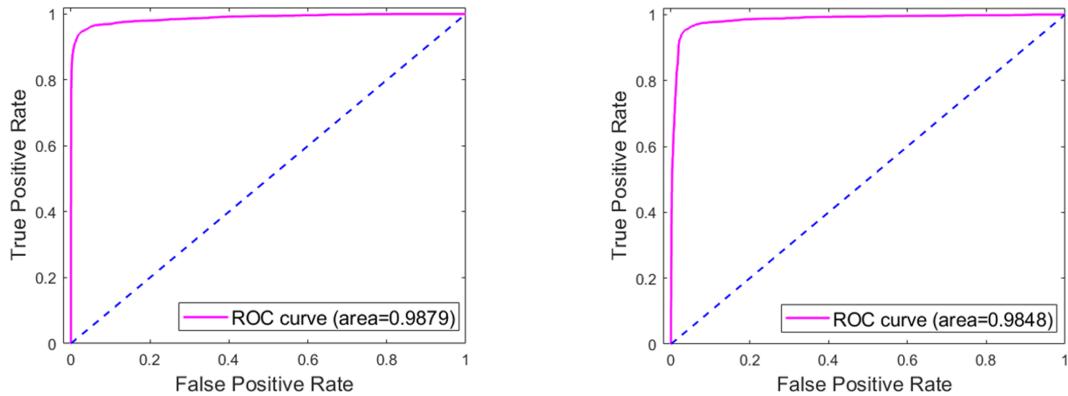


FIGURE 5.7: ROC curves using test segment 2 PPG signals (left) and test segment 1 PPG signals (right).

have used several entropy functions such as fuzzy entropy, sample entropy and approximate entropy. Results show that degree of irregularity and noise decreases using reconstructed PPG signals.

TABLE 5.10: Evaluation of Cycle-GAN for different passes with entropy functions.

Types	Fuzzy Entropy	Sample Entropy	AP Entropy
Raw Original	0.0202	0.1869	0.2316
Pass 1	0.0173	0.01815	0.2124
Pass 2	0.0169	0.1802	0.2169

CHAPTER 6

CONCLUSION AND FUTURE WORKS

This study presented a novel approach for classifying atrial fibrillation signals without using any specific information from patients. We have employed state-of-the-art pre-processing techniques and proposed a robust 1D Self-MobileNet model for AF classification. ECG act as a gold standard for reliable atrial fibrillation detection. Since it has short monitoring cycles and data collection issues, wrist PPG signals non-invisibly collected by smart watches can be extensively used to detect AF. We have run two experiments for evaluation to provide comparative analysis for robust AF detection in real world applications. We have employed several Self-ONN layers in 1D Self-MobileNet model and conducted experiments with different Maclaurin approximations (i.e. $q=1, 3, 5$ and 7). Here, each generative neuron in an operational layer has the ability to enhance each kernel's nodal operator function. Heterogeneous diversity at neuron-level subsequently improves the network diversity and classification performance. We have achieved the best performance on both ECG and wrist PPG segments using order of $q=3$. This is the first study attempting to remove motion artifacts and baseline drifts and restore wrist PPG signals with the aid of 1D-Cycle-GAN. Results demonstrate that performance of our AF classification pipeline significantly improves using reconstructed PPG signals with Cycle-GAN and it closely matches to the ECG performance of AF detection task. 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. It is a very challenging task to acquire 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 many signal processing techniques were extensively used to eliminate MA and signal cut-off conditions, the results were not always satisfactory in some circumstances. We can use 1D version of Cycle-GAN for reconstruction of PPG signals to eliminate the motion- artifacts and noises of PPG signals. Then we can create a robust pipeline to extract morphological features for measuring blood pressures, respiratory rate, and blood oxygen levels. In addition, we can efficiently extract information about cardio-vascular system and detect several arrhythmia diseases.

BIBLIOGRAPHY

- [1] 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.
- [2] G. Lippi, F. Sanchis-Gomar, and G. Cervellin, “Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge,” *Int. J. Stroke*, vol. 16, no. 2, p. 217–221, 2021.
- [3] B. A. Schoonderwoerd, M. D. Smit, L. Pen, and I. C. V. Gelder, “New risk factors for atrial fibrillation: causes of ‘not-so-lone’ atrial fibrillation.”
- [4] 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,” *Am. J. Cardiol*, vol. 103, no. 11, p. 1572–1577, 2009.
- [5] M. Myrstad *et al.*, “Effect of years of endurance exercise on risk of atrial fibrillation and atrial flutter,” *Am. J. Cardiol*, vol. 114, no. 8, p. 1229–1233, 2014.
- [6] E. Guasch *et al.*, “Atrial fibrillation promotion by endurance exercise: demonstration and mechanistic exploration in an animal model,” *J. Am. Coll. Cardiol*, vol. 62, no. 1, p. 68–77, 2013.
- [7] S. S. Barold, “Willem Einthoven and the birth of clinical electrocardiography a hundred years ago,” *Card. Electrophysiol. Rev*, vol. 7, no. 1, p. 99–104, 2003.
- [8] C. Bruser, J. Diesel, M. D. H. Zink, S. Winter, P. Schauerte, and S. Leonhardt, “Automatic detection of atrial fibrillation in cardiac vibration signals,” *Health informatics*, vol. 17, no. 1, p. 162–171, 2012.

- [9] C. Yang, C. Veiga, J. J. Rodriguez-Andina, J. Farina, A. Iniguez, and S. Yin, “Using PPG signals and wearable devices for atrial fibrillation screening,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 11, p. 8832–8842, 2019.
- [10] G. H. Tison *et al.*, “Passive detection of atrial fibrillation using a commercially available smartwatch,” *JAMA Cardiol.* vol. 3, no. 5, p. 409–416, 2018.
- [11] N. Singh, S. Chun, D. Hadley, and V. Froelicher, “Clinical implications of technological advances in screening for atrial fibrillation,” *Prog. Cardiovasc.*, vol. 60, p. 550–559, 2018.
- [12] 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.
- [13] A. L. Goldberger *et al.*, “Physiobank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.”
- [14] N. Tajbakhsh *et al.*, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” *Imaging*, vol. 35, no. 5, p. 1299–1312, 2016.
- [15] Q. Wang, J. Wan, and X. Li, “Robust hierarchical deep learning for vehicular management,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, p. 4148–4156, 2018.
- [16] R. Assabumrungrat *et al.*, *Ubiquitous Affective Computing: A Review.* IEEE Sens, 2021.
- [17] S. Purushotham, C. Meng, Z. Che, and Y. Liu, “Benchmarking deep learning models on large healthcare datasets,” *Inform.*, vol. 83, p. 112–134, 2018.
- [18] S. Liu, A. Wang, X. Deng, and C. Yang, “Mgnn: A multiscale grouped convolutional neural network for efficient atrial fibrillation detection,” vol. 148, p. 105863, 2022.
- [19] M. U. Zahid, S. Kiranyaz, and M. Gabbouj, “Global ECG classification by self-operational neural networks with feature injection,” *Eng.*

- [20] 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, 2021.
- [21] T. Pereira *et al.*, “Pulse pressure waveform estimation using distension profiling with contactless optical probe,” *Eng. Phys.*, vol. 36, no. 11, p. 1515–1520, 2014.
- [22] 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.
- [23] 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, p. 57–61, 2016.
- [24] A. Reiss, P. Schmidt, I. Indlekofer, and K. V. 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.
- [25] 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, p. 172692–172706, 2020.
- [26] S. K. Bashar *et al.*, “Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches,” ” *Sci. Rep.*, vol. 9, no. 1, p. 1–10, 2019.
- [27] F. Mohagheghian *et al.*, “Optimized signal quality assessment for photoplethysmogram signals using feature selection,” *Eng.*, vol. 69, no. 9, p. 2982–2993, 2022.
- [28] I. Goodfellow *et al.*, “Generative adversarial networks,” *Commun. ACM*, vol. 63, no. 11, p. 139–144, 2020.

- [29] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” Mar. 2017.
- [30] 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, p. 294–308, 2021.
- [31] J. Malik, S. Kiranyaz, and M. Gabbouj, “Self-organized operational neural networks for severe image restoration problems, Neural Networks,” vol. 135, p. 201–211, 2021.
- [32] T. Ince *et al.*, “Early bearing fault diagnosis of rotating machinery by 1d self-organized operational neural networks, IEEE Access,” vol. 9, p. 139260–139270, 2021.
- [33] O. C. Devecioglu, J. Malik, T. Ince, S. Kiranyaz, E. Atalay, and M. Gabbouj, “Real-time glaucoma detection from digital fundus images using self-onns, IEEE Access,” vol. 9, p. 140031–140041, 2021.
- [34] 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.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] L. Friberg and L. Bergfeldt, “Atrial fibrillation prevalence revisited,” *Journal of internal medicine*, vol. 274, no. 5, pp. 461–468, 2013.
- [39] 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 lékařství*, vol. 58, no. 9, pp. 633–639, 2012.
- [40] 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.
- [41] G. Borianì and L. Padeletti, “Management of atrial fibrillation in bradycardias,” *Nature Reviews Cardiology*, vol. 12, no. 6, pp. 337–349, 2015.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] S. H. Kim, K. R. Lim, J.-H. Seo, D. R. Ryu, B.-K. Lee, B.-R. Cho, and K. J. Chun, “Higher heart rate variability as a predictor of atrial fibrillation in patients with hypertension,” *Scientific reports*, vol. 12, no. 1, pp. 1–8, 2022.
- [47] 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.
- [48] 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.
- [49] 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.
- [50] 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.
- [51] 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.
- [52] 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.

- [53] 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.
- [54] 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.
- [55] 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.
- [56] 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.
- [57] 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.
- [58] 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.
- [59] 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.
- [60] 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.

- [61] 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.
- [62] 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.
- [63] 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.
- [64] 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.
- [65] 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.
- [66] 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.
- [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] 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.

- [69] 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.
- [70] K. Aschbacher, D. Yilmaz, Y. Kerem, S. Crawford, D. Benaron, J. Liu, M. Eaton, G. H. Tison, J. E. Olgin, 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.
- [71] G. Lippi, F. Sanchis-Gomar, and G. Cervellin, “Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge,” *International Journal of Stroke*, vol. 16, no. 2, pp. 217–221, 2021.
- [72] 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.
- [73] S. Kiranyaz, T. Ince, A. Iosifidis, and M. Gabbouj, “Operational neural networks,” *Neural Comput. Appl.*, vol. 32, no. 11, p. 6645–6668, 2020.
- [74] J. Malik, S. Kiranyaz, M. Yamac, E. Guldogan, and M. Gabbouj, “Convolutional versus Self-Organized Operational Neural Networks for Real-World Blind Image Denoising,” *arXiv Prepr. arXiv2103*, p. 03070, 2021.
- [75] M. A. Yílmaz, O. Keleş, H. Güven, A. M. Tekalp, J. Malik, and S. Kíranyaz, *Self-Organized Variational Autoencoders (Self-Vae) For Learned Image Compression, ” in 2021 IEEE International Conference on Image Processing (ICIP)*. 2021, 2021.
- [76] D. T. Tran, S. Kiranyaz, M. Gabbouj, and A. Iosifidis, *Knowledge transfer for face verification using heterogeneous generalized operational perceptrons, ” in 2019 IEEE International Conference on Image Processing (ICIP)*, 2019.
- [77] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C., Chen, *Mobilenetv2: Inverted residuals and linear bottlenecks, in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.

- [78] J. Bacevičius *et al.*, *Long-term electrocardiogram and wrist-based photoplethysmogram recordings with annotated atrial fibrillation episodes.* doi: 10.5281/ZENODO.5815074.