Atrial Fibrillation Detection Service for Stroke Prevention

Murtadha Kamil Kareem Kareem

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

September 2021

Candidate Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.

2. None of the material contained in the thesis has been used in any other submission for an academic award.

3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

4. The work undertaken towards the thesis has been conducted in accordance with the SHU Principles of Integrity in Research and the SHU Research Ethics Policy.

5. The word count of the thesis is 47897.

Name	Murtadha Kareem
Date	20/09/2021
Award	PhD
Faculty	Science Technology and Art
Director(s) of Studies	Dr Oliver Faust

DEDICATION

This study is passionately dedicated to my beloved parents, siblings, and my Director of Study Dr. Oliver Faust. Special gratitude to my father-Kamil Buraihi who died from heart failure due to a COVID-19 infection in 2020. He was in the hospital in IRAQ and kept asking me to return to the UK and complete my research. From my heart, I would say that he was a genuine and an extraordinary father from all prospective. The beloved ones have been the source of inspiration, spiritual and strength to achieve the success.

Abstract

This study aims to develop a cost-effective atrial fibrillation detection service that improves outcomes for patients. The service offers continuous Atrial Fibrillation (AF) detection which might help to address the real-world problem of stroke prevention. AF is the most common sustained heart rhythm disorder in adults. AF is either intermittent (paroxysmal) or permanent. Both types increase the risk ischemic stroke around fivefold. An accurate diagnosis of AF is mandatory for treatment initiation. AF treatment reduces the stroke risk and for individual patients prevent stroke. Unfortunately, current AF detection methods often fail to detect paroxysmal AF cases, because the observation duration is too short. We propose to address this problem with real time monitoring and artificial intelligence for AF detection. We developed two distinct deep learning models to detect irregular heartbeats. For the first experiment, a Long Short-Term Memory (LSTM) classifier was used to detect and differentiate AF beats and Normal beats. The data were collected from MIT-BIH Atrial Fibrillation Database. This database incorporates 10-hour Electrocardiogram (ECG) signals from 23 participants. The second experiment was based on using a ResNet algorithm to detect common arrhythmias, namely, AF, and Atrial Flutter (AFL), as well as Normal Sinus Rhythm (NSR). The algorithm was trained with data from 4051 subject. The LSTM model achieved 98.51% accuracy with 10-fold cross-validation (20 subjects) and 99.77% with blindfold validation (3 subjects). Whilst, the ResNet model achieved, the following results: accuracy = 99.98%, sensitivity = 100.00%, and specificity = 99.94%. In addition, the LSTM model was validated with five independents benchmark databases to establish the robustness and maturity through more and more varied datasets. With the LSTM validation, we established trust which enabled us to conduct a clinical trial study with Sheffield Teaching Hospital. As part of this work, an AF detection service validation tool was built for hybrid decision support where machine learning decisions are verified by a stroke consultant. This detection method makes economic sense because Heart Rate (HR) signals are cost-effective to measure, transmit, and process. Having such a cost-effective solution might lead to widespread long-term observation, which can help detecting arrhythmia earlier. Detection improves the outcomes for patients and reduces healthcare cost.

Acknowledgement

Foremost, it is a genuine pleasure to express my deep sense of thanks and gratitude to my DoS Dr. Oliver Faust. His dedication and keen interest above all his overwhelming attitude to help his first PhD student had been solely and mainly responsible for completing my work. His timely advice, meticulous, scrutiny, scholarly advice and scientific approach have helped me to a very great extent to accomplish this award.

In addition, I would like to thank my supervisory team Prof Reza and Prof Marcos for their kind support and guide. From PGRT team, I would like to thank Dr. Xu for her valuable advice and direction during the drop-in sessions. Grateful for all staff whom being a matter of help and support at Sheffield Hallam university.

I owe a deep sense of gratitude to Dr. Ali Ali, from Sheffield Teaching Hospital. His previous time is much appreciated at every stage of our clinical trial study.

I would like to thank Grow MedTech for funding our cutting-edge research and translate it in the practical settings. Besides that, I thank profusely all Grow MedTech manager for supporting the project and make it happen.

It is my privilege to thank Keith Errey the CEO of Isansys for his dedicated support in providing the innovative sensors and technology for Proof of Feasibility study.

Finally, I would like to express the deepest appreciation for my parents and family members and friends. Extremely thankful to my father who was substantially supported my study financially and encourage me to broaden my knowledge in the UK.

List of abbreviations

ACC	Accuracy
AF	Atrial Fibrillation
AFDB	Atrial Fibrillation Database
AFL	Atrial Flutter
ANN	Artificial Neural Network
API	Application Programming Interface
AUC	Area Under Curve
BLE	Bluetooth Low Energy
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DBN	Deep Belief Network
DCT	Discrete Cosine Transform
DNN	Deep Neural Network
DL	Deep Learning
ECG	Electrocardiogram
EMD	Empirical Mode Decomposition
FD	Fractal Dimension
FN	False Negative
FP	False Positive
FT	Fourier Transform

GOE	Global Observatory for eHealth
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HOS	Higher Order Spectra
HR	Heart Rate
HRA	Health Regulatory Approval
HRV	Heart Rate Variability
HRVAS	Heart Rate Variability Analysis Software
ICA	Independent Component Analysis
ICT	Information and Communication Technology
IIoMT	Intelligent Internet of Medical Things
IoT	Internet of Thing
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LLE	Largest Lyapunov Exponent
LMNN	Levenberg–Marquardt Neural Network
LSTM	Long Short-Term Memory
MCOT	Mobile Cardiac Outpatient Telemetry
ML	Machine Learning
NB	Naive Bayes
NSR	Normal Sinus Rhythm

NSRDB	Normal Sinus Rhythm Database
PCA	Principle Component Analysis
PNN	Probabilistic Neural Network
PPG	Photoplethysmogram
PPIEP	Public Patient Involvement, Engagement and Participation
PPV	Positive Predictive Value
PSD	Power Spectral Density
PSE	Patient Status Engine
ReLU	Rectified Linear Unit
RF	Random Forest
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
RQA	Recurrence Quantification Analysis
SEN	Sensitivity
SNR	Signal to Noise Ratio
SPE	Specificity
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
SWT	Stationary Wavelet Transform
TIA	Transit Ischemic Attack
TN	True Negative
TP	True Positive

TVCF	Time-Varying Coherence Function
WHO	World Health Organisation
WPD	Wavelet Packet Decomposition

List Publications and awards

Journal Publications

1. O. Faust, A. Shenfield, M. Kareem, T. San, H. Fujita, and U. Acharya, "Automated detection of atrial fibrillation using long short-term memory network with RR interval signals," *Comput. Biol. Med.*, vol. 102, 2018, doi: 10.1016/j.compbiomed.2018.07.001

2. O. Faust, M. Kareem, A. Shenfield, A. Ali, and U. R. Acharya, "Validating the robustness of an internet of things based atrial fibrillation detection system," *Pattern Recognit. Lett.*, vol. 133, 2020, doi: 10.1016/j.patrec.2020.02.005.

3. N. Lei, M. Kareem, S. K. Moon, E. J. Ciaccio, U. R. Acharya, and O. Faust, "Hybrid decision support to monitor atrial fibrillation for stroke prevention," *Int. J. Environ. Res. Public Health*, vol. 18, no. 2, 2021, doi: 10.3390/ijerph18020813.

4. M. Kareem, N. Lei, A. Ali, E. J. Ciaccio, U. R. Acharya, and O. Faust, "A review of patient-led data acquisition for atrial fibrillation detection to prevent stroke," *Biomed. Signal Process. Control*, vol. 69, no. March, p. 102818, 2021, doi: 10.1016/j.bspc.2021.102818

5. O. Faust, M. Kareem, A. Ali, E. J. Ciaccio, and U. R. Acharya, "Automated Arrhythmia Detection Based on RR Intervals," *Diagnostics*, vol. 11, no. 8, pp. 1–18, 2021.

6. O. Faust, M. Kareem and N. Lei, Atrial fibrillation detection service validation tool, Software Impacts (2021).

Conference Publications

1. M. Kareem and O. Faust, "Establishing the safety of a smart heart health monitoring service through validation," 2019, doi: 10.1109/BigData47090.2019.9006478USA.

2. M. Kareem and O.Faust," Validating a deep learning model for atrial fibrillation detection"2021, Condition Monitoring/UK.

Awards

1. 20th International Conference on Wireless Healthcare Systems outstanding work entitled Design and Implementation of a “Wireless Heart Health Monitoring System for Stroke Risk Mitigation”.

2. Secondment with Isansys £5k funded by Grow MedTech.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

Stroke, also known as cerebrovascular accident, is the second leading cause for mortality and the third leading cause of disability worldwide [1], [2]. Prevalence in the general population and severity of outcomes are the main drivers behind these statistics. Annually, stroke affects around 15 million adults globally [3]. About one in five of the individuals who experience a stroke die during the first 30 days [4], and over 40% of stroke survivors are functionally dependent after 6 months [5]. Around 5.7 million people die annually from stroke worldwide.

Ischemic stroke occurs when the bloodstream to any part of the brain is blocked by blood clots. Once this occurs, brain tissue is damaged due to a lack of oxygen. Therefore, the brain needs a steady supply of blood, delivering essential oxygen and nutrients, to maintain normal levels of functionality. The Framingham study showed that there is a connection between Atrial Fibrillation (AF) and stroke, by examining 5,070 participants [1]. The researchers found a positive correlation between nonrheumatic AF and stroke. Clinical studies showed that AF, either permanent or intermittent (paroxysmal), increases the risk of cardio-embolic stroke fivefold. Around 25% of all strokes happen in people aged 65 years or younger. In the UK, the estimated rate of having a stroke, in people aged 45 years and below, is approximately 20,000 cases every year [6]. There are 150,000 people getting stroke in the UK each year [7], and 53,000 people died [8]. The overall cost for stroke treatment is expected to exceed £ 9 billion in the UK [9]. In addition, stroke ranks third in the annual mortality statistics in the UK, after heart attacks and cancer [10]. Stroke prevalence differs according to age and sex. Statistics indicate that death caused by stroke in men is 9% in the UK, whereas 13% female fatalities were accounted to stroke. The severity of strokes in people with AF is higher and a stroke event has worse outcomes

when compared to people who have strokes without AF [11]. Therefore, AF treatment can reduce the stroke risk.

Continuous treatment monitoring can be achieved through automated AF detection and wireless communication technology. In the past, remote health monitoring was difficult because appropriate mobile communication technology was not available [12]. Wireless networks can play a main role in continuously monitoring patient's health. The purpose of utilising wireless networks is to establish the continuous patient monitoring in a cost-effective way. Wearable health care systems monitor various vital signs from a patient, such as blood pressure, Heart Rate (HR), heartbeats, respiratory rate, body temperature and body orientation [13]. For instance, nurses or caregivers might not be available 24/7 in hospitals or at the home environment for monitoring patients' status constantly. Therefore, there are periods of insufficient care. During these periods, the patient condition might significantly worsen. The only way to avoid the care gaps is to monitor the patient continuously. However, nurses or other human caregivers are inefficient in monitoring patient health. A much better setup is to involve care giving only if it is needed and not spend time and attention when the patient is not in a critical phase.

We propose a real time AF detection service for stroke prevention. This service comprises of a commercial HR patch, central server to access patient HR data, automated AF detection algorithm, physician support tool to review and verify the machine decision. The monitoring duration of the proposed service is not limited. That means our service can support a long observation duration, which might help to detect paroxysmal AF cases. The proposed value for the healthcare providers is twofold. From a medical perspective, a long observation duration has the potential to generate a higher AF detection rate in patients who use the monitoring service. Furthermore, the unlimited observation duration enables a cardiologist to monitor the AF treatment's efficacy indefinitely. To achieve this task, we introduce the concept of patient-led data acquisition where HR data travels from the point of measurement (patient) to a central storage point, i.e., a cloud server. As such patient led data acquisition is ideal for long-term monitoring. The second value proposition comes from hybrid decision support, which leads to efficiency in terms of both time and cost. A physician gets involved only if a deep learning algorithm detected a sequence of AF beats; at all other times, human expert intervention is not required. Furthermore, the combination of continuous machine analysis and human oversight creates a cost-effective system for hybrid decision support. Hence, the AF detection

service reduces the time a medical doctor spends on routine screening tasks. Once the estimated AF probability is established, the service provides information extraction tools to analyse critical sections of the HR trace effectively. The physician can combine the extracted information with other information sources, such as patient records, to reach a safe and reliable diagnosis. These diagnosis results can be transmitted via a feedback channel to the patient. Implementing the AF detection algorithm for real-time monitoring loads a current Central Process Unit (CPU) core by about 50%. This translates into low processing cost if the algorithm runs on a cloud server. Furthermore, HR has a low data rate and high information content when compared with ECG signals. As such, the low data rate implies that the wireless HR sensors have a low energy consumption, which keeps both size and cost down. The value propositions focus on the healthcare provider. The patient benefits from the AF detection service through patient-led data acquisition, unobtrusive HR measurement and peace of mind through real-time HR monitoring and diagnosis.

1.2 Aims and objectives

This study aims to develop a cost-effective atrial fibrillation detection service that improves outcomes for patients. The main objectives of this study are as follows:

1. Measure the electrical activity through RR intervals signal.
2. Deploy the RR intervals for AF detection to prevent a stroke.
3. Use a computer-aided-diagnosis that represented by a deep learning algorithm to detect AF episodes continuously.
4. Detect more AF events based on long-term monitoring.
5. Validate a deep learning algorithm with more and more datasets.
6. Develop service directions that improve patient outcomes.

1.3 Research Questions

1. Can we detect atrial fibrillation from RR intervals?
2. Does validation setup increase the confidence?
3. Does the automated atrial fibrillation detection service reduce the risk of stroke?
4. Do the service directions help patients and healthcare provider?
5. How do we continue to improve deployed medical decision support systems?
6. How do we improve the safety of a machine decision?

In the literature review, we will support and detail this research question.

1.4 Tasks and Steps taken

The following list details the steps taken to achieve certain objectives:

1. Develop techniques based on neural computing deep learning for AF diagnosis. Evaluate the system's performance with benchmark data from freely available databases (i.e., PhysioNet).
2. Design, build and evaluate a HR monitoring system that incorporates the latest wireless technology to be record RR intervals related health data.
3. Develop cloud computing techniques to facilitate communication between the signal monitor and facilities to store information.
4. Patient led data acquisition through real-time transmission.
5. Validate the LSTM model with unknown data from varied databases.
6. Validate the proposed AF detection service for stroke prevention in the clinical trial.
7. Validate the deep learning classification results by an experienced cardiologist.
8. Establish the feasibility of hybrid decision support through results analysis formulation of the clinical study.

1.5 Expected Contributions to Knowledge

This research will contribute to the literature substantially in terms of human factor, technology, and cost [9].These factors can establish an effective environment to real-time patient-led data acquisition. To be specific, that scope focuses on monitoring with only non-invasive devices which do not require surgical procedures. As such, they offer technical solution which can help to detect more AF. We propose HR signal measurement as an ideal technology for long-term AF detection [14]. HR signal is a cost-effective method to measure, communicate, store and processing [15]. Distinct deep learning algorithms were developed to detect AF, AFL, and normal rhythm from HR measurement [14], [16], [17]. Intelligent Internet of Medical Things (IIoMT) were developed to facilitate data distribution, processing through integrated DL algorithm in the central cloud system, data storage and visualisation [18]. The combination of IIoMT, DL system, and physician involvement form the service platform concept [19]–[22]. Therefore, these technologies influence the human factors which determine the product success. The goal for all patient-led data acquisition systems is to become a product; hence, cost pressures exist early in the design process. Even research is influenced by arguments about cost, at least when it involves selecting a particular topic or conducting a specific study.

Thus, this research has a novel approach that contains in-depth understanding and knowledge of identifying unmet needs as well as solving the real-world problem which can benefit both patients and healthcare providers. In addition, this research attempts to fill the gaps by using advanced technological approaches that overcomes the challenges of traditional methods.

1.6 Research methodology overview

This research fellows a quantitative approach that demonstrate latest technical solutions in complex computational of digital signal processing. The quantitative approach based on deep learning algorithms rather than traditional machine learning. This is due to the fact that deep learning does not require feature engineering, feature selection and information reduction [23]. All the features can be extracted automatically during the learning phase. In addition, the quantitative methods involve training and testing the DL

model with benchmark data. This data is freely available for research on large database known as PhysioNet. The LSTM model was established to detect AF from RR intervals. Whereas, the ResNet model was created to detect AFL, AF and NSR from around 4050 participants. Both models achieved such a promising result that has been compared with the state of the art. To be specific, the LSTM model was validated with varied datasets to show the potential of an accurate detection as well as evaluating the performance measures.

The safety of machine decision can be improved through human verification [19], [24]. From these projective, we conducted a practical study that based on AF detection in clinical setting. The sample size includes 20 participants from two different cohorts, 10 participants from known AF group and the other 10 from normal group.

1.7 Thesis outline

This thesis structured from seven chapters. Chapter 1 incorporates a background on stroke risk and its relevance to AF. In chapter 2, we conducted an informative literature review on AF occurrence, technology based on signal measurement and devices for this emerging healthcare field that allows us to formulate research gaps which might be addressed in future projects. Chapter 3 introduces the methods used for automated arrhythmia detection. Chapter 4 demonstrates the additional validation setup for a deep learning model to reach the robustness and maturity with varied databases. Chapter 5 based on hybrid decision support to reach an accurate diagnosis. Chapter 6 documents the clinical trial study procedures, data collection and results analysis. The chapter 7 summarizes the thesis findings, and it delivers a conclusion, limitation, and future work.

Chapter 2 Literature Review

In this section, we introduce AF as a heart rhythm disorder, and we discuss various AF detection methods. Furthermore, we make an attempt to frame the research question better by introducing state of the art systems and frameworks that can be used to detect more AF. Having reviewed signal acquisition, communication, and processing for automated AF detection, we are in a position to put forward a number of research gaps.

2.1 Atrial Fibrillation

AF is the most common serious irregular heart rhythm associated with rapid HR in adults [25]. Atrial means to the atria (pleural of atrium), which describes the locations at the top two chambers of the heart. Fibrillation refers to irregular, rapid and unsynchronised contraction of muscle fibres. Sinus rhythm represents the normal beat of the heart, which is managed by a sophisticated electrical control system. This system controls the timing of the heart pump. When the electrical system is functioning correctly, it maintains a normal HR rhythm. Problems with this electrical system can cause an arrhythmia. Two arrhythmia types can be identified which are associated with abnormalities such as tachycardia and bradycardia. Tachycardia describes a situation when the heart beats too fast whereas bradycardia indicates that the heart beats too slow.

The sinus node consists of a cluster of special cells, which acts as the heart's natural pacemaker. The sinus node controls the rate at the atria while the heart muscles contract and relax. In AF, disordered electrical activity progresses in the walls of the atria, exceeding the sinus node. As a result, the rhythm will change from normal to abnormal when the atria begin to fibrillate; that leads to a rapid rhythm as their muscular walls fail to contract with coordination and regularity. 0.4% of adults are affected by this disease, and prevalence increases with age. Less than 1% of people are affected by AF in the age

group of 60- years or younger. The prevalence increases to 6% for those aged 80-years and older [26]. It is anticipated that the occurrence of AF increases, because of the aging population. In addition, AF incidence is related to a significant increase in stroke, heart failure, poor mental health, diminished life quality, as such it is a leading cause of death [27]. This disease is associated with various types of symptoms, such as chest pain, shortness of breath, fainting, fatigue, palpitation, and light-headedness [28]. AF can be treated by a procedure known as cardioversion, which tackles the electrical problem in the heart. that can be subjected to the physicians who attempt to treat the rate and rhythm of the heartbeat, or they can return the heart to normal sinus rhythm.

In 2019, National Health Service (NHS) England defined their ambition to enhance outcomes for stroke patients by increasing the AF detection rate from 79% to 85%. Recent analysis from Public Health England and NHS reveals that, over 3 years, accomplishing optimum treatment for people diagnosed with AF could prevent up to 14,220 strokes across England [9]. This could save £ 240 million. To accomplish that ambition, cost-effective technological solutions for AF detection are needed. Currently, 12-lead Electrocardiogram (ECG) recording and prolonged Holter ECG monitoring are the most common methods used to screen for AF. 12-lead ECG recordings require specialized medical facilities to do the recording. Holter monitoring systems can be applied inpatient wards; however, most of them are bulky, burdensome, and they have a limited monitoring duration. This limits their utility in clinical practice. In other words, resource constraints restrict the use of the measurement equipment, and the observation duration is insufficient to detect all rhythm abnormalities. The medical framework around AF was shaped by these limitations. That framework unfolds with all patients belonging to the ‘first detected AF’ category [29]. If a first detected episode ceases on its own in less than seven days and then another episode begins later, the category changes to paroxysmal AF. Although people in this category have episodes lasting up to seven days, in most cases of paroxysmal AF, the episodes will stop in less than 24 hours. If the episode lasts for more than seven days, it is unlikely to stop on its own and is then known as persistent AF. In this case, cardioversion can be applied to restore the rhythms back to normal. If cardioversion fails, and the episode continues for a long time (i.e., a year or more), the person's AF is classified as permanent. Labelling all confirmed cases as ‘first detected AF’ implies that the arrhythmia might have been present before it was discovered. Hence, it is expected that more measurements will lead to a larger number of first detected AF

cases. However, the signal acquisition period for current AF detection methods is small when compared to the symptomless periods of paroxysmal AF [30], [31]. Consequently, a patient might be symptomless during the recordings. That will result in a false negative diagnosis. Thus, current recording techniques will increase the chance of misdiagnosis. That ambiguity can diminish the number of first detected AF cases which can be increased by extending the recording duration [32], [33]. Detecting more AF might lead to more and earlier treatment which can reduce the stroke severity and in some cases prevent a stroke altogether [34]. That potential to reduce the disease burden is significant because estimates indicate that one-third of all patients who suffered a stroke have undiagnosed AF. Apart from this direct effect on enhancing outcomes for patients, more and longer measurements also aid our understanding [35]. The Framingham heart study has also found that AF is an independent risk factor for stroke [1]. However, the sample size and the analysis methods were inadequate to determine the degree of danger from untreated AF. A better understanding is needed to establish the stroke risk for a specific patient with paroxysmal AF. Treatment is another aspect that could benefit from statistical analysis. Current treatment methods are invasive, and they carry the risk of death. With long-term measurements it might be possible to show the efficacy of less invasive treatment, such as diet, exercise, and lifestyle changes.

2.2 Photoplethysmogram

Photoplethysmogram (PPG) is an optical method that measures the HR from blood volume changes. PPG signals are established by measuring light reflected from human tissues [36]. The main feature of PPG signals is a shape known as peripheral plus, which corresponds to R waves in the ECG. Once the signal is captured, each consecutive beat is quantified and maintained as an RR interval [37]. Hence, the methods used for AF detection are related to those that process RR intervals extracted from ECG. The pulse measurement can be achieved through mobile devices, such as wristwatches [38]. Therefore, the PPG functionality is often being adapted as an add-on feature into mobile devices, which is cost effective. Sensors and associated measurement technology is likely to become widespread for HR signal acquisition in the patient environment.

For user comfort, these devices are designed to be movable which is likely to cause activity specific artefacts in the measurement [39]. These artefacts include mandatary pre-processing algorithms that filter the signal before it is fed to the AF detector. Salesmen et al. [40], addressed the lack of labelled PPG data by medical professional that found from publicly MIT-BIHAF databases repository to generate a PPG detection model. This model designed to consider cardiovascular system; hence, there might be changes by optimising how well that model works for a specific patient. That uncertainty also impacts the AF detection approaches that were established with the PPG model. Another method is to measure PPG alongside ECG. Once the measurement is completed, the labels generated by an experienced cardiologist, from ECG to PPG signals [41], [42]. Dual measurement helps to overcome the limitations that arise from the fact that PPG signals cannot be examined directly through visual inspection by cardiologists. In other words, the cardiologist identifies AF episodes based on ECG signals. Yet, PPG signal has not been used in clinical practice as a reliable diagnostic tool.

2.3 Electrocardiogram

ECG is a physiological signal that measures the electrical activity of the heart. ECG is a basic non-invasive measurement method. This measurement can be achieved through placing several electrodes on the chest and on other parts of the human body. The outcome is a time domain signal. The initial diagnosis of heart conditions is normally made through monitoring both characteristic ECG variations and observed symptoms during the clinical visit. In AF detection, the standard observation is used 12-lead ECG [43]. To reach an effective diagnosis, the signal is analysed by a well-trained clinical physiologist [44].

Figure 1 demonstrates two different segments of ECG signals that were taken from MIT-BIH Atrial Fibrillation Database [45], [46]. Normal Sinus Rhythm (NSR) indicates a normal ECG cycle of healthy heart. ECG comprises from three main components which are P wave, QRS-complex and T wave.

Figure 1 provides an example NSR signal. In that plot, the first component P-wave is marked and annotated. Medically, AF can be identified as a collapse of organised atrial electrical activity [47]. This collapse is indicated by the absence of a P wave in the ECG signal. If that happens, the R peak becomes irregular [48]. The second plot in Row1 of

Figure1 shows an example of an ECG signal that appears with AF signs. The graphical representation of AF is related to irregular time intervals of R peaks due to the absence of P waves. The R peak, in the ECG plots, indicates ventricular depolarization, for example, the time of the heartbeat, assume a and b, is indicated as the RR interval. That time duration forms the amplitude, and the time location of the second beat b is the time location of an RR interval sample. The P wave, labeled in the NSR ECG plot, indicates atrial depolarization.

The morphology of ECG signals varies from person-to-person [49]. ECG-based AF detection means to diagnose the disease-associated with variations. The main difficulty is that these morphology variations might not be exclusive, which results in reducing the specificity in diagnosing the disease that causes those changes. To solve the sensitivity issue, extending the observation duration might work effectively [50]. Moreover, the relationship between AF mechanisms, recorded by surface ECGs, and atrial activity, is not yet well comprehended [51]. This uncertainty translates into imprecise detection algorithms. AF detection based on the symptomatic rhythm disorder could be a helpful option for moving forward. Despite ECG being a non-invasive method, which facilitates that measurement, the ECG measurement setup is complex because the electrical signals have low amplitude. Furthermore, ECG signals have a high data rate, which makes them difficult to distribute and process in real-time.

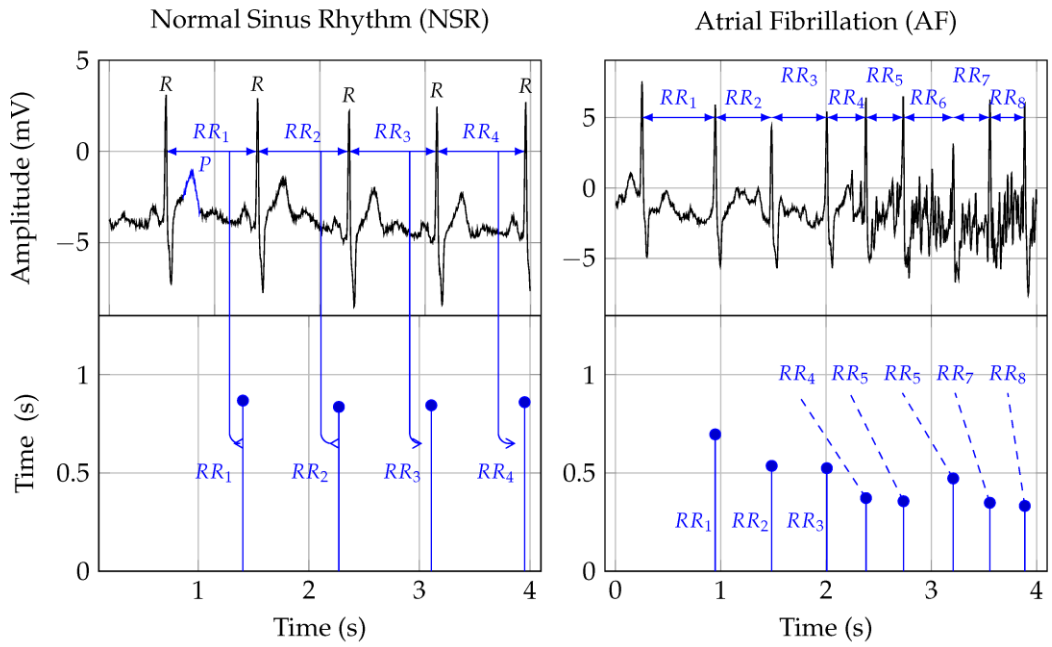


Figure 1: The first row shows the plotting of two different ECG signals and the second row the subsequent RR interval traces. The first column in the plots show NSR, and the second column show AF symptoms [22].

2.4 Heart Rate

AF episodes can be also detected through HR signals [52]. The lower plot in

Figure 1 shows the RR intervals that correspond to the ECG signals shown in the upper plot. As such, beat-to-beat signals are formed from RR intervals. These RR intervals refer to the time measured from one R peak in the ECG signal to the next. As shown in

Figure 1, this concept is indicated by the arrows that showed the RR intervals signal extracted from the ECG. Even for an experienced cardiologist, it is difficult to diagnose the subtle linear and nonlinear disorders in the RR interval trace that related-to AF. Therefore, manual interpretation of HR measurements might result in inter-and intra-observer variability. Algorithmic decision support synchronised with digital biomarkers can assist to overcome these difficulties and consequently enhance the diagnostic quality [53], [54]. In intermittent AF, the incidence of symptomatic episodes is ambiguous [55], [56]. Some AF-events might end more than 48 h [55]. Numerical analysis proof that during AF events, the RR intervals have a shorter correlation length and a greater standard

deviation than those during NSR. Hence, these measures can be applied as digital biomarkers for detecting AF occurrence. However, the measures are not good enough when it involves distinguishing AF from other arrhythmias.

2.4.1 Intelligent internet of medical things based on M-health

Patient-led data acquisition systems must identify patient demands and data specific requirements. Physiological signal recording is a main factor of such systems. Once the signal is recorded and digitized, the samples become data with specific requirements for transmission, storage, and analysis. Patients require improvement of their outcomes and better management of their stroke risk. At the same time, a system needs to be convenient enough to ensure patient compliance [9]. Mobile health, also known as m-health, aims to provide the technical solutions which highlights addresses both data requirements and patient needs [57]. Thus, m-health was defined by the Global Observatory for eHealth (GOE) as medical and public health practice that operates through portable devices, such as mobile phones, patient monitoring tablets, personal digital assistants, and other wireless devices [58]. The core idea of m-health systems is that the data rather than the patient travels. More specifically, close to the patient the data is interconnected through wireless channels to increase convenience for long-term recordings. M-health delivers hospital-in-house services that overcome geographical as well as organizational impediments [59], [60] . More advanced m-health services provide automated diagnosis support through built-in AI algorithms [20]. Furthermore, the cost-effective nature of m-health services makes them attractive for developing countries [61], [62], where health-care facilities are sometimes inaccessible. Hence, the m-health approach can be used to implement patient-led data acquisition for AF detection to prevent stroke [63]–[65]. The m-health concept outlines a general framework of components. that can be used for patient-led data acquisition. That concept is flexible enough to accommodate the rise of big data which established itself as an independent goal for healthcare applications.

To highlight that substantial change in healthcare technology, the Intelligent Internet of Medical Things (IIoMT) improves m-health to capture the data centric nature of patient-led data acquisition systems which transmit signals from a patient to a central cloud server

for processing and storage [66]. This approach comes from the fact that wireless technology enables us to exchange healthcare data in real time [67]. The real time, or more specifically the statistical real time, aspect is crucial, because there is only finite buffering in the system, and data loss will occur if one component in the processing chain cannot maintain the required speed. To address that problem, IIoMT systems use packet-based transmission protocols which make the data exchange very robust [68], [69]. Linking the advancements of AI technology, telemedicine, and wireless body sensor networks, facilitates IIoMT systems to automate clinical decision-making. Human decision empowered AI can provide better support in terms of learning and discovering new knowledge from big healthcare data [70], [71]. This can be achieved through human validation of AI decisions. For example, in cardiology departments, human experts can work cooperatively with deep learning algorithms to establish a medical diagnosis [69], [72]. Patient-led data acquisition for AF detection to prevent stroke can be established with mobile health systems that measure and process physiological signals in real time. During the systems design, individual requirements that arise from stakeholder needs must be refined into a specification which governs the implementation [73]. As such, this improvement process has many degrees of freedom, and resolving design choices is based on available technologies, organization strategies, and human factors.

2.4.2 The internet of things technology for healthcare settings

IoT is an advanced method used to reshape the traditional healthcare domain, as such promising technological, economic, and social aspects [74]. IoT based healthcare provides various solutions which can facilitate the clinical missions in terms of monitoring, data acquisition, processing, and data analysis for a subject. The medical applications significantly increase their performance by tapping into the potential of IoT systems. These apps facilitate remote health monitoring of chronic diseases, sports programs, and elderly care. In addition, taking the medication and treatment on the specified time at home is another possible way IoT systems can be used in healthcare applications. The idea is that such systems give an alert to the patient when it is needed. Therefore, IoT services are projected to diminish the healthcare costs, enhance the subject's experience and increasing the quality of life. Furthermore, IoT based healthcare is an essential trend which can be updated through wireless technologies that predicted to

support medical emergencies, real-time monitoring, and early diagnosis. At this stage, cloud platform plays a main role of delivering and storing patient records on demand health services. The databases must be established and transmitted from medical server, gateways, and other health database resources. Drew [75] demonstrated that IoT techniques can be deployed for ECG monitoring in clinical settings. Their system facilitates HR measurement and the determination of heart rhythms as well as the diagnosis of multi-faced arrhythmias, myocardial ischemia, and prolonged QT intervals. The application IoT based ECG monitoring has the capability to provide the maximum information that could be extended to the longest period needed [76].

2.5 Automated diagnosis support architectures

An applicable IoT infrastructure can facilitate the cognitive prospective in terms of forming automated AF detection as a service [22]. These cognitive prospective are associated with improving the information that streams from the patient to the cardiologist. The diagram, shown in

Figure 2: Streaming the information from patient to cardiologist [22].

shows that the patient represents information acquisition source and cardiologist as the information sink. There are two separate directions for the information flow. The first direction refers to information extraction through deep learning. The second direction indicates to information extraction through using the digital biomarker algorithm. The obtained information can support as a measurable disease indicator for the physician during the examination process. Regardless the direct use, the obtained information can also apply as input for the classical Machine Learning (ML). The ML algorithms enhance that input into a single score. In the next section, we provide a review on the digital biomarker, deep learning concept, and classical ML algorithms used to reach the information refinement.

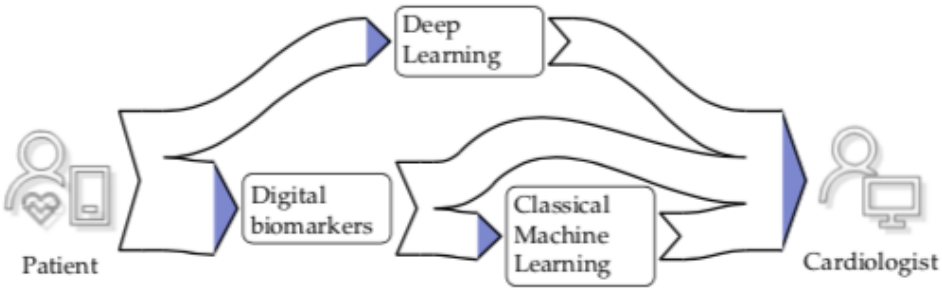


Figure 2: Streaming the information from patient to cardiologist [22].

2.5.1Digital biomarkers

The conception of digital biomarker is a significant finding factor in the review. Digital biomarkers refer to identify the information of interest that was picked up from the original measurement [77]. To do this, algorithms can be used to extract these biomarkers [77]. In the past, digital biomarker was understood as complementary to the traditional methods, such as cardiologist interpretation of ECG [78]. Since there is a possibility of establishing excellent digital biomarkers quality by using statical methods, and machine classification, they have begun to acquire acceptance [79].

Digital biomarkers for AF detection are often clustered into frequency, time, nonlinear and time frequency.

1. Figure 3 shows a taxonomy of individual groups and the following points provide more details on the techniques applied to obtain the digital biomarkers: Time-domain can generate biomarkers from the algorithms that analysis signals over time. From that prospective, they can be readily utilised to model the dynamic performance of the human heart. These are the most understandable digital biomarkers because they can measure some of the annotations that could be obtained through visual examination of the signal. However, these biomarkers are not successful to describe the non-linear features of heart's fluctuation process. Therefore, they do not provide all the existing information. To give an examples on the digital biomarkers algorithms, such as Principle Component Analysis

(PCA), differential equations [80], Linear Discriminant Analysis (LDA) [81], the Hadamard transform [82], and Independent Component Analysis [ICA].

2. Heart rhythm is related to the frequency component of the ECG signal. From the frequency domain, we can extract the digital biomarkers measures that refer to the features of the rhythm variations. These biomarkers are sensitive to rhythm variations initiated by AF. To be specific, frequency analysis techniques have linear characteristic. Therefore, the digital biomarker obtained from the frequency-domain fail to reflect non-linear features that is probable having a related source of information. The theory of Fourier Transform (FT) and Power Spectral Density (PSD) [83] are used as algorithms to extract the digital biomarker.
3. Hybrid-domain digital biomarkers are based on the concept of mixing the analytic power of multiple domains [84]. The practical understandings are relied on the fact that the spectrum changes are leading from the time progresses [85]. It is applicable to track spectral variations over time [86]. Based on the evidence that a rhythm irregularity will appear in the frequency domain, time resolution allows us to predict the time occurrence of AF event [87]. These digital biomarkers are an excellent tool for describing the arrhythmia positions. However, the analysis does not go beyond linear associations. There is list of examples that refer to hybrid domain transformation algorithm are: Short-Time Fourier Transform (STFT) [88], Empirical Mode Decomposition (EMD) [89], Wavelet Packet Decomposition (WPD) [90], Stationary Wavelet Transform (SWT), Discrete Cosine Transform (DCT) [91], Continuous Wavelet Transform (CWT) [92], Wavelet Transform (DWT) [84], and Time-Varying Coherence Function (TVCF) [93].
4. Nonlinear digital biomarkers target to consider the nonlinearity nature of the human heart [94]. Data tests replacement [95] show that the nonlinearity cannot be avoided in the physiological signals such as ECG and RR intervals. To test the performance of nonlinear digital biomarkers, with statistical and classification approaches, indicating that they are an independent source of the information. Herein some examples of algorithms that supply parameters for such biomarkers Higher Order Spectra (HOS) [96], Fractal Dimension (FD) [97], entropy [98], energy, Recurrence Quantification Analysis (RQA) [99] and Largest Lyapunov Exponent (LLE) [100]. RQA describes a set of digital biomarkers that generate

from examining the recurrence plots. Measuring the duration and number of occurrences in the physiological signal can assist us to comprehend the phase space route. The presence of AF might be related to changes in the phase space route [99]. The HOS refers to a third order of the spectral representation, and higher instants and cumulants. It harvests nonlinear digital biomarker that compute the nonlinear correlation between numerous frequency contents[101]. These measures are important to rhythm variations, which make them valuable for detecting arrhythmias. FD and LLE are disorder measures in terms of the signal complexity [102]. Entropy scales the information content of a signal [103]. In general, signals with fewer structure have more information content such as RR intervals signals, compared to signals have more structure, i.e., ECG. Energy biomarkers measure the rhythm's regularity of physiological signals.

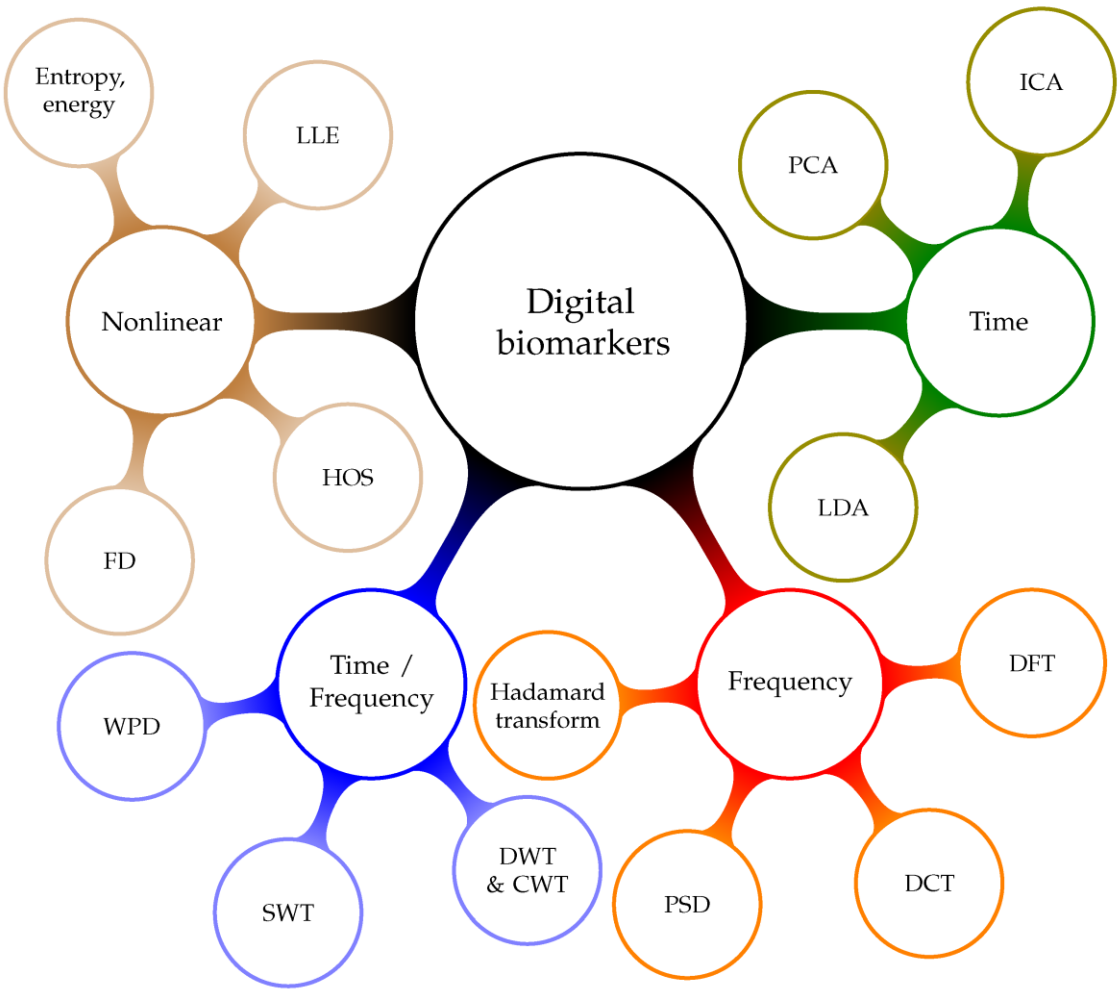


Figure 3: Digital Biomarkers taxonomy [22].

2.5.2 Artificial Intelligence

Artificial Intelligence (AI) incorporates from sets of computational algorithms that help in creating the input signal class. To be specific, classification algorithms are categorised into three main types: (a) supervised learning (b) unsupervised learning, and (c) reinforcement learning [104]. These algorithms are required a specific design step as well as tailored algorithms. Figure 4 illustrates a taxonomy graph of AI system management. Unsupervised learning structures the data as sets of clustering. To establish the cluster, a new data vector is used to group the objects which it most probable belongs. K-means clustering [105] and self-organising map [106] are examples of classification algorithms that based on unsupervised learning. While in the reinforcement learning approach, the algorithm requires continuous training by taking the feedback into the considerations for improving the quality of the decision. To give examples of the reinforced learning, hidden Markov models [and generative adversarial networks are used as algorithms for decision making.

A student-tutor relationship reflects the idea of supervised learning. Annotated data can be applied to train and test the supervised algorithms. The typical technique to verify the decision reliability, is based on using 10-fold cross validation [107]. Holdout dataset is indicating into blindfold cross validation which establishes the classifier performance through practical conditions [108]. There are some examples of the classical ML algorithms that relies on supervised learning as shown in Table 1.

Table 1: Shows lists of examples of classical ML algorithms.

Classical ML algorithms examples	References
1. Naive Bayes (NB)	[109]
2. Probabilistic Neural Network (PNN)	[110]
3. Support Vector Machine (SVM)	[111]
4. Random Forest (RF)	[112], [113]
5. Levenberg–Marquardt Neural Network (LMNN)	[114]
6. K-Nearest Neighbor (K-NN)	[115]

7. Decision Tree (DT)	[156]
8. rule-based	[116]

The core idea of Deep Learning relies on multi-hidden layers that shapes a complex structure and deals with complex scenario. A Deep Neural Network (DNN) imitates the human intelligence that is performed by the brain in terms of the complexity and functionality. In the past decades, the Artificial Neural Network (ANN) algorithms was designed with maximum of two hidden layers, which has a small number of interconnected neurons [104]. These intelligent systems have the capability of learning through updating weight parameters within the individual neurons. However, ANN has less complexity; therefore, these algorithms cannot easily process high-dimensional data.

Whereas deep learning algorithm deals with complexity to extract the knowledge from data structured with high-dimensional [23]. For instance, it is applicable to feed RR intervals data directly into a deep learning. system. In contrast, ANN algorithms require information reduction through its dimensionality when they extract the digital biomarkers and before processing the data. Deep learning approach was effectively implemented to a wide range of applications, such as image analysis [117], sound classification [118], and signal analysis. DNNs are mainly comprised of three types of layers which are convolutional, pooling and fully connected. The first layer (convolutional) is an adaptive filer, which can update its weights during the training phase. The pooling layer achieves a result directed that involves the dimension reduction of the values that spreads via the network. A common approach performs dimension reduction which is known as Max pooling [23]. The fully connected layers come at the end of DNN processing series. The data is combined from the previous layers, which could be more focused at the end of third layer, such that individual labels appear [119]. However, there are numerous types of variant DNN algorithms, i.e., Recurrent Neural Network (RNN) [120], Convolutional Neural Network (CNN) [121], Deep Belief Network (DBN) [122], and Long Short-Term Memory (LSTM) [123]. Figure 4 shows the taxonomy of AI. Faust et al. [22] referred to ANN and other algorithms as classical ML. These algorithms, plus the DNN algorithms belong to set of supervised learning algorithms. The main difference between deep learning and ML algorithms are that the previous set of algorithms require extracting the

digital biomarker, while the latter set can handle the physiological signal directly without feature extraction and information reduction.

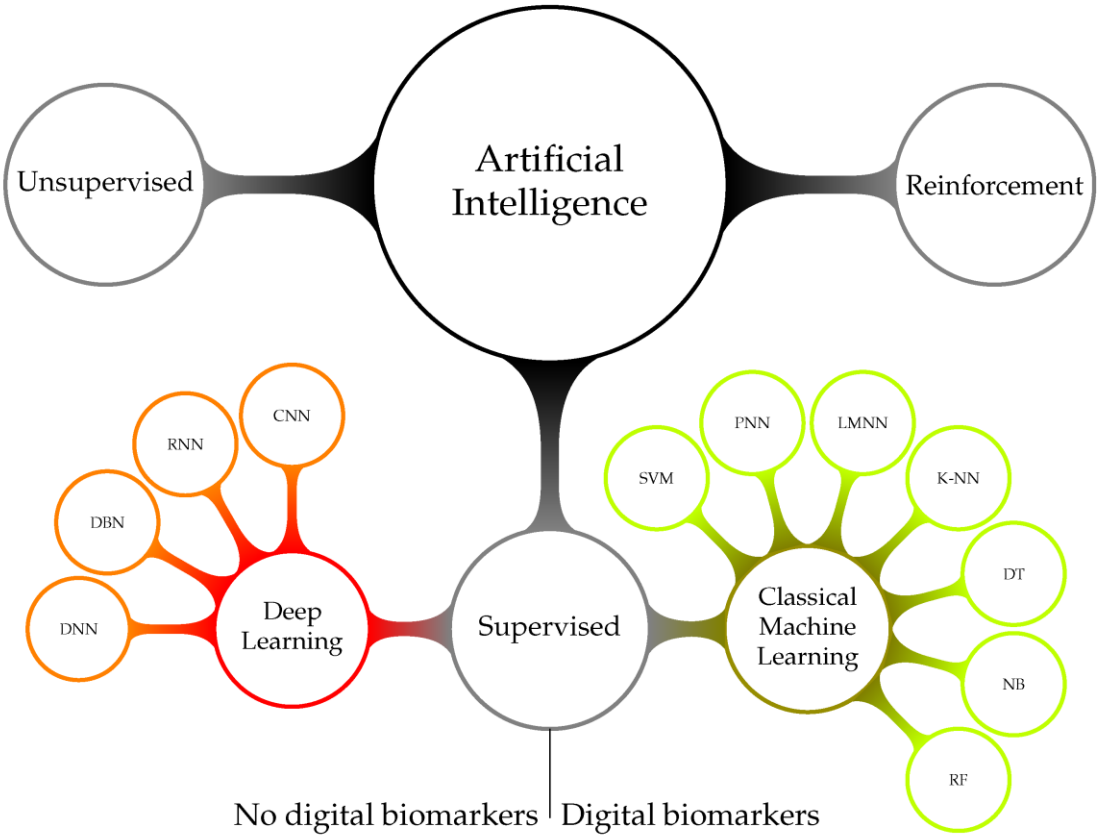


Figure 4: Categorisation of Artificial Intelligence [22].

2.6 Automated Arrhythmia detection based on ECG and RR intervals

Enormous number of studies that have been conducted by researchers to differentiate AF, Atrial Flutter (AFL) and NSR [16], [124], [125]. This differentiation based on automated arrhythmia intelligent detection algorithms through extracted knowledge from ECG and RR intervals. AF and AFL are types of arrhythmias that impact an increasing number of patients. AF is more common than AFL due to AF is a leading cause for stroke that increase its risk five-folds. In addition, AF increases the risk of death with double when compared to healthy people from the same age. In contrast, AFL is less common than AF with a prevalence of 0.09% of population, but whose anticipated disease burden is also subjected to increase by at least 2050 [17]. Approximately, over half of AFL subjects

have AF; however, AFL also indicate to independently increased risk factor of stroke and heart failure compared to those with a NSR [126].

In the clinical practice, the treatment management of symptoms from these both arrythmias does differ, but they have the same symptoms which might lead to misdiagnosis through human expert. This means that the discrimination between them is a significant clinical step based on computer aided diagnosis. For example, while the stroke prophylaxis decisions in the form of anticoalition does not differ between AF and AFL. Treatment decisions for managing the symptoms such as a shortage of breath and palpitations differ greatly. In respect to the risk of effective treatment is based on using the medications (drugs) or cardioversion treatment (to restore the heart rhythm to normal), which has been significantly reported as higher in patients with AF and AFL [127].

Looking beyond the literature, a rich context is available about automated arrhythmias detection based on wide ranges of machine learning algorithms. Faust et al. [17], introduced a distinct deep learning algorithm for detecting arrhythmias by using RR intervals. This algorithm is based on ResNet/detrending 10-folds cross validation approaches that achieved promising results, indicating to 99.98% of accuracy, 99.94% of sensitivity and 100% of specificity. Fujita et al. [128], used CNN with normalisation LSTM to detect arrhythmias from ECG signals that they achieved the following results, 98.45% of accuracy, 99.87% of specificity and 99.27% of sensitivity. These results represent the top best findings available for arrhythmias discriminations.

2.6.1 A symbolic dynamic transformation and MSSE algorithm used for detecting AF based on RR intervals

Zhang [129] have done significant research on specific algorithms for detecting AF based on RR intervals. Their method extracted the features from RR intervals in two stages. The initial stage began with pre-processing the raw RR interval by deploying symbolic dynamic transformation and multiscale Shannon entropy. Therefore, that step measured the variability of the beat-to-beat interval and hence it obtained a preliminary detection result. The second stage established the difference in the distribution curves of Delta RR intervals. These results were used to adjust the boundary between AF and normal beats, so that acquiring more stable and accurate results. The experimental work achieved

similar results when compared with the accuracy of other algorithms. The proposed method achieved the best performance of sensitivity and specificity with approximate results of 98.81% and 96.53% respectively. Consequently, this achievement is appropriate for ECG comprehensive analysis in clinical usage which could detect AF after gathering long-term ECG recordings. The detection algorithm can assist the doctors to enhance their work efficiency. In addition to that, the MIT/BIH- AF database has been collected from PhysioNet. The database includes 25 long- term ECG recordings of human subject with AF incidents (mostly paroxysmal), two leads signals sampled in 250 Hz has been used for each recording with duration of 10 hours. Using 23 ECG recording of the raw data due to the first two recordings are not available now. Signals are labelled with annotations which could identify the beats as normal (annotated "N") and AF (annotated "AFIB") as test data.

2.6.2 A deep learning method based on atrial fibrillation detection by using RR intervals

Faust et al. [14], introduced a cutting-edge method for detecting AF beats based on HR signals with a deep leaning system. The proposed system incorporates a combination of two deep learning algorithms: LSTM and deep RNN with 6 layers. The data was pre-processed before feeding it to the system. The data was partitioned with a sliding window of 100 beats having a step length of 1. The bidirectional LSTM (1-3) layers represent the learning and extracting of features from the input data sequence of HR. In the second step these features are sent to the fully connected layers these layers classify whether AF is present or not. The proposed system used Computer Aided Diagnosis (CAD) for long term monitoring of the subject's heart. As a result, this system is the first to comprise deep learning method for AF detection. The data were collected from PhysioNet, which publishes the MIT-BIH Atrial Fibrillation Database [46] . This database consists of 23 long-term ECG Holter recordings from different subjects. The data is recorded for a duration of 10 hours with two leads. The signals are sampled at 250 Hz, and they come with AF annotations. The data from 20 subjects was used for 10-fold cross- validating the model. The 3 remaining subject's data were employed for blind-fold validation. With 10-fold cross validation to system achieved a sensitivity of 98.51% and 99.77% specificity.

2.7 Improving the safety of atrial fibrillation monitoring systems through human verification

Faust et al. [24] and Kareem et al. [19], introduced the concept of improving the safety of machine decision based on human verification in the clinical settings. From the assumption tested is shown that a deep learning system have the potential to detect the AF episodes in the real-time, also monitoring a prevalent of other types of heart arrhythmias, and an experienced cardiologist will confirm the outcomes to reach a diagnosis. This verification stage can add the essential checks and balance to improve the safety of computer-aided-diagnosis.

The hybrid decision support approach has been tested through creating a prototype of AF monitoring service [21]. This treatment monitoring service incorporates using HR sensors for data acquisition as well as IoT technology for data transmission and storage. These high-tech facilitate transfer the HR data from patient worn sensor to universal cloud server. An AI algorithm task is to analysis the data in real-time, by them it would be reviewed by a human cardiologist once abnormal signal is detected. This human specialist contributes then for verifying the deep learning results by relying on HR data and additional knowledge acquired through patient records such as the history of clinical diagnosis for a patient. To establish a prerequisite for safety in any computer expert system, the purpose of the decision-making process should be clear and identified. Healthcare providers can register patients into the system of AF monitoring service [19]. The service provides real-time detection support by delivering a timely warning messages and HR analysis. Therefore, the safety prospective is a critical decision which needs the intervention of human practitioner. Chip integrated into wristwatch with little extra cost, might result in high volume of data. Getting the correct detection algorithms, that data can assist to determine AF periods.

2.8 E-health

The World Health Organisation (WHO) defines e-health as a wide-ranging set of activities that involves using electronic means to provide information and services related to health. This is mainly manifest on Information and Communication Technology (ICT) deployment for health. E-health incorporates a variety of standards, tools and events that use the electronic devices to deliver virtual instead of face-to-face contact medical service [130]. The core idea of implementing and embedding e-health is to have a vision of enhancing the quality of health-related-information, underpinning national health systems by ensuring that it is accessible and having good quality of healthcare for all. The usage of e-health became more common over the last decades due to benefiting from several factors, such as, it is straightforward to educate non-lay person. In addition, a vast majority of elderly people became familiar with internet technology. However, the privacy problems were not indicating to any concern for the senior adults [131]. They were concerned more about how to achieve good health value. Moreover, cost effectiveness was also a top priority of elderly people for selecting the e-health application. Many patients had related skills to ICT through using apps, visiting websites regularly and email correspondence. A research study showed that a number of senior adults that were involved in the use of internet technology which marked with high percentages, especially in the developed countries, and the education played a vital role as an effective factor in e-health deployment [132]. One of the most important point was noticed from weaknesses, is the lack of evidence in terms of supporting its relevance when the participants begun involving in the trial which did not show the actual condition of elderly people. Another study emphasised that the physical barriers referred to a huge weakness factor. In addition, the lack of e-health education in some countries is a major cause for those limitations.

Furthermore, Thingspeak from MathWorks [74] is a cloud storage solution that can be used for e-health applications. It is an open-source application associated with IoT technology and API. Storing and retrieving information from things such as smartphones, tablets, sensors, through HTTP protocol over internet network or via Local Area Network (LAN). Herein, enhancing IoT technology by expanding the power of the site to communicate to a social network of things with status updates in high level of remote monitoring, collecting, handling, and analysing the data. This technique will provide great

accuracy during visualizing the information in real-time. Meanwhile, IoT will play the significant role in the domain of healthcare. IoT systems have several advantages for health care projects [66], which can be used by health care projects, these positive points are cost reduction when Caregivers utilise healthcare connectivity solution, subject monitoring can be achieved in the real-time, which significantly decreases unnecessary visits to clinicians. Data storage functionality is to transmit information to the could platform where it is stored in either a private or public channel, and HR sensors can be enabled through Thingspeak application. The private channel is used to store Thingspeak data by default; however, public channels can be employed to share data with others. Once data is in a Thingspeak channel, we can analyse and visualize it, calculate new data, or interact with social media, web services, and other devices. Using Thingtweet application to link a Twitter account to a Thingspeak account. As such, Things, sensors, devices, and channels can be updated though Twitter by using Tweeter-Control API. For instance, we can make a device tweet us by giving alerts notification when the AF disease detected.

2.8.1 Intelligent internet of medical things based on M-health

Patient-led data acquisition systems must identify patient demands and data specific requirements. Physiological signal recording is a main factor of such systems. Once the signal is recorded and digitized, the samples become data with specific requirements for transmission, storage, and analysis. Patients require improvement of their outcomes and better management of their stroke risk. At the same time, a system needs to be convenient enough to ensure patient compliance [9]. Mobile health, also known as m-health, aims to provide the technical solutions which addresses both data requirements and patient needs [57]. Thus, m-health was defined by the Global Observatory for eHealth (GOE) as medical and public health practice that operates through portable devices, such as mobile phones, patient monitoring tablets, personal digital assistants, and other wireless devices [58]. The core idea of m-health systems is that the data rather than the patient travels. More specifically, close to the patient the data is interconnected through wireless channels to increase convenience for long-term recordings. M-health delivers hospital-in-house services that overcome geographical as well as organizational impediments [59], [60]. More advanced m-health services provide automated diagnosis support through built-in

AI algorithms [20]. Furthermore, the cost-effective nature of m-health services makes them attractive for developing countries [61], [62], where healthcare facilities are sometimes inaccessible. Hence, the m-health approach can be used to implement patient-led data acquisition for AF detection to prevent stroke [63]–[65]. The m-health concept outlines a general framework of components. that can be used for patient-led data acquisition. That concept is flexible enough to accommodate the rise of big data which established itself as an independent goal for healthcare applications.

IoT is an advanced method used to reshape the traditional healthcare domain, as such promising technological, economic, and social aspects [74]. IoT based healthcare provides various solutions which can facilitate the clinical missions in terms of monitoring, data acquisition, processing, and data analysis for a subject. The medical applications significantly increase their performance by tapping into the potential of IoT systems. These apps facilitate remote health monitoring of chronic diseases, sports programs, and elderly care. In addition, taking the medication and treatment on the specified time at home is another possible way IoT systems can be used in healthcare applications. The idea is that such systems give an alert to the patient when it is needed. Therefore, IoT services are projected to diminish the healthcare costs, enhance the subject's experience and increasing the quality of life. Furthermore, IoT based healthcare is an essential trend which can be updated through wireless technologies that predicted to support medical emergencies, real-time monitoring, and early diagnosis. At this stage, cloud platform plays a main role of delivering and storing patient records on demand health services. The databases must be established and transmitted from medical server, gateways, and other health database resources. Drew [75] demonstrated that IoT techniques can be deployed for ECG monitoring in clinical settings. Their system facilitates HR measurement and the determination of heart rhythms as well as the diagnosis of multi-faced arrhythmias, myocardial ischemia, and prolonged QT intervals. The application IoT based ECG monitoring has the capability to provide the maximum information that could be extended to the longest period needed [76].

To highlight that substantial change in healthcare technology, the IIoMT improves m-health to capture the data centric nature of patient-led data acquisition systems which transmit signals from a patient to a central cloud server for processing and storage [66]. This approach comes from the fact that wireless technology enables us to exchange healthcare data in real time [67]. The real time, or more specifically the statistical real

time, aspect is crucial, because there is only finite buffering in the system, and data loss will occur if one component in the processing chain cannot maintain the required speed. To address that problem, IIoMT systems use packet-based transmission protocols which make the data exchange very robust [68], [69]. Linking the advancements of AI technology, telemedicine, and wireless body sensor networks, facilitates IIoMT systems to automate clinical decision-making. Human decision empowered AI can provide better support in terms of learning and discovering new knowledge from big healthcare data [70], [71]. This can be achieved through human validation of AI decisions. For example, in cardiology departments, human experts can work cooperatively with deep learning algorithms to establish a medical diagnosis [69], [72]. Patient-led data acquisition for AF detection to prevent stroke can be established with mobile health systems that measure and process physiological signals in real time. During the systems design, individual requirements that arise from stakeholder needs must be refined into a specification which governs the implementation [73]. As such, this improvement process has many degrees of freedom, and resolving design choices is based on available technologies, organization strategies, and human factors.

2.9 State of the art atrial fibrillation detection systems

Patient led data acquisition is based on data travelling from point of measurement (patient) to the universal central processor (Cloud Computing). In this review, Kareem et al. [9] introduce the state-of-the art of patient led data acquisition for automated AF detection to prevent stroke. According to Framingham heart study showed that there is a link between AF and stroke, which increase the risk prevalence probability by five-folds. Hence, there is a significant need to monitor AF for long-term with reliable detection methods that requires potentially unrestricted measurements. Enormous studies showed that patients with paroxysmal AF might have hidden symptoms during the measurement [9], [17], [133], [134]. That implies with undiagnosed patients which could lead for further cardiac complication.

The latest measurement approaches will leave a significant gap when it comes to AF detection. This concern can be reduced as well as the number of first detected AF cases increase by extending the data acquisition duration. Furthermore, detecting more AF episodes could lead to manage the treatment earlier, which can decrease the stroke severity, and in some cases, provide stroke prevention altogether.

Practical restrictions mandate that these measurements are achieved in the patient environment. The measurement technology must be organised in line of day-to-day activities of the patient. Health economics require that the required resources for prolong measurements and monitoring should be minimal. To replicate these requirements, we have examined the current technology as providing real solutions in terms of benefits for patients, a detection quality, long-term observation duration and cost. The optimum patient led data acquisition system does not require face-to-face interaction during normal monitoring procedures. The sensor can be delivered to patient by post and the data connection setup can be achieved online.

In this review, we found two significant measures for detecting AF, which are ECG and RR-interval signals. Decision support algorithms can extract relevant knowledge for medical diagnosis from both measures. However, RR-interval signals are straight forward to measure and inexpensive to communicate, store as well as to analyse. Yet, the current clinical settings and indeed cardiologist are well trained towards AF detection based on ECG signals. Looking beyond large body of literature, we anticipate that there will be a change from the traditional measurement to RR-interval based systems where will become widely applicable for AF detection. Initially, smart devices can be applied to create a suspicion that AF is detected, and the diagnosis is achieved through conventional with ECG gold standards measurement. Once there is enough trust in RR-interval based system, they might be used to even support the AF diagnosis.

In the meantime, healthcare givers depend on costly ECG based systems for AF detection that opens an avenue for commercial solutions which might be not projected to medical device regulation. In other words, having altered measurement method-based detection might be driven by commercial companies which collect and process physiological signals to provide an automated AF diagnosis service. This service concept will be directed by patient led data acquisition which enables big data business strategies to resolve or at least to highlights healthcare demands from patients.

In this study, we review patient led data acquisition as smart system for indefinite observation of AF events. Throughout the investigation, we found that all there viewed

systems can capture AF events based on physiological measurements [9]. However, these systems designed with excellent aspects in terms of technology applied, human factor design and cost. Our focus is related to the technical concepts which support the patient led data acquisition systems. From that review, we obtained two fundamentally different approached those results in designing very distinct patient led data acquisition systems. The first approach refers to signal recordings, which is responsible on data handling. The second approach is related to event triggering, which leads to information handling. In general, the obtained information about when a specific AF event occurred has a significantly lower data rate as same as the physiological signal data, specifically the HR signal. This is such a great benefit when it comes to sensor battery lifespan as well as data storage or communication channel requirements. Whereas the event trigger systems cannot maintain the evidence which triggered at that event. As such, it is impossible to verify independently the outcomes of the event trigger algorithm. Moreover, the event trigger systems do not produce signal data that could be utilised to enhance patient data acquisition systems. In contrast, signal recording systems generate raw data that requires processing. That processing is normally achieved in the most appropriate place, i.e., compute clusters linked to cloud server. Such a setup can be continuously enhanced by refinement on the detection algorithms. This enhancement can come from the validated signal partitions. To be specific, according to the available data, medical professionals can independently validate the decision made by a machine algorithm [18], [19], [21]. Regardless the verified results being more reliable, such results can also be applied as labelled data to re-train the detection algorithms. Due to the fundamental differences between signal recording and event triggered systems, we have used these approaches as level 1 differentiation in our taxonomy for patient led data acquisition. Figure 5 shows a taxonomy of patient-led data acquisition devices for AF detection. In the remainder of this section, we introduce the individual concepts in detail. This provides the framework for the device reviews. Each reviewed device is represented as a leave node in the taxonomy.

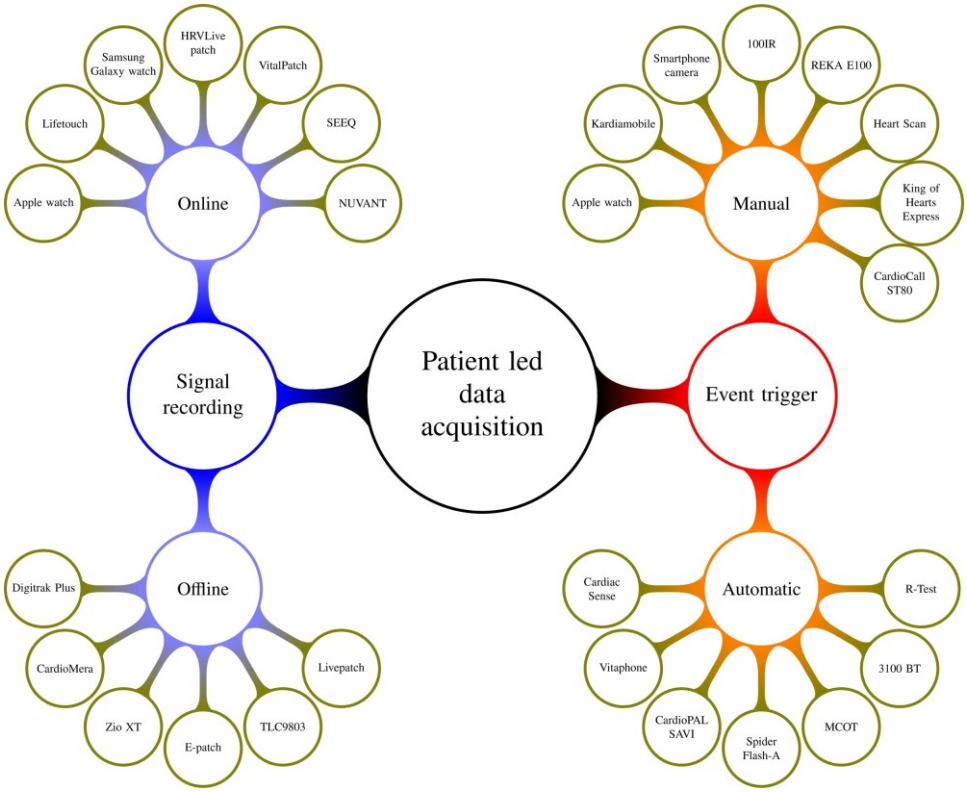


Figure 5: Taxonomy of patient led data acquisition concept [9].

2.9.1 Signal recordings

The main concept behind signal recordings systems is to maintain the original cardiac measurements in a digital format. Hence, the stored signals keep the track record of the evidence upon which a specific diagnosis was based [9]. This evidence can be used for identifying a root cause analysis where there is a suspicion of misdiagnosis. Moreover, continuous data recording implies to huge data dimensionality with reasonable trustworthy labels. By using real-time data acquisition offers the possibility of improving the decision support algorithms. These improvement leads to establish deeper understanding of the link between AF and stroke risk. To be specific, monitoring patient over long duration could help to recognise the signal pattern that are more related and sensitive to predict stroke risk than AF rhythms itself. Another technical benefit comes from the fact that data volume is foreseeable, for ECG recordings is completely true and for RR interval is also true from a statistical perspective. Online and offline capabilities differentiate signal recording systems for patient led data acquisition.

2.9.1.1 Offline signal recording

Offline recording is achieved by storing the recorded cardiac data in the sensor itself. That sensor can only maintain a restricted amount of digital storage- normally in form of removeable memory stick once the patch is returned for analysis [135], as such, the recording time is limited. That result in either one-off measurement or interval. Traditionally, there were many systems classed as the first ones that allowed physiological data acquisition in the patient environment [136]. Meanwhile, having a built-in storage capability was an important technological step that enabled data acquisition anywhere and anytime.

Offline signal recording has the following advantages:

1. It has lower energy requirements comparing to online-signal recording. That implies to mini sensors and longer duration of the cardiac measurement with one battery
2. Independent of network coverage.
3. Most trustworthy proven technology.
4. Data is recorded – all evidence is preserved

Drawbacks:

1. No patient compliance feedback.
2. The storage capacity limits the recording duration.
3. The storage medium must travel to and from the patient environment. In most cases, that means the patient must travel to initiate and conclude the measurement.
4. Long time (weeks) between a AF episode and the diagnosis.
5. Labor intensive data handling at the medical facility.

There are a wide range of offline sensors as presented in signal recording taxonomy in Figure 5, but we will mention a few of them as follows:

1. TLC9803 Dynamic ECG Monitor- This sensor has 5 electrodes, 3 channels for chest ECG measurements. Recording duration is up to 24 hours. Expert electrode placement required such as qualified nurse. Local data processing after the measurement. In clinical studies, the device was used to acquire standard ECG signals [137]

2. Zio XT patch cardiac arrhythmia monitor Single-lead display and two electrodes skin placed to record ECG for up to two weeks. The device is sent by post to patients, and once the measurement is completed, it is also returned by post. The data are analysed by iRhythm, the technology provider. The device was validated in a clinical pilot study [138]. The results show that the Zio XT Patch was accepted by patients, and long observation durations of two weeks were achieved. Subsequently, the diagnostic utility of this device was confirmed with three separate clinical studies [30]. In a clinical study with 2659 participants, the device was utilized to establish the efficacy of home-based wearable continuous ECG monitoring for AF detection [139], [140]. A literature review established that long-term monitoring with the Zio XT patch results in higher cardiac arrhythmia detection rates when compared with traditional Holter monitoring [33]. The device was also used to establish incidence and timing of potentially high-risk arrhythmias. The device was also used to establish incidence and timing of potentially high-risk arrhythmias. However, the disadvantage of this patch is based on offline recording aspect.

3. E-patch extended Holter monitor Up to three channels. Five days continuous chest ECG recording augmented with manually triggered events. The patient can activate and apply the patch. After the measurement, the sensor is mailed back to Biotelemetry, where the data is analysed.

4. Digitrak Plus That sensor consist of 5 electrodes, 3 channels for chest ECG measurements. Up to one week recording duration. Expert electrode placement required. A clinical study found that the device provided accurate assessment of atrial beats and rhythm diagnosis [141].

5. CardioMera ECG Holter Contains 7 electrodes, configurable for 1, 2, 3, or 5-channel chest ECG measurement. Measurement duration is up to 24 hours. Once the ECG is measured, the ECG data is transferred by either disconnecting a memory card or via an optical transmission link with a local workstation. The device comes with analysis software for arrhythmia detection. The device was used in a multi-sensor study to determine general health [142].

2.9.1.2 Online signal recording

Online signal recording can also be defined a medical telemetry that facilitates transmitting data from the point of recordings to a central monitoring unit, such as, cloud storage [143]. This approach is possible to achieve with wireless technology in the patient environment. Hence, online signal recordings operate in real-time, which means that within a specific timeframe all the recorded data must be transferred to the cloud location. Failure to meet the real-time requirements will lead to data loss. To reduce the real-time requirements, the sensors are designed with a built-in buffer which can be used as a spare storage in case the wireless link fails for connections. Another critical problem for measuring the online signal is energy costing. Energy is needed to establish the electromagnetic waveform that conveys the physiological data from the biosensor to a relay station. In general, that energy requirements rely upon the wireless channel condition. As such, this will make the real energy requirement difficult to predict. Despite that uncertainty, on average the energy requirement for an online system is higher when compared to offline systems.

Online signal recording offers the following benefits:

1. Data travels and patient can stay at home. For an ideal system arrangement that might mean the smart sensor plus the setup instructions are delivered to patient's home via mail carrier, and the patient's registration can be done via a web-portal on the internet.
2. Offering Real-time detection as well as decision support. That might lead to: a) Stroke physician or cardiologist gets alerted within minutes of detecting an AF episode. b) Timely diagnosis.
3. Feedback channel for patient compliance.

Drawbacks:

1. Online signal recordings require higher energy requirements when compared to offline systems. With current technology, the battery might need to be replaced as it runs out quickly or the device must be charged.
2. Wireless Network or data cellular coverage required.
3. Data plan. A network provider needs to be selected and there will be a continuous operation cost.

The next paragraphs present specific patient led data acquisition systems that are based on online signal communication technology.

Herein a list of an online devices that are commonly used for online detection:

1. Isansys Lifetouch Two electrodes, single lead display ECG and RR interval measurements. Wireless connection based on Bluetooth to a smart tablet which relays the data to a cloud server. The smart sensor was used in clinical studies on cirrhosis that showed a significant reduction in HRV for 90 days and predict the mortality [144] , vital sign monitoring after major abdominal surgery [145], and chronic obstructive pulmonary disease [146].

2. Apple watch Wrist PPG. The device was implemented for a proof of feasibility study in a clinical setting where patients can be observed and overseen by healthcare providers. A built-in diagnostic algorithm has been used to detect AF based on PPG signals [147]. The accuracy of the AF detection model is validated with two clinical studies [148], [149]. Apart from that, the device was also used for mental state detection [150].

3. Medtronic SEEQ – Two electrodes, one lead, up to 30 days ECG recording. Wireless telemetry based on an uplink station which transmits the data via satellite or terrestrial channels to the Medtronic network. Clinical studies were used to validate the device [151], [152].

4. Corventis NUVANT – Two electrodes, single lead which records ECG from a patient chest for up to 30 days chest ECG recording. Wireless telemetry based on an uplink station (zLink) which communicates the data via terrestrial channels to a Corventis monitoring centre. Disposable patch sensor. Clinical studies were used for validation [153], [154].

5. Vital-Patch-Vital-Connect RTM– includes 2 electrodes one channel chest for one week ECG recording. Data connection via Bluetooth to an Android mobile device with VistaTablet software. The mobile device transmits the data to a cloud server. Both device wearability and low impact on day-to-day patient activities was verified in a clinical study [155].

6. HRVLive Monitor –Two electrodes recording ECG from chest or finger PPG. Unrestricted recording duration. RR interval extraction and analysis software. Standalone offline PC-based solution.

7. Samsung Galaxy Watch3 – Wrist ECG with unrestricted observation duration. In May 2020 the device was cleared by the Korean Ministry of Food and Drug Safety for health monitoring.

2.9.2 Event trigger

Event trigger systems make the promises based on their ability to identify a specific segment of the physiological signal that shows AF symptoms. That diagnosis is reached in the real-time. The core concept behind having the signal analysis within the measurement device is the demand for data reduction. To be more specific, event monitoring systems store only signal segments of interests. on a fundamental level, each sensor type has its own specifications that involves the energy requirements for recordings, including data storage, data analysis and data communication. Event trigger systems prosper on the sense that the energy kept by communicating or storing the physiological data is larger than the amount of energy required to achieve the processing which extracts the events from the measurement data. The trigger mechanisms, i.e., manual, or automatic, can be used to distinguish event trigger systems.

2.9.2.1 Manual event trigger

This type of device accomplishes the ultimate energy saving through pushing the trigger decision button manually by patient. In terms of AF detection, this could be possible due to many patients feel the difference once their heart rhythm becomes irregular. However, such manual trigger devices cannot be utilised by patients with special requirements such as patient suffer from physical or mental ability to operate the trigger manually. In addition, service users or patients are not able to trigger the devices whilst they are asleep. There might be even psychological consequences growing from being responsible to trigger the device. From a positive perspective, patients might realise their active engagement in the recordings process as empowerment, for example, they are considered as a significant active part in the disease monitoring process. whereas, on the negative

side, patients' responsibility may generate the anxiety caused by the fear of missing AF events. To summarize, manual event trigger systems offer the following benefits

1. Minimum energy requirements are needed. That can lead to: a) mini sensors and b) greater measurement duration with one battery.
2. Lower data rate. That can lead to: a) Lower storage requirement and b) Lower requirements for the communication channel – better wireless connectivity.

Drawbacks:

1. Reliance on the patient to trigger the measurement
2. Measurement during sleep is not possible.
3. Evidence, in the form of signal data, is discarded.

Patient compliance can only be established on a local level. That means, compliance cannot be managed from a central location which might have more knowledge of the patient.

The examples of manual event trigger are listed in the following:

1. Tele-EKG-Card 100IR
2. REKA E100
3. Smartphone camera – Finger PPG
4. Omron W7720 HCG-801
5. HeartScan
6. King of Hearts Express
7. CardioCall ST80
8. Kardia mobile

2.9.2.2 Automatic trigger

Automatic event trigger devices demonstrate the capability to decide whether AF symptoms are present. The decision information is either saved in the device itself or communicated in a real-time channel to a central cloud. Information about the suspected AF episode might include the signal snippet which caused the AF detection algorithm to trigger. The trigger algorithm is executed close to the point of measurement - usually in the sensor itself. Hence, these algorithms must operate in the resource constraint environment of a deeply embedded system. This makes it much harder to update the algorithm once better methods are available. To summarize, automatic event trigger systems have the following advantages:

1. Low data rate, when compared to data recording systems.

Drawbacks:

1. Increased energy requirements for the analysis algorithm.
2. Black box decision-making. Humans are not involved in the decision-making process.
3. Difficult to update the decision-making algorithm.
4. Evidence, in the form of signal data is discarded.

The examples of Automatic event trigger are listed in the following:

1. R-Test Evolution, Novacor, France
2. Mobile cardiac outpatient telemetry (MCOT)
3. 1-channel Holter monitor 3100 BT Loop Recorder
4. Cardiac sense
5. Vitalphone
6. Spyder Flash A
7. CardioPal SAVI

2.10 Research Gaps

All the reviewed scientific studies stopped at a one-time evaluation of the learning model. In other words, the implied premise is that the developed model will work for unseen data. However, without independent support, there is doubt that this premise holds. The assumption might be wrong and hence the claim that the methods are relevant for practical setting is invalid. We identify that as a research gap because the lack of follow up verification results in a lack of trust in the computer aided diagnosis systems. Our plan is to establish that trust, at least to some extend, by validating the existing DL system with new data. Another gap identified that there was no scientific work on HR analysis with DL before 2017. Our paper documents the first scientific work on analysing and detecting AF from HR signals by using a DL algorithm [14]. We believe that such work is beneficial for scientific progress, because DL is a very promising analysis method and HR is a good health indicator. To fill that research gap, we propose to analyse HR signals from normal and AF subjects with a DL system.

In addition, state of the art Holter monitors operate at least with two-electrodes, which must be setup by a specialist i.e., nurses [156]. Another limitation of the standard care routine Holter monitor that it is an offline monitoring device, and the data can only be analysed once the monitoring session is completed which might take weeks to do so. From patients' point of view, they identified Holter monitors as being bulky, and uncomfortable when their feedback sought during a Public Patient Involvement, Engagement and Participation (PPIEP) event for stroke survivors. Moreover, the PPIEP community addressed their preference of wearing a re-attachable heart patch rather than wearing a strap with HR sensor. Therefore, we selected a certified device by the NHS trust. The proposed monitor to be used in the study is a lightweight real-time Lifetouch sensor that records both ECG and RR intervals where the RR intervals forms the HR. These signal recordings are measured by placing a sensor on the chest, and that can be done by a patient. The third gap, we found that there is insufficient monitoring duration of AF for both inpatient and outpatient. Hence, we propose a service platform and directions that can extend the observation duration for AF detection indefinitely. To achieve that extension, we use RR intervals as an ideal measurement to record, process and store in central cloud storage. As such, long-term observation requires having patient-led data acquisition to underpin regular service directions, i.e., the stroke risk monitoring

service can be directed to either healthcare providers or to patient environments, as proposed in our publication [9]. Figure 6 shows the graphical representation of the service directions which presents some of the advantages and drawbacks.

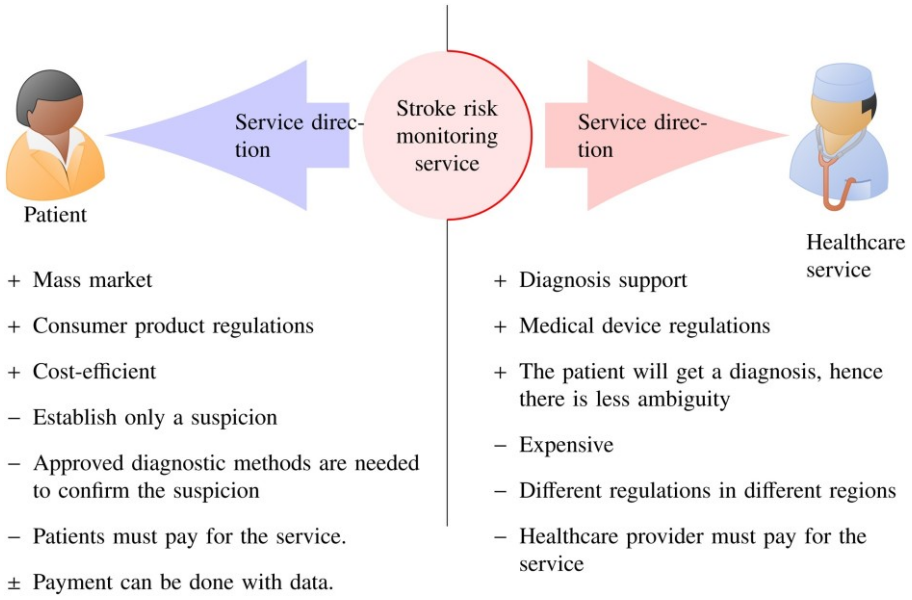


Figure 6: Service directions for long duration of AF detection based on patient.-led data acquisition [9].

2.11 Summary

This chapter involves a large body of literature that has been investigated various aspects such as signal recording methods of human's heart, varied medical devices and DL algorithms. In addition, providing the relevant approach of how to improve the safety of machine's decision through human verification. That approach helps to increase the reliability of using the DL algorithms in the clinical practice as main diagnostic tool. This step is significant which adds value for research and practical deployment. Another important point was addressed in that chapter is introducing the concept of patient-led data acquisition. The proposed concept tends to solve the problem of monitoring duration from short to long-term. To conclude, this chapter also attempt to cover the research questions that presented in chapter 1.

Chapter 3 Automated arrhythmia detection with deep learning based on RR intervals

3.1 Introduction

Heart rhythm irregularities, known as arrhythmias, are a leading cause of mortality and morbidity. AF and AFL are common types of arrhythmias that affect an increasing number of patients. These arrhythmias can be measured either by ECG or RR intervals. In this study, we use RR intervals as alternative measurement method to standard ECG recordings. RR intervals are an ideal approach to communicate, process and store patients' data in the cloud due to having low data rate and high information content. From that prospective, these conditions require long-term monitoring. RR intervals can facilitate extending the observation duration. Automated detection can be achieved through using a DL system. This chapter introduces two different algorithms that were used to detect common heart arrhythmias. The directional-LSTM was used to classify AF and normal beats while ResNet was applied to accomplish the discrimination of AF, AFL, and NSR.

In Chapter 2, we identified a research gab which highlights the problem of detecting AF episodes with computer-aided-diagnosis technologies that help with extending the observation duration, including medical devices and signal recordings. This chapter describes the design of an AF detection service as a proposed solution of detecting cardiac arrhythmias for those who based in hospital and their home environment. First, the chapter describes the project 's design methodology Section 3.2, and material and methods Section 3.3. A proposed solution is then discussed Section 3.4, followed by the proposed prototype.

3.2 Design a decision support system based on physiological signal

The main aim of designing the DL algorithm is to outperform the established methods. The design structure is based on the concept of having an offline and online system, as shown in Figure 7. The offline system is applied to shape the required algorithm arrangement based on labelled data. Therefore, applying the knowledge of the designed algorithm in the online system to process live measurement data. That system incorporates three sequential processing stages: 1) downloading and pre-processing the data from the source 2) creating DL model 3) validating the DL model [14]. The first two stages establish the analysis system which extracts the required information from the ECG signal. The third stage evaluates the model with unseen data [23].

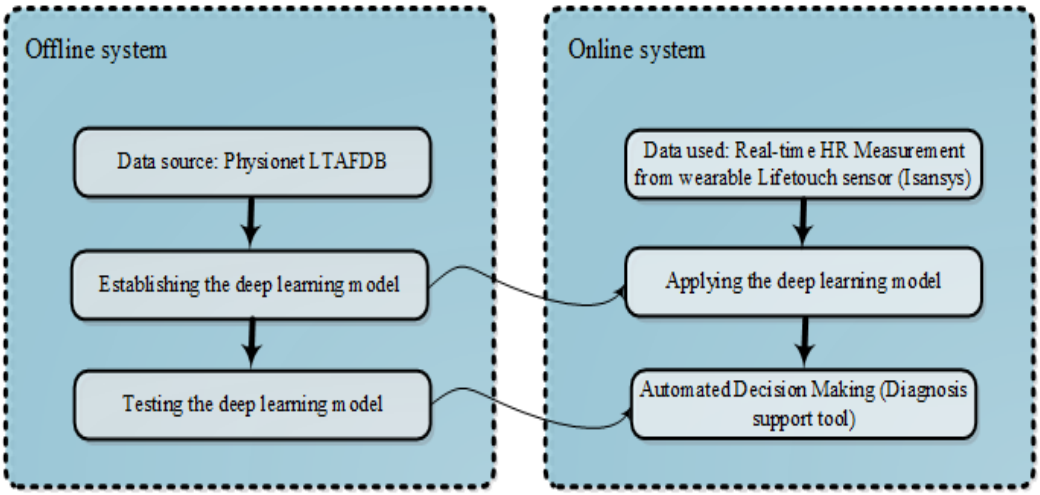


Figure 7: Block-diagram illustrating the design of decision support algorithm based on data source.

3.2.1 Offline system

The block diagram outlines the offline system which comprises from three sequential steps, as shown in Figure 7:

1. Data for AF disease detection was collected from PhysioNet databases (LTAFDB).
2. MATLAB software has been used to download ECG recordings as Excel file for each individual subject. Installing ECG-toolkit was required to facilitate the PhysioNet data visualisation. Signal processing involves the analysis, interpretation, and manipulation of the physiological signals. To be specific, building as such explicit algorithm is to extract the beat-to-beat interval from ECG signal, and this aspect based on Pan-Tompkins algorithm.
3. The final stage was associated with partitioning the data into sliding window of 100 beats for the whole time series. A deep learning system was applied to detect AF beats in HR signals, and the resulting signal blocks were directly fed into a RNN with LSTM.

3.3 Materials and Methods

This section presents the data used for training and testing the algorithm as well as the processing methods that were used to design the AF detection system. The discussion begins with describing the data and the pre-processing methods. The pre-processed data is directly fed into a DL system. As such, that system contains all the computational complexity. Hence, this section focuses on the DL algorithm and the design decisions which led to the proposed Computer-Aided Diagnosis (CAD) system.

3.3.1 Data used for training and testing the deep learning algorithm

The experimental work were conducted based on data collected from MIT-BIH Atrial Fibrillation Database (AFDB) which is available on PhysioNet [45], [46]. This database includes 23 long-term ECG Holter recordings of patients with paroxysmal or sustained AF. Each dataset contains two ECG signals that were recorded in parallel for 10 hours and sampled at 250 Hz with AF annotation. These recordings incorporate also beat

annotations and rhythm annotations accomplished manually by experienced cardiologists. Moreover, the R peaks are labelled, and the RR interval sequence was extracted based on these labels. The RR intervals sequences have been partitioned with overlapping windows into sequences of 100 beats for each HR trace. A sequence is labelled as AF if it contains one or more beats that were classified as showing signs of AF, all other sequences are labelled as normal or non-AF. Data from 20 subjects has been utilised for 10-fold cross validation of the model. This means the proposed methods can be used to generalise not only unknown data, but to unknown patients as well. The remaining data of 3 patients was held out for the usage in a blind-fold validation phase after achieving the training and validation stages of the model and adjusting the model with so-called hyper-tuning parameters. This concept ensures that the proposed approach is applicable not only to unseen data, but to unseen patients as well. The diagram in

Figure 8 provides an overview of the DL design concept.

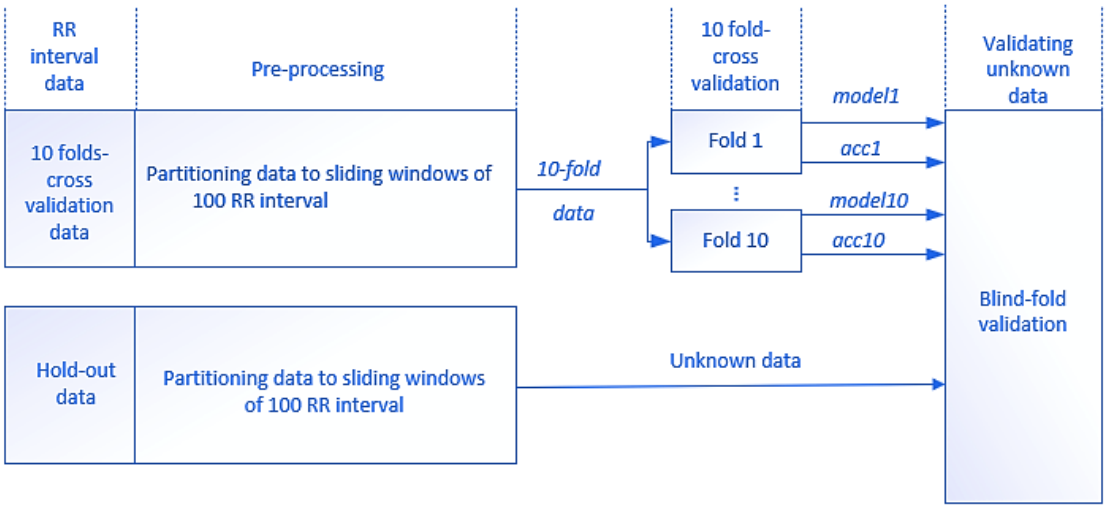


Figure 8: Shows a block diagram of training and validating the DL model.

3.3.2 10-folds cross validation

10-fold cross-validation aims to reduce the impact of selecting test samples from an available dataset. Kohavi et al. [157] advises to use this approach for model selection. As

such, that performance measure is related to compare the classification models; as shown in table. The fundamental concept is to partition the dataset into 10 segments. Each segment contained varied data from the cross-validation dataset. To be specific, tuning model method were commonly followed in the practice within bioinformatics and machine learning community [158]. Once the data partitioning is done, the segments are used to generate 10 folds to be fed into the deep learning model for training and testing. For instance, in fold 0, segment location 0 is applied to test and the remaining 9 segments are utilised to train the network. Similarly, the idea is replicated for the following folds through shifting the segments that corresponds to fold number by allocating one segment for testing and the remaining 9 segments for training. The model fitting process is set to 80 epochs. For each epoch, the bidirectional LSTM network is trained and tested. The training stage creates a model which is expressed as a set of weights. The LSTM network testing step can validate the prediction quality of the model. Prediction quality means selecting the best model's block that decides which is the best model for specific fold. The data for the next fold is loaded once all the epochs are processed. The algorithm yields once all the folds processing are completed and the K best models, together with their accuracy (acc) are established. The right segments in the flow chart depicts the epoch-based fold processing.

3.3.3 Bidirectional Long-Short Term Memory network

DL algorithms attempt to improve the model by using all the existing information from the input [23]. Getting this information creates the knowledge which supports the robust decision process. Therefore, the deep learning method is more feasible than conventional machine learning, such as SVM [159]. RNN models have acquired a growing popularity in recent years because they resolve some of the key limitations of the traditional machine learning algorithms. The hypothesis is that the inputs and outputs to a model are completely independent from each other [160]. This hypothesis is falsely claimed when compared to natural language processing as one of many problems. For instance, to classify an expression with a sentence, it is essential to add the individual words of the sentence into context. RNN models achieve this by enabling the network to maintain and

use state information, such as information about what has occurred in the prior time-steps/inputs. However, Bengio [161] presented that, whilst standard RNNs can theoretically process input dependences over long interval. Training such networks with gradient descent becomes more ineffective when the time span of the input sequence increases. Hence, it is difficult to trains RNNs successfully.

Bidirectional LSTM overcomes some of the limitations of standard RNN models by including a gates mechanism which enhances the processing of time step information from long interval input sequences [162]. That mechanism governs the amount of information from the prior time steps, that contributes to the current output The LSTM gating mechanism comprises of three gates: 1) input-gate, 2) forget- gate and 3) output-gate. The training algorithm determines which information is remembered and which information is forgotten.

Schuster and Paliwal [163] suggested to utilise the bidirectional RNN for problems where the complete input sequence is available. To be specific, the bidirectional RNN mechanism from an input sequence uses the past and future data to train both a forward state RNN (working in the positive time dimension) and a backward state RNN (working in the negative time dimension). This enables the network to make more precise predictions, because of the increased context provided. Recently, bidirectional LSTM models have demonstrated excellent results in fields such as speech recognition. Graves and Schmidhuber [164] showed that the bidirectional-LSTM networks can be substantially more effective than the unidirectional LSTM architecture.

The number of neurons can be determined by the input layer which is equal to the number of the features space (explanatory variables). However, the number of neurons in the output layer represents the output space. The main features of LSTM network are storing the information in so-called memory cells. Each memory cell has three gates to maintain and tuning its cell state S_t, which are forget gate F_t, an input gate i_t and an output gate O_t. The schematic diagram illustrated the structure of a memory cell in Figure 9.

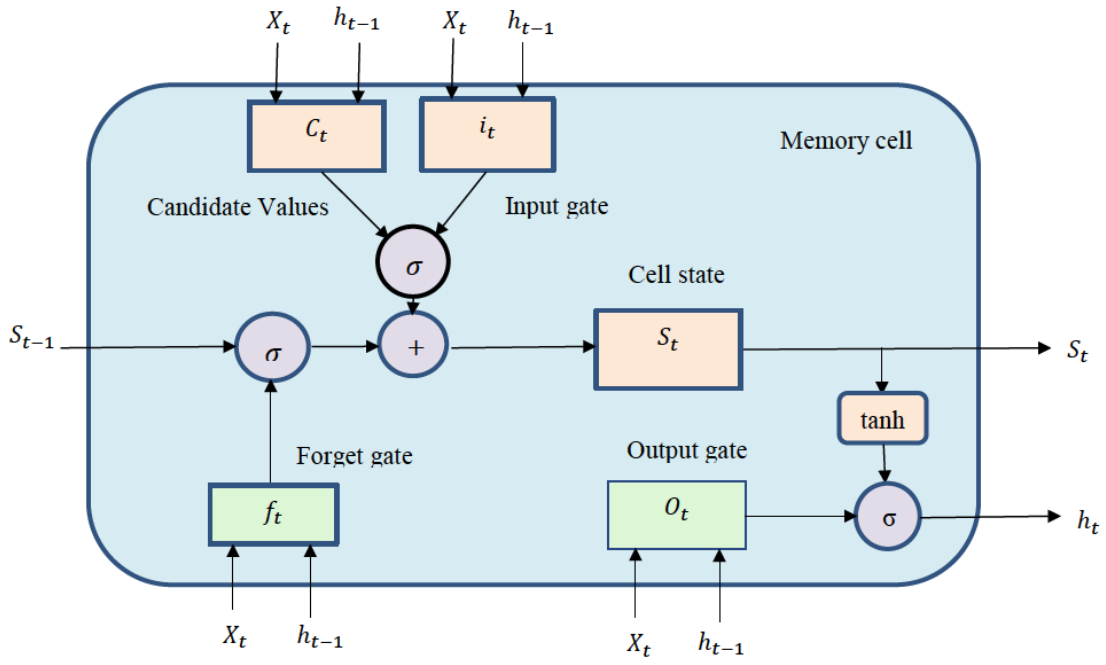


Figure 9: Shows memory cell structure

These three gates function as follows:

1. Input gate: define which information require to add to the memory (cell state)
2. Forget gate: define which information need to be removed from the memory (cell state)
3. Output gate: in this gate decide which information from the memory (cell state) to be used as output.

The structure of memory cell demonstrated in the equations below are described the update of the memory cells in the LSTM layer at every timestep t. Hereby, the following notation:

1. xt Represents the input vector at timestep t.
2. W_f, $W_{f,h}$, $W_{c,x}$, $W_{c,h}$, $W_{i,x}$, $W_{i,h}$, $W_{o,x}$ and $W_{o,h}$ are weight matrices.
3. b_f, b_c , b_i and b_o are bias vector.
4. f_t, i_t and o_t represent the victors for the activation values of the respective gates.
5. $S(t)$ and $C(t)$ are vectors for the cell states and candidate values.
6. h_t represents a vector for the output of the LSTM layer.

At the first stage, the LSTM layer decides which should be replaced from its previous cell states S_{t-1}. Therefore, the activation valuesf_t of the forget gates at timestep t are calculated by using the current input x_t, the output h_{t-1} of the memory cells at timestep$(t-1)$, and the bias terms b_f of the forget gates. The sigmoid function measures finally all activation values into the range between Zero (completely forget) and One (completely remember). Equation 1 forms the sigmoid function with inputs parameters.

$$f(t) = sigmoid(W_{f,x}\, Xt + W_{f,h}\, h(t-1) + bf\,) \qquad 3.1$$

In the second stage, the LSTM layer controls which information should be added to the network's cell state (s_t). This procedure comprises two operations: first, candidate values(C), which could potentially be added to the cell states, are computed. Second, the activation values it of the input gates are calculated:

$$C(t) = tan\, h\, (W_c, Xt + W_{c,h}\, h(t-1) + \, b_c) \qquad 3.2$$

$$I(t) = \, sigmoid\big(W_{i,x}\, X_t + W_{i,h} h_{t-1} + b_i\big) \qquad 3.3$$

In the third stage, multiplying the two previous stages which resulted in computing the new cell states *S(t),* as illustrated in the following equation:

$$S(t) = (f(t)\, S_{t-1} + I(t)C(t)) \quad 3.4$$

In the last step, the output h_t of the memory cells is derived as denoted in the following two equations:

$$O(t) = sigmoid(W_{o,x}\, X_t + W_{o,h}\, h_{t-1} + b_0) \quad 3.5$$

$$h(t) = O(t) \tanh(S(t)) \quad 3.6$$

3.3.4 Proposed system architecture

The parameter details of the proposed bidirectional LSTM model are presented in both Table 2 and Figure 10. To be specific, the number of LSTM cells in each forward/backward layer was determined to twice the input sequence length. This has been empirically demonstrated to perform well on a range of time series classification and natural language recognition problems. To produce the final classification, we used two fully connected layers arranged as the top model. A single dimension of Global max pooling was applied between bidirectional LSTM Layers and fully connected layers so that it can reduce the features of the output sequences generated by the LSTM layers through selecting the highest values with a matrix. These LSTM layers perform effectively to learn and extract the knowledge represented by features from the input of RR interval data sequence, before forwarding these features to the fully connected top model to classify whether AF signs are detected or not. The proposed model was implemented using both TensorFlow and Keras [165]–[167].

Table 2: Bidirectional-LSTM architecture [14].

Layer number	Type	Output shape	Number of parameters
Layer-1	Input	100,1	0
Layer-2 a	LSTM feed (forward)	100,400	161600
Layer-2 b	LSTM feed (backward)	100,400	161600
Layer-3	Global ID max pooling	400	0
Layer-4	Fully connected Rectified Linear Unit (ReLU)	50	20050
Layer-5	Dropout	50	0
Layer-6	Fully connected (Sigmoid)	1	51

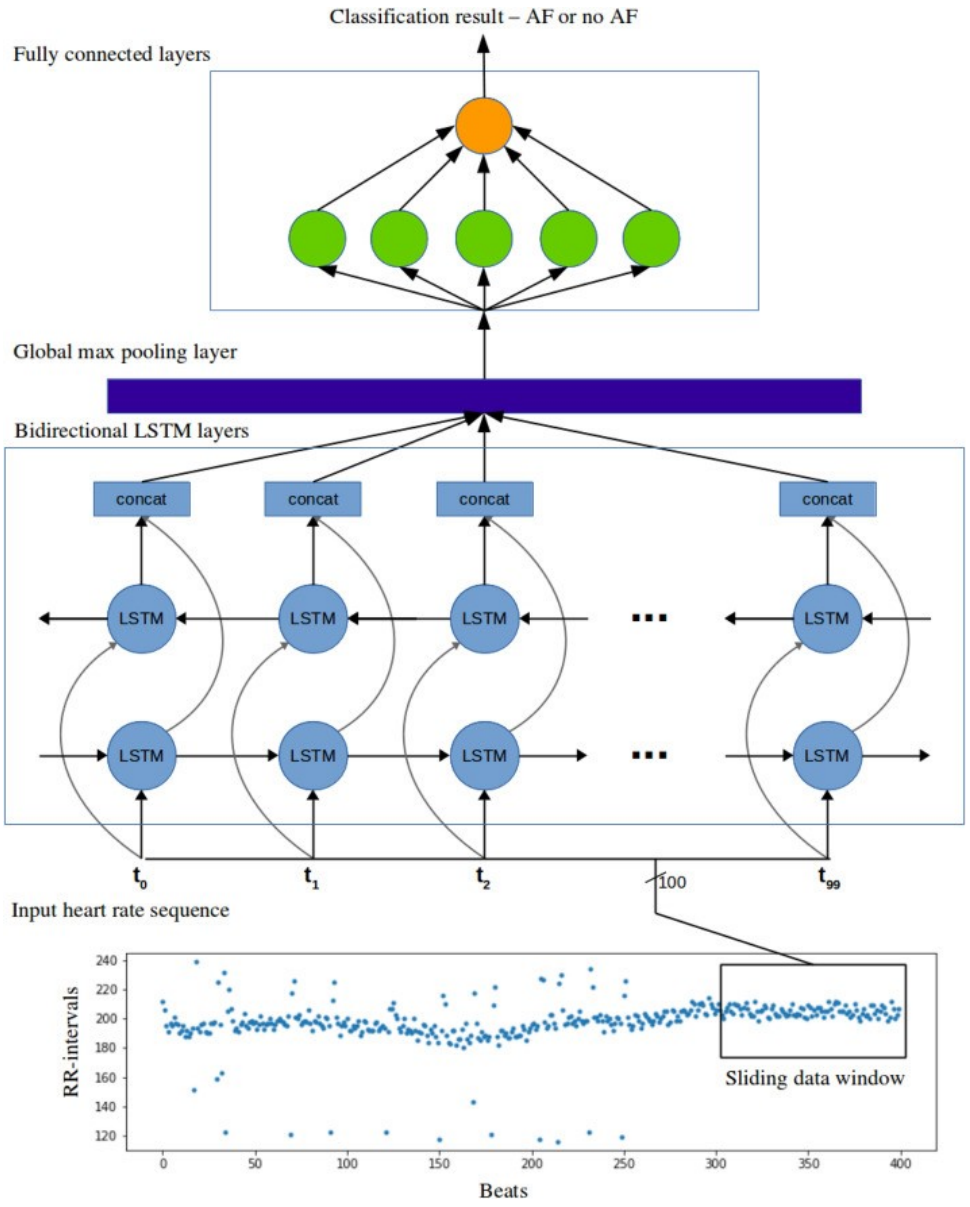


Figure 10: The architecture of bidirectional-LSTM used for AF classification [14].

3.3.5 Model training

Xavier [168] introduced the initialisation aspect of model training in terms of initialising all the weights of a bidirectional LSTM model as well as gradient descent backpropagation, through applying the Adam optimiser method [169]. We used that method to update the weights. The Adam optimiser was set to $1e^{-3}$ and binary cross

entropy was used to assess the network loss. In this experiment, a small batch size of 1024 input sequences were used throughout the learning phase, providing a good trade-off between the available memory capacity in Graphical Processing Unit (GPU) and training speed. The recurrent dropout [170] was set with a probability of 0.1 during training the model to both inputs and hidden states of the bidirectional-LSTM cells while the standard dropout [171] was used among the fully connected layers, also with a probability of 0.1 in order to decrease model overfitting and enhance model generalisation.

In addition, the binary cross-entropy function was used to evaluate the training performance of the proposed model, and this approach provides a better understanding of the model performance within a range of operating conditions which is different from the classification accuracy that only indicates to the performance of a model at one point. That entropy function can compare two probability distributions which are the predicted distribution and true distribution so that it can bring more details about the basis of search landscape. A layered 10-fold cross validation approach was used to evaluate the model performance and tuning both the hyperparameters and model architecture. The layered cross fold validation was essential to ensure that each fold represented the full dataset.

3.3.6 Performance measures

The confusion or error matrix summarizes the prediction results of a classification problem [18]. The layout allows visualization of the performance of a decision-making algorithm. To be specific, the matrix has two rows and two columns that represent the classes of actual and predictive analysis. Therefore, these performance measures report the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Most medical test results refer to a positive case (classifying the subject having the disease) and a negative case (classifying the patient not having the disease). In addition, from these terminologies, we can calculate the accuracy, sensitivity and specificity as shown in the following equations:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad 3.7$$

$$Sensitivity = TP/(TP + FN) \quad 3.8$$

$$Specificity = TN/(TN + FP) \quad 3.9$$

The ROC curve is a method used to evaluate the diagnostic accuracy of tests in modern medicine. It is widely used to illustrate how well a diagnostic model can distinguish between the presence and absence of disease and works equally well with data sets that exhibit class imbalance. The ROC represents the TP rate (sensitivity) plotted against the FP rate (1-specificity) for various cut-off points [172]. Each point on the ROC graph indicates the sensitivity/specificity corresponding to a specific decision threshold. The Area Under Curve (AUC) is a summary metric indicating the discriminatory power of a classifier.

3.3.7 LSTM model results

This section presents the outcomes for training and testing according to 10-fold cross validation. We also present blind-fold validation results. In this study, the bidirectional LSTM model was trained by using a Nvidia Quartro M5000 system. That system has a Graphical Processing Unit (GPU) with 8GB of GDDR5 graphics RAM. The average time required to train one epoch of this model was about 215 seconds. The primary experiments showed that the model was set to 80 epochs during training and model fitting. As such, applying 80 epochs can meet the requirement of generating an excellent model and it limits opportunities for overfitting.

3.3.8 10-fold cross validation results

Figure 11 and Figure 12 depict the training and validation set performance versus the number of epochs. These figures illustrate the average of the performance as a solid line

for each of the 10-folds, and the standard deviation of the performance represented as the shaded region.

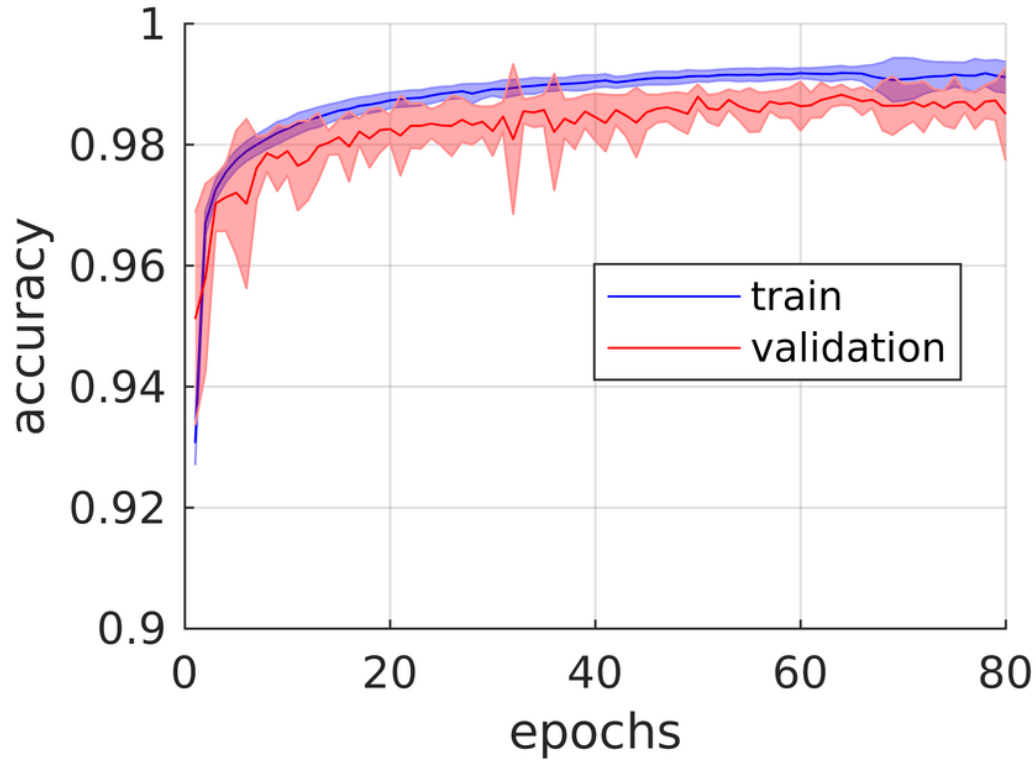


Figure 11: Training and validation accuracy curve over 80 epochs. The shaded area with red marker indicates to the variance. The solid line with blue maker refers to the mean of 10-folds cross validation [14].

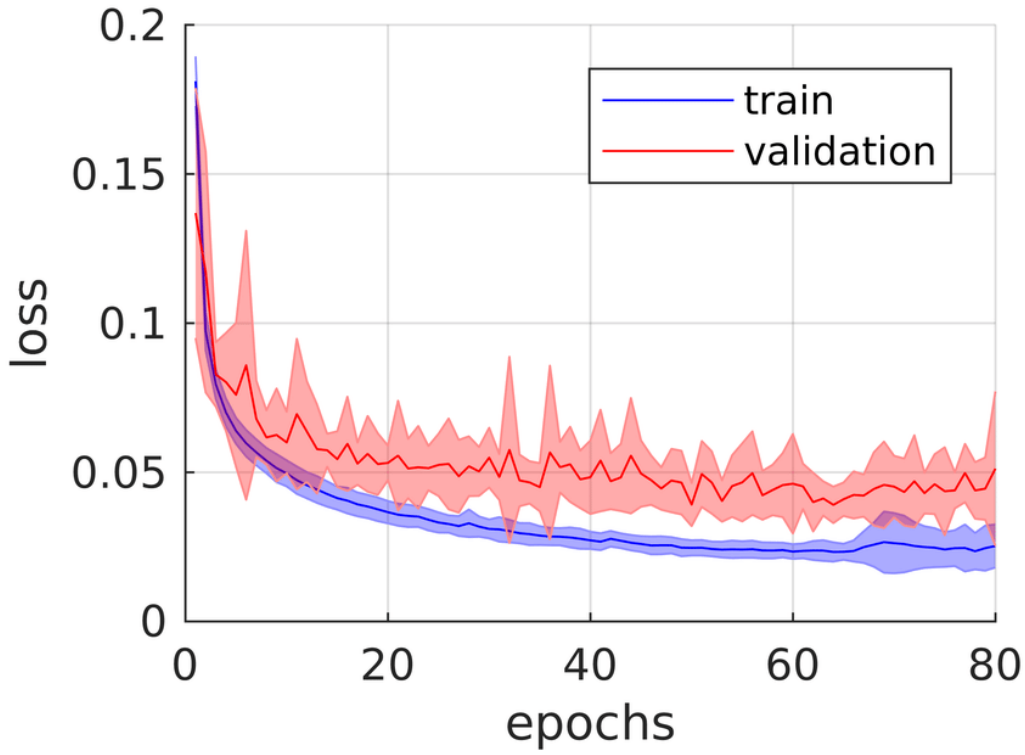


Figure 12: Training and testing loss function over the 80 epochs. The shaded area indicates the variance whilst the solid line is the mean [14].

These figures indicate that measuring the performance on the training set is slightly better than the validation set. The model was tuned to a stable value and there are no signs of overfitting, i.e., the training performance shows ongoing improvement while the validation declines. The confusion matrix for the 10-fold cross-validation process is shown in Figure 13 and the Receiver Operating Characteristic (ROC) curve is shown in Figure 14. The results, shown in these figures, were achieved by employing the model to the validation sets and aggregating the results from all 10 folds. The mean performance of all 10-folds is shown in Table 3. Based on the outcomes of Figure 13 and 14 as well as

Table 3: Overall 10-folds cross-validation performance outcomes of LSTM model.

TN	FP	FN	TP	Accuracy	Sensitivity	Specificity	AUC
523,241	7,040 7	7,040 7	430,615	98.51%	98.32%	98.67%	0.9986

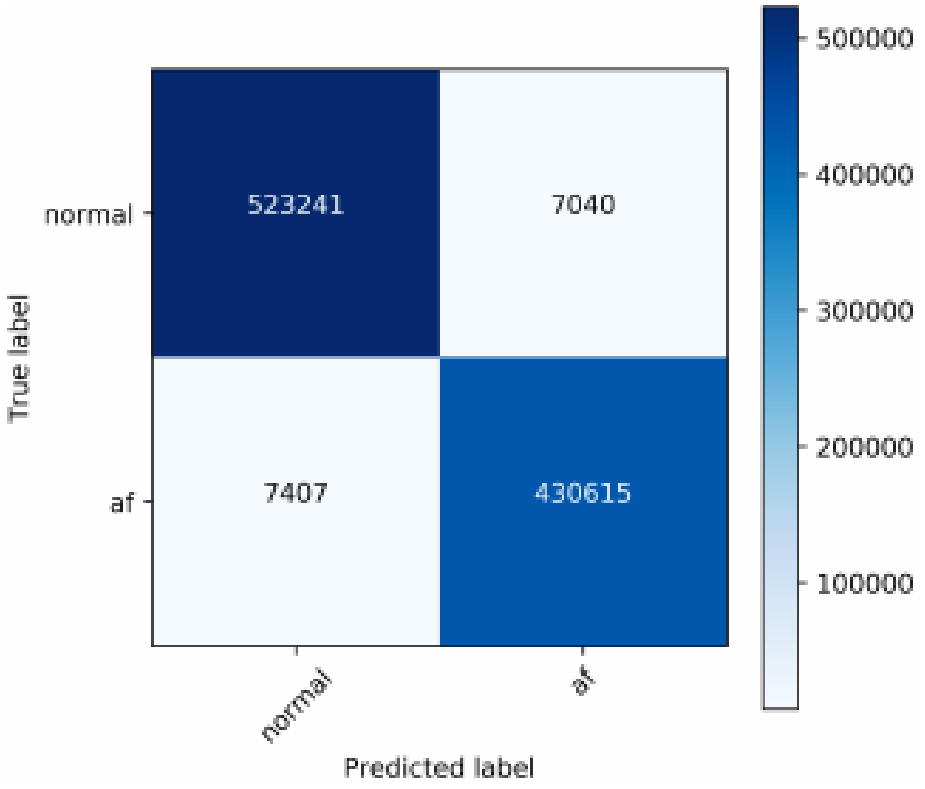


Figure 13: Confusion Matrix plot of 20 subjects from 10-folds cross-validation process [14].

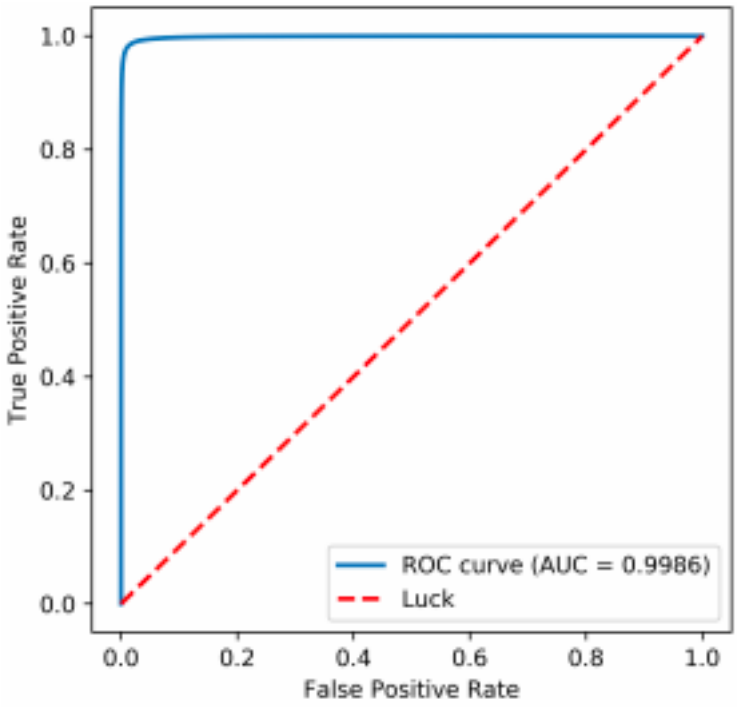


Figure 14: ROC curve from the stratified 10-folds cross validation [14].

3.3.9 Blind-fold validation results

After 10-fold cross validation we performed a blind-fold evaluation of the proposed model. That was undertaken by using normal and AF HR sequences from 3 subjects that were not used during 10-fold cross validation. The outcomes from this holdout test set are presented in Figure 15 and Table 17 as well as in Table 4. It can be noticed from the results that the proposed LSTM classifier accomplishes an overall accuracy of 99.77% on the tested set of completely unseen subjects- identifying correctly 99.61% of normal HR sequences and classifying 99.87% of HR sequences that correctly showed the presence of AF signs.

Table 4: Overall blind-fold validation performance results of LSTM classifier.

TN	FP	FN	TP	Accuracy	Sensitivity	Specificity	AUC
65,699	255	116	91,888	99.77%	99.87%	99.61%	1

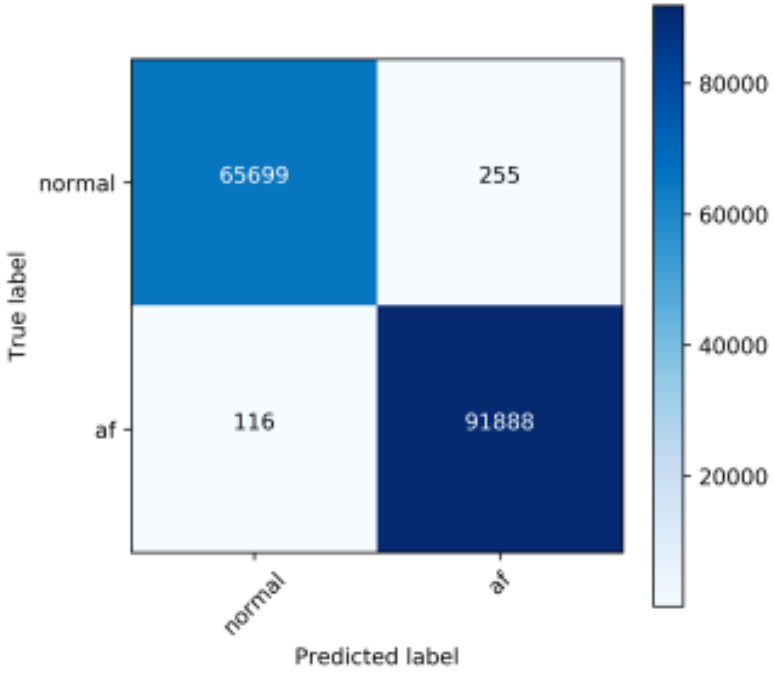


Figure 15: Confusion Matrix of blind-folds validation for 3 holdout subjects [14].

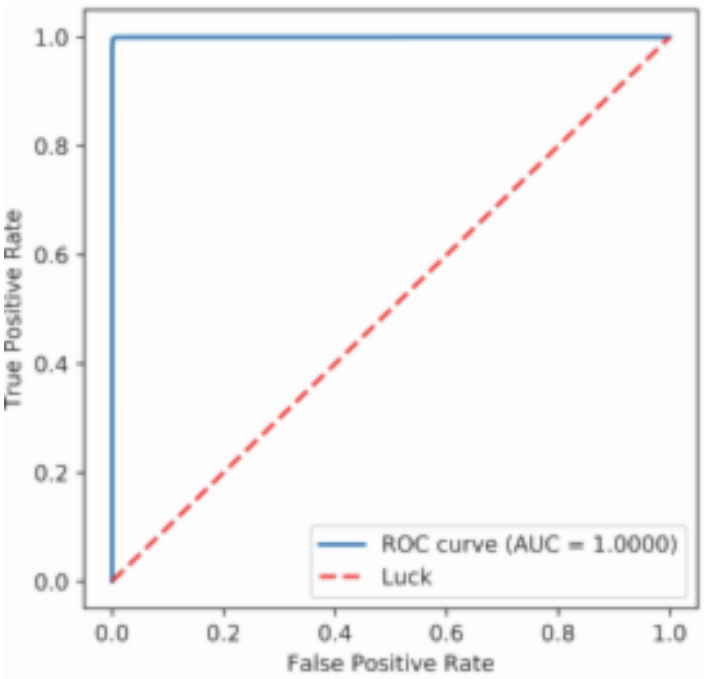


Figure 16: ROC curve of blind-folds validation for 3 holdout subjects [14].

3.4 Online system

The proposed solution incorporates establishing an automated AF detection service based on RR intervals.

Figure 17 shows a graphical representation of the proposed technology that aims to prevent stroke. This technology involves wearing a heart patch sensor that operates in real time, and the measured HR data is communicated to smart phone or tablet via Bluetooth Low Energy (BLE). The data will be transmitted from the point of measurements (patient) to a central point known as cloud server through a Wi-Fi or a mobile connection. Once the data has reached the cloud server, the processing phase beings by using the integrated DL algorithm for automated AF detection. The machine's decision quality will be verified and evaluated by a physician to improve the safety aspect. In case any emergencies are noticed, the healthcare provider will inform the patient through a feedback channel either by sending an email or a text message.

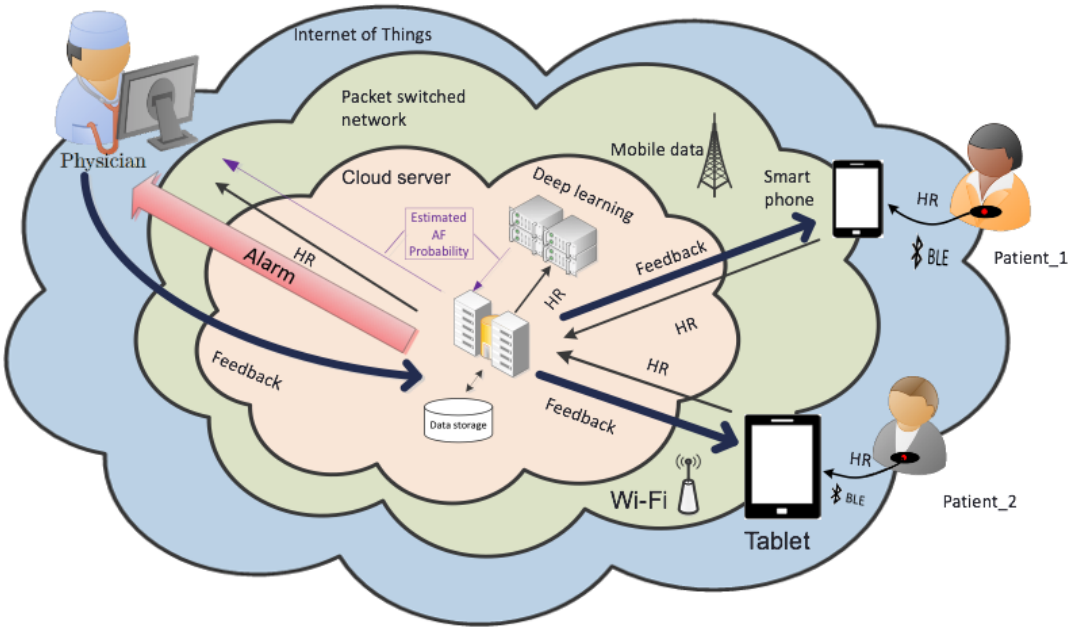


Figure 17: Architecture of the proposed AF monitoring system.

There is a clear demand from NHS trusts for establishing automated detection and smart management systems of AF patients [173], [174]. The main objective of the service concept is to prolong the observation duration. With a service timeline we show how the proposed system can achieve that through patient-led data acquisition. The collected data from the patient is either ECG or RR intervals measured by a wearable patch. The resulting signals were continuously analysed, and AF signs are detected. An urgent medical intervention is only required when AF is detected. To document the service approach, we designed a timeline that involves a sequence of synchronised actions that provide the desired functionality.

Figure 18 shows the timeline that refers to certain actions for healthcare givers, patient, cardiologist, and AF detection service. The timeline begins with the activity of signing up a patient through a nurse for the AF detection service. This is when the patient-led data acquisition process starts. Data are measured, transmitted, stored, and analysed via IoMT and cloud technology [175]. This creates the continuous AF monitoring functionality. The cardiologist is being alerted once the DL algorithm detects AF events in the stored data. A cardiologist can examine the available data and combine this information with

knowledge gained from standard care to reach an accurate and effective diagnosis [24]. In case the cardiologist rejects the classification outcomes from the AF detection service, the monitoring continues to operate as per normal. However, if AF is detected, the patient is informed, and treatment can start. The AF detection service is then used for treatment monitoring.

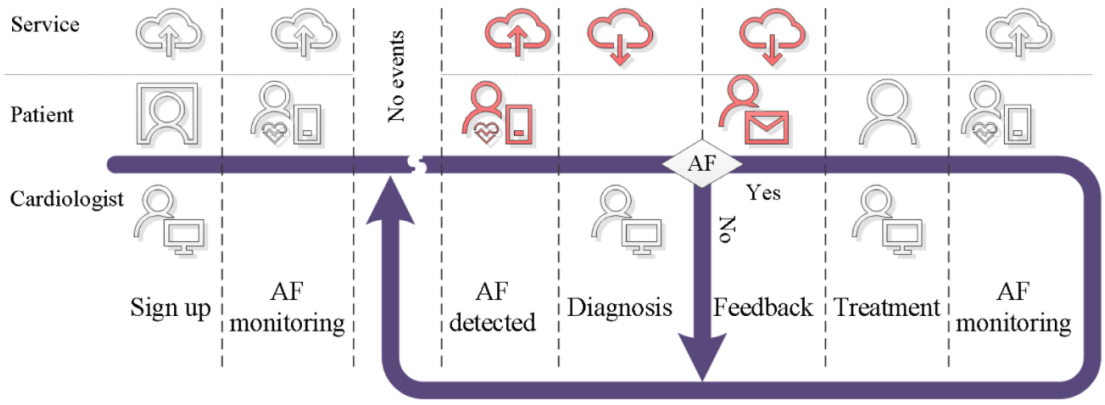


Figure 18: Timeline of AF detection service [22].

3.5 ResNet Model

This section describes the methods used to underpin our statement that automated arrhythmias detection in RR interval signals is applicable. These methods were employed to create a signal processing system which can be used to train and test a ResNet DL algorithm with benchmark data. There were two targets of augmenting [176] and balancing [177] the dataset as directed by the design strategy. Balancing a dataset indicates to establish the same amount of training data for each class. From the benchmark data, we found that AFL had the least number of RR intervals. Therefore, we applied a scrambling approach to augment the dataset. To achieve the augmentation, round robin windowing technique was used to increase the amount of data for all signal classes. Ultimately, puncturing was utilised to balance the dataset [17].

Figure 19 provides a data processing overview block diagram. The processing starts with classifying the available ECG datasets, from benchmark database, into three different classes, which are AFIB, AF, and NSR. The RR interval signal was extracted from the available ECG signals. These RR interval signals were processed to train and test the ResNet model. The model was evaluated by using the performance measures known as ROC curve and confusion matrix. The following sections introduce both data and processing steps in more detail.

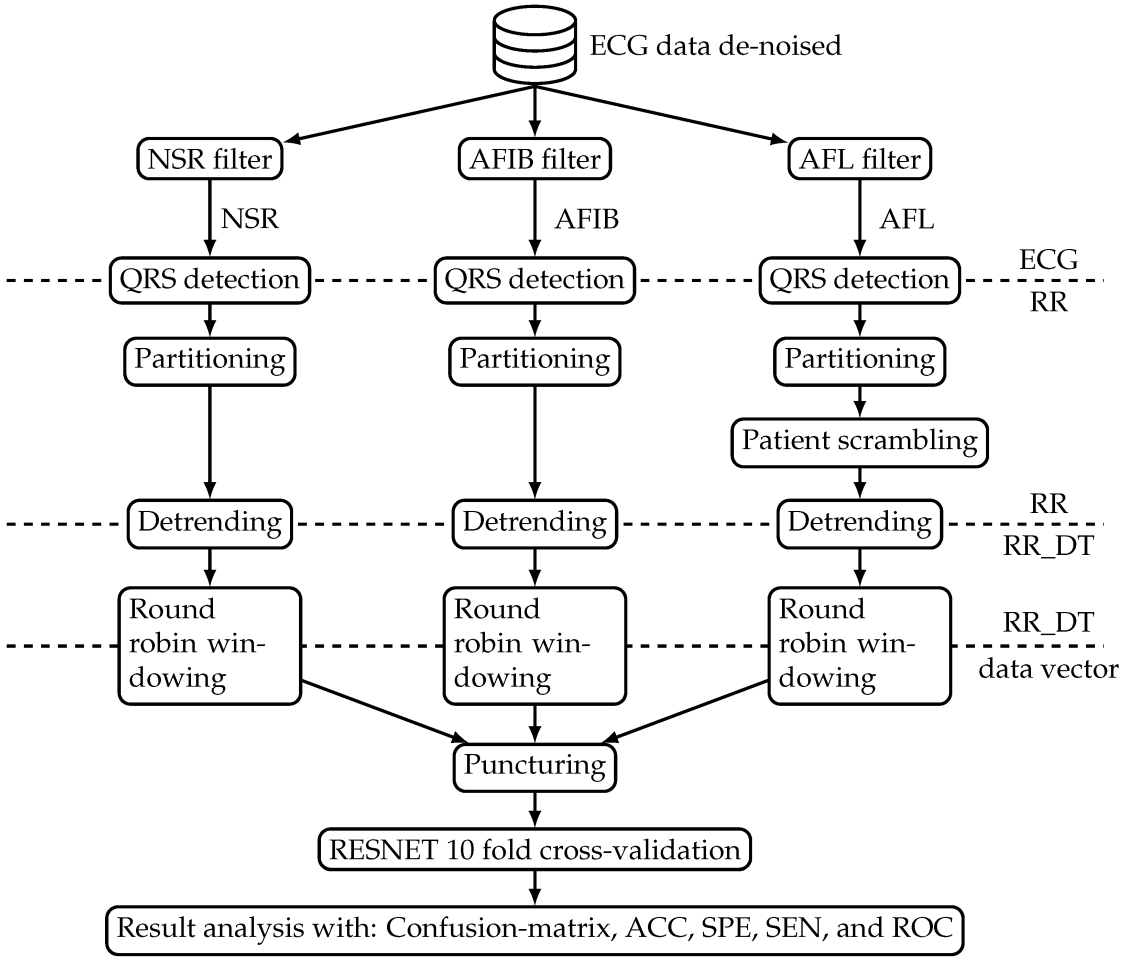


Figure 19: Block diagram mapping the overview study setup [17].

3.5.1 ECG benchmark data

Figure 19 shows that one ECG database is used to provide benchmark data to train and test the ResNet algorithm. As a perquisite processing step, the ECG data requires denoising which contained recording information of 12-lead ECG signals from 10,646 patients. These signals have been sampled with 500 HZ for a recording duration of 10 seconds. The data were collected at Chapman University as well as Shaoxing people's hospital (Shaoxing Hospital Zhejiang University School of Medicine) [178]. Each recorded signal was annotated by a cardiologist to identify one of 11 common rhythms. The annotation has been provided as a table that relates disease labels and ECG signal file name. From the given table, we selected all the files of interest that were labelled as AFIB, AFL, NSR. Table 5 indicates to the number of subjects for each individual signal class and the accumulated ECG duration (over the individual patient within a class).

Table 5: Data properties for the three signal classes. The 'ECG Duration (s)' column provides the time duration of all ECG signal blocks for each individual class. The following two columns to the right provide the number of RR intervals and the number of RR DT.

Property / Class	ECG duration	RR intervals	RR_DT Samples	Number of blocks	Number of participants
AFIB	17,800	25,995	25,995	1780	1780
AFL	4450	7536	7536	445	445
NSR	18,260	33,976	33,976	1826	1826
Total	40,510	67,507	67,507	4051	4051

The table contains the entries for ECG duration that shows that all ECG signals recorded with the length of 10 seconds. The ECG recording for each patient compromises of one data block. Each block of ECG data includes an array of 12 × 5000 samples, where 12 refers to the number of leads and 5000 indicates to the samples captured within 10 seconds. Data block term refers to describe the subsequent processing stages to express the data from one patient.

Figure 20 shows three examples of signal recordings namely, NSR, AFIB, and AFL. There are three different signals for each signal class. The first of these signals depicts a 10 second ECG signal.

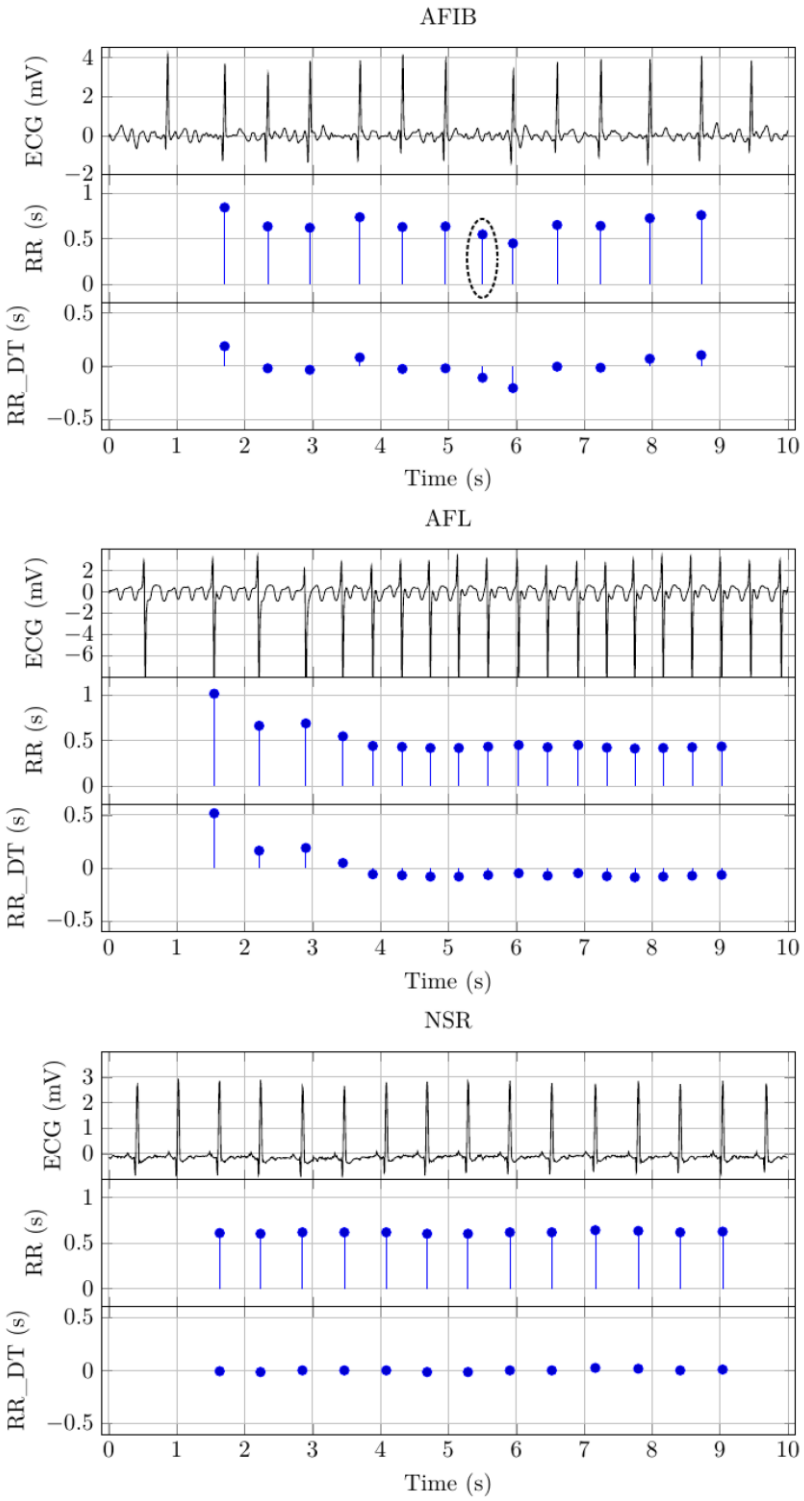


Figure 20: Example plots from AFIB, AFL, and NSR signal classes. The ECG signal was recorded with the aVL lead. The RR-intervals, plotted as RR-intervals over time,

were extracted from the ECG via QRS detection. The detrended RR-intervals were plotted as RR DT [17].

3.5.2 QRS detection

The QRS detection is an important step that based on extracting the RR intervals from the ECG data blocks [17]. Hence, QRS represents the main component element in an ECG signal. It is generated by the ventricular depolarisation that happens when the heart's muscles contract during a heartbeat. To be specific, within the QRX complex combination, the R wave indicates to the peak and the time location of that wave which symbolises the time location of the heartbeat. One RR interval means the distance between two consecutive R peaks. The well-known ECG kit for MATLAB was used to process the ECG [179]. The ECG kit framework has the wave detect algorithm that was implemented by Martínez et al [180]. Once the RR interval sequences are generated, these sequences can be saved and maintained as the block structure. From Table 5, it can be seen that there is a column of RR intervals for each signal class. As such, this step comprises a substantial data reduction. A better description of data reduction is illustrated with the following example. The number of NSR recorded as ECG data blocks involved 109560000 samples. These samples reduced to only 33976 RR intervals after QRS detection. Therefore, the compression ratio accomplished by the QRS detection phase was 3224.6291.

Figure 20 shows the example of each class signals of the extracted RR interval. The x-axis scale refers to the location of RR interval such as the time location where RR interval ends. The y-axis scale indicates to the duration of RR interval. From the visual inspection, it can be recognised that the QRS detection algorithm has detected an additional beat for the AFIB as example signal. We have addressed that RR interval with a black circle in AFIB plot.

3.5.3 Fold generation and Patient scrambling

Figure 20: Example plots from AFIB, AFL, and NSR signal classes. The ECG signal was recorded with the aVL lead. The RR-intervals, plotted as RR-intervals over time, were extracted from the ECG via QRS detection. The detrended RR-intervals were plotted as RR DT 10-fold cross validation involves partitioning the RR intervals data into 10 folds, of approximately equal size [181]. This approach has been described in detail in Section 3.3.2. To achieve the fold generation, we have divided the data along RR interval blocks. That strategy is the same as generating the fold along subjects. In other words, the data related to specific subject can only be found in one-fold. Table 6 demonstrates this activity by presenting the number of RR intervals for NSR, AFIB, and AFL.

Table 6 Number of RR intervals per signal class for each fold. AFL_{SC} denotes the scrambled AFL dataset.

Fold / Class	F-1	F-2	F-3	F-4	F-5	F-6	F-7	F-8	F-9	F-10
NSR	2015	1980	1980	2029	2020	1973	1992	2017	1975	1975
AFIB	2651	2667	2584	2566	2633	2649	2594	2512	2604	2535
AFL	742	759	786	784	762	766	727	702	721	787
AFL_{SC}	2226	2277	2358	2352	2286	2298	2181	2106	2163	2361

From the table above, it can be noticed that in all folds, the number of AFL RR intervals is lower by more than three times when compared to AFIB RR intervals. To correct that imbalance, we have implemented patient scrambling to increase the AFL data. The core concept of patient scrambling was established with a fold generation algorithm that uses the sequence in which the RR interval block appears in the dataset to create fold data. This sequence would affect the data vectors, which were generated via round robin windowing due to the window length is greater than the number of RR intervals in any

specific data block. Each data vector contains 100 detrended RR intervals from different patients. In the scrambling step, we use this feature to create more AFL data. To do this, we established 3 variations of the sequence in which each patient data emerged in the training and testing datasets-for each individual fold. Table 6 indicating to the number of the augmented RR intervals for AFL_{SC} is exactly three times greater than the number of AFL RR intervals for the same fold.

3.5.4 Detrending

With detrending approach, the DC offset can be removed from the RR intervals signal [17]. This processing method was applied to aid the learning phase by reducing the training time as well as the network complexity [182]. In this experiment, we have utilised detrending and low-pass filter introduced by Fisher et al. [183]. The filter combination is based on an Ornstein-Uhlenbeck third-order Gaussian process which acts on the RR-interval signal directly [184], [185]. Once the detrending step is completed, the datasets contain RR DT samples. Table 5 contains the column of processed RR DT samples. As such, the detrending does not increase the amount of data; therefore, the number of RR DT samples corresponds to RR intervals.

Figure 20 shows the detrended plot of RR signals for each signal class. The signal graphs show that the DC bias is significantly reduced.

3.5.5 Round robin windowing and Puncturing

Applying a round robin windowing approach can augment the data through generating a data vector with 100 elements for each RR DT sample. That approach increases the data dimensionality 100-fold [17]. These vectors indicate to the specific class data for each fold that arranged in a window with 100 elements. This window was slid over the RR DT signal one sample at a time. Round robin method refers that first 100 RR DT samples for

each dataset were copied at the end, before applying the window. That extension would link one data vector for each RR DT sample.

Puncturing is the subsequent step used after windowing which can adjust data size for AFIB and AFL_{SC} datasets. The puncturing algorithm removes equidistant data vectors. This technique ensures that the number of training data, for each of the three classes in a fold, is equal. Table 7 shows that NSR has the least number of data vectors for the 10-folds when compared with AFIB and AFL_{SC}. As a result, we have selected the number of NSR data vectors as a target for puncturing AFIB and AFL_{SC}.

Table 7: The number of data vectors per signal class within each fold. AFL_P and $AFIB_P$ denote the punctured datasets for NSR and AFIB respectively.

Fold Class	F-1	F-2	F-3	F-4	F-5	F-6	F-7	F-8	F-9	F-10
AFIB	2651	2667	2584	2566	2633	2649	2594	2512	2604	2535
AFL	742	759	786	784	762	766	727	702	721	787
NSR	2015	1980	1980	2029	2020	1973	1992	2017	1975	1975
AFL_p	2015	1980	1980	2029	2020	1973	1992	2017	1975	1975
AFIB p	2015	1980	1980	2029	2020	1973	1992	2017	1975	1975

The puncturing algorithm will decrease the number of data vectors until reaching the number of NSR data vectors in the same fold. For instance, the number of NSR data vectors in Fold-1 has 2015 samples. After puncturing processing, the number of $AFIB_P$ and AFL_P is equal to the number of NSR data vectors. Once the pre-processing is completed, the new data vectors NSR, $AFIB_P$, and AFL_P were used to train and test the ResNet model.

3.5.6 ResNet 10-fold cross-validation

Overfitting is a main issue for classifying the physiological signals with DL. This term indicates to the fact that ResNet network can remember the signals itself rather than the signal properties which are related to disease symptoms. In a practical implementation, overfitting happens when the DL algorithm classifies the training datasets correctly but fails to achieve that with testing datasets. There are several approaches to avoid or at least reduce overfitting. Model selection plays a main role in that process. In this study, we followed the findings by Fawaz et al. [186], in their review on DL for time series classification they found that ResNet Network outperforms all the other tested models. Figure 21 shows the ResNet model was established by using the data low structure.

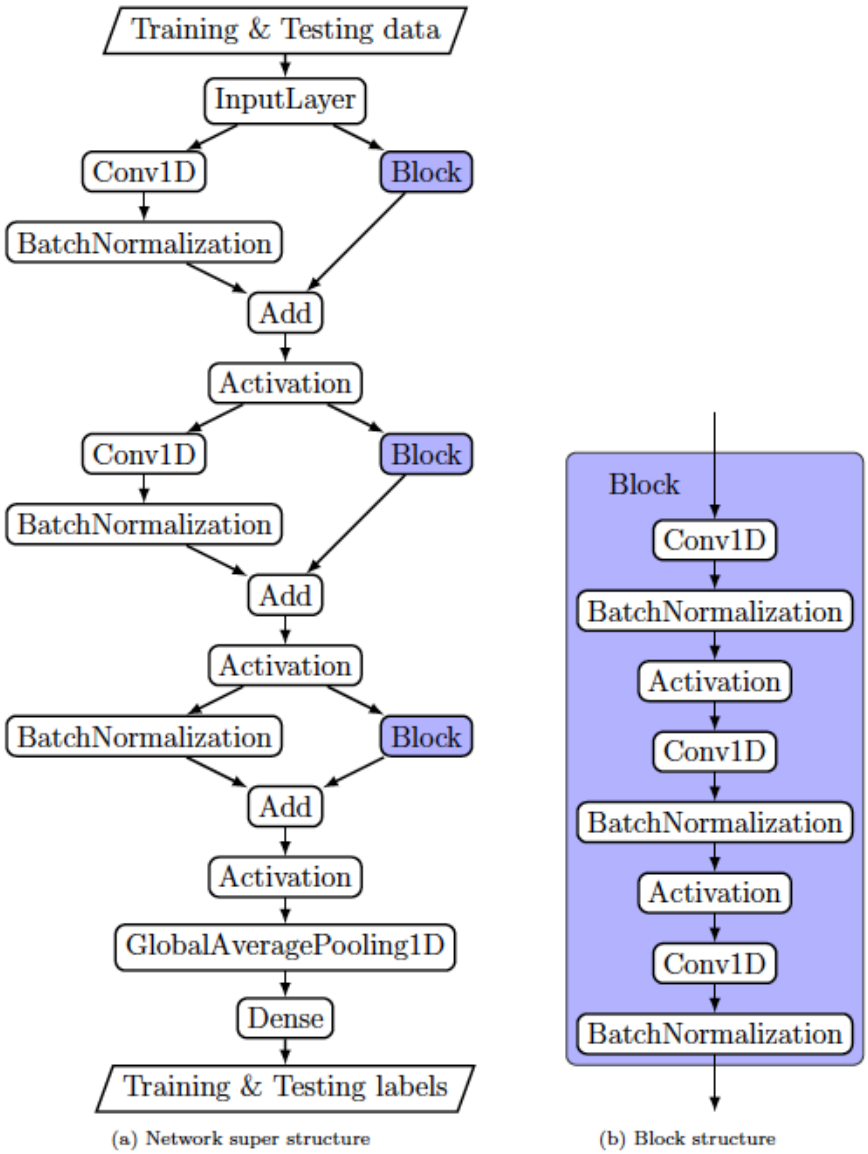


Figure 21: ResNet architecture used for training and testing [121].

The data flow diagram is comprised from standard components which have a direct correspondence with the support of python Application Programming Interface (API) Keras [187] for the DL framework TensorFlow [188]. The structure of data flow indicates that there are three shortcut connections which enable the information to skip the processing block. This structure is known as residual block. That structure can address another limitation of the particular DL model which is the so-called vanishing/exploding gradient problem [189]. From a practical aspect, this problem occurs when increasing the

number of network layers results in lower training accuracy. Hence this kind of problem is distinct from the general overfitting problem. The hyperparameters require tuning once the network is selected. We have applied a trial-and-error technique to narrow down the optimal parameters. To be precise, we implemented a collaborative process which was controlled by an increasing understanding of the interaction between signal processing and classification model. Table 8 provides the number of data vectors used to train and test the ResNet model. In a final stage, these data vectors were used to shape both training and testing sets. Table 6 provides the parameters for these datasets. They were arranged by choosing one-fold for testing and using the data vectors in the remaining folds for training. That process is repeated until each fold was used for testing. This comes from the fact that the number of data vectors for AFLSC, AFLP, and AFIBP is equal for each fold, see Table 7, results in a perfectly balanced training dataset. That implies for any given test fold, the number of vector data for NSR, AFIB and AFL has the same amount. To achieve this process, Column1 in refers to the test fold and the remaining columns on the right indicate to the number of training and testing data vectors. For instance, if fold 1 was used for testing, the data vector categories are provided in Row1 Table 8. Hence, the network was trained for all the data vectors (53823) starting from fold-1 to fold-10 (53943). The network was tested with all data vectors (6892) from fold-1, including NSR, AFIB and AFL.

Table 8 shows the number of data vectors used for training and testing during the 10-fold cross validation.

Training datasets					Testing datasets			
N Fold	NSR	AFIB	AFL	Total	NSR	AFIB	AFL	Total
1	17,941	17,941	17,941	53,823	2015	2651	2226	6892
2	17,976	17,976	17,976	53,928	1980	2667	2277	6924
3	17,976	17,976	17,976	53,928	1980	2584	2358	6922
4	17,927	17,927	17,927	53,781	2029	2566	2352	6947

5	17,936	17,936	17,936	53,808	2020	2633	2286	6939
6	17,983	17,983	17,983	53,949	1973	2649	2298	6920
7	17,964	17,964	17,964	53,892	1992	2594	2181	6767
8	17,939	17,939	17,939	53,817	2017	2512	2106	6635
9	17,981	17,981	17,981	53,943	1975	2604	2163	6742
10	17,981	17,981	17,981	53,943	1975	2535	2361	6871

3.5.7 Results analysis methods

The result analysis starts with creating a confusion matrix based on validating the classification outcomes of the DL model. Table 9 describes the confusion matrix in the form of the number of beats with a true and a predicted label: N predicted, true label. The predicted label was produced with the ResNet classification model. The combination of three predicted labels and three true labels forms the confusion matrix dimensionality with 3×3=9 beat labels. Table 9 shows the arrangement of these beat labels in the confusion matrix. AFIB and AFL are both arrhythmias. Hence, it is logical to combine AFIB and AFL beats to shape an arrhythmia class. The NSR beats constitute a non-arrhythmia class. Table 9 shows the confusion matrix that reflects these considerations. Based on the confusion matrix (cm), we define TP, TN, FP, FN.

Table 9: Confusion matrix for AFIB,AFL, and NSR.

		AFIB	**AFL**	**NSR**
True label	**AFIB**	$N_{AFIB, AFIB}$	$N_{AFL, AFIB}$	$N_{NSR, AFIB}$
	AFL	$N_{AFIB, AFL}$	$N_{AFL, AFL}$	$N_{NSR, AFL}$
	NSR	$N_{AFIB, NSR}$	$N_{AFL, NSR}$	$N_{NSR, NSR}$

Predicted label

Table 10: Confusion matrix for Arrhythmia and Non-arrhythmia.

True label \ Predicted label	**Arrhythmia**	**Non-arrhythmia**
Arrhythmia	$N_{\text{AFIB, AFIB}} + N_{\text{AFL, AFIB}} + N_{\text{AFIB, AFL}} + N_{\text{AFL, AFL}}$	$N_{\text{NSR, AFIB}} + N_{\text{NSR, AFL}}$
Non-arrhythmia	$N_{\text{AFIB, AFL}} + N_{\text{AFL, AFL}}$	$N_{\text{NSR, NSR}}$

Based on the confusion matrix, we define TP, TN, FP, and FN for a specific class (cl) as follows:

$$\text{TP}_{\text{cl}} = N_{cl,cl} \tag{3.10}$$

$$\text{TP}_{\text{cl}} = \left(\sum_{i \in \, class\ set} N_{i,i}\right) - N_{cl,cl} \tag{3.11}$$

$$\text{FP}_{\text{cl}} = \left(\sum_{i \in \, class\ set} N_{i,cl}\right) - N_{cl,cl} \tag{3.12}$$

$$\text{FN}_{\text{cl}} = \left(\sum_{i \in \, class\ set} N_{cl,i}\right) - N_{cl,cl} \tag{3.13}$$

Where cl ∈ (Class set) and Class set is either (AFIB; AFL; NSR) or (Arrhythmia; Non-arrhythmia).

These definitions were used to establish the performance measures of ACC, SPE, and SEN for the individual class:

$$ACC_{cl} = \frac{TP_{cl}+TN_{cl}}{TP_{cl}+TN_{cl}+FP_{cl}+FN_{cl}} \tag{3.14}$$

$$SEN_{cl} = \frac{TP_{cl}}{TP_{cl}+FN_{cl}} \tag{3.15}$$

$$SPE_{cl} = \frac{TN_{cl}}{TN_{cl}+FP_{cl}} \tag{3.16}$$

10-fold cross-validation results in 10 different performance measures. These individual performance measures were averaged to determine the overall performance. For the confusion matrix, that combination implies of the accumulation the matrix classes. Table 11 shows the considerations of the confusion matrix. The overall performance measures can be determined by using Equations 3.10 to 3.16.

A ROC curve demonstrates how the threshold level affects the diagnostic capability of a binary classifier [190]. AUC indicates the general performance of the classifier, for example, if the area is nearer to 1 that is an indication of achieving a better classification performance. Both Equations 3.15 and 3.16 were utilised to compute the class specific True Positive Rate (TPR) and False Positive Rate (FPR), respectively. The micro-average is the mean of the individual class results for AFIB, AFL, and NSR. The macro-average is computed by accumulating all the FP. Table 11: Average cross-validation confusion matrix. The summation symbol for test fold indicates to sum over all test folds.

True label	AFIB	AFL	NSR
AFIB	$\sum\{N_{AFIB,AFIB}\}$ (Test Fold)	$\sum\{N_{AFL,,AFIB}\}$ (Test Fold)	$\sum\{N_{NSR,,AFIB}\}$ (Test Fold)
AFL	$\sum\{N_{AFIB,AFL}\}$ (Test Fold)	$\sum\{N_{AFL,,AFL}\}$ (Test Fold)	$\sum\{N_{NSR,AFL}\}$ (Test Fold)
NSR	$\sum\{N_{AFIB,NSR}\}$ (Test Fold)	$\sum\{N_{AFL,NSR}\}$ (Test Fold)	$\sum\{N_{NSR,NSR}\}$ (Test Fold)

Predicted Label

[1]https://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html

3.5.8 ResNet network classification outcomes

The outcomes, presented in this section, demonstrate the classification performance achieved by the ResNet model. In order to establish that performance, 10-fold cross validation approach was applied to train and test the model. Table 8 describes the properties of the training and testing data. Hence, the table presents 10 Test Folds requiring 10 individual training and testing iterations. The training was achieved in 50 epochs with batch size of 16. Categorical cross-entropy was used as loss function and Adam [191] was used as optimizer. 50 epochs were specified for each training and testing iteration, and the obtaining the greatest testing accuracy that was used to form the confusion matrix. The confusion matrix structure is described in Table 9 . Having created the individual confusion matrices, as such, we are able to determine the overall confusion matrix as presented in Table 11. Equations 3.8 to 3.10 were used to compute the accuracy (ACC_{cl}), Sensitivity (SEN_{cl}), and Specificity (SPE_{cl}) where cl ∈ {AFIB, AFL, NSR}, calculating the performance measures for each class in matrix dimensionality 3x3 for each Test Fold. Table 12 documents the quality of the classification measures ACC, SEN, SPE, as well as the confusion matrix results for the 10 separate Test Folds and aggregating the overall test folds. Hence, the average of the overall performance measures for 10 folds highlighted in the last Row of table 10, are over the 95%. This reflects how perfect the proposed ResNet model was able to classify AFIB, AFL, and NSR using the RR interval signal.

From a medical perspective, the binary problem classification of Arrhythmia vs Non-Arrhythmia is also significant. Hence, we have used the terms, provided in Table 10, to refine all Test Fold confusion matrixes, presented in the last row of Table 12. In that step, we have established the two-class results, shown in Table 12. Figure 22 shows the ROC curve which provides the graphical representation of the classification outcomes. The large area under curve is a direct result of the excellent classification performance followed by the performance measures presented in Table 12 and Table 13.

Table 12: Analysis results for each fold and overall aggregated folds.

Test Fold	*cl*	ACC $_{cl}$ (%)	**SEN**$_{cl}$ (%)	**SPE**$_{cl}$ **(%)**	Confusion matrix		
1	AFIB	97.16	92.72	99.27	2064	162	0
	AFL	97.16	98.72	99.18	34	2617	0
	NSR	100.00	100.00	100.00	0	0	2015
2	AFIB	99.87	99.60	100.00	2268	9	0
	AFL	99.87	100.00	99.79	0	2667	0
	NSR	100.00	100.00	100.00	0	0	1980
3	AFIB	95.81	87.70	100.00	2068	290	0
	AFL	95.81	100.00	93.31	0	2584	0
	NSR	100.00	100.00	100.00	0	0	1980
4	AFIB	96.95	91.11	99.93	2143	209	0
	AFL	96.95	99.88	95.23	3	2563	0
	NSR	100.00	100.00	100.00	0	0	2029
5	AFIB	98.83	96.98	99.74	2217	69	0
	AFL	98.83	99.54	98.40	12	2649	0
	NSR	100.00	100.00	100.00	0	0	1970
6	AFIB	100.00	100.00	100.00	2298	0	0
	AFL	99.96	100.00	99.93	0	2649	0
	NSR	99.96	99.85	100.00	0	3	1970
7	AFIB	96.81	90.10	100.00	1965	216	0
	AFL	96.81	100.00	94.82	0	2594	0
	NSR	100.00	100.00	100.00	0	0	1992
8	AFIB	94.32	83.05	99.56	1749	357	0
	AFL	94.32	99.20	91.34	20	2492	0
	NSR	100.00	100.00	100.00	0	0	2017
9	AFIB	98.28	95.42	99.63	2064	99	0
	AFL	98.15	99.35	97.39	17	2587	0
	NSR	99.86	99.54	100.00	0	9	1966
10	AFIB	100.00	100.00	100.00	2361	0	0
	AFL	100.00	100.00	100.00	0	2535	0
	NSR	100.00	100.00	100.00	0	0	1975
All	AFIB	97.82	93.76	99.81	21197	1411	0
	AFL	97.80	99.67	96.66	86	25909	0
	NSR	99.98	99.94	100.00	0	12	19944

Table 13: Overall classification outcomes. Where cl=Arrhythmia

ACC $_{cl}$ (%)	SEN$_{cl}$ (%)	SPE$_{cl}$ (%)	Confusion matrix	
99.98	99.94	100.00	48603	0
			12	19944

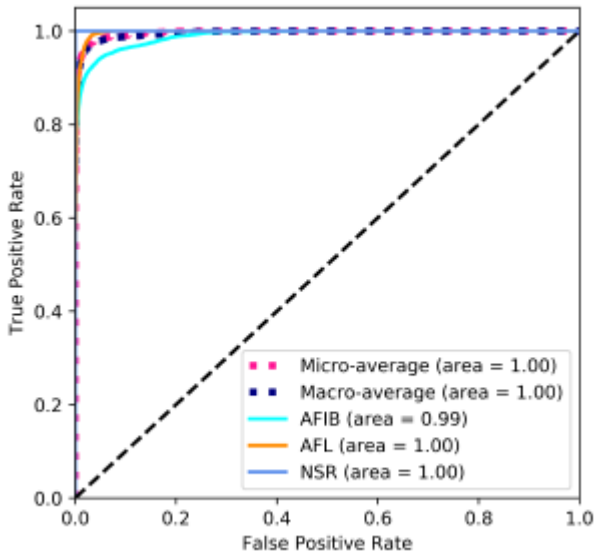


Figure 22: The ROC curve shows the diagnostic quality results for 10 folds.

3.6 Discussion

Data acquisition is always a concern for learning systems. The proposed automated AF detection system is no exception. The database from PhysioNet MIT-BIH AF were used to enable collaboration and competition. Competition means that the used data is well known, and classifications outcomes are available from other research study. Collaboration is doable due to the data being publicly accessible. There was no need to feature extraction in the processing structure, all the required information exciting in the training data set is fed into the deep learning system. Therefore, knowledge can be extracted implicitly which allows it to make good decisions even for unseen data. Looking

beyond the research, we found that there are two main methods to measure the electrical activity of human's heart. The LSTM method based on classifying RR interval from HR signal and detects the abnormality through computer-aided-diagnosis. However, the standard measurement method is relying on the ECG and the AF detection can be done either professional cardiologist or by machine learning.

The ResNet method investigates the problem of classification NSR, AFIB and AFL by using the RR intervals. This problem has been highlighted with various research studies which extracted the information from ECG signals. The morphology of ECG signals has principal structural elements, such as the QRS complex, which aids the classification efforts. Cardiologists use changes in ECG morphology for arrhythmia diagnosis. These structures are removed during QRS detection which is used to extract the RR-interval sequence. The RR-interval reflects only the heartbeat rhythm. That rhythm is distinct for NSR and AFIB. Therefore, arrhythmia research based on RR-interval sequences has focused on differentiating AFIB and AFL. Only Ivanovic [125] address the three-class problem of AFIB, AFL, and NSR. Direct competition with this study is difficult because the authors have used a private dataset. To be specific, we could not apply the ResNet algorithm to their dataset and therefore we can only compare the performance results achieved with different datasets. A statistical comparison reveals that the LSTM based detection method, proposed by [125] has a ≈10% lower accuracy when compared to our ResNet approach.

Table 14 provides an overview of arrhythmia detection studies based on ECG and RR interval signals.

In general, ECG based arrhythmia detection achieves better accuracy values when compared to RR interval-based detection. We believe that this holds true, even though a direct comparison is not possible because different datasets were used to establish the performance results. ECG contains all the information about the electrical activity of the human heart. As such, the RR interval is part of this information. Hence, during the process of extracting the RR intervals we lose all the additional information contained in the morphology of the ECG. However, when we compare the accuracy performance reported by Fujita [192], with the ResNet accuracy, we find that our performance is just 0.49% lower. The small performance benefit might not justify the increased measurement

effort and significantly higher data rate of ECG signals when compared to RR interval signals. The increased measurement effort results in the fact that ECG monitors require expert instrumentation, for example the sensors must be attached by a specialist nurse. In contrast, RR intervals can be measured with sensors that were placed by patients. State-of-the-art ECG sensors deliver 250 samples per second. However, the heart beats around once a second, generating about one RR interval value per second. The fact that RR interval signals have a 250 times lower data rate, when compared to ECG signals, leads to significant cost savings when it comes to communication, storage, and processing. RR interval-based arrhythmia detection becomes even more important when we move away from the electrical activity of the human heart and consider RR intervals extracted from pulse signals [38]. Pulse sensors are less expensive and more readily attached when compared to RR interval sensors that measure the electrical activity of the human heart [65]. Therefore, pulse sensors can be used in wearable devices, such as smart watches. Coupled with the low data rate of RR interval signals, wearable technology may facilitate low barrier and low-cost arrhythmia detection systems. Such systems are governed by the laws of big data, where individual beat classifications become less significant when compared to accumulated evidence. Furthermore, big data helps to diversify and to improve classification results. This may lead to a better understanding and detection of early-stage arrhythmia.

Table 14: Selected arrhythmia detection studies using RR intervals and ECG from varied databases.

Authors list	**Classifier**	**Data**			**Performance**		
		Signal type	**DB**	**Rhythm**	**ACC%**	**SPE%**	**SEN%**
Our proposed	Bidirectional-LSTM	RR	MIT-AFIB	AF and normal	98.51	98.67	98.32

methods: [14], [17]	Detrending, ResNet		ECG-DB	NSR, AFIB, AFL	99.98	100.00	99.94
Ivanovic et al., 2019 [125]	CNN, LSTM	RR	Hospital	NSR, AFIB AFL	88	–	87.09
Fujita et al., 2019 [192]	CNN with normalisation	ECG	AFDB, MIT-DB, VFDB	NSR, AFIB AFL, VFIB	98.45	99.87	99.27
Faust et al., 2018 [14]	LSTM	RR	AFDB	NSR, AFIB	98.32	98.32	98.51
Acharya et al., 2017 [193]	CNN with Z-score	ECG	AFDB, MITDB, VFDB	AFIB, AFL, VFIB, NSR	92.50	98.09	93.13
Henzel et al., 2017 [194]	Statistical features with generalized Linear Model	RR	AFDB	AFIB NSR	93	95	90

Desai et al., 2016 [195]	RQA with Decision Tree, Random Forest, Rotation Forest	ECG	AFDB, MITDB, VFDB	AFIB, AFL, VFIB, NSR	98.37		
Acharya et al., 2016 [196]	Thirteen nonlinear features with ANOVA with KNN and DT	ECG	AFDB, MITDB,VFDB	AFIB, AFL, VFIB, NSR	97.78	99.76	98.82
Hamed et al., 2016 [197]	Thirteen nonlinear features with ANOVA with KNN and DT	ECG	AFDB	AFIB, AFL, NSR	98.43	96.89	98.96
Xia et al., 2018 [198]	DWT, PCA and SVM	ECG	AFDB	AFIB	98.63	98.79	97.87
Petrenas et al., 2015 [199]	STFT/SWT with CNN	RR	NSRDB, AFDB	AFIB NSR		98.3	97.1
Zhou et al., 2014 [200]	Median filter with threshold	RR	LTAFDB, AFDB, NSRDB	AFIB NSR	96.05	95.07	96.72

Muthuchudar et al., 2013 [201]	Median filter & Shannon entropy with threshold UWT NN	ECG	AFDB	AFIB, VFIB, NSR	96		
Yuan et al., 2016 [202]	Unsupervised autoencoder NN SoftMax regression	ECG	AFDB, NSRDB, ltdb, hospital	AFIB	98.18	98.22	98.11
Dinakarrao et al., 2018[203]	Daubechies-6 with counters Anomaly detector	ECG	MITDB	AFIB, VFIB	99.19	98.25	78.70
Salem et al., 2018 [204]	Spectrogram with CNN	ECG	AFDB NSRDB VFDB EDB	AFIB, AFL VFIB NSR	97.23		

3.7 Summary

This chapter described design, implementation, and testing of arrythmia detection systems. The chapter started by describing the design approach which mainly based on offline and online system. From the offline system, we extracted the knowledge from small datasets, by generating an AI model. In the online system this model is applied to larger datasets. The proposed algorithms for AF detection and arrhythmia detection were discussed in detail in Section 3.6. Validation of a mature deep leaning system is covered in detail in Chapter 4 which involves testing the LSTM algorithm for AF detection with more and more databases to establish a feasibility.

Chapter 4 Validating the robustness of a mature deep learning system

4.1 Introduction

This study describes the validation of a mature deep learning model for IoT based healthcare applications. The original deep learning model was established to detect AF epoxides using RR intervals. The initial LSTM model was trained with 20 subjects, collected from the publicly available AFDB database, known as PhysioNet. This model achieved an AF detection accuracy of 98.51% with 10-fold cross-validation. In this chapter, we describe how we validated the initial outcomes by testing the developed DL model with more unknown datasets. To be specific, we used these databases for testing the model with a so-called blind-fold validation method. The blind fold validation approach has been done independently for each database which they have its own performance measures and results representation. Physicians can compare the classification results of the machine decision with their diagnosis. These results showed that the LSTM model can extract the feature maps from unseen data which led to accurate detection of AF events. Testing the model with blindfold validation we contravened a well-known design rule for learning systems which states that more data should be used for training rather testing. Therefore, we have established that the tested DL model is fit for practical applications. To be specific, we found that the DL model can apply knowledge which was extracted from a small training data set to a HR trace from a patient.

4.2 Background on intelligent internet of things

Intelligent Internet of Things (IIoTs) transmit the measurement of heart data to a central storage location for centralised decision making [205]. In the clinical domain, these

measurement data are usually indicated to physiological signals, such as HR signals [66]. Figure 23 demonstrates a use case scenario for a medical IIoTs. Monitoring session starts with session initialization by attaching the HR patch to patients, which presented in the figure as a black oval with a red dot. That real-time patch communicates the HR signal to a tablet or a smartphone through Bluetooth Low Energy (BLE), which then transmits the data from the smartphone to the cloud server via IoT protocol. The LSTM based deep learning model, was used for validation in this study, can analyse the transmitted HR signals in real-time and act as decision support tool. Once AF event is detected, a cardiologist will be notified. The cardiologist can examine the HR trace and reach a diagnosis. That diagnosis can be interpreted to the patients in form of a simple traffic light scheme, as shown in the Figure 23.

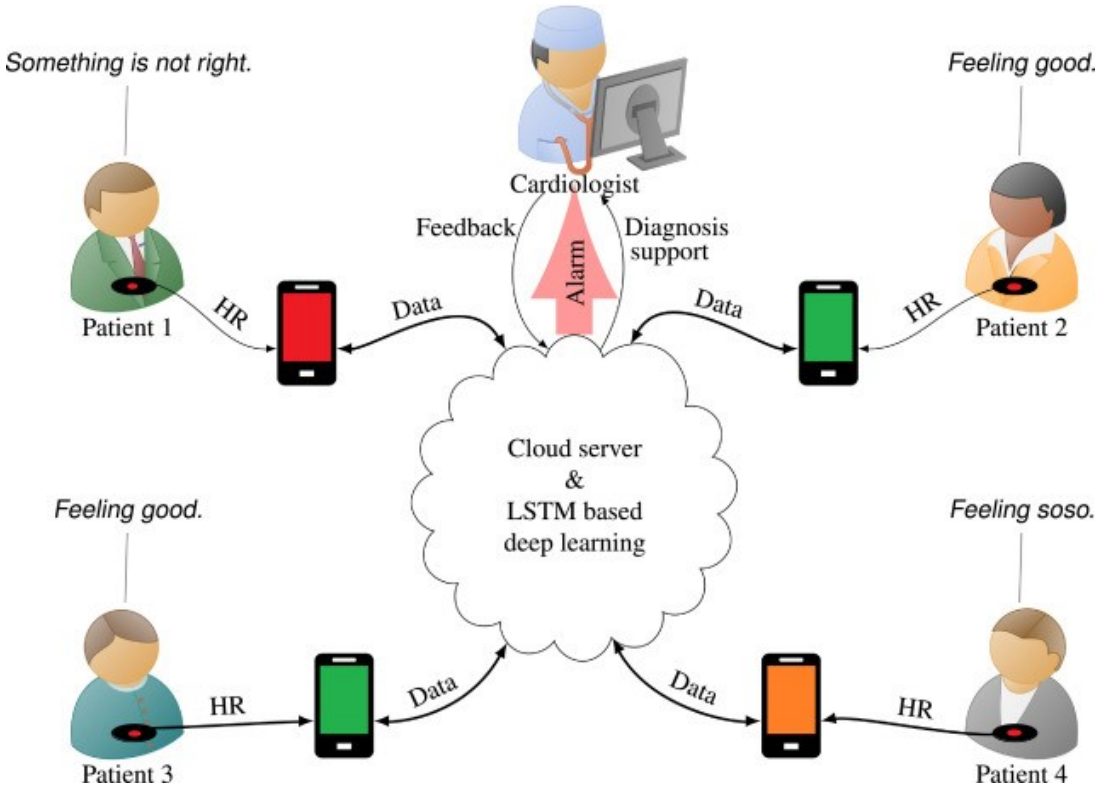


Figure 23: The diagram presents the use case technology of an IoT based decision support system. The colour on the smartphone changes in line with diagnosis outcomes. Red screen refers to serious event detected and requires an urgent intervention, orange indicate to cautions, and green no need for attention [18].

4.3 Methods

In Chapter 3 we outlined the implementation of a LSTM model based on DL to detect the AF episodes in RR interval signals [14]. The model provides the required intelligence analytic tool for the state-of-the-art IoT based diagnosis support systems. The current study setup tests the same model with more and more distinct data acquired from different database sources. Figure 24 shows a block diagram of the study setup. The upper phase of the graph illustrates the initial study setup. The LSTM based DL system was trained and validated with Labelled RR interval signal data from 20 subjects obtained from PhysioNet' AFDB. The model was tested with completely hold-out 3 subjects. Training the deep learning network means generating good weight values of the individual neurons. The weight vectors generated from the initial study were applied in the validation setup. However, more and more varied data were used for the blind-fold validation. The second phase of the figure shows the sources of different databases used for the validation arrangement. The following sections detail the validation setup by introducing the data sets and the processing steps in more detail.

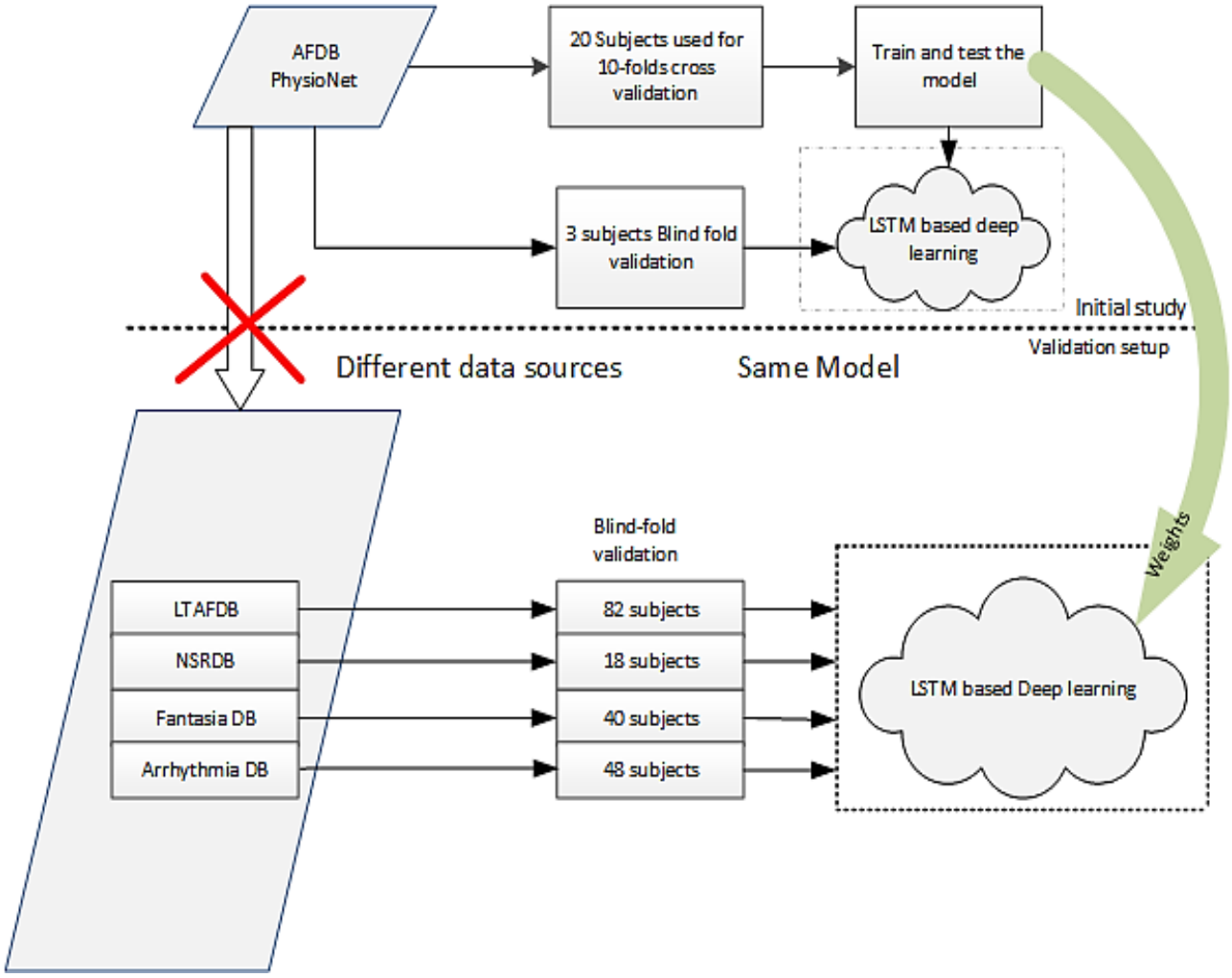


Figure 24: Overview diagrame of the validation setup.

4.3.1 Data used for validation setup

The data acquisition obtained for the initial study from MIT-BIH AFDB, which is available on PhysioNet [45], [46]. This database involves 23 subjects for 10 hours recordings with ECG. These recordings were labelled with so-called rhythm annotation files, the content of which was arranged manually. There are different types of rhythm annotations found in these files, such as AFIB (atrial fibrillation), AFL referring to (Atrial Flutter), J: AV junctional rhythm), and N used to indicate al other rhythms. Furthermore, the database contains beat annotation files that were used to extract the HR signals.

For the blindfold validation, we used the data from multiple database sources. These databases employed for validation setup, including Long-Term Atrial Fibrillation Database (LTAFDB) with 82 subjects [30], Normal Sinus Rhythm Database (NSRDB)

with 18 subjects, Fantasia database has 40 subjects and Arrhythmia database involved 48 subjects. A brief description of LTAFDB, all subjects in that database were identified with either paroxysmal or sustained AF. Each dataset is constituted of two ECG signals where each signal has a recording duration of 24 to 25 hours with the sample frequency of 128 HZ. In addition, the datasets also contain rhythm and beat annotations, which were prepared by experienced cardiologists. Furthermore, the R Peaks are labelled, and the RR interval sequence was extracted based on these labels. The RR interval sequence indicates to either AF or non-AF based on rhythm annotations.

However, the data were collected from the NSRDB [46] investigated by experienced cardiologists which showed that clean datasets with no signs of arrhythmias or AF. The signals were measured from 5 males, aged between 26 to 45, and 13 females aged from 20 to 50. This dataset is composed from 18 long-term ECG recordings sampled at 250 Hz. Similarly, all datasets from Fantasia database do not contain AF or arrhythmia episodes, all datasets are normal according to the clinical information published on the PhysioNet website. That database includes 20 young participants with age group between 21-34 years old and other 20 participants from elderly group (68-85 years old). Each subgroup of subjects includes equal numbers of men and women. The monitoring duration underwent continuously for 2 hours of the healthy subjects with ECG recordings. The ECG signal sampled with 250 HZ. Each heartbeat was labelled using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection. In contrast, the MIT-BIH Arrhythmia Database contains 48 subjects monitored with two-leads ambulatory ECG recordings for 24 hours. ECG measurements were collected from a mixed population of inpatients wards around 60% and outpatients about 40% at Boston's Beth Israel Hospital. These recordings were digitized with sampling frequency 360 HZ per second.

: Provide the recordings parameters that were used to collect the data from five databases.

Table 15: Comparison between databases sources for validation setup study.

Database category	Recording duration	Sampling frequency	Subjects Number	Voltage range	ADC Resolution
AFDB	10h	250 Hz	3	± 10 mV	12-bit
LTAFDB	24h	128 Hz	82	± 20 mV	12-bit

NSRDB	24-25h	250 Hz	18	± 10 mV	12-bit
Fantasia DB	2h	250 Hz	40		
Arrythmia DB	24h	360 Hz	48	± 10 mV	11-bit

The fact that the sampling frequencies were varied among these databases. The pre-processing of the data in the validation setup follows closely the initial study. A sliding window was used to partition the HR signals into overlapping sequences of 100 beats. Each sequence of 100 beats was labelled as AF if one of the beats, within the sequence was labelled AF, similarly repeated with arrhythmia, all other sequences were labelled non-AF, indicating normal or other cardiac disease. The overall data from 188 subjects were used with the blindfold validation strategy to evaluate the performance of the robust deep learning system. This means the proposed methods can be used to generalize not only unknown data, but also the unknown patients as well.

4.3.2 Performance analysis

The performance measures can be established by using both confusion matrix and ROC curve. The confusion matrix is also known as error matrix which summarise the prediction outcomes of a classification problem [206]. The arrangement of such a matrix can visualise the performance of a decision-making algorithm. To be specific, the matrix comprises of two rows and two columns that represent the classes of actual and predictive analysis. Therefore, these performance measures report the number of TP, TN, FP, FN. TP in most medical test results refer to a positive case (patient correctly identified with a disease) while TN a negative case (subject correctly identified without having a disease). The ROC curve is a technique used to evaluate the diagnostic ac- curacy of tests in modern medicine [172]. It is widely used to demonstrate how well a diagnostic model can differentiate between the presence and absence of the disease and works equally well with data sets that exhibit class imbalance. The ROC represents the TP rate (sensitivity) plotted against the FP rate (1-specificity) for various cut-off points [207]. Each point on the ROC graph indicates the sensitivity/specificity corresponding to a specific decision threshold.

The Area Under Curve AUC) is a summary metric indicating the discriminatory power of a classifier.

4.3.3 Initial study setup

In the initial study, we designed a developed deep RNN with LSTM model to detect AF based on RR interval [14]. The resulting data blocks from the pre-processing method were directly fed into the model without doing any feature engineering. Table 16 provides the 10-folds cross validation results achieved for the initial model that was trained with data of 20 subjects from AFDB database.

Table 16: 10-fold cross-validation performance measure outcome established during the initial training and testing of the LSTM based deep learning model.

TN	FP	FN	TP	Accuracy	AUC
523,241	7040	7407	430,615	98.51%	0.9986

4.4 Validation Results

This section forms the blind-fold validation results of the LSTM based DL model which was initially established by Faust et.al [14]. The test results for LTAFDB datasets based on processing RR intervals signals for each individual patient's file with overall 82 subjects. The radar plot, shown in Figure 25, provides a graphical representation of the model performance in terms of accuracy, precision, recall and F1 score. The labels placed around the radar plot, correspond to the subject ID as found in the LTAFDB. These labels were arranged in order in terms of having the highest accuracy. The signal with patient ID 10, reflects a position at 12 o'clock, has the greatest accuracy. The accuracy performance of the classification model reduces its values when it moves clockwise, for example, the accuracy values were fluctuating from label 12 to label 22. The model achieved the lowest accuracy of just under 70% for the HR signal from patient 22. Therefore, we have evaluated the classifier results with the performance measures such as the confusion matrix and ROC curve to validate the classification of AF and normal cases in the HR sequences processed from 82 subjects. Figure 26 shows the results

obtained using the new test data: (a) confusion matrix used for blindfold validation from 82 subject's data, (b) ROC curve of the classification model and diagnostic quality measures.

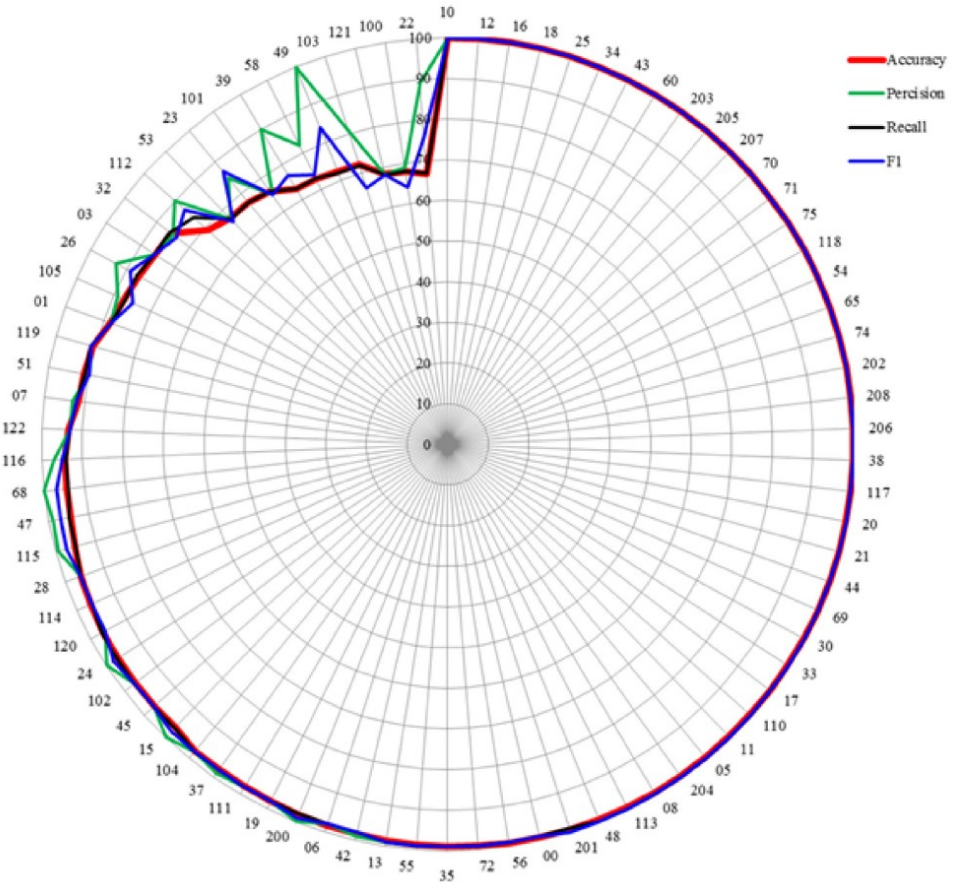
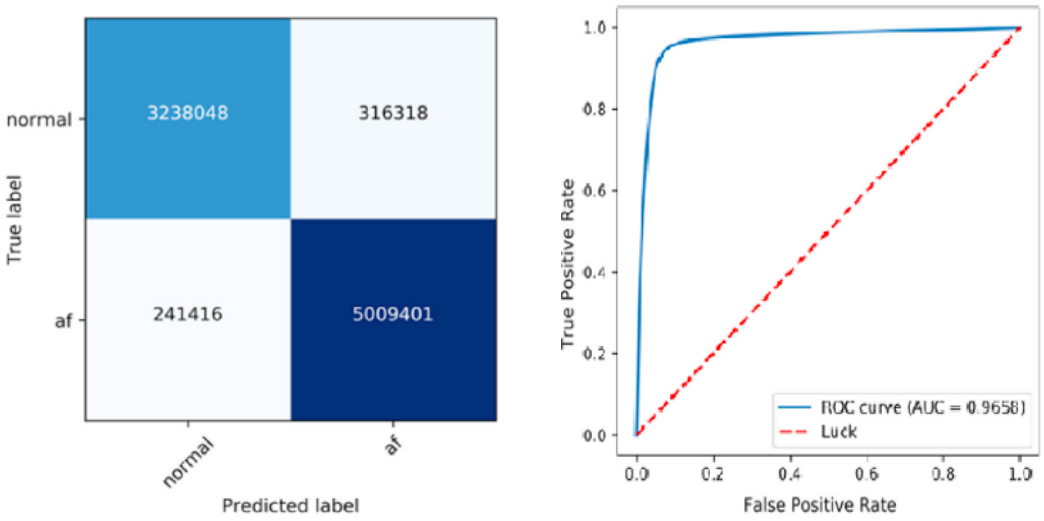


Figure 25: Illustration of model performance for testing each individual subject measurements from LTAFDB achieved through blindfold validation .



(a) (b)

Figure 26: Results obtained using the new test data: (a) confusion matrix by blindfold validation, (b) ROC of the model.

It can be observed from the results in Table 17 that the LSTM based DL model achieved 94% of the overall classification accuracy on the LTAFDB blindfold validation datasets - classifying 95% of normal and 95% of HR sequences showing signs of AF correctly, the overall accuracy detection in the NSRDB 97% as well as 97.57% accuracy for Fantasia database, but with arrhythmia database the model achieved 88% of accuracy despite of the model has not been trained with this datasets. For the ROC curve plot, the DL classifier achieves 96.58% of an AUC, which is close to a perfect score of 1, as such, the classifier discriminates well between the presence and the absence of the AF episodes. The classifiers values threshold is set between 0 and 1 where zeros represent true negative values and 1.0 refers to true positive. The closer curve to 1 in ROC plot, the higher overall diagnostic accuracy of TPR against FPR is achieved. Figure 26b shows the ROC curve indicating the diagnostic performance of AF. Table 17 provides the blindfold validation results for 3 subjects of AFDB, as documented during the initial study, and 82 subjects from LTAFDB, 18 subjects from NSRDB, 40 subjects of Fantasia database and 48 subjects of Arrhythmia database. The overall subjects that have been validated the model with blindfolds validation model are 188 subjects. From the matrix parameters TN, FP, FN and TN, we can calculate the accuracy, sensitivity and specificity. AUC curve is plotted once AF beats are detected represented by TPR.

Table 17: Blindfold validation obtained from the LSTM deep learning model for five databases.

Databases	TN	FP	FN	TP	Total beats No	Accuracy %	Sensitivity %	Specificity %	AUC
AFDB	65,699	255	116	91,888	157,958	99.77	100	100	1
LTAFDB	3238048	316318	241416	5009401	8,808,183	94	95	95	96.58
NSRDB	1617117	45936	0	0	1,663,053	97	97	100	—
Fantasia DB	269343	6630	0	0	275,973	97.57	98	100	—
Arrhythmia DB	77941	11365	1232	12067	102,605	88	88	92	

The tested model achieved with blindfold validation for all databases above than 80% which reflects excellent results according to the state-of-art. Figure 27 provides the confusion matrix outcomes for NSRDB. Figure 28 shows the performance measure results of the classification model for Fantasia database. Figure 29 depict confusion matrix results for arrhythmia database.

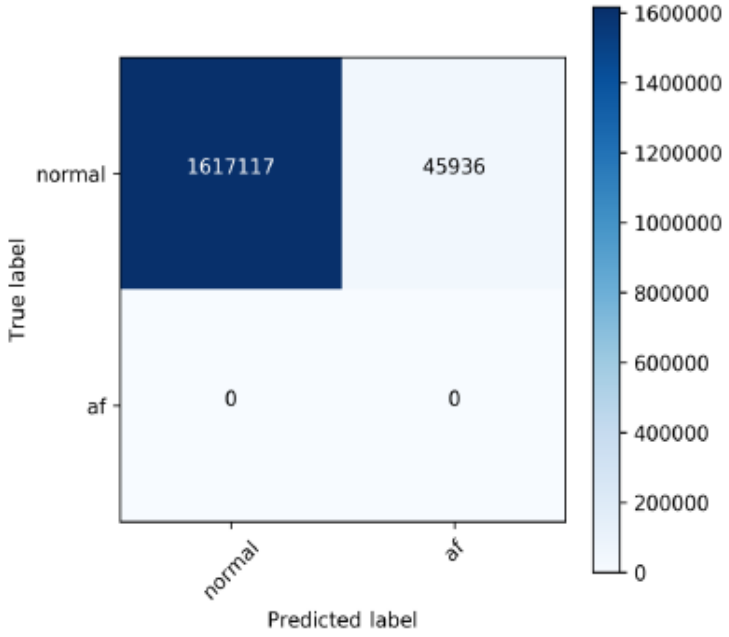


Figure 27: Confusion matrix for NSRDB.

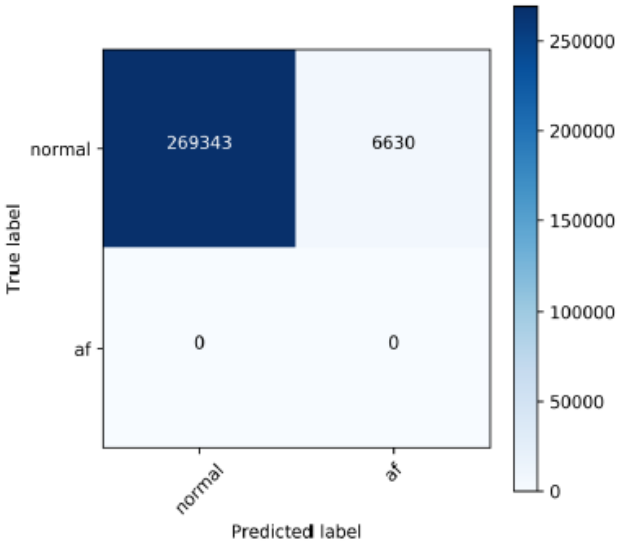


Figure 28: Confusion matrix outcome for Fantasia.

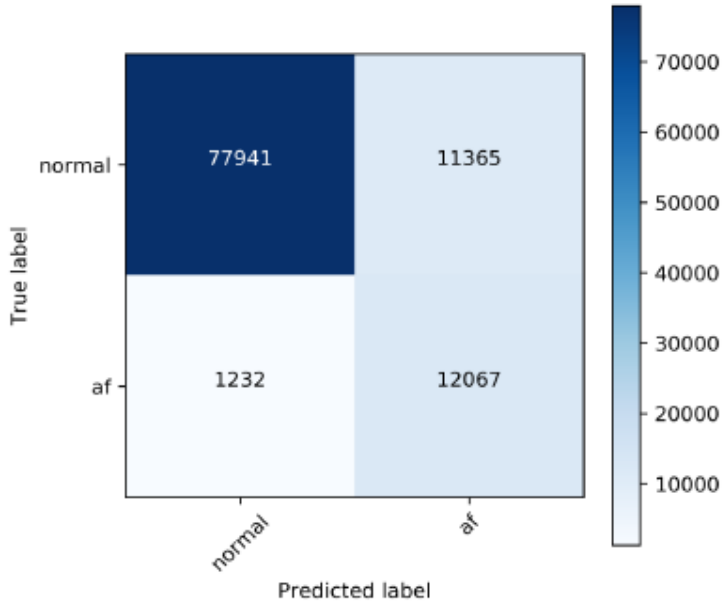


Figure 29: Confusion matrix for Arrhythmia database.

4.5 Discussion

This validation study demonstrates that the mature DL model can mimic human decision-making, when it comes to extracting relevant knowledge from a labelled data set. Training the DL algorithm establishes the distinct model with the knowledge initiation which transpires during the learning of a cardiologist. The blindfold validation, conducted in this study, reflects the clinical practice where the decision-making system is based on physiological signals from a wide range of patients. Validating the DL with unknown data that was collected with a different measurement setup that builds the trust in AF detection model. Moreover, the results might be transferable to other areas of Computer-Aided Diagnosis and beyond. Blindfold validation should become a standard method to evaluate deep learning and other decision support algorithms.

The fact that the DL system was trained with HR signals has particular importance for transferring the healthcare to patient home environment. HR signals capture the RR intervals of the human heart [208]. The R peak is the most prominent signal feature when the electrical activity of the human heart is recorded. That prominence is reflected in the high Signal-to-Noise Ratio (SNR). Hence, recording the RR intervals are significant to

noise. Therefore, the measurement setup is simple, when compared to other physiological signals, such as ECG [209]. Another benefit of HR is that the time from one R peak to the next can be encoded with a 2-bytes. Hence, a digital HR signal consists of approximately 2 bytes a second. In contrast, to sample the complete electrical activity of the human heart requires around 256 samples a second, each of which is encoded with 2 bytes. Therefore, ECG signals have a 256 times higher data rate. The lower data rate has practical benefits, in terms of data communication, storage and processing. For instance, HR signals can be communicated via BLE whereas ECG signals require a broader wireless channel, such as radio frequency. Applying BLE instead of radio frequency has beneficial implications for battery lifespan powered sensors. The low data rate and patient led signal acquisition makes HR signals is an ideal choice for IoT based healthcare applications. For such applications the HR data transmits from the point of measurement (patient) to a central cloud service over communication infrastructure. Having the data at a central location has several advantages. DL can be used to detect AF in real time. That is a significant advantage over ECG measurements with Holter monitors, because in the clinical practice Holter data can only be analysed after the measurement period is completed with a month timeframe. Validating the deep learning model for AF detection has paved the way for an IoT based AF diagnosis support tool. The knowledge extracted from a small, labelled dataset can be used to provide real-time decision support. To reach the diagnosis, the deep learning outcome should be validated by a cardiologist. The basis for this validation can be the HR data which has been flagged as showing the subtle waveform alterations caused by AF.

4.6 Summary

This chapter described the validation of the mature DL model, which was based on an initial design described in Chapter 3. The validation setup was introduced in the method section. Our DL model understands the HR signals in such way that it can discriminate AF episodes from non-AF affected signals. This classification method differs from the classical machine learning approach that is based on manual feature extraction.

To be specific, the classical approach discriminates AF from non-AF signals relied on parameters. It is possible that these parameters vary when more and more varied data are processed. As such, for the studies based on classical methods., it is common practice to

state this as a limitation, for example, more and more various data is required as well as retraining the machine learning model to enhance the diagnostic quality. In this study, we tested the model with five different databases without retraining the DL model. Varied data from 188 subjects were used for this blindfold validation. The accuracy of the classification model varies from 88% to 99.77% for the five different databases. Hence, these outcomes indicate that the DL model has less limitations when compared to classical machine learning methods. We demonstrated that the knowledge was extracted from small training dataset which can be applied to a larger and more diverse validation dataset. Our findings have significant implications for practical diagnostic support, because applying knowledge obtained during the limited training period is exactly what an experienced cardiologist achieves in the practice. Therefore, a practical IoT service based clinical decision support system must use the extracted knowledge during the training phase, to samples, for example, patient data, from a different dataset.

Chapter 5 Intelligent decision support for Atrial Fibrillation detection through human verification

5.1 Introduction

In this chapter, we present hybrid decision support for long term AF monitoring to prevent stroke. Hybrid decision support means that an experienced cardiologist works cooperatively with machine algorithms to reach an accurate diagnosis. The human expert can verify the machine decisions through visual data inspection together with knowledge obtained from interacting with patients. This links with stroke prevention because patients with AF have a five times higher stroke risk. Early diagnosis, which results in adequate AF treatment, can reduce the risk of stroke by approximate 66% and thus it prevents stroke occurrence. The monitoring service can be achieved through measuring the HR signals in real-time. These signals are transmitted and stored with IoT technology. The DL algorithm automatically processes the estimated AF probability. from a technological perspective, we propose four different services to healthcare providers:

1. Accessible universal location to patient data such as IoT Cloud based.
2. Automated AF detection service supported with patient alert.
3. Cardiologist support tool.
4. Feedback channels.

These four services establish an environment where cardiologists can interact symbiotically with machine algorithms to generate and discuss a high-quality AF diagnosis.

5.2 Materials and methods

The principles of service design have been used to analyse and structure the AF detection problem [22]. First, we contemplated the demands of all stakeholders impacted by the proposed service [210]. This leads to a better understanding for the requirements on the AF detection service. The next step was to interpret the requirements and form a system specification. A prototype was implemented to test the validity of this specification, which involved hybrid decision support. The following sections provide further details on the individual steps that led to the AF detection service.

5.3 Demand definition

To establish a demand definition, it is essential to present the link between AF detection in relation to stroke prevention in further detail. For ischemic stroke, a lack of oxygen causes a blockage of the arteries that supply oxygen-rich blood to the brain. In most cases, that cut-off is due to plaque debris in the bloodstream. The heart pumps blood, and the debris might travel to the brain through arteries with a shrinking diameter. If the blood vessel diameter is not large enough for the debris to pass, it will block the artery, and that will stop the oxygen supply to the associated brain tissue [21]. The incidence of plaque debris is related to the fluid dynamics of the bloodstream, which is ruled by the HR variability. From that prospective, the first service design step was to identify the key stakeholders and their requirements. We found that there are four key stakeholders in the AF detection service, as described in Table . The only reason for constructing the service is the fact that AF presents in patients and its prevalence increases. Hence, this group subjects to early diagnosis when it comes to AF detection for stroke prevention.

Healthcare providers aim to identify the unmet need by creating an appropriate framework. That infrastructure requires investment based on the cost and the expected benefits for patients. From a medical perspective, cardiologists are key stakeholders. Their input is vital when it comes to determining the benefits of a proposed service. Hence, innovators who design the proposed service must consider the needs of cardiologists to generate a successful service. However, the effort spent in addressing

these needs must correspond with the required profitability of a practical problem solution.

Table 18: Stakeholders' beneficiaries from AF detection service based on hybrid decision support.

Stakeholders	Identify unmet needs and wants
Patients	Reduction in the cases of stroke risk, less clinical admissions, flexibility in mobility, safety
Cardiologist	Reduced the burden on the medics, enhanced the medical outcomes, high diagnostic quality, increased safety.
Healthcare givers	High efficacy and quality, cost effectiveness, improved productivity, and outcomes
Innovators of stroke risk monitoring service	Enhanced the outcome, profitability

5.4 Requirements evaluation

From the demand classification, we determined the required service elements and the associated value proposition. provides an overview for both service requirements and value proposition. Cost effectiveness and decision support quality are the most significant requirements because they determine if the proposed service element can be used to enhance and extend the current infrastructure. All remaining service requirements are practical requirements that answer the question: What service do we need to create and develop? An alert message should be only sent once AF event is present. This requirement leads to the information improvement and management nature of the service. The functional design highlights the main requirement for workload reduction on cardiologist and healthcare provider [21]. To be specific, the task to generate a suspicion that AF is present, has transferred from humans to machines. The AF detection service is a diagnosis support tool, which means all diagnostic decisions align with cardiologist verification. To

underpin that decision, the AF detection service must give support evidence that indicates to the suspicion of a disease is present. This step can help us to ensure having excellent evaluation for quality and safety aspect of the diagnosis tool. Furthermore, that service can provide additional evidence even if there is no alarm message raised, such as detecting normal rhythms. As such, it assists during the root cause analysis and to enhance the service. For instance, the proposed service misdiagnosed the AF episodes in a specific patient. Having the capability to retrieve the evidence in the shape of raw signals might help to investigate what caused that error. That root cause analysis leads to first step to develop the algorithms specification that provide hybrid decision support. Moreover, the proposed service should also provide a feedback channel that allows the service provider to contact with the patient. That channel can be used to circulate diagnosis results and send as messages that help with patient compliance.

Table 19: Service requirement based on the proposed value.

Service Requirement	Value proposal
1. Cost effectiveness and decision support quality.	More e-health facilities to help a greater number of patients
2. Notification with an alert once AF is present	Detecting and transmitting a suspicion alarm that AF is present in real time.
3. Show the evidence for raising the alert	Establishing an overview of the estimated AF probability; this can be used to review the DL results that established a suspicion and triggered an alarm message
4. Enable a time interval of interest; subsequently, the corresponding HR trace can be processed	Download the HR trace that matches to the selected time interval of interest, and calculate features from that HR trace
5. Provide a feedback channel to the patient	Act on the diagnosis by providing appropriate and timely feedback to the patient; act on meta data, such as data

	stream interruptions, to ensure patient compliance

To get a better understanding of the functional requirements of the proposed service, we visualized the service requirements as a sequence of related actions; see

Figure 30. These actions were arranged along a timeline to establish a relevant structure that orders the individual events. The timeline begins with the healthcare provider, represented by a qualified nurse which can register a new patient with the AF detection service. Once registered, the patient session initiates capturing the HR measurements, which are displayed via either via smartphones or an android tablet to a cloud server [211]. In the cloud server, the data are stored and processed by deep learning model [212]. When the analysis outcomes indicate that estimated AF periods were detected in the HR data, the cloud logic will notify through an alarm message to the appointed physician. That message is sent within 5 min of the AF event. In relation to the alarm message, the cardiologist will review the evidence contained in the HR trace and fuse this information with further knowledge and experience concerning the patient, to reach a diagnosis. If the diagnosis is negative, i.e., the physician found that AF does not present in a patient record, monitoring for AF continues. Once AF is detected, treatment can be instructed. The treatment efficiency can now be monitored with the same system setup. If AF is diagnosed again, treatment can be changed, and the monitoring continues. The next section details the functional specification that was created to meet the system requirements.

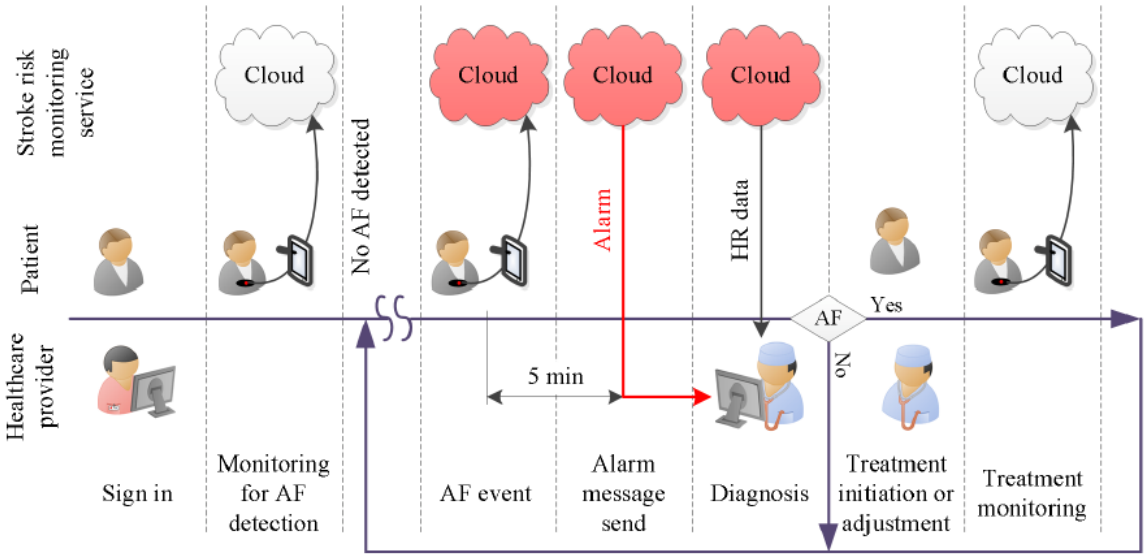


Figure 30: Provides the timeline of all activities for AF detection service over requested duration by an experienced cardiologist [19].

5.5 Specification

The specification determines how the AF detection service is constructed. This is achieved by improving the requirements and thereby increasing both the clarity and rigour of the documentation. The AF monitoring is launched by identifying a disease related variation in HR signals. These signals are ideal to record, cost efficient to transmit, as well as resource efficient to store and analysis. Hence, this refinement tackles the cost efficacy requirement for the proposed service [22]. The requirement for using HR signals provides the foundation for the functional specification. We structured the functional specification into six service components. The following list details how to build these service components:

1. Smartphone activation app: This service allows a patient's phone to activate and sign up an account with the healthcare provider. At the start of the service subscription, the healthcare provider records the patient with the database on a cloud server. The unique account contains patient information. The necessary fields are: patient ID, appointed cardiologist, service start date, service end date.

The registration will provide the cloud server login key. This login key is used for both user authentication and data acquisition setup.

2. Cloud server storage: The patient's HR data and the deep learning classification results are stored in the cloud server. This service allows the authorized users to retrieve the data anytime and anywhere.
3. Real-time HR monitoring service: The patient wears a heart patch sensor that records both ECG and HR signal. The sensor collects both signals. However, for continuous monitoring, the HR signal is more ideal for communicating the data with cloud server. Hence, these real-time data are displayed on patient smartphones. The patient co-creates value by providing and integrating the data into the AF detection service.
4. Automated AF detection and alarm service: The deep learning algorithm processes patient real-time HR data and classifies the data as AF or non-AF. Once an AF sequence is detected, the system will send an alarm message to the assigned cardiologist. The DL algorithm generates the core value for the system.
5. Cardiologist diagnosis support service: The cardiologist support service includes algorithm support in the form of deep learning results and diagnosis support tools. It helps the professional medics to verify the deep learning outcomes and to reach a diagnosis. The value of this diagnosis is twofold. First and foremost, it aids to start treatment, which might enhance the outcomes for the patient. A secondary use for an established diagnosis arises when we preserve refining the deep learning algorithm. To be specific, a diagnosis becomes the ground truth, which can be used to continuously retrain the deep learning model. That sustained retraining has the potential to increase the detection quality of the algorithm.
6. Feedback and intervention service: Once the cardiologist has reached a diagnosis, the feedback service can be used to contact the result to the patient. Social media such as Twitter, email and personal phone calls can be used to provide feedback. Timely appropriate intervention can be carried out to boost the outcomes for patients. Another use for the feedback service is the dissemination of patient compliance messages. For example, through data analytics, it is likely to establish if there is a signal interruption. A compliance message over the feedback channel might help to re-establish the data flow.

5.6 Results

This section describes how we translated the specification into an implementation. The service modules were interpreted into software analysis, executed by standard machine architectures, and transferred over available infrastructure. Figure 31 demonstrates the data stream between different functional entities of the service. The arrangement of the data flow diagram indicates the central role of the cloud storage. The HealthCare app transmits the sensor data to the cloud storage. The cluster computing sources the data from the cloud server and, once the data are analysed, puts the result back. The processes are managed based on information from the real-time database. This information is particularly useful to establish the conditions when and to whom an alarm message is sent. This functionality is essential to create the hybrid decision support, which allows medical experts to work efficiently with smart machines. The following sections introduce the functional entities in more detail.

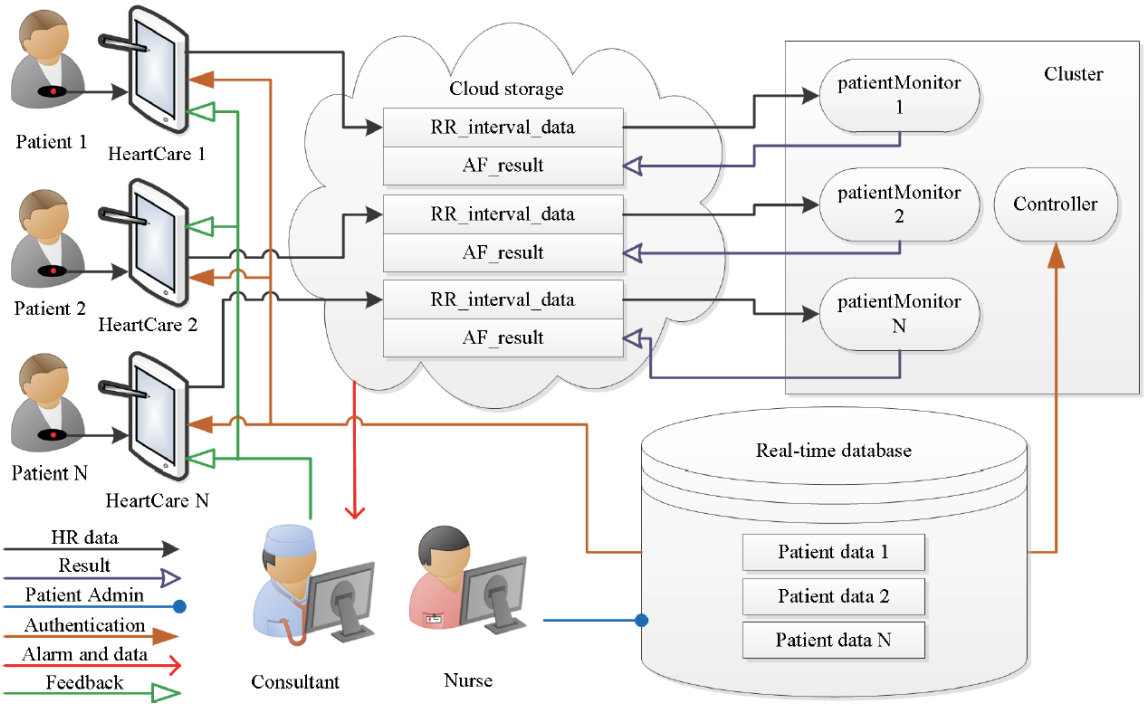


Figure 31: A diagram shows the main combinations of AF detection service for hybrid decision support [21].

5.6.1 Real-time database

The real-time database entries facilitate the patient information management. Throughout the initial registration process, a representative of the healthcare professionals generates a patient record. That record incorporates patient-personal information, such as the patient ID and password, as well as system-specific information like a cloud server key, which unlocks dedicated data channels. Once the initial registration is completed, a patient can apply the patient ID and password to login to the HeartCare app. This authentication ensures that the HR recordings are communicated to the patient-specific cloud server channels. The controller node in the cluster uses the patient records to set up the patient monitors, which processes the HR data in real time. The patient information is also used to manage the alarm message distribution.

5.6.2 HeartCare mobile app

For the pre-validation study in the clinical, we have designed an android application that receives, displays, and disseminates HR data. BLE technique establishes the connection between an android smart device and HR biosensor (Polar H10). As such, BLE transmits data in small packets which require less power when compared to normal Bluetooth packets [213]. In addition, there are some specific data types were added into Bluetooth standards body for supporting the Bluetooth version 4.0 specification: The Health Heart Rate Profile.

The AF detection service enables patient-led data acquisition. Figure 32 shows a snapshot of the HeartCare app login interface. The graph illustrates a sequence of HR trace recorded with a polar H10 sensor. In that state, the app transfers the HR data to the Thingspeak cloud server [214]. Each patient has provided with a unique API key. Once logged in, the HeartCare app transmits the HR data from the sensor to the patient-specific

RR_interval_data channel on the cloud server. Both the patient and authorized cardiologist can access the patient's data anywhere using the same API key.

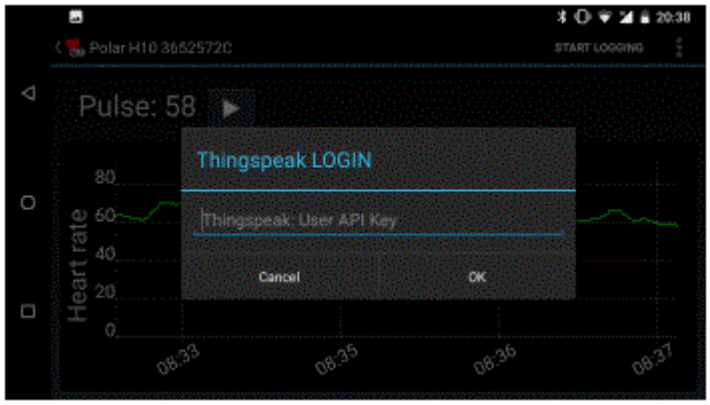


Figure 32: HeartCare app login interface.

5.6.3 Thingspeak account setup

Thingspeak is an open cloud data platform where the collected data can be stored retrieved, processed, and visualised. Creating new Thingspeak account is by using an email address and password for logging into the account. Channel establishments can be achieved automatically by using MATLAB program. Each channel has a capacity of storing and retrieving the information with maximum of 8 fields. Moreover, Users have unique channel ID for identifying the required channel. Two channels were created for storing and visualising the received data, channel-1 so called Murtadha HR data which performs delivering live HR measurements from wearable sensor via smart phone (android). In contrast, channel 2 was allocated for disseminating the deep learning results. In addition, Application Programming Interface (API) keys have the main role of accessing the channel. For instance, API keys is considered as the password to access the channel. API write key is used to update channel or logging data from the source. While the read key utilises for retrieving the data from the channel. API keys contain part of the transmitted link which formed as following:

URL(**"https://api.thingspeak.com/channels/476872/bulk_update.json"**);

GUI allows the users to interact with Thingspeak webpage, as shown in Figure 33.

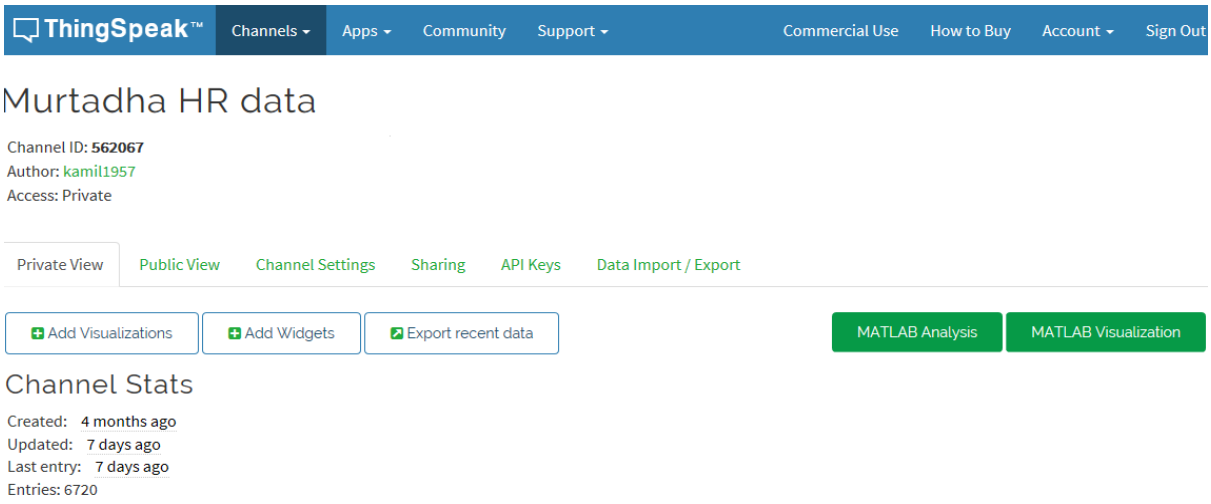


Figure 33: Visualises the GUI of Thingspeak application.

5.6.4 Cloud storage

Once the Thingspeak account has been created successfully, each patient account has two channels for cloud storage. The first channel, called RR_interval_data, store the HR measurements. The data is updated when the Heart-Care app sends new HR traces to the cloud server with 100 beats per minute. The second channel, called AF detection result, keeps the deep learning classification results. The result channel data is updated after the patient monitor produces a new result. Figure 34 shows a patient's HR data on the Thingspeak cloud server. Plotting RR intervals values and pulses duration which visualised as a graph varied with number of beats. The sequence of 1000 RR interval and each dot represent the position of beat measured in second from the heart rate signal. Figure 35 displaying the Pulse in beats per minute overtime and these measurements accomplished within 12 minutes.

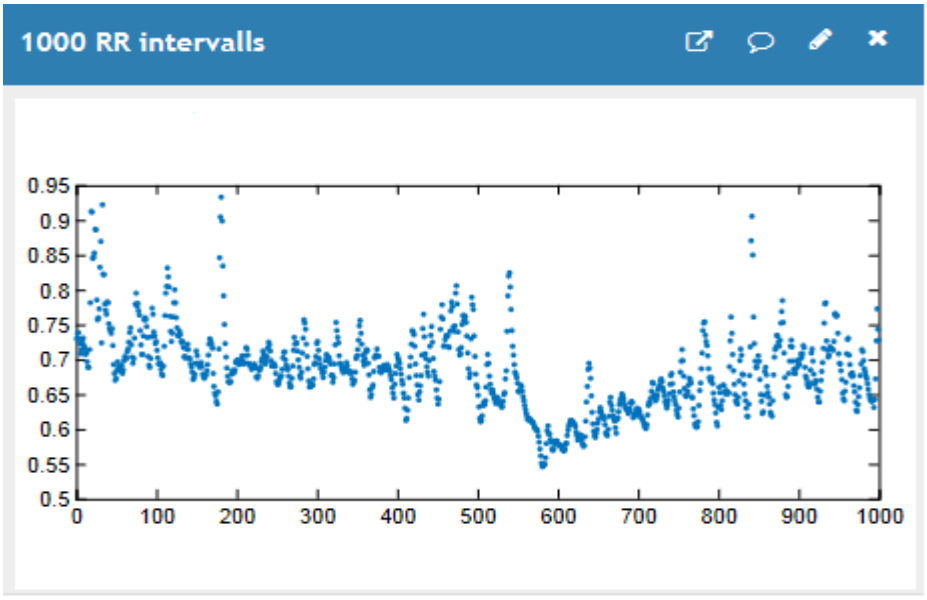


Figure 34: Heartbeats visualisation over RR_interval_data channel.

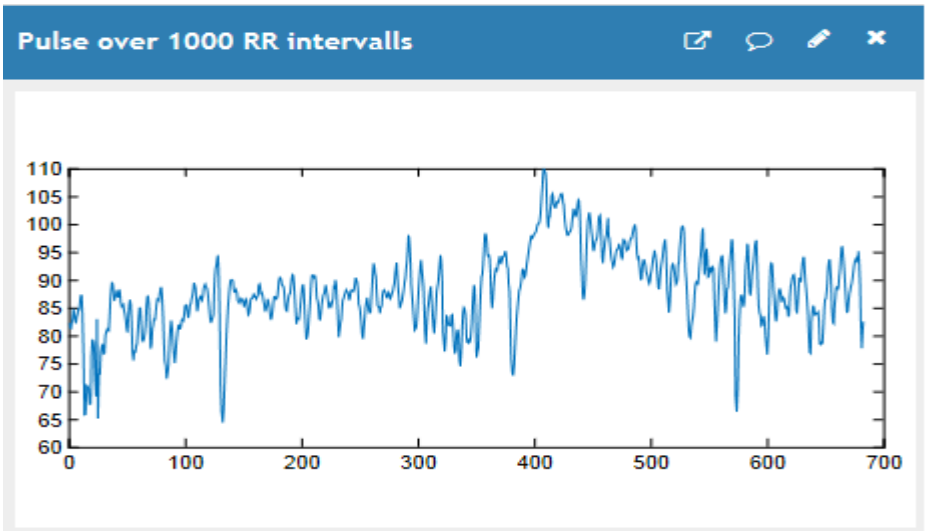


Figure 35: Shows pulse in beats over 1000 RR intervals.

The HR measurements transmit to channel-1 based Thingspeak that can be partitioned into five fields to solve the issue of timing. Uploading a block of 100 RR intervals sampled per second. The time delay between each block takes 15 second in accordance with time scheduling that specified for this setup. In addition to that, each field receives a packet of 20 beats per seconds which implies to be delivered to 5 fields at the same time. As a result, our channel-data distributed equally for each field, a packet of 20 beats

times 5 fields equal to block of 100 RR intervals. shows HR live measurements visualised on Thingspeak channel 1.

Figure 36: Channel-1 comprises fives fields for storing and visualising HR measurements in real-time.

The last step is to query these measurements from Thingspeak-cloud computing platform to deep learning system. Fetching the data based on using key-Read of channel-1 that facilitate processing the information. Deep learning system performs classifying the HR measurements with two classes of AF and Normal. Consequently, binary 0 sequences classify the pattern as un-diseased subject. Whilst binary 1 episode indicate to presence of AF detected and correctly identified the subject of having the disease. However, our

proposed system aims to monitor patients heart health for long-term in the real time. Short period of AF might not be subjected to the disease. Medical intervention is needed to validate the acquired test outcomes. Figure 37 indicates if the HR pattern is AF or Normal.

Figure 37: Deep learning classification results of HR measurements visualised on Thingspeak for decision support.

When an AF episode is detected by the deep learning algorithm, the Thingtweet feature based on cloud logic will send an alert to the assigned physician. Sending the alert message can be facilitated with a range of communication channels, such as email, Twitter, and instant messages. The message alerts the physician that a dangerous condition has occurred, i.e., AF event was detected. The physician decision support and diagnosis service can be used to review the available evidence and to reach a diagnosis.

5.6.5 Patient HR data processing in the cluster

The cluster performs a patient monitor procedure for each patient. That process network enables a real-time data processing [215]. To accomplish that task, each patient monitor compromises of three processes. The first process verifies if there is new HR data in the RR_interval_data channel on the cloud. The new data are shifted to the second node, which executes a deep learning model. The deep learning results are handed to the third process, which transmits them to the AF_detection_result channel on the cloud server. Processes 1 and 2 of the patient monitors handle the data exchange between the cluster and the cloud server. The main role for the patient monitor and certainly for the AF detection service is real-time HR analysis and visualisation. We recognized this

functionality with an LSTM-RNN of DL model. That model was initially trained, tested from AFDB database, and validated with unknown data from different database sources. For more information about the algorithm chapter 4 describe in detail the validation process of the LSTM based deep learning model. Figure 38 shows the design structure of the proposed deep learning system. The deep learning algorithm is constituted of three layers, known as bidirectional LSTM layer, global max pooling layer and fully connected layer. The simple structure leaves little space for design errors [216]. Furthermore, the implemented deep learning algorithm does not require feature engineering. Hence, there is no information reduction due to feature selection, which improves both the accuracy and robustness of the performance results [23].

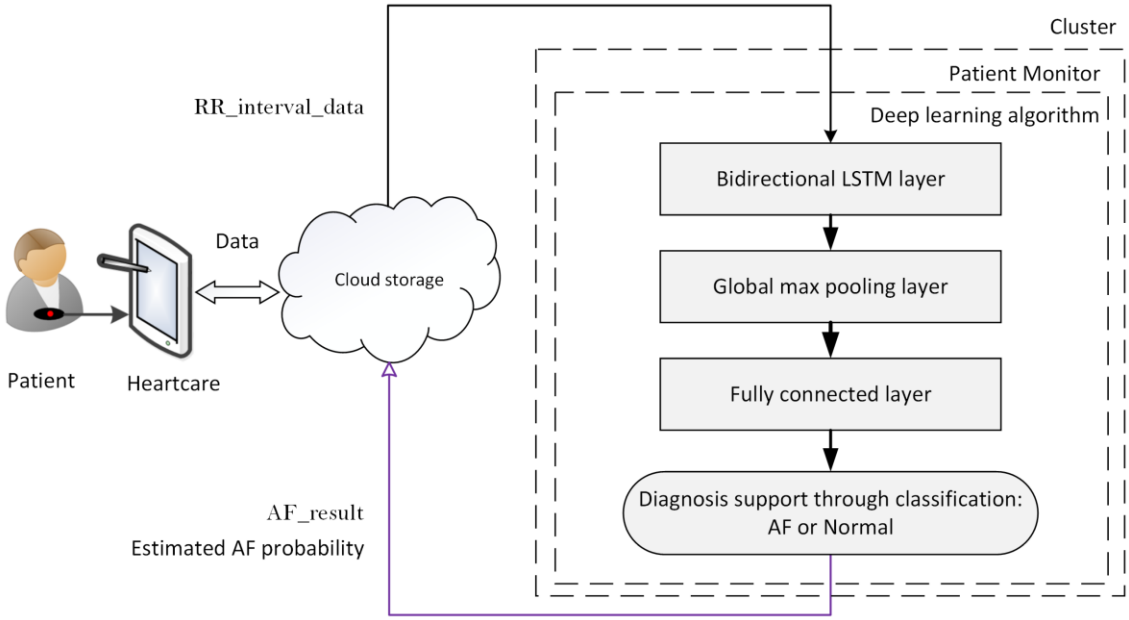


Figure 38: Flowchart shows data flow and classifcation processes system [21].

5.6.6 Cardiologist support tool

Cardiologist diagnosis support tool is an important service component. The implementation of this service module directs the data available on the cloud server. The service module creates an interface that allows a medical doctor to confirm the automated diagnosis results. In other words, the appointed stroke consultant can visually inspect the data and either accept or reject the decision achieved by the deep learning system. We implemented that service component by extending a current HR analysis and

visualization tool. The tool is called the Heart Rate Variability Analysis Software (HRVAS) program, originally developed by Ramshur [217], and published under the GNU public license (https://github.com/jramshur/HRVAS). We developed the program with the capability to download both HR data and the estimated AF probability from the cloud server. Having both, the raw data, and the DL results, allows a physician to review the HR measurements which contains available evidence either through visual inspection or through using the digital biomarkers. For instance, visual inspection might show fundamental data problems, such as all RR samples having the same value. Digital biomarkers can assist to confirm the deep learning decision result.

The ability to create independent human verification of the machine learning results is a significant factor for the proposed hybrid decision making process [24]. Figure 39 shows a snapshot of the developed HRVAS program. A drop-down menu allows the user to select the HR signal from a specific patient. The snapshot shows that the signal from Murtadha HR file was selected. As such, the signal for Murtadha record was streamed form HR sensor in real-time to cloud storage. The GUI of HRVAS program displays the deep learning outcomes in the upper graph on the left. Presenting the deep learning results gives an overview of the estimated AF probability, i.e., the examination through physician can determine at what time the patient had an increased AF probability. Based on that examination, the physician can select a region of interest and view the HR signal, which links to that region in the second window. The HR signals trace is coloured in accordance with the estimated AF probability.

Apart from visual signal inspection, the main purpose of the HRVAS program is to visualize digital biomarkers. The workflow unfolds as follows. The physician selects a region of interest on the estimated AF probability graph. Once the region is selected, the matching AF trace is displayed, and the digital biomarkers for this region are calculated. The biomarker values are displayed in the right part of the HRVAS GUI. The snapshot in Figure 39 shows time domain biomarkers. The HRVAS documentation provides more details on the available digital biomarkers [217]. These biomarkers are designed to help physicians during the process of validating the deep learning results and establishing a diagnosis.

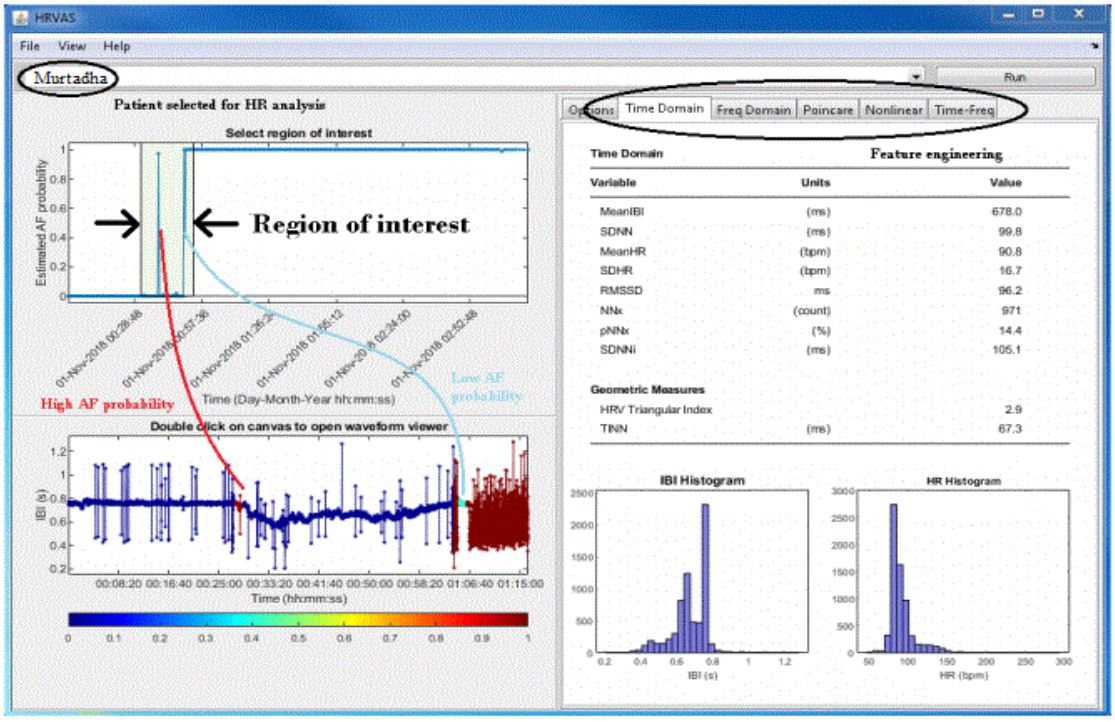


Figure 39: A snapshot of the developed HRVAS program.

5.6.7 Feedback and intervention

Once the physician has reached a diagnosis, the feedback and intervention service communicate with the concerned patient. Social media, email and personal phone calls can be used to provide feedback. One way to structure the feedback content is a simple traffic light system: green, all is well; orange, take predetermined precautionary action; red, see your physician immediately.

5.7 Discussion

A hybrid decision making process accomplishes an accurate diagnosis [24]. The hybrid process proposes three main benefits: 1) safety through human professional's verifications and balances, 2) reduced Stroke consultant workload, and 3) increased effectiveness, which facilitates real-time diagnosis. The hybrid decision making process is based on analysis results, which are directed to an independent first opinion on the data

[19]. Specifically, we propose a system where an AI algorithm involves in analysing the available data in real time, and a human expert only becomes involved if a suspicion is established. However, that design option is only valid if the AI algorithm is very sensitive when it comes to the detection of AF in HR signals. Another central requirement is cost efficiency. Furthermore, irrelevant decision making is not cost effective, because a human expert receives an alert frequently, and the machine decisions are routinely overruled. Such unnecessary involvement of human expertise would be ineffective, and indeed, it would be wasteful in terms of time spent rejecting the machine decision, which translates into additional cost for the healthcare provider. Hence, we require the decision support algorithm to have both high Specificity (SPE) and high Sensitivity (SEN)). In effect, that leads to a high Accuracy (ACC). Table 20 summarise research work for the automated detection of AF based on ECG and RR intervals measurement.

The performance measures, reported in the three columns on the right of the table, indicate two points:

1. There is no performance difference between studies based on ECG and RR inter signals.
2. Both the SEN and SPE values are very high. Therefore, these algorithms are sufficiently potent to justify large-scale AF detection in a practical service environment.

The proposed AF detection service is based on hybrid decision support, which uses advanced AI for automated AF detection. The high accuracy of this algorithm sets it apart from other solutions currently on the market. The following paragraphs provide some background on current solutions. An Apple Watch and iPhone combination can be used to detect an irregular pulse. The Apple Watch measures the pulse. Once the signal is captured, an algorithm chain analyses the data. The user receives an alarm message if an irregular pulse is detected. During hold-out validation with benchmark data, that system achieved a positive predictive value of 71% (for example, only 71% of AF detections by the Apple Watch were actual AF detections; the remaining 29% were not). Based on the same measurements, researchers found that 84% of the participants that received irregular pulse messages. In a subsequent open study, 400,000 users were enrolled. 50% percent of the participants notified with irregular pulse messages. Apart from those pulse-based studies, the Apple Watch also features a finger ECG sensor with an AF detection function. However, this only works for as long as the user holds their fingers on the sensor. This

might not be adequate duration to detect AF. All Apple Watch-based health applications are consumer gadgets, which can create a suspicion that AF might be present. This suspicion would need to be verified by a medical doctor using a heart rate monitoring system. KardiaMobile with KardiaPro were employed to detect AF at home. The system is based on two electrodes sticked on behind the smartphone that measure the ECG by pressing both finger indexes. Based on these signals, the device decides if AF is present. In a study with 51 participants, the device had an 8% AF yield, i.e., four people were subsequently diagnosed with AF. Like the Apple Watch and iPhone combination, KardiaMobile is a gadget that establishes a suspicion that AF is present. However, the measurement is not continuous: 30 s ECG snippets are acquired whenever a patient activates the device. Based on such ad hoc measurements, the AF detection algorithm might miss an AF period. If an AF period is detected, the device raises an alarm, and it is up to the patient to interpret that information.

Table 20: Shows the results of arrhythmia detection based on RR intervals and ECG signals.

			Performance		
Author	Classifier	Signal type	ACC%	SPE%	SEN%
Salem et al., 2018[204]	Spectrogram with CNN	ECG	97.23		
Pudukotai Dinakarrao & Jantsch, 2018 [203]	Daubechies-6 with counters Anomaly	ECG	99.19	98.25	78.70
Yuanet et al., 2017 [202]	Unsupervised autoencoder NN SoftMax regression	ECG	98.19	98.22	98.11
Muthuchudar and Baboo, 2013 [201]	UWT NN	ECG	96		

Zhou et al., 2014 [200]	Shannon entropy with threshold	RR interval	96.05	95.07	96.72
Faust & Acharya,2021 [16]	ResNet	RR interval	98.55	94.30	99.40
Keidar et al, 2021 [218]	Modified Entropy Scale& decision tree	RR interval	97.8	97.4	98.1
Wang,2020 [219]	CNN and modified Elman neural network	ECG	97.4	97.1	97.9
Petrenas et al., 2015 [199]	Median filter with threshold	HR		98.3	97.1
Xia et al., 2018 [198]	STFT/SWT with CNN	ECG	98.63	98.79	97.87
Henzel et al., 2017[194]	Statistical features with generalized Linear Model	RR interval	93	95	90
Faust et al., 2018 [14]	Bidirectional LSTM	RR interval	98.39	98.32	98.51
Fujita and Cimr, 2019 [192]	CNN with normalization	ECG	98.45	99.87	99.27
Ivanovic et al., 2019 [125]	CNN, LSTM	RR interval	88		87.09

Our proposed method [17]	Detrending, ResNet	RR interval	99.98	100.00	99.94

5.8 Summary

In this study, we propose that a hybrid decision support approach for stroke prevention based on automated AF detection tool in HR signals. To achieve this task, data acquisition is accomplished by using Commercial HR sensors. The sensor data is transmitted via mobile phone to a cloud server for data storage. A DL model evaluates the HR data in real time. The real-time evaluation results represent the estimated AF probability. The physician can use that result as a second opinion which might refine the AF diagnosis, which ultimately leads to a stroke risk stratification. To support physicians during the diagnosis, we have integrated the analysis results from the LSTM classifier and displaying a digital biomarker in the proposed GUI to provide two independent categories of the analysis results. Having two options has the benefit that there is additional verification of the result analysis, and the digital biomarkers can be used to validate the DL outcomes. Real-time AF monitoring and detection systems are of great interest because they enable an early diagnosis, which might enhance patient quality of life, and provide a promising alternative to current healthcare processes. The value propositions focus on the healthcare provider. Therefore, the hybrid decision support for stroke prevention plays the key role in reduction of patient admission as well as decrease the workload on both healthcare provider and physician.

Chapter 6 Practical Atrial Fibrillation detection in a clinical setting

6.1 Study overview

We have developed a DL algorithm that can detect symptoms of AF in RR interval signals. With benchmark data from five different publicly accessible databases, the algorithm achieves excellent performance measures through a validation setup. To test the algorithm even further, we have conducted a clinical study which aims to validate the detection accuracy and to establish the diagnostic relevance. As part of this study, we will recruit 20 patients who have had stroke or Transient Ischaemic Attack (TIA): 10 who are known to have AF, and 10 who are not known to have the disease. We measured the electrical activity of the patients with two sensors. A Holter monitor was employed to measure the ECG signals and the Lifetouch sensor was allocated to record RR intervals. Once all recordings are collected, an experienced cardiologist or stroke physician will analyse the ECG signal to establish whether AF is present in a specific region of interest. We plan to compare this ground truth with the deep learning results. The study will be considered successful if the accuracy of the deep learning prediction is above 80%.

6.2 Methods

We have chosen the Isansys heart rate sensor because it is a medical grade device with a CE mark for clinical use. Further, the Isansys measurement setup is non-obtrusive and easy to apply. Such a device is ideal to be used alongside a Holter monitor as is the aim in this study.

We have opted to minimise wireless patient data transmission. During measurement the data is stored in the Isansys sensor; hence there is no wireless data transmission in the patient environment. Wireless patient data transmission happens only in the clinic once the patient has returned the sensor. From that time onwards, the data stays in the clinic. To be specific, the patient data is stored on a laptop which is in the clinic. Adopting this method limits the opportunity for an external entity to access and / or temper with patient data. A sample of patients with known AF (n=10) and another without known AF (n=10) has been chosen to enhance the chances of capturing episodes of AF to allow a comparative analysis between the deep learning system and Holter ECG recording.

6.3 Setting

The study has been undertaking on the stroke unit at the Royal Hallamshire Hospital and the Stroke Pathway Assessment and Rehabilitation Centre, Beech Hill, both part of Sheffield Teaching Hospitals NHS Foundation Trust.

6.3.1 Inclusion criteria

1. Adult (age > 18) patients who have suffered an ischaemic stroke or TIA.

2. Ability to provide written informed consent.
3. Ability to comply with study procedures in the opinion of the treating physician.

6.3.2 Exclusion criteria

1. Haemorrhagic stroke.
2. Disability preventing adherence to study procedures.
3. Clinically unstable.
3. Premorbid bed-bound state, Modified Rankin Scale (mRS = 5).

6.3.3 Sample size

A sample size of 20 has been chosen for practical reasons (ensure practice at obtaining the measures that we are aiming to measure, resource and time constraints). We have chosen to include a sample of patients with known AF to optimise our chances of picking up episodes of AF for comparison of the 2 monitoring methods (Deep learning system and ECG data analysis software).

6.3.4 Recruitment

Recruitment to this study have several prospective. We will recruit 20 patients from the stroke unit at the Royal Hallamshire Hospital. The stroke service in Sheffield receives approximately 900 stroke patients a year, approximately 300 of whom will require inpatient rehabilitation for longer than 1 week. The stroke unit currently houses 40 beds, and the large majority of these occupancies can be eligible for inclusion into the trial.

Patients eligible for recruitment into the study will be highlighted by the clinical treating team, who will then introduce the study to the patients. Principal Investigator (PI) identify from the impatient bed base on the stroke unit from daily ward rounds. Eligible patients will be asked if they would like an information leaflet regarding the study (an aphasia version if appropriate) and given at least 24 hours to consider the information. Patients who have read the information leaflets and are keen on being involved in the study can then inform the research team who will arrange a mutually convenient time to meet the patient and answer any further questions they may have about the study. If they are still keen on participating consent can be taken. Capacity to consent to the study would be determined by the treating clinical team and confirmed by the research team according to the principles of the mental capacity act.

6.4 Outcome measures

6.4.1 Primary outcomes

1. Safety of using the Isansys sensor and deep learning system. This will be assessed by review of the clinical data at baseline and follow up visit and review of the side effect diary during the observation period. Safety will be defined as:

2. NO Society of Automotive Engineer (SAE's) related to the Isansys monitor.

3. Less than 4 patients with any SAE.

4. Less than 8 adverse events across all 20 participants.

5. Patient acceptability: This will be assessed by review of acceptability of intervention diaries utilized throughout the intervention period. Acceptability will be defined as:

    - Less than 1/3 of participants reporting moderate or greater discomfort.

6. Compliance with the HR monitor. This will be assessed using the patient monitoring diaries. Compliance will be defined as:

    - More than 80% of the intended duration of wear.

7. Prediction of an AF risk score by the DL system. We anticipate that the DL system will produce an AF risk prediction score throughout the period of RR interval recording. Success will be defined as:

    - DL system produces an AF risk score for 100% of participants.

6.4.2 Secondary outcomes

6.4.2.1 Accuracy of the deep learning system

The accuracy of the deep learning system will be compared to the ground truth established by a human expert: review of ECG recording data by a stroke consultant, such that the

onset and offset of any periods of AF can be recorded. These time stamps can then be compared to the RR interval analysis produced by the deep learning system. If the algorithm produces a risk score of greater than or equal to 0.5 for these periods of AF then this will demonstrate evidence between the 2 methods. Success criteria for accuracy will be defined as:

1. >95% of the ECG identified periods of AF also identified as high risk by the deep learning system.

6.4.3 Study protocol

Potential participants have been identified from the TIA clinical and stroke unit at Sheffield Teaching Hospitals. Any patients meeting the inclusion criteria and interested in the study will be referred to the research team by their clinician. They will be given an information sheet and sufficient time to consider the study. If they wish to take part a baseline visit will be arranged.

Baseline visit 1

During the baseline visit the following will occur:

1. Eligibility criteria will be assessed and confirmed.
2. Informed consent will be obtained.
3. Socio-demographic (age, sex, marital status, pre-morbid function) and clinical (co-morbid history, medications, type, and severity of stroke, measured impairment, laboratory (including CRP and albumin) and radiology results, electrocardiogram, current mobility status) details will be recorded.
4. One of the researchers will apply the Holter monitor and an Isansys HR sensor (Lifetouch), and explain that these will be kept on for 24 hours. The participants will be taught how to take the sensors off for showering and will be asked to demonstrate this to the researchers.
5. The participants will be given a side effect diary to record any untoward effects and how acceptable they found having each type of HR monitor. The Holter monitor will record an ECG signal for 72 h. The Lifetouch sensor will record the HR signal for 24

h also. The participants will be allowed home or back on the ward following completion of the first visit.

Visit 2

After the 72-h measurement duration, the patient will return to the clinic for the second study visit. During the second visit, a researcher (Murtadha Kareem) will take off both Holter monitor and Lifetouch sensors. After that. There will be a short interview. A researcher will ask several structured questions aim to assess your experience as participant of the study. There will be feedback on the measurement results in the form of a diagnosis by the reading cardiologist and the stroke physician in the usual manner via letter sent to the participant. Following this, the participants involvement in the study will be complete. Figure 40 Shows a story board for the study protocol. Sign up and sensor placement will happen during Visit 1. Sensor return, which includes taking the questionnaire, will happen during Visit 2.

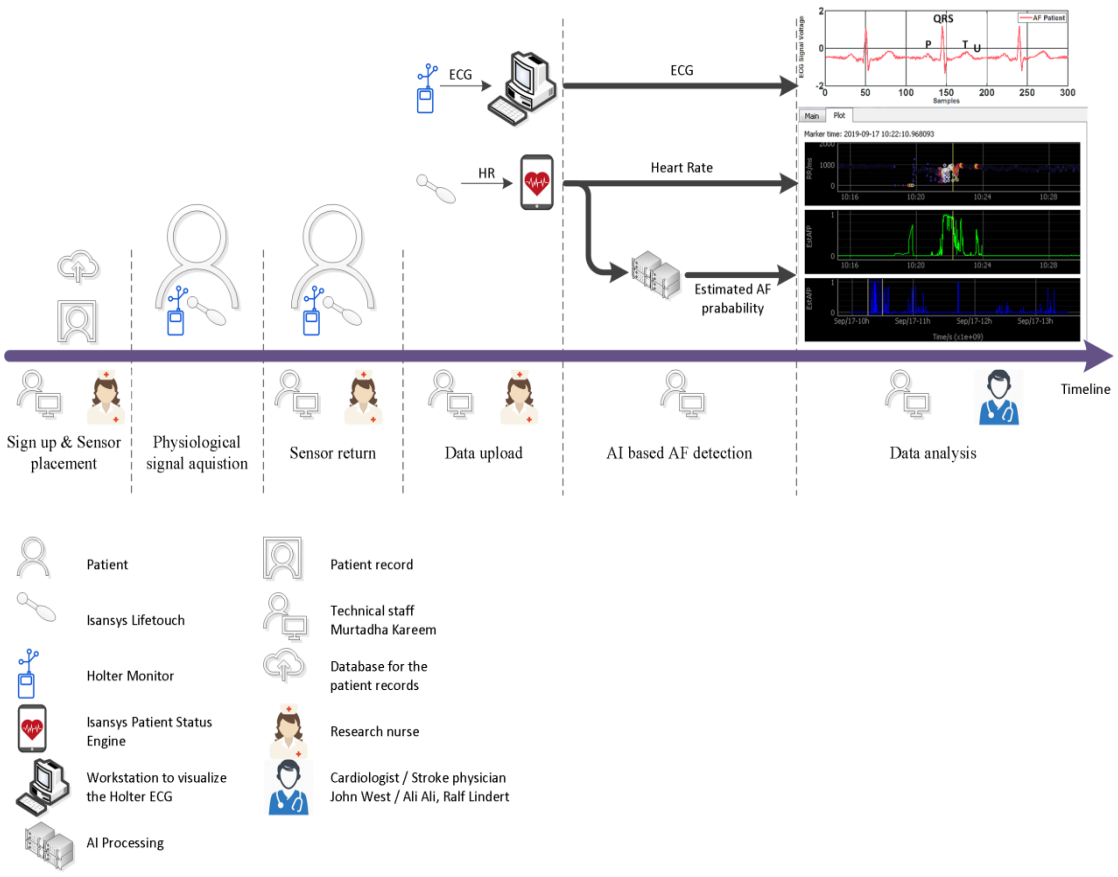


Figure 40: A timeline which shows the overall activities conducted at Sheffield Teaching hospital.

Software/Hardware overview

6.5 Software: Atrial fibrillation service validation tool

We developed a software tool to validate a deep learning algorithm for an atrial fibrillation detection service with heart rate data from a clinical study. The DL algorithm analyses the measurement data and establishes an estimated AF probability for each heartbeat. The software tool displays both data and DL analysis results. Furthermore, the graphical user interface can be used by medical experts to detect AF periods in the data and establish a reference result which will be treated as ground truth in subsequent result analysis steps. Once both DL and expert results are available, a confusion matrix is produced and the algorithm performance is validated by establishing accuracy, sensitivity, specificity, and f1-score. The software tool was created in Python and the software incorporated a graphical user interface as well as functional elements for data display and deep learning. To establish the required functionality, we used three different parallel processing methods for: 1) user interface processing, 2) data handling, and 3) deep learning. This highlights the need for parallel processing methods even for projects with a low or mid-range complexity. We have learned that the functionality of individual components can be expressed elegantly in Python. To be specific, there are four important activities that needs following to enable us visualising the graph and select the region of interest as follows:

1. **Control through Graphical User Interface (GUI)**
    - Load the Heart Rate (HR) data obtained from Lifetouch sensor as an excel file.
    - Process the data in the Deep Learning System (DLS).
    - Plot the HR measurements
    - Label all AF regions.
    - Results analysis
    - Save results as excel files and graph

2. **Graphical representation that shows four sections:**

- subplot 1: HR amplitude.
- subplot 2: RR intervals processed via DLS.
- subplot 3: Estimated probability of Atrial Fibrillation Events.
- subplot 4: Movable window to Select the area of interest.

3. **Add region of interest**
   - In case the DLS does not detect AF event, then an experienced cardiologist can identify and add AF region.

4. **Inference**

   - Confusion matrix graph: including accuracy, sensitivity, specificity, and F1-score.

   - ROC curve.

Figure 41: Visualisation of AF validation tool functionality as graphs and control panel.

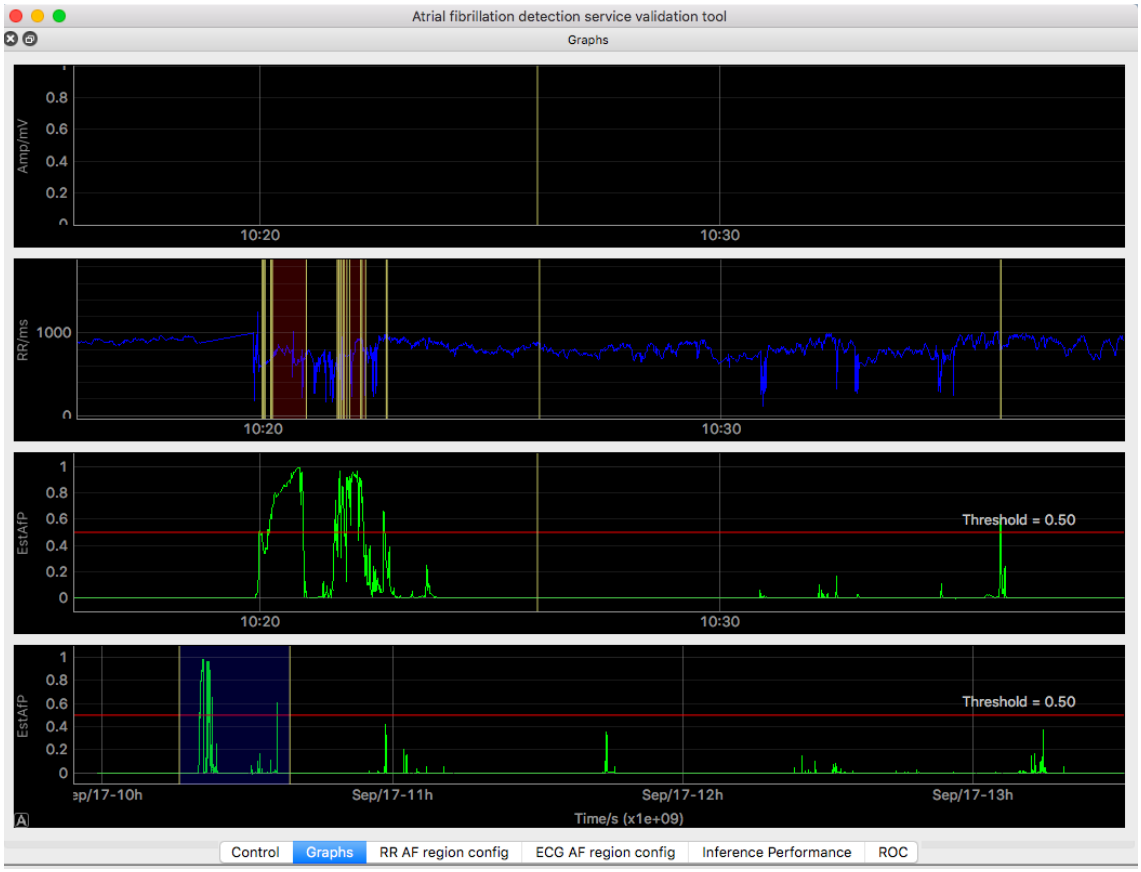


Figure 41: Demonstrates the graphical representation of the AF service validation tool.

6.5.1 Requirements

The following requirements describe what software we should build [220]. The software tool should automate the comparison between human expert and deep learning results. The measurements should be processed offline with the previously developed deep learning algorithm. The software should allow human experts to input the ground truth analysis results. The output of the program should be a confusion matrix and classification quality measures [221].

6.5.2 Specification

The requirements are refined into a specification which defines how we build the software [222]. The software takes excel sheets as input. These excel sheets are produced by the Lifeguard server from Isansys and they contain HR and ECG signals amongst other information. The HR signals are analysed with the LSTM deep learning algorithm, which produces an estimated AF probability score for each heartbeat. The analysis results are displayed alongside the signal data in two dimensional graphs. Navigation through these graphs is established with crosshair functionality. These graphs can be used to review and annotate the ECG signal with regions of AF [223]. The annotated regions are treated as ground truth with which the deep learning result is compared. To be specific, every beat that falls within the annotated region is treated as AF and every beat outside the region is non-AF [224]. A threshold is used to generate regions of estimated AF. This is done by comparing the threshold value with the estimated AF probability for each heartbeat. Whenever the estimated AF probability value is larger than the threshold, that beat belongs to a region of estimated AF. As a result, each beat has two labels: one from a human expert and one from the deep learning algorithm. Based on these labels a 2×2 confusion matrix is established. The matrix elements are used to calculate the performance measures of accuracy, sensitivity, specificity, and f1-score. These performance results, together with the estimated AF probability as well as the expert and algorithmic regions are saved in a separate excel file [225].

6.5.3 Implementation

The implementation was a meandering journey between learning the Python language and establishing the specified functionality [225]. Despite the relative inexperience with the language itself, the need for parallel processing became apparent early in the implementation cycle. To start, the inference functionality of Keras, which utilizes our DL model, incorporates parallel processing to establish the estimated AF probability [18]. Fortunately, this functionality is very well abstracted and indeed hidden from the user. Engaging with parallel processing libraries was required to realize a speedup for the signal display processing. We have used pyscp to compose ECG, HR, and estimated AF probability data vectors in parallel. Composing these vectors and the inference processing has high and very high computational complexity, respectively. This translates into waiting times for the program user. A progress spinner was implemented which indicates the processing of a potentially long task. This required us to use the QT multithreading functionality.

We have successfully established the specified functionality with three different parallel processing methods. However, the lack of debug support for parallel processing in the Python development environment Spyder made that task unnecessary hard. At times we resorted to trace messages and sometimes we bypassed the Qt multithreading functionality to inspect variables in code with is normally executed as a Qt thread. Furthermore, there is also some scope for formalizing the design and standardizing both code as well as file structure. This might lead to improved code quality.

6.6 Hardware

6.6.1 Sheffield Hallam University laptop

This laptop contains local Lifeguard server as well as AF service validation tool. If the devices are connected to the server, the HR data can be visualised in real-time on Patient Status Engine (PSE) interface. This enables the researcher and stroke physician to monitor the HR changes over the entire session duration. Figure 42 visualise the patient status engine page.

Figure 42: Sheffield Hallam Laptop displays the patient status engine interface.

6.6.2 Patient Gateway

The patient Gateway is a technologically advanced, all wireless patient monitoring system. A complete ready-built configurable platform, the gateway as part of PSE is also a fully certified Class IIa CE-marked, and Class II 510(k) cleared medical device that monitors patients automatically, continuously and in real-time. It combines sensors, connectivity and digital biomarker in a medical platform which collects, analyses, and transforms vital sign data into actionable clinical insights. By automating the basic process of taking patient observations, the PSE offers significant efficiency gains and

higher quality, continuous data for making better clinical decisions, as shown in Figure 43. The system uses trend data combined with the integrated Early Warning Scores to provide alerts to healthcare professionals, enabling timely interventions that enhance the care and safety of patients.

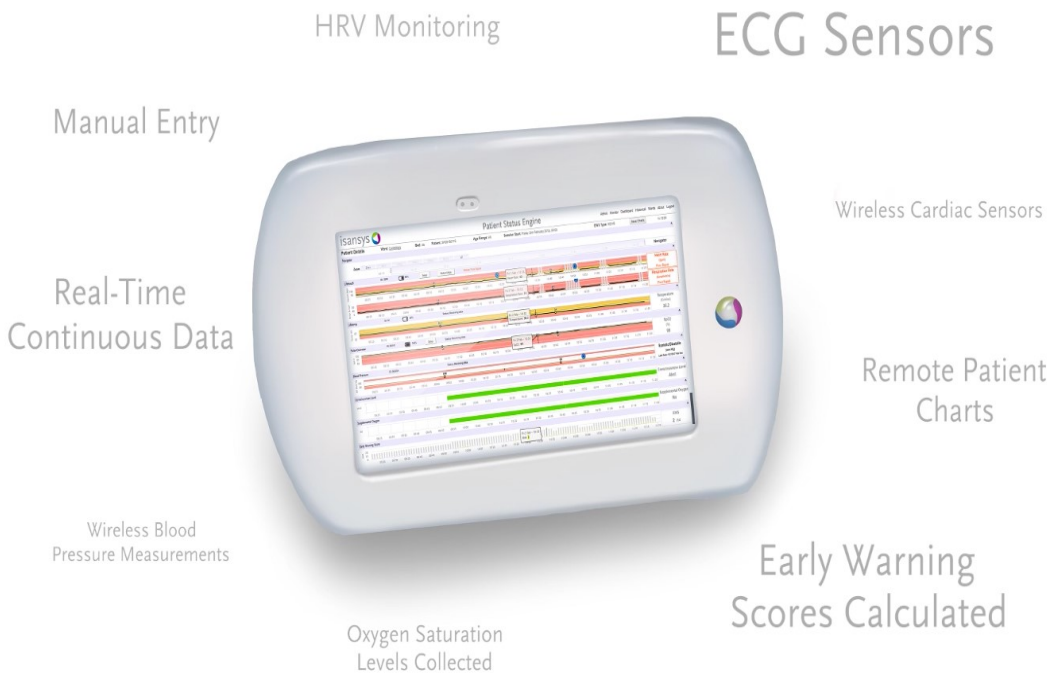


Figure 43: Patient Gateway functionality utilisation.

6.6.3 Local Area Network router

For data protection reason, the laptop is disconnected from Wi-Fi so that data cannot be communicated with the hospital network. The connection is made through local router that arrange transmission between patient Gateway and SHU Laptop. Figure 44 documents the connections.

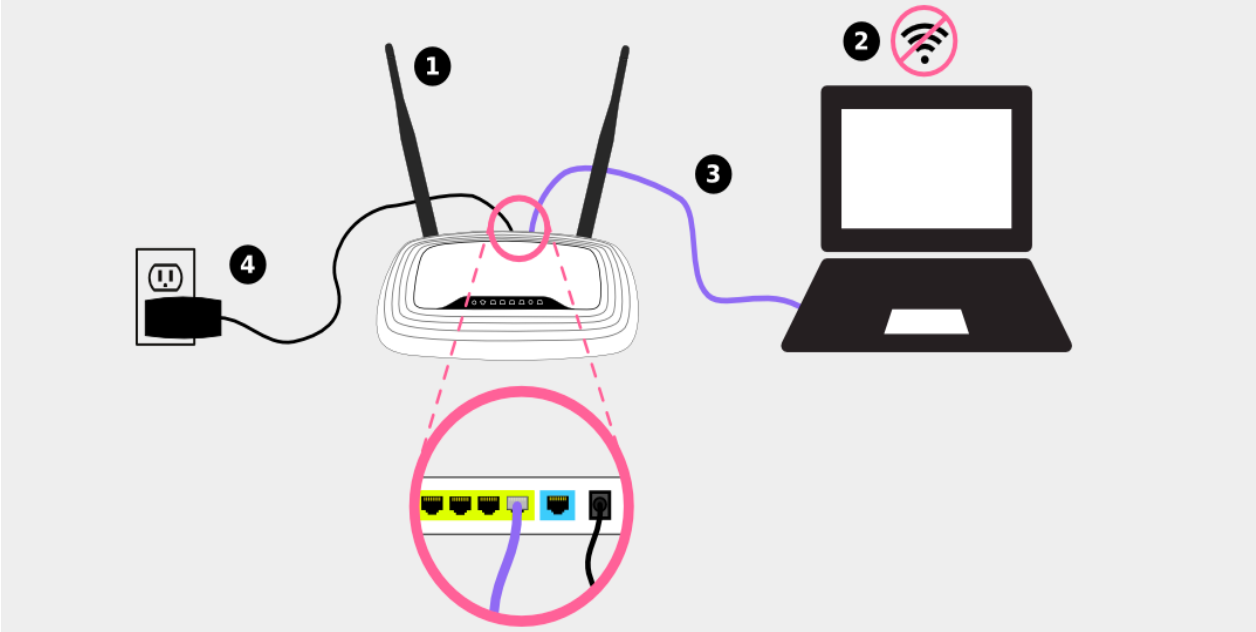


Figure 44: TP link router to activate the the local area network.

6.6.4 Lifetouch Sensor

The Lifetouch intelligent wearable biosensor is a validated and clinically proven wireless medical device that has already been used to collect over 150,000 hours of data from patients in acute care settings. It has been developed in association with diverse clinical teams, particularly nurses, to ensure ease of use, reliable operation, and seamless integration into nursing workflows.

With its lightweight design and ultra-low power operation, this state-of-the-art technology continuously samples and analyses the ECG signal in the sensor itself to extract the parameters of clinical interest. Providing continuous operation for 4 to 5 days the Lifetouch generates the data from which HR, respiration rate and HRV are calculated, while offering a real-time ECG streaming functionality. The device also includes a three-axis accelerometer to provide information on patient orientation, activity, and motion. Figure 45: Shows an example of Lifetouch sensor.

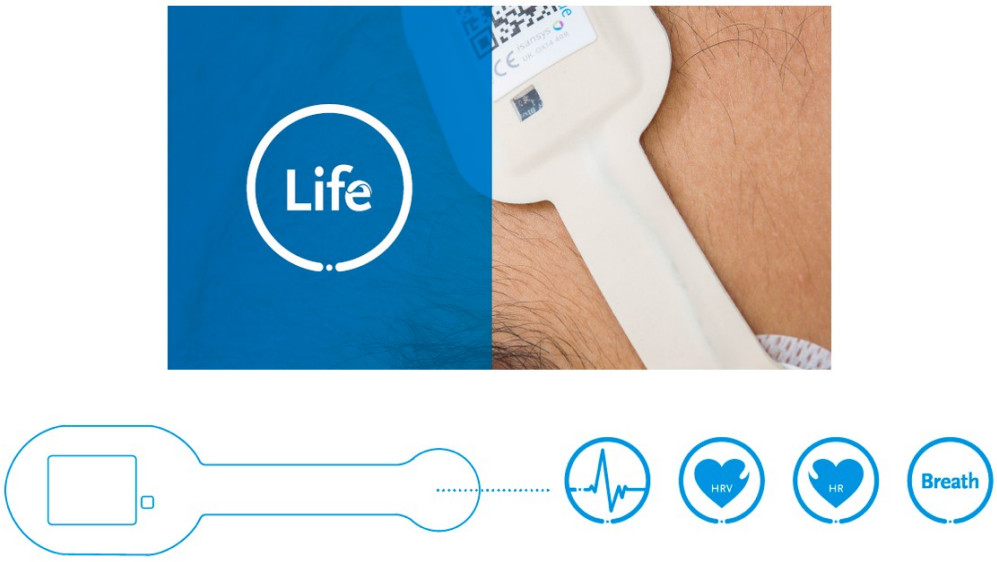


Figure 45: Lifetouch sensor applied on a patient chest[2].

Figure 46 illustrates the complete PSE setup for continuous monitoring. For the clinical trial study, our strategy for inpatient and outpatient participants are to transfer the monitoring session from the patient gateway which allows the sensor detecting heartbeats and store the data in the sensor itself. This option enables us to collect data from multiple participants in one patient Gateway device. Once the monitoring session is completed, sensor removal can be applied to download the data wirelessly from the sensor to patient gateway point and then directly to be transmitted to Local Lifeguard server through TP link router.

[2]https://www.isansys.com/en/Wearable-Sensor

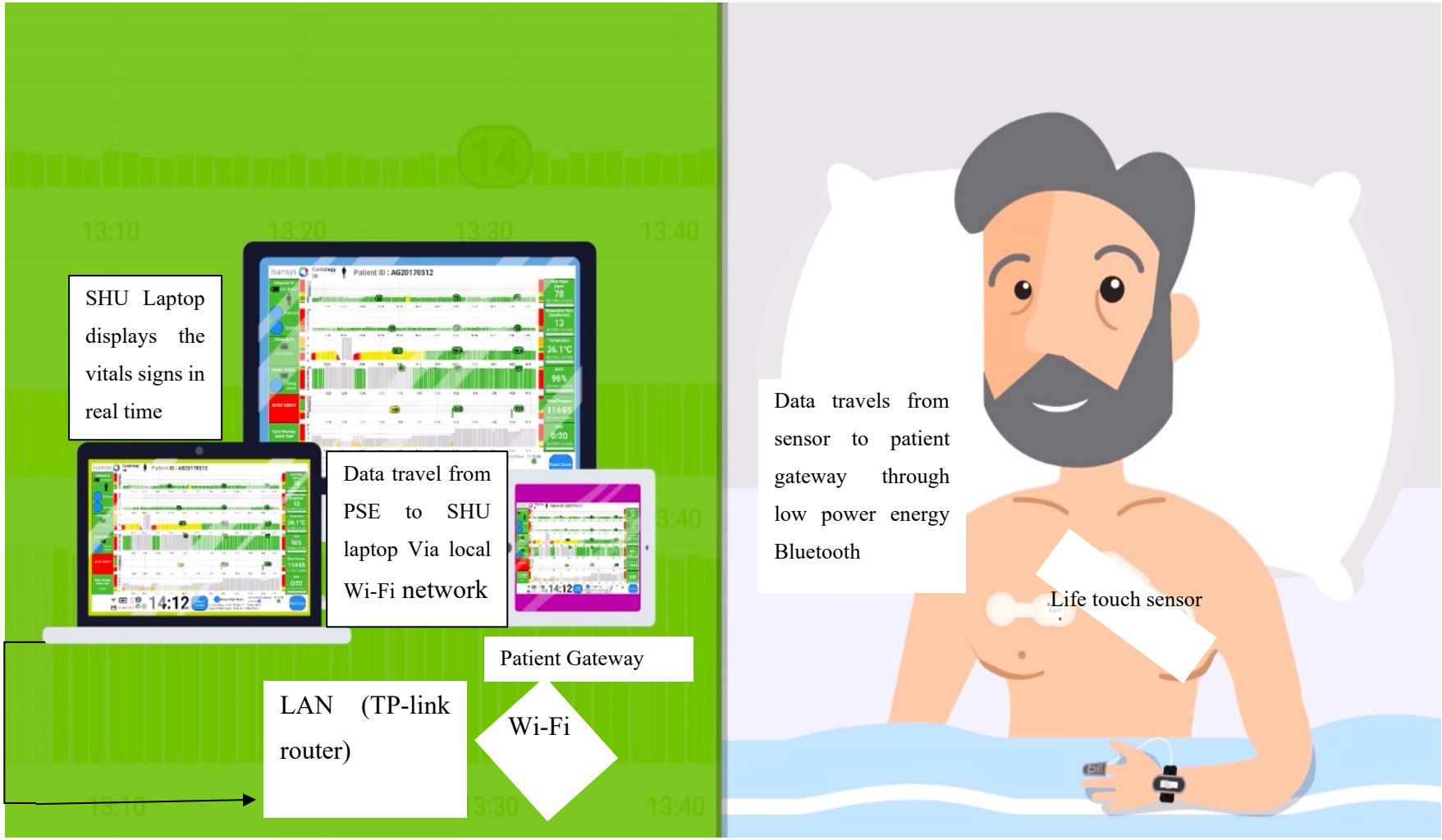


Figure 46: PSE setup.

6.7 Data collection

Currently, we have collected data form 13 participants and there will more participate in the following weeks. The recruitment strategy has been carried out either sending around 100 invitation letters for patients who have been discharged from the hospital or recruiting some of those who are based in the stroke ward and, willing to take part in the study. This step results in receiving great response from participants. However, the main challenge is having limited numbers of the available Holter monitor in the cardiology department. These devices are mainly allocated to the patients rather than being used for research. Therefore, this takes some time to arrange as the participant should be contacted for the scheduled appointment. In addition, the monitoring duration is determined by the principal investigator Dr. Ali, who can assess the necessities to have longer, or shorter monitoring period based on diagnosis needs. Table 21 shows the details of the data collected from two samples.

Table 21: Details of the data collected from 13 participants.

Sample No	**Participant ID**	**Heart monitor type**	**Recording duration**	**Cohort group**
Sample-1	B6048	Holter monitor + Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	Normal
Sample-2	B6049	Holter monitor +Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	Normal

Sample-3	B6047	Holter monitor +Lifetouch sensor	24 h with Holter & 24 h for Lifetouch sensor	AF
Sample-4	B6050	Holter monitor +Lifetouch sensor	24 h with Holter & 24 h for Lifetouch sensor	AF
Sample-5	B6051	Holter monitor +Lifetouch sensor	24 h with Holter & 24 h for Lifetouch sensor	Normal
Sample-6	B6052	Holter monitor +Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	AF
Sample-7	B6053	Holter monitor +Lifetouch sensor	24h with Holter & 24 h for Lifetouch sensor	AF, Stroke
Sample-8	B6054	Holter monitor +Lifetouch sensor	24 h with Holter &	AF

			24 h for Lifetouch sensor	
Sample-9	B6055	Holter monitor +Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	AF, Stroke
Sample-10	B6056	Holter monitor +Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	AF
Sample-11	B6159	Holter monitor +Lifetouch sensor	72 h with Holter & 24 h for Lifetouch sensor	AF, Stroke
Sample-12	B6160	Lifetouch sensor only	30h using Lifetouch sensor only, recording HR and ECG in one sensor	Normal
Sample-13	B6161	Lifetouch sensor only	48h Lifetouch sensor, recording HR and ECG in one sensor	Normal

6.8 Atrial fibrillation detection service validation tool description

In this phase, there are multiple steps to plot the Heart Rate (HR) measurements alongside with ECG recordings. These steps are listed as follows:

1. **Download the participant data from the Isansys lifeguard server**

   Participant data can be downloaded from the historical page once logging in into local server built-in Sheffield Hallam University laptop for data protection prospective.

2. **Load the Excel file into AF service validation tool**

   The excel file can be loaded from the control panel which allow us to process both HR and ECG, as shown in Figure 47.

3. **Process the Participant data**

   In this step, we use the deep learning technology to process the data once the process button is pressed. The measurement from 24 hours takes almost 2-3 hours processing by using Python compiler so-called Spyder. This duration could be less if we use the GPU. However, using computer-aided-diagnosis is still far better compared with the classical processing made in the hospital which almost take weeks to complete the results analysis. Figure 48 visualise data processing through Spyder software.

4. **Plot data**

   The graphical representation of the AF service validation tool shows that there are four subplots indicating to data acquisition and the estimated probability of AF that was detected by deep learning model. The first subplot refers to ECG measurements obtained by Holter where the stroke consultant can draw a region of AF with start and end timestamp. Whereas the second subplot refers to RR intervals signal recorded by Blue Lifetouch sensor. The last two subplots indicate to the classification outcomes of the estimated AF probability. In the third subplot, the threshold was set to 0.5 due to binary classification varies between zeros and ones. The values under the 0.5 relates to normal beats whereas the

values equal or above 0.5 refer to AF beats detected by deep learning algorithm. This approach is used as standard measure to evaluate the classifier performance. Figure 49: Depicts plotting data as graphs.

5. Label all AF regions

Labelling all AF regions detected through RR interval can be done automatically once this button is selected. It labels all the region of interest which is related to AF episodes based on the start of event and ending the event. These events are determined by using the timestamp for each region that specify the condition occurrence within certain period. Furthermore, we have added the option for cardiologists to label the other AF regions manually in case the classifier misclassifies some regions. In contrast, the human expert examines AF events through visual inspection of the entire ECG trace to correlate the machine classification with human diagnosis. In other words, the human specialist then works cooperatively with algorithmic tool to verify the deep learning results based on the HR data and additional knowledge obtained through patient records or by personal interaction with the patient. This approach would Improving the safety of AF monitoring systems through human verification[19], [21].In addition, the stroke physician can also confirm the AF suspicion generated by the algorithm which represents false positive classification outcome. Figure 50 shows some regions of AF detected by deep learning algorithm.

6. Results analysis

From the results analysis icon, we can generate both the confusion matrix and ROC curve. Confusion matrix has true label classes represented by a cardiologist when identify normal and abnormal rhythms whilst the deep learning algorithm performs the predicted label classes. Both classes [True Negative, True Positive] represent the correctly identified beats normal and AF respectively. However, classes [False Positive, False Negative] are incorrectly identified by algorithmic machine and the cardiologist respectively. In terms of ROC curve, the curve can be plotted the discrimination between TPR and FPR. For the initial analyse, deep learning algorithm processes the performance measure without stroke

consultant intervention. Once data acquisition stage will be completed, we will arrange a meeting for result formulation to reach a feasibility. Therefore, there were no ROC curve that plotted during the initial analysis, but certainly, cardiologist intervention will help refining the classification outcomes for both confusion matrix and ROC curve.

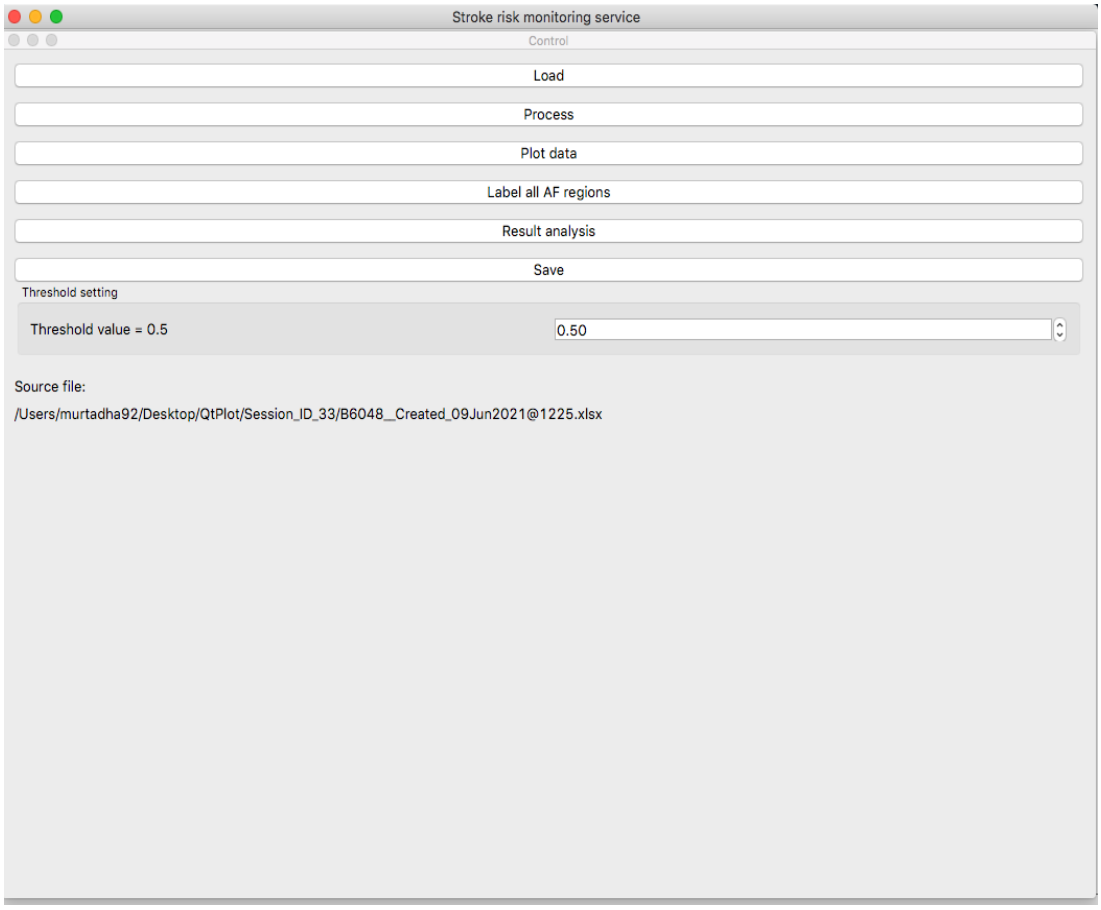


Figure 47: Load function performs data loading to the deep learning algorithm from the control panel.

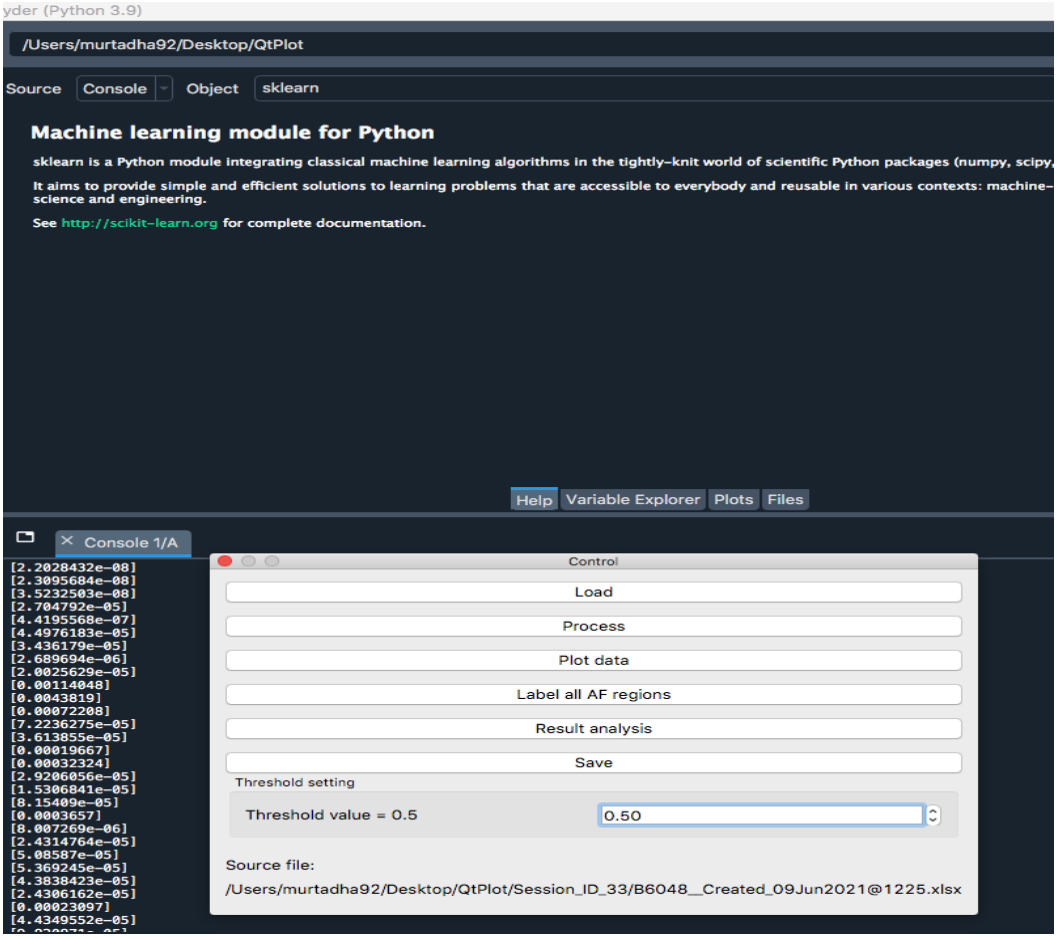


Figure 48: Process option enables us to start estimating the AF probability.

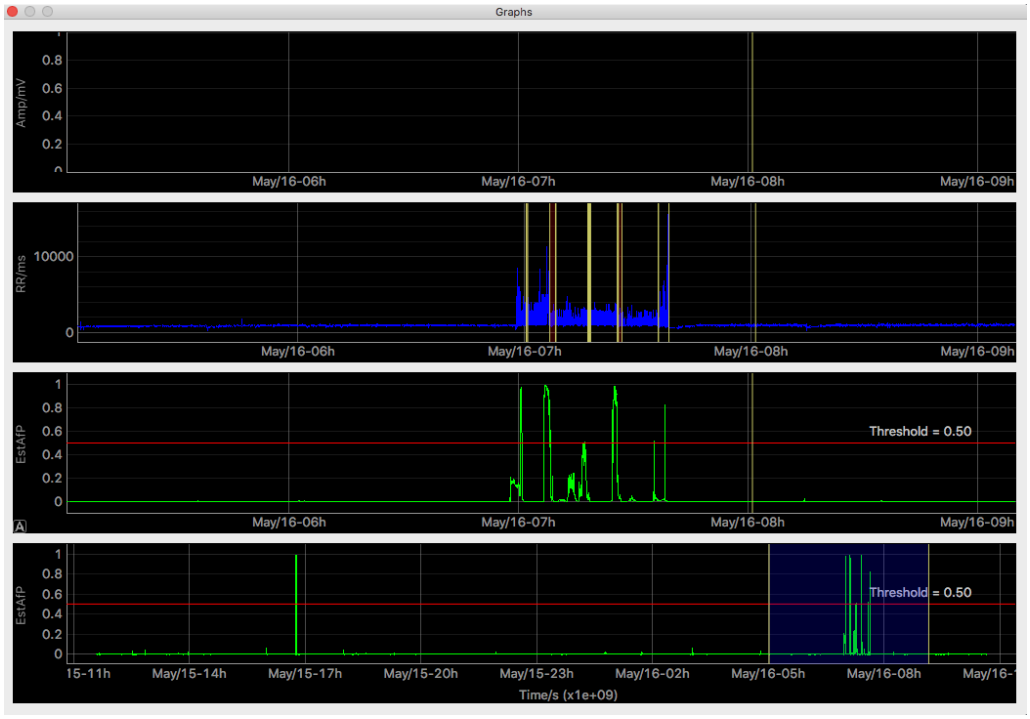
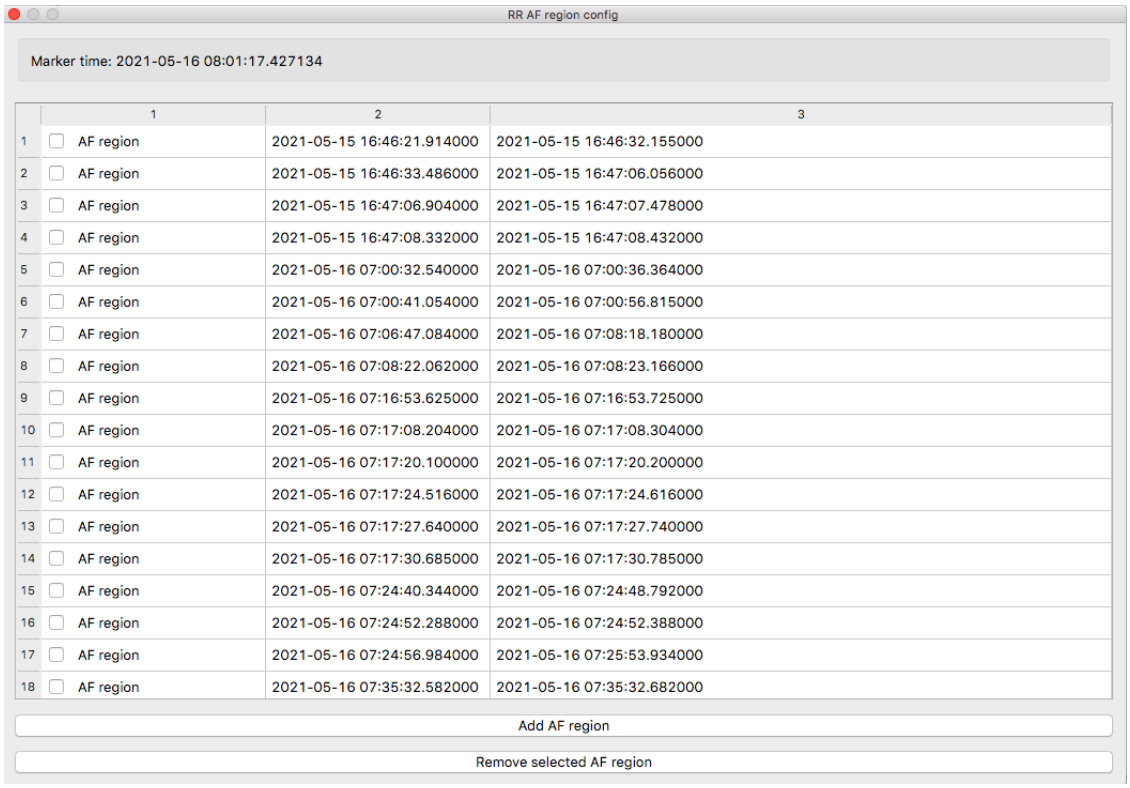


Figure 49: Graphical representation of data plot.



	1	2	3
1	AF region	2021-05-15 16:46:21.914000	2021-05-15 16:46:32.155000
2	AF region	2021-05-15 16:46:33.486000	2021-05-15 16:47:06.056000
3	AF region	2021-05-15 16:47:06.904000	2021-05-15 16:47:07.478000
4	AF region	2021-05-15 16:47:08.332000	2021-05-15 16:47:08.432000
5	AF region	2021-05-16 07:00:32.540000	2021-05-16 07:00:36.364000
6	AF region	2021-05-16 07:00:41.054000	2021-05-16 07:00:56.815000
7	AF region	2021-05-16 07:06:47.084000	2021-05-16 07:08:18.180000
8	AF region	2021-05-16 07:08:22.062000	2021-05-16 07:08:23.166000
9	AF region	2021-05-16 07:16:53.625000	2021-05-16 07:16:53.725000
10	AF region	2021-05-16 07:17:08.204000	2021-05-16 07:17:08.304000
11	AF region	2021-05-16 07:17:20.100000	2021-05-16 07:17:20.200000
12	AF region	2021-05-16 07:17:24.516000	2021-05-16 07:17:24.616000
13	AF region	2021-05-16 07:17:27.640000	2021-05-16 07:17:27.740000
14	AF region	2021-05-16 07:17:30.685000	2021-05-16 07:17:30.785000
15	AF region	2021-05-16 07:24:40.344000	2021-05-16 07:24:48.792000
16	AF region	2021-05-16 07:24:52.288000	2021-05-16 07:24:52.388000
17	AF region	2021-05-16 07:24:56.984000	2021-05-16 07:25:53.934000
18	AF region	2021-05-16 07:35:32.582000	2021-05-16 07:35:32.682000

Figure 50: Labelled RR AF regions.

6.9 Results analysis

The results analysis for the 13-participants showed excellent classification outcomes which has been processed by using the deep learning support. Each participant results have been processed independently where patient data loaded into the AF service validation tool. To be specific, accessing patients' RR interval data can be retrieved from the built-in local cloud server at SHU laptop, known as lifeguard server. The Researcher, Murtadha Kareem, has only the granted access to that Patient Status Engine as greed in study protocol and NHS-Ethics approvals. For the time being, the cardiology department at Sheffield Teaching Hospital has provided the ECG analysis of Holter monitor for those participants as confirmed sheets by cardiologist. Initially, these sheets help the stroke consultant to undertake the diagnosis based on digital biomarkers, such as, Min/Max HR, mean, as well as visually inspecting the region of interest where the AF events present. However, for the POF study, we would need to have the start timestamp and the end timestamp of AF episodes provided from the cardiologist for all the events with complete duration. This can allow us to correspond the labelled region by the cardiologist to the

labelled regions processed by the deep learning algorithm. Once these labelled regions are symmetric, it indicates to the accuracy and the efficiency of the validated deep learning model based on professional human verification. Alternatively, the researcher can combine both the RR interval measurement from Lifetouch sensor and ECG recordings from Holter monitor used gold standard device for recording the paroxysmal AF in the hospital. Both Sensors are wearable, and patients can be freely sent home for designated duration by the Stroke Consultant. In the meantime, the data analysis phase has been partially completed only for 13 participants with HR data and some ECG episodes captured by Lifetouch sensor. The performance measure of the binary classification showed that the algorithm achieved such a promising result in terms of accuracy and specificity.

Table 22: Performance measure results for each participant with average.

Sample No	ACC_{cl} (%)	SEN_{cl} (%)	SPE_{cl} (%)	Confusion matrix	
1	99.72	Nan	99.72	91135	252
				0	0
2	99.85	Nan	99.85	84225	128
				0	0
3	99.51	Nan	99.51	104159	514
				0	0
4	98.70	Nan	98.70	80223	1049
				0	0
5	99.23	Nan	99.23	90160	100
				0	0
6	33.12	Nan	33.12	41000	82785
				0	0
7	98.99	Nan	98.99	95646	250
				0	0
8	88.96	Nan	88.96	84134	10433
					0
9	99.80	Nan	99.80	110600	200
				0	0
10	99.51	Nan	99.51	92300	400
				0	0
11	85.47	Nan	85.47	82521	14660
				0	0
12	99.22	Nan	99.22	131620	104
				0	0
13	99.80	Nan	99.80	133918	255
				0	0
Average	92.45	Nan	92.45	90644	9241

	0	0

Table 22 shows the results of accuracy, sensitivity, and specificity for each participant as well as the confusion matrix outcomes. The average is accumulated for the overall entities. We observed from the results analysis that there was unexpected accuracy and specificity of 33.12% respectively for sample 6. This is due to the suspicion of generating high false positive by deep learning algorithm. However, the overall false positive beats are produced because of the absence of the cardiologist input during the analysis phase. For the preliminary analysis, the stroke consultant has confirmed that participant has several heart problems and other conditions such as AFL, AF, stroke, Asama, and heart failure. As such, all these conditions are synchronised with rhythm abnormalities which correctly identified as irregular beats. In addition, the consultant stated that there were around 90,000 beats irregular detected from the Holter monitor. Therefore, this claim can confirm the deep learning findings and would improve the performance measures results once the experienced cardiologist provides the annotation of AF regions and establish identical diagnosis. Another observation found from the above table is the sensitivity values are not calculated due the true positive values are equal to zero. The link between the absence of true positive values and sensitivity's values will be refined as soon as the clinical interpretation of ECG measurements are combined into the AF validation tool and corresponded with the estimated AF probability result. This leads to results analysis refinements in terms of confusion matrix outcomes, ROC curve, accuracy, specificity, and sensitivity.

6.10 Discussion

The atrial fibrillation detection service validation tool automates the comparison between stroke consultant and classification results of deep learning algorithm [19], [225]. Moreover, relevant plotted signals on graphs allow the medical professional to review the analysis results which might lead to a deeper understanding of AI for AF detection. To be specific, for medical professionals it is significant to visualize the deep learning the results analysis, because they tend to establish what misdiagnoses are happening and the extend of overreporting and underreporting of events. Based on this visualization, we can start a discussion with medical professionals on what action to perform as a result of a specific scenario. For instance, we might be able to answer questions like: How long can we automatically detect AF justifies potentially life-threatening intervention, such as anticoagulation. Currently, this is where we draw the line between machine and human work. The machine provides an estimated AF probability over time and human experts must interpret and verify that result. The interpretation should be done by combining the result information with history record about the patient to reach a diagnosis.

Establishing a hybrid environment where humans work with machine algorithms is an important goal for future deployment in the practice [19], [21]. The current AF detection service validation tool can only support as an initial attempt with which we can study interaction patterns. These patterns might indicate a direction for further automatization. Currently, we are considering about rules and regulations to create a notification system. Understanding the estimated AF probability signal shape for a treatable case might lead to the automated generation of notification messages. For example, a notification message is sent once the estimated AF probability is equal or above a 50% threshold for more than 5 min within one hour. Calculating that and transmitting the alarm message is straight forward, but significantly more research is needed to establish a useful amount of alarm cases. The focus on establishing alarm message conditions might seem like a minor point, but this is what lies at the heart of all IoMT devices that provide diagnosis support by measuring, transmitting, and processing patient data in real time. These systems are capable to extend the observation duration indefinitely which holds the promise of detecting diseases earlier and that detection is largely independent from whether there are long asymptomatic episodes [9]. Having the long observation duration together with the

alarm functionality is likely to improve outcomes for patients through an early diagnosis which will lead to less intrusive interventions.

6.11 Summary

At the time of writing, the clinical study is still ongoing as most of the work has been delayed the COVID-19 pandemic. We continue to maintain the study moving forward until the completion of data collection and analysis. Once these phases are done, the Chief Investigator, Principal Investigator and Co-investigator will discuss the final outcomes of the study to reach a feasibility. So far, the proposed algorithm a very promising performance, but these results need to be confirmed. To be specific, a cardiologist will establish the ECG outcome which is compared with HR classification. If there are any undetected events captured by the ECG monitor, we added an option for stroke cardiologist to label the additional region of interest which helps to improve the safety and accuracy of diagnostic tool. This approach can increase the learning phase and extracting useful features for future diagnosis.

Chapter 7 Conclusion and further work

7.1 Conclusion

This study incorporates developing an LSTM based DL model to detect AF episodes by using RR intervals for stroke prevention. A large body of literature has been reviewed. Current technology was used in the market and clinical practice, and signal recording types through medical devices. ECG signals are most widely used for AF diagnosis. We found that ECG recordings yield a significantly higher data rate when compared with RR intervals. Apart from a significant data reduction, the RR interval measurement setup is also less complex. This makes RR interval signal acquisition less expensive. We proposed patient led data acquisition for continuous treatment monitoring, because with current technology this is a practical way to extent the observation duration for those who have an identified AF risk or stroke survivors. To achieve this task, a real time sensor should be attached to patients' chest, and data travels from the point of measurement to the central cloud processing unit. DL algorithms should be used to automate the analysis and they might provide accurate AF detection functionality.

Initially, the algorithm was trained with 20 subjects and tested with completely held out of 3subjects. In addition, the LSTM model has been validated with varied datasets obtained from publicly accessible databases for researchers. The overall data used for validation was around 188 subjects obtained from LTAFDB, NSRDB, Fantasia database and arrythmia database. The validation outcomes showed promising results which implies that the algorithms extracted the knowledge from small datasets and apply the knowledge on larger datasets.

We propose the hybrid decision support tool for stroke prevention based automated AF detection in HR signals. From this tool, the stroke physician reviews and annotates the region of interest which relates to deep learning classification. The physician can use that

result as a second opinion, which might improve the AF diagnosis, which ultimately leads to a stroke risk stratification. To support physicians during the diagnosis, we incorporate deep learning results and digital biomarkers in the proposed GUI to provide two independent analysis results. Having two independent results has the advantage that there is no single point of failure, and the digital biomarkers can be used to validate the deep learning result.

Based on the distinct results obtained from the validation study setup, we conducted a clinical trial study in collaboration between Sheffield Hallam University and Sheffield Teaching Hospital. This study involves collecting data from overall 20 participants, classified into two groups. These groups are 10 participants from normal cohort and the other 10 participants from AF and Stroke unit. However, we only collected data for 13 participants in the meantime, and recruitment strategy will continue until completion of the full sample size. AF detection service support tool has been designed to helps stroke consultant in the analysis phase. The initial analysis of deep learning algorithm compared to gold standard showed an excellent results and promising approach for future deployment in the clinical practice.

7.2 Summary of Work Done

This section summarises the work done throughout the PhD project. The first step started with identifying the research gaps and showing the potential in filling that gaps scientifically. Two distinct algorithms were developed to detect arrhythmias. The results of these algorithms were successfully published in peer-reviewed journals. Collaborating with translate MedTech in Sheffield and Leeds helped to broaden the network with academic, clinical and industry partners. To be specific, translate the technology offers greater opportunity that identify the unmet needs. Attending local and international conferences in which contribute the undertaken research with the relevant community. In addition, participating in the clinical trial study that aimed to validate the proposed algorithms in the ground truth.

7.3 Contributions to Knowledge

The objectives of this thesis are to contribute to knowledge in theory and practice. Through our literature review, some of the research gaps were identified, namely that there was an absence of using RR intervals in AF detection based on distinct DL model in clinical settings. In addition, we are the first who introduced the concept of patient-led data acquisition that can extend the observational duration of AF [9]. The service directions play the main role to form patient-led data acquisition. We have critically investigated and evaluated the chosen topic. This section shows how the results from this research fill the related gaps in knowledge, thereby contributing to theory and practice. The achieved objectives summary our key contributions to knowledge as follows:

1. We used HR measurements as a main monitoring and detection method. This comes from the fact that HR is a cost-effective signal to measure, distribute, process and store in the cloud. Furthermore, HR is good indicator of human health. Hence, we demonstrated that AF detection can be done through using RR intervals which were extracted from the ECG measurement obtained of publicly benchmark data known as PhysioNet.
2. We designed, implemented, and tested a DL model that trained with 20 subjects and tested with 3 subjects from MIT AFDB databases. The LSTM model was evaluated with the performance measures, namely confusion matrix and ROC curve. These performance measures showed promising results that were achieved in both 10-folds-cross validation and blind folds validation. Therefore, it is a valid claim to argue that RR intervals is an applicable approach for automating AF detection. Moreover, we established a ResNet model to detect AF, AFL and NSR that was trained with 4051 participants. This model achieved great results in terms of accuracy, sensitivity and specificity.
3. We developed a cloud computing technique based on Thingspeak. A smart app so-called Heartcare mobile app was developed to communicate the HR measurement from wireless sensor polar 10 to a smartphone through BLE connection. The HR data will be distributed to Thingspeak server via Wi-Fi network. We integrated our DL algorithm into Thingspeak which can detect AF episodes in real-time.

4. We demonstrated patient-led data acquisition approach where the data travel from the point of measurement(patient) to the central location point (cloud server). This approach is relevant to AF detection service for stroke prevention which helps to extend the observational duration.
5. The LSTM model was validated from five different benchmark databases, namely, AFDB, LTAFDB, NSRDB, Fantasia database and Arrhythmia database. These databases are completely unknown to the model. To be specific, the bi-directional LSTM model achieved greater results in the validation setup study. Therefore, we established the maturity and robustness of the LSTM model through validation.
6. The proposed AF detection service for stroke prevention has been validated in the clinical trial at Sheffield Teaching Hospital. All ethics approvals completed before conducting the study. The sample size of the study is 20 participants which were divided into two cohorts, 10 participants from normal cohort and other 10 participants from AF cohort.
7. We applied the concept of hybrid decision support in both research and practical setting. The human experts can work cooperatively with AF detection service validation tool to verify the classification outcome of the DL model through knowledge obtained of the patient interaction. This step improves the safety of machine decision through medical professionals' verifications.

7.4 Limitations and challenges

1. Retrain the deep learning model during the validation and learning features. To be specific, a cardiologist learns while doing the job. The proposed deep learning model is static, i.e., it did not learn during the validation. At one point the knowledge, extracted from 20 patients, will be insufficient to cope with practical scenarios. In the future, we have to find a way to model that continuous training in order to improve the diagnostic quality of the proposed AF detection system. One way of providing this continuous learning is to retrain the deep learning model with measurement data. A prerequisite for such a methodology is to have the HR data stored in a central location. Hence, streaming the HR data to a central

cloud server might prove to be an advantage when the continuous learning problem is tackled.

2. An alarm message is sent when a dangerous situation arises. Initially, what forms a dangerous condition could follow Holter monitoring protocols. For example, an AF event is detected when the estimated AF probability is above 0.5 for at least 30 s. However, it is not known if such an approach is sensitive and indeed specific enough to capture the stroke risk for patient.

3. Obtaining necessary regulatory approvals for accessing the NHS cloud to facilitate uplink transmission.

4. The current algorithm is not integrated into a medical device due to further Health Regulatory Approval (HRA) is required.

7.5 Future work

1. We plan to collect more data from multiple participant sites. We will recruit participants with both known and unknown aetiology to get deeper insights into the link between HR and the nature of embolisms, which might lead to stroke.

2. Extending the observation for indefinite duration so that it increases AF detection probabilities and lead to stroke prevention.

3. Using the proposed AF detection service for many patients over long time periods leads to big data with reliable labels.

4. Integrates our deep learning algorithm into a medical device after the approval is granted.

5. Establish a service platform that can monitor multiple diseases, such as, arrhythmias, diabetes, sleep apnea, and congestive heart failure.

REFERENCES

[1] P. A. Wolf, R. D. Abbott, and W. B. Kannel, "Atrial fibrillation as an independent risk factor for stroke: The framingham study," *Stroke*, vol. 22, no. 8, pp. 983–988, 1991, doi: 10.1161/01.STR.22.8.983.

[2] A. G. Thrift *et al.*, "Global stroke statistics," *Int. J. Stroke*, vol. 9, no. 1, pp. 6–18, 2014, doi: 10.1111/ijs.12245.

[3] C. Wang, J. Redgrave, M. Shafizadeh, A. Majid, K. Kilner, and A. N. Ali, "Aerobic exercise interventions reduce blood pressure in patients after stroke or transient ischaemic attack: A systematic review and meta-analysis," *Br. J. Sports Med.*, vol. 53, no. 24, pp. 1515–1525, 2019, doi: 10.1136/bjsports-2017-098903.

[4] R. Grysiewicz, K. Thomas, D. P.-N. Clinics, and U. 2008, "Epidemiology of ischemic and hemorrhagic stroke: incidence, prevalence, mortality, and risk factors," *Elsevier*, vol. 26, no. 4, pp. 871–895, 2008, Accessed: Aug. 09, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0733861908001047.

[5] K. B. Slot, E. Berge, P. Dorman, S. Lewis, M. Dennis, and P. Sandercock, "Impact of functional status at six months on long term survival in patients with ischaemic stroke: Prospective cohort studies," *Bmj*, vol. 336, no. 7640, pp. 376–379, 2008, doi: 10.1136/bmj.39456.688333.BE.

[6] K. Daniel, C. D. A. Wolfe, M. A. Busch, and C. Mckevitt, "What are the social consequences of stroke for working-aged adults?: a systematic review," *Stroke*, vol. 40, no. 6, pp. 431–440, 2009, doi: 10.1161/STROKEAHA.108.534487.

[7] M. Colin, GretchenStevens, and M. Maya, "GLOBAL HEALTH RISKS Mortality and burden of disease attributable to selected major risks," 2009. [Online]. Available: http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf.

[8] K. Carroll, S. Murad, J. Eliahoo, and A. Majeed, "Stroke incidence and risk factors in a population-based prospective cohort study," *Heal. Stat. Q.*, vol. 12, pp. 18–26, 2001.

[9] M. Kareem, N. Lei, A. Ali, E. J. Ciaccio, U. R. Acharya, and O. Faust, “A review of patient-led data acquisition for atrial fibrillation detection to prevent stroke,” *Biomed. Signal Process. Control*, vol. 69, no. March, p. 102818, 2021, doi: 10.1016/j.bspc.2021.102818.

[10] P. Scarborough *et al.*, “Stroke statistics 2009 edition Contents,” 2009.

[11] A. N. Ali and A. Abdelhafiz, “Clinical and economic implications of AF related stroke,” *J. Atr. Fibrillation*, vol. 8, no. 5, pp. 85–92, 2016, doi: 10.4022/jafib.1279.

[12] U. Varshney, “Pervasive healthcare and wireless health monitoring,” *Mob. Networks Appl.*, vol. 12, no. 2–3, pp. 113–127, 2007, doi: 10.1007/s11036-007-0017-1.

[13] A. Pantelopoulos and N. G. Bourbakis, “A survey on wearable sensor-based systems for health monitoring and prognosis,” *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 40, no. 1, pp. 1–12, 2010, doi: 10.1109/TSMCC.2009.2032660.

[14] O. Faust, A. Shenfield, M. Kareem, T. . San, H. Fujita, and U. . Acharya, “Automated detection of atrial fibrillation using long short-term memory network with RR interval signals,” *Comput. Biol. Med.*, vol. 102, 2018, doi: 10.1016/j.compbiomed.2018.07.001.

[15] E. Hoefman, H. C. P. M. van Weert, and P. J. E. Bindels, “Efficacy of diagnostic tools for detecting cardiac arrhythmias: Systematic literature search,” *Netherlands Hear. J.*, vol. 18, no. 11, pp. 543–551, 2010, doi: 10.1007/s12471-010-0831-0.

[16] O. Faust and U. R. Acharya, “Automated classification of five arrhythmias and normal sinus rhythm based on RR interval signals,” *Expert Syst. Appl.*, vol. 181, no. March, p. 115031, 2021, doi: 10.1016/j.eswa.2021.115031.

[17] O. Faust, M. Kareem, A. Ali, E. J. Ciaccio, and U. R. Acharya, “Automated Arrhythmia Detection Based on RR Intervals,” *Diagnostics*, vol. 11, no. 8, pp. 1–18, 2021.

[18] O. Faust, M. Kareem, A. Shenfield, A. Ali, and U. R. Acharya, “Validating the robustness of an internet of things based atrial fibrillation detection system,” *Pattern Recognit. Lett.*, vol. 133, 2020, doi: 10.1016/j.patrec.2020.02.005.

[19] M. Kareem and O. Faust, “Establishing the safety of a smart heart health monitoring service through validation,” 2019, doi: 10.1109/BigData47090.2019.9006478.

[20] O. Faust, N. Lei, E. Chew, E. J. Ciaccio, and U. R. Acharya, “A smart service platform for cost efficient cardiac health monitoring,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 17, pp. 1–18, 2020, doi: 10.3390/ijerph17176313.

[21] N. Lei, M. Kareem, S. K. Moon, E. J. Ciaccio, U. R. Acharya, and O. Faust, “Hybrid decision support to monitor atrial fibrillation for stroke prevention,” *Int. J. Environ. Res. Public Health*, vol. 18, no. 2, 2021, doi: 10.3390/ijerph18020813.

[22] O. Faust, E. J. Ciaccio, and U. R. Acharya, “A review of atrial fibrillation detection methods as a service,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 9, 2020, doi: 10.3390/ijerph17093093.

[23] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, “Deep learning for healthcare applications based on physiological signals: A review,” *Computer Methods and Programs in Biomedicine*, vol. 161. Elsevier Ireland Ltd, pp. 1–13, Jul. 01, 2018, doi: 10.1016/j.cmpb.2018.04.005.

[24] O. Faust, E. J. Ciaccio, A. Majid, and U. R. Acharya, “Improving the safety of atrial fibrillation monitoring systems through human verification,” *Saf. Sci.*, vol. 118, no. May, pp. 881–886, 2019, doi: 10.1016/j.ssci.2019.05.013.

[25] E. J. Benjamin, P. A. Wolf, R. B. D’Agostino, H. Silbershatz, W. B. Kannel, and D. Levy, “Impact of atrial fibrillation on the risk of death: The Framingham Heart Study,” *Circulation*, vol. 98, no. 10, pp. 946–952, 1998, doi: 10.1161/01.CIR.98.10.946.

[26] J. Adamson, A. Beswick, and S. Ebrahim, “Is stroke the most common cause of disability?,” *J. Stroke Cerebrovasc. Dis.*, vol. 13, no. 4, pp. 171–177, 2004, doi: 10.1016/j.jstrokecerebrovasdis.2004.06.003.

[27] S. Stewart, C. L. Hart, D. J. Hole, and J. J. V. McMurray, “A population-based study of the long-term risks associated with atrial fibrillation: 20-Year follow-up of the Renfrew/Paisley study,” *Am. J. Med.*, vol. 113, no. 5, pp. 359–364, 2002, doi: 10.1016/S0002-9343(02)01236-6.

[28] G. Thrall, G. Y. H. Lip, D. Carroll, and D. Lane, “Depression, anxiety, and quality of life in patients with atrial fibrillation,” *Chest*, vol. 132, no. 4, pp. 1259–1264, 2007, doi: 10.1378/chest.07-0036.

[29] S. Lévy, “Classification system of atrial fibrillation,” *Curr. Opin. Cardiol.*, vol. 15, no. 1, pp. 54–57, 2000, doi: 10.1097/00001573-200001000-00007.

[30] M. P. Turakhia *et al.*, “Diagnostic utility of a novel leadless arrhythmia monitoring device,” *Am. J. Cardiol.*, vol. 112, no. 4, pp. 520–524, 2013, doi: 10.1016/j.amjcard.2013.04.017.

[31] M. Pankaala *et al.*, “Detection of atrial fibrillation with seismocardiography,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2016-Octob, pp. 4369–4374, 2016, doi: 10.1109/EMBC.2016.7591695.

[32] F. Enseleit and F. Duru, “Long-term continuous external electrocardiographic recording: A review,” *Europace*, vol. 8, no. 4, pp. 255–266, 2006, doi: 10.1093/europace/euj054.

[33] M. Yenikomshian *et al.*, “Cardiac arrhythmia detection outcomes among patients monitored with the Zio patch system: a systematic literature review,” *Curr. Med. Res. Opin.*, vol. 35, no. 10, pp. 1659–1670, 2019, doi: 10.1080/03007995.2019.1610370.

[34] J. Anczykowski, S. Willems, B. A. Hoffmann, T. Meinertz, S. Blankenberg, and M. Patten, “Early Detection of Symptomatic Paroxysmal Cardiac Arrhythmias by Trans-Telephonic ECG Monitoring: Impact on Diagnosis and Treatment of Atrial Fibrillation,” *J. Cardiovasc. Electrophysiol.*, vol. 27, no. 9, pp. 1032–1037, 2016, doi: 10.1111/jce.13025.

[35] L. G. Y, F. B, D. C. R, and P. T.S, “Stroke prevention in atrial fibrillation: Past, present and future. Comparing the guidelines and practical decision-making,” *Thromb. Haemost.*, vol. 117, no. 7, pp. 1230–1239, 2017, doi: 10.1160/TH16-11-0876.

[36] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiol. Meas.*, vol. 28, no. 3, 2007, doi: 10.1088/0967-3334/28/3/R01.

[37] A. G. Bonomi *et al.*, “Atrial fibrillation detection using photo-plethysmography and acceleration data at the wrist,” *Comput. Cardiol. (2010).*, vol. 43, no. 2016, pp. 277–280, 2016, doi: 10.22489/cinc.2016.081-339.

[38] C. Yang, C. Veiga, J. J. Rodríguez-Andina, J. Fariña, A. Iñiguez, and S. Yin, “Using PPG Signals and Wearable Devices for Atrial Fibrillation Screening,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 11, pp. 8832–8842, 2019, doi: 10.1109/TIE.2018.2889614.

[39] S. Nemati *et al.*, “Monitoring and detecting atrial fibrillation using wearable technology,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2016-Octob, pp. 3394–3397, 2016, doi: 10.1109/EMBC.2016.7591456.

[40] A. Sološenko, A. Petrėnas, V. Marozas, and L. Sörnmo, “Modeling of the photoplethysmogram during atrial fibrillation,” *Comput. Biol. Med.*, vol. 81, no. December 2016, pp. 130–138, 2017, doi: 10.1016/j.compbiomed.2016.12.016.

[41] A. Hochstadt, E. Chorin, S. Viskin, A. L. Schwartz, N. Lubman, and R. Rosso, “Continuous heart rate monitoring for automatic detection of atrial fibrillation with novel bio-sensing technology,” *J. Electrocardiol.*, vol. 52, pp. 23–27, 2019, doi: 10.1016/j.jelectrocard.2018.10.096.

[42] A. Aliamiri and Y. Shen, “Deep learning based atrial fibrillation detection using wearable photoplethysmography sensor,” *2018 IEEE EMBS Int. Conf. Biomed. Heal. Informatics, BHI 2018*, vol. 2018-Janua, no. March, pp. 442–445, 2018, doi: 10.1109/BHI.2018.8333463.

[43] T. Sanna *et al.*, “Cryptogenic Stroke and Underlying Atrial Fibrillation,” *N. Engl. J. Med.*, vol. 370, no. 26, pp. 2478–2486, 2014, doi: 10.1056/nejmoa1313600.

[44] K. Harris, D. Edwards, and J. Mant, “How can we best detect atrial fibrillation?,” *J. R. Coll. Physicians Edinb.*, no. 42, pp. 5–22, 2012.

[45] M. George, “A new method for detecting atrial fibrillation using RR intervals,” *Comput. Cardiol.*, pp. 227–230, 1983, Accessed: Mar. 26, 2021. [Online]. Available: https://ci.nii.ac.jp/naid/80014474876/.

[46] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.,” *Circulation*, vol.

101, no. 23, 2000, doi: 10.1161/01.cir.101.23.e215.

[47] A. J. Camm *et al.*, “Guidelines for the management of atrial fibrillation,” *Eur. Heart J.*, vol. 31, no. 19, pp. 2369–2429, 2010, doi: 10.1093/eurheartj/ehq278.

[48] M. Budeus, M. Hennersdorf, C. Perings, and B. E. Strauer, “Der nachweis atrialer spätpotentiale mittels P-wellen-signalmittelungs-EKG bei patienten mit paroxysmalem vorhofflimmern,” *Z. Kardiol.*, vol. 92, no. 5, pp. 362–369, 2003, doi: 10.1007/s00392-003-0921-8.

[49] H. Khorrami and M. Moavenian, “A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification,” *Expert Syst. Appl.*, vol. 37, no. 8, pp. 5751–5757, 2010, doi: 10.1016/j.eswa.2010.02.033.

[50] F. Censi, G. Calcagnini, E. Mattei, A. Gargaro, G. Biancalana, and A. Capucci, “Simulation of monitoring strategies for atrial arrhythmia detection,” *Ann. dell’Istituto Super. di Sanita*, vol. 49, pp. 176–182, 2013, doi: 10.4415/ANN.

[51] S. Petrutiu, J. Ng, G. M. Nijm, H. Al-Angari, S. Swiryn, and A. V. Sahakian, “Atrial fibrillation and waveform characterization: A time domain perspective in the surface ECG,” *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 6, pp. 24–30, 2006, doi: 10.1109/EMB-M.2006.250505.

[52] K. Tateno and L. Glass, “A method for detection of atrial fibrillation using RR intervals,” *Comput. Cardiol.*, vol. 27, pp. 391–394, 2000.

[53] M. M. A. Hadhoud, M. I. Eladawy, and A. Farag, “Computer aided diagnosis of cardiac arrhythmias,” *2006 Int. Conf. Comput. Eng. Syst. ICCES’06*, pp. 262–265, 2006, doi: 10.1109/ICCES.2006.320458.

[54] Y. Hagiwara *et al.*, “Computer-aided diagnosis of atrial fibrillation based on ECG Signals: A review,” *Inf. Sci. (Ny).*, vol. 467, pp. 99–114, 2018, doi: 10.1016/j.ins.2018.07.063.

[55] C. W. Israel, G. Grönefeld, J. R. Ehrlich, Y. G. Li, and S. H. Hohnloser, “Long-Term Risk of Recurrent Atrial Fibrillation as Documented by an Implantable Monitoring Device: Implications for Optimal Patient Care,” *J. Am. Coll. Cardiol.*, vol. 43, no. 1, pp. 47–52, 2004, doi: 10.1016/j.jacc.2003.08.027.

[56] R. L. Page, W. E. Wilkinson, W. K. Clair, E. A. McCarthy, and E. L. C. Pritchett, “Asymptomatic arrhythmias in patients with symptomatic paroxysmal atrial fibrillation and paroxysmal supraventricular tachycardia,” *Circulation*, vol. 89, no. 1, pp. 224–227, 1994, doi: 10.1161/01.CIR.89.1.224.

[57] S. R. Steinhubl, E. D. Muse, and E. J. Topol, “The emerging field of mobile health,” *Sci. Transl. Med.*, vol. 7, no. 283, pp. 1–7, 2015, doi: 10.1126/scitranslmed.aaa3487.

[58] B. M. C. Silva, J. J. P. C. Rodrigues, I. de la Torre Díez, M. López-Coronado, and K. Saleem, “Mobile-health: A review of current state in 2015,” *J. Biomed. Inform.*, vol. 56, pp. 265–272, 2015, doi: 10.1016/j.jbi.2015.06.003.

[59] P. Zha *et al.*, “Utilizing a Mobile Health Intervention to Manage Hypertension in an Underserved Community,” *West. J. Nurs. Res.*, vol. 42, no. 3, pp. 201–209, 2020, doi: 10.1177/0193945919847937.

[60] A. B. del Río-Lanza, A. Suárez-Vázquez, L. Suárez-Álvarez, and V. Iglesias-Argüelles, “Mobile health (mhealth): facilitators and barriers of the intention of use in patients with chronic illnesses,” *J. Commun. Healthc.*, vol. 13, no. 2, pp. 138–146, 2020, doi: 10.1080/17538068.2020.1777513.

[61] K. Källander *et al.*, “Mobile health (mhealth) approaches and lessons for increased performance and retention of community health workers in lowand middle-income countries: A review,” *J. Med. Internet Res.*, vol. 15, no. 1, 2013, doi: 10.2196/jmir.2130.

[62] C. Dèglise, L. S. Suggs, and P. Odermatt, “Short message service (SMS) applications for disease prevention in developing countries,” *J. Med. Internet Res.*, vol. 14, no. 1, pp. 1–20, 2012, doi: 10.2196/jmir.1823.

[63] Y. Guo *et al.*, “Mobile Health Technology to Improve Care for Patients With Atrial Fibrillation,” *J. Am. Coll. Cardiol.*, vol. 75, no. 13, pp. 1523–1534, 2020, doi: 10.1016/j.jacc.2020.01.052.

[64] Y. Guo, Y. Chen, D. A. Lane, L. Liu, Y. Wang, and G. Y. H. Lip, “Mobile Health Technology for Atrial Fibrillation Management Integrating Decision Support, Education, and Patient Involvement: mAF App Trial,” *Am. J. Med.*, vol. 130, no.

12, pp. 1388-1396.e6, 2017, doi: 10.1016/j.amjmed.2017.07.003.

[65] Y. Guo *et al.*, "Mobile Photoplethysmographic Technology to Detect Atrial Fibrillation," *J. Am. Coll. Cardiol.*, vol. 74, no. 19, pp. 2365–2375, 2019, doi: 10.1016/j.jacc.2019.08.019.

[66] D. V. Dimitrov, "Medical internet of things and big data in healthcare," *Healthc. Inform. Res.*, vol. 22, no. 3, pp. 156–163, 2016, doi: 10.4258/hir.2016.22.3.156.

[67] and A. S. Shelke, Yogesh, "INTERNET OF MEDICAL THINGS," *Technol. Intell. IP Res. aranca*, 2016.

[68] M. Usak, M. Kubiatko, M. S. Shabbir, O. V. Dudnik, K. Jermsittiparsert, and L. Rajabion, "Health care service delivery based on the Internet of things: A systematic and comprehensive study," *Int. J. Commun. Syst.*, vol. 33, no. 2, pp. 1–17, 2020, doi: 10.1002/dac.4179.

[69] V. L. Patel *et al.*, "The coming of age of artificial intelligence in medicine," *Artif. Intell. Med.*, vol. 46, no. 1, pp. 5–17, 2009, doi: 10.1016/j.artmed.2008.07.017.

[70] B. Farahani, M. Barzegari, F. Shams Aliee, and K. A. Shaik, "Towards collaborative intelligent IoT eHealth: From device to fog, and cloud," *Microprocess. Microsyst.*, vol. 72, p. 102938, 2020, doi: 10.1016/j.micpro.2019.102938.

[71] A. Sarraju and M. P. Turakhia, "Wearables for arrhythmia care: Challenges and futureprospects," *Cardiovasc. Digit. Heal. J.*, vol. 1, no. 2, pp. 56–58, 2020, doi: 10.1016/j.cvdhj.2020.09.001.

[72] S. E. Dilsizian and E. L. Siegel, "Artificial intelligence in medicine and cardiac imaging: Harnessing big data and advanced computing to provide personalized medical diagnosis and treatment," *Curr. Cardiol. Rep.*, vol. 16, no. 1, 2014, doi: 10.1007/s11886-013-0441-8.

[73] O. Faust, U. R. Acharya, and T. Tamura, "Formal design methods for reliable computer-aided diagnosis: A review," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 15–28, 2012, doi: 10.1109/RBME.2012.2184750.

[74] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K. S. Kwak, "The internet

of things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015, doi: 10.1109/ACCESS.2015.2437951.

[75] B. J. Drew *et al.*, "Practice standards for electrocardiographic monitoring in hospital settings," *Circulation*, vol. 110, no. 17, pp. 2721–2746, 2004, doi: 10.1161/01.CIR.0000145144.56673.59.

[76] B. J. Drew, "Electrocardiogram monitoring," *AACN Clin. Issues Crit. Care Nurs.*, vol. 46, no. 4, pp. 251–260, 2002, doi: 10.1097/00044067-199302000-00003.

[77] M. A. Hlatky *et al.*, "Criteria for evaluation of novel markers of cardiovascular risk: A scientific statement from the American heart association," *Circulation*, vol. 119, no. 17, pp. 2408–2416, 2009, doi: 10.1161/CIRCULATIONAHA.109.192278.

[78] A. J. Atkinson *et al.*, "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," *Clin. Pharmacol. Ther.*, vol. 69, no. 3, pp. 89–95, 2001, doi: 10.1067/mcp.2001.113989.

[79] U. R. Acharya *et al.*, "An integrated diabetic index using heart rate variability signal features for diagnosis of diabetes," *Comput. Methods Biomech. Biomed. Engin.*, vol. 16, no. 2, pp. 222–234, 2013, doi: 10.1080/10255842.2011.616945.

[80] N. A. Abdul-Kadir, N. Mat Safri, and M. A. Othman, "Dynamic ECG features for atrial fibrillation recognition," *Comput. Methods Programs Biomed.*, vol. 136, pp. 143–150, 2016, doi: 10.1016/j.cmpb.2016.08.021.

[81] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Comput. J.*, vol. 7, no. 2, pp. 149–154, 1964, [Online]. Available: http://dx.doi.org/10.1016/j.biotechadv.2010.07.003%0Ahttp://dx.doi.org/10.1016/j.scitotenv.2016.06.080%0Ahttp://dx.doi.org/10.1016/j.bbapap.2013.06.007%0Ahttps://www.frontiersin.org/article/10.3389/fmicb.2018.02309/full%0Ahttp://dx.doi.org/10.1007/s13762-.

[82] J. W. Manz, "A Sequency-Ordered Fast Walsh Transform," *IEEE Trans. Audio Electroacoust.*, vol. 20, no. 3, pp. 204–205, 1972, doi: 10.1109/TAU.1972.1162377.

[83] T. K. Boehme and R. Bracewell, "The Fourier Transform and its Applications.,"

Am. Math. Mon., vol. 73, no. 6, p. 685, 1966, doi: 10.2307/2314845.

[84] P. S. Addison, "Wavelet transforms and the ECG: A review," *Physiol. Meas.*, vol. 26, no. 5, 2005, doi: 10.1088/0967-3334/26/5/R01.

[85] O. Faust, R. Acharya U, S. M. Krishnan, and L. C. Min, "Analysis of cardiac signals using spatial filling index and time-frequency domain," *Biomed. Eng. Online*, vol. 3, pp. 1–11, 2004, doi: 10.1186/1475-925X-3-30.

[86] P. R. Gomes, F. O. Soares, J. H. Correia, and C. S. Lima, "ECG data-acquisition and classification system by using Wavelet-domain Hidden Markov Models," *2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC'10*, pp. 4670–4673, 2010, doi: 10.1109/IEMBS.2010.5626456.

[87] B. Weng, J. J. Wang, F. Michaud, and M. Blanco-Velasco, "Atrial Fibrillation Detection Using Stationary Wavelet Transform Analysis," *30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, doi: 10.22489/CinC.2017.210-084.

[88] G. P. Nason and B. W. Silverman, "The Stationary Wavelet Transform and some Statistical Applications," *Springer*, pp. 281–299, 1995, doi: 10.1007/978-1-4612-2544-7_17.

[89] S. Pal and M. Mitra, "Empirical mode decomposition based ECG enhancement and QRS detection," *Comput. Biol. Med.*, vol. 42, no. 1, pp. 83–92, 2012, doi: 10.1016/j.compbiomed.2011.10.012.

[90] B. Bradie, "Wavelet packet-based compression of single lead ECG," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 5, pp. 493–501, 1996, doi: 10.1109/10.488797.

[91] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete Cosine Transform," *IEEE Trans. Comput.*, vol. C–23, no. 1, pp. 90–93, 1974, doi: 10.1109/T-C.1974.223784.

[92] M. Unser and W. Aldroubi, "A Review of Wavelets i Biomedical Ap," *Rev. Lit. Arts Am.*, vol. 84, no. 4, 1996.

[93] H. Zhao, S. Lu, R. Zou, K. Ju, and K. H. Chon, "Estimation of time-varying coherence function using time-varying transfer functions," *Ann. Biomed. Eng.*, vol. 33, no. 11, pp. 1582–1594, 2005, doi: 10.1007/s10439-005-7045-4.

[94] U. R. Acharya *et al.*, "Linear and nonlinear analysis of normal and CAD-affected

heart rate signals," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 55–68, 2014, doi: 10.1016/j.cmpb.2013.08.017.

[95] T. Schreiber and A. Schmitz, "Improved surrogate data for nonlinearity tests," *Phys. Rev. Lett.*, vol. 77, no. 4, pp. 635–638, Sep. 1996, doi: 10.1103/physrevlett.77.635.

[96] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Process. Mag.*, no. July, 1993.

[97] B. B. Mandelbrot and J. A. Wheeler, "The Fractal Geometry of Nature," *American Journal of Physics*, vol. 51, no. 3. pp. 286–287, 1983, doi: 10.1119/1.13295.

[98] E. Czogała and J. ŁĘSki, "Application of entropy and energy measures of fuzziness to processing of ECG signal," *Fuzzy Sets Syst.*, vol. 97, no. 1, pp. 9–18, 1998, doi: 10.1016/S0165-0114(96)00359-4.

[99] J. P. Eckmann, O. Oliffson Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *World Sci. Ser. Nonlinear Sci. Ser. A*, vol. 16, pp. 441–446, 1995, doi: 10.1209/0295-5075/4/9/004.

[100] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "A practical method for calculating largest Lyapunov exponents from small data sets Michael," *Phys. D Nonlinear Phenom.*, vol. 65.1, no. 2, pp. 117–134, 1993.

[101] R. J. Martis, U. R. Acharya, C. M. Lim, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Application of higher order cumulant features for cardiac health diagnosis using ECG signals," *Int. J. Neural Syst.*, 2013, doi: 10.1142/S0129065711002912.

[102] R. J. Martis, U. R. Acharya, H. Prasad, C. K. Chua, C. M. Lim, and J. S. Suri, "Application of higher order statistics for atrial arrhythmia classification," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 888–900, 2013, doi: 10.1016/j.bspc.2013.08.008.

[103] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," *Comput. Biol. Med.*, vol. 48, no. 1, pp. 133–149, 2014, doi: 10.1016/j.compbiomed.2014.02.012.

[104] C. M. Bishop, “Pattern recognition,” *Springer*, vol. 9, no. 128, 2006, doi: 10.1109/TIT.1963.1057854.

[105] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967, vol. 1, no. 14, pp. 281–297, [Online]. Available: http://books.google.de/books?hl=de&lr=&id=IC4Ku_7dBFUC&oi=fnd&pg=PA281&dq=MacQueen+some+methods+for+classification&ots=nNTcK1IdoQ&sig=fHzdVcbvmYJ-lTNHu1HncmOFOkM#v=onepage&q=MacQueen some methods for classification&f=false.

[106] Y. Wang, Y. S. Zhu, N. V. Thakor, and Y. H. Xu, “A short-time multifractal approach for arrhythmia detection based on fuzzy neural network,” *IEEE Trans. Biomed. Eng.*, vol. 48, no. 9, pp. 989–995, 2001, doi: 10.1109/10.942588.

[107] I. Goodfellow *et al.*, “Generative adversarial nets,” *Adv. Neural Inf. Process. Syst.*, 2014.

[108] L. Devroye and T. Wagner, “Distribution-Free Performance Bounds for Potential Function Rules,” vol. I, no. 2, pp. 601-604., 1979.

[109] R. O. Duda and P. E. Hart, “Pattern Classification and Scene Analysis,” *J. R. Stat. Soc. Ser. A*, vol. 3, 1973, doi: 10.2307/2344977.

[110] D. . Specht, “Probabilistic Neural Networks,” *Nat. Hazards*, vol. 3, p. 109 118, 1990, doi: 10.1007/s11069-015-1595-z.

[111] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene selection for cancer classification using DCA,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 46, no. 1, pp. 389–422, 2002, doi: 10.1007/978-3-540-88192-6_8.

[112] L. Breiman and A. Cutler, “State-of-the-Art Data Mining Using Random Forests State-of-the-Art Data Mining Using Random Forests,” in *Proceedings of the Salford Data Mining Conference*, 2012, pp. 24–25.

[113] J. J. Rodríguez, L. I. Kuncheva, and C. J. Alonso, “Rotation forest: A New classifier ensemble method,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1619–1630, 2006, doi: 10.1109/TPAMI.2006.211.

[114] D. W. Marquardt, “An Algorithm for Least-Squares Estimation of Nonlinear Parameters,” *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[115] T. M. Cover and P. E. Hart, “Nearest Neighbor Pattern Classification,” *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[116] F. Sufi and I. Khalil, “Diagnosis of cardiovascular abnormalities from compressed ECG: A data mining-based approach,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 33–39, 2011, doi: 10.1109/TITB.2010.2094197.

[117] A. Poznanski and L. Wolf, “CNN-N-Gram for HandwritingWord Recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2305–2314, 2016, doi: 10.1109/CVPR.2016.253.

[118] A. Graves, A. Mohamed, and G. Hinton, “Speech Recognition with Deep Recurrent Neural Networks , Department of Computer Science, University of Toronto,” in *IEEE international conference on acoustics, speech and signal processing*, 2013, vol. 3, no. 3, pp. 45–49, [Online]. Available: https://ieeexplore.ieee.org/stampPDF/getPDF.jsp?tp=&arnumber=6638947&ref=aHR0cHM6Ly9pZWVleHBsb3JlLmllZWUub3JnL2Fic3RyYWN0L2RvY3VtZW50LzY2Mzg5NDc/Y2FzYV90b2tlbj1OQUo1VFJxWk5JRUFBQUFBOmtPZmdDbS00NGhqaGI2N3dMd2JrU3lSaEdJREhBWnpMSkxoT201Um5YMXR0S0poUDAtM2hkbT.

[119] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, “Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2015, vol. 2015-Augus, pp. 4580–4584, doi: 10.1109/ICASSP.2015.7178838.

[120] M. Tomas, “Recurrent neural network based language model ´ s Mikolov Introduction Comparison and model combination Future work,” in *Eleventh annual conference of the international speech communication association*, 2010, pp. 1–24.

[121] Y. Zhang and B. Wallace, “A Sensitivity Analysis of (and Practitioners’ Guide to) Convolutional Neural Networks for Sentence Classification,” 2015, [Online].

Available: http://arxiv.org/abs/1510.03820.

[122] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, “Unsupervised learning of hierarchical representations with convolutional deep belief networks,” in *In Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 609–616, doi: 10.1145/2001269.2001295.

[123] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, “Long Short Term Memory Networks for Anomaly Detection in Time Series,” in *ESANN 2015 proceedings, 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2015, pp. 89–94.

[124] A. L. Waldo, “Atrial fibrillation and atrial flutter: Two sides of the same coin,” *Int. J. Cardiol.*, vol. 240, pp. 251–252, 2017, doi: 10.1016/j.ijcard.2017.02.146.

[125] M. D. Ivanovic, V. Atanasoski, A. Shvilkin, L. Hadzievski, and A. Maluckov, “Deep Learning Approach for Highly Specific Atrial Fibrillation and Flutter Detection based on RR Intervals,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 1780–1783, 2019, doi: 10.1109/EMBC.2019.8856806.

[126] V. Fuster *et al.*, *ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation: Executive summary a report of the american college of cardiology/american heart association task force on practice guidelines and the european society of cardiology committee* , vol. 38, no. 4. 2001.

[127] B. Taha, S. Reddy, Q. Xue, and S. Swiryn, “Automated discrimination between atrial fibrillation and atrial flutter in the resting 12-lead electrocardiogram,” *J. Electrocardiol.*, vol. 33, pp. 123–125, 2000, doi: 10.1054/jelc.2000.20303.

[128] H. Fujita and D. Cimr, “Decision support system for arrhythmia prediction using convolutional neural network structure without preprocessing,” *Appl. Intell.*, vol. 49, no. 9, pp. 3383–3391, 2019, doi: 10.1007/s10489-019-01461-0.

[129] Y. Zhang and L. Zhang, “Research on Algorithm for Detecting Atrial Fibrillation Using RR Intervals,” *2018 Int. Conf. Audio, Lang. Image Process.*, pp. 205–209, 2018.

[130] R. S. H. Istepanian and T. AlAnzi, *Mobile health (m-health)*. Elsevier Inc., 2020.

[131] A. Abbas and S. U. Khan, “A review on the state-of-the-art privacy-preserving approaches in the e-Health clouds,” *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 4, pp. 1431–1441, 2014, doi: 10.1109/JBHI.2014.2300846.

[132] C. de Grood, A. Raissi, Y. Kwon, and M. J. Santana, “Adoption of e-health technology by physicians: A scoping review,” *J. Multidiscip. Healthc.*, vol. 9, pp. 335–344, 2016, doi: 10.2147/JMDH.S103881.

[133] K. Sekido *et al.*, “Exploring Hidden Atrial Fibrillation in Patients with Type 2 Diabetes Mellitus Admitted to Shinshu University Hospital,” *Shinshu Med. J.*, vol. 69, no. 2, pp. 75–81, 2021, doi: 10.11441/shinshumedj.69.75.

[134] Y. S. Baek, S. C. Lee, W. Choi, and D. H. Kim, “A new deep learning algorithm of 12-lead electrocardiogram for identifying atrial fibrillation during sinus rhythm,” *Sci. Rep.*, vol. 11, no. 1, pp. 1–10, 2021, doi: 10.1038/s41598-021-92172-5.

[135] S. M. Al-Khatib *et al.*, “2017 AHA/ACC/HRS Guideline for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death,” 2018. doi: 10.1016/j.jacc.2017.10.054.

[136] H. L. Kennedy, “The history, science, and innovation of Holter technology,” *Ann. Noninvasive Electrocardiol.*, vol. 11, no. 1, pp. 85–94, 2006, doi: 10.1111/j.1542-474X.2006.00067.x.

[137] C. M. Mba *et al.*, “Short term optimization of glycaemic control using insulin improves sympatho-vagal tone activities in patients with type 2 diabetes,” *Diabetes Res. Clin. Pract.*, vol. 157, 2019, doi: 10.1016/j.diabres.2019.107875.

[138] M. A. Rosenberg, M. Samuel, A. Thosani, and P. J. Zimetbaum, “Use of a noninvasive continuous monitoring device in the management of atrial fibrillation: A pilot study,” *PACE - Pacing Clin. Electrophysiol.*, vol. 36, no. 3, pp. 328–333, 2013, doi: 10.1111/pace.12053.

[139] P. M. Barrett *et al.*, “Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring,” *Am. J. Med.*, vol. 127, no. 1, pp. 95.e11-95.e17, 2014, doi: 10.1016/j.amjmed.2013.10.003.

[140] S. Pradhan, J. A. Robinson, J. K. Shivapour, and C. S. Snyder, “Ambulatory

Arrhythmia Detection with ZIO® XT Patch in Pediatric Patients: A Comparison of Devices," *Pediatr. Cardiol.*, vol. 40, no. 5, pp. 921–924, 2019, doi: 10.1007/s00246-019-02089-0.

[141] A. Janata *et al.*, "Quality of ECG monitoring with a miniature ECG recorder," *PACE - Pacing Clin. Electrophysiol.*, vol. 31, no. 6, pp. 676–684, 2008, doi: 10.1111/j.1540-8159.2008.01070.x.

[142] E. Vavrinsky, J. Subjak, M. Donoval, A. Wagner, T. Zavodnik, and H. Svobodova, "Application of modern multi-sensor holter in diagnosis and treatment," *Sensors (Switzerland)*, vol. 20, no. 9, 2020, doi: 10.3390/s20092663.

[143] J. P. Tsang and S. Mohan, "Benefits of monitoring patients with mobile cardiac telemetry (MCT) compared with the Event or Holter monitors," *Med. Devices Evid. Res.*, vol. 7, no. 1, pp. 1–5, 2013, doi: 10.2147/MDER.S54038.

[144] C. Jansen *et al.*, "Significant reduction in heart rate variability is a feature of acute decompensation of cirrhosis and predicts 90-day mortality," *Aliment. Pharmacol. Ther.*, vol. 50, no. 5, pp. 568–579, 2019, doi: 10.1111/apt.15365.

[145] C. L. Duus *et al.*, "Continuous vital sign monitoring after major abdominal surgery—Quantification of micro events," *Acta Anaesthesiol. Scand.*, vol. 62, no. 9, pp. 1200–1208, 2018, doi: 10.1111/aas.13173.

[146] M. Elvekjaer *et al.*, "Physiological abnormalities in patients admitted with acute exacerbation of COPD: an observational study with continuous monitoring," *J. Clin. Monit. Comput.*, vol. 34, no. 5, pp. 1051–1060, 2020, doi: 10.1007/s10877-019-00415-8.

[147] K. Chon and D. M. Cardiology, "Detection of atrial fibrillation using a smartwatch," *Nat. Rev. Cardiol.*, vol. 15, no. 11, pp. 657–658, 2018, Accessed: Aug. 18, 2021. [Online]. Available: https://www.nature.com/articles/s41569-018-0057-1.

[148] T. Inui *et al.*, "Use of a Smart Watch for Early Detection of Paroxysmal Atrial Fibrillation: Validation Study," *JMIR Cardio*, vol. 4, no. 1, pp. 1–13, 2020, doi: 10.2196/14857.

[149] G. H. Tison *et al.*, "Passive detection of atrial fibrillation using a commercially

available smartwatch," *JAMA Cardiol.*, vol. 3, no. 5, pp. 409–416, 2018, doi: 10.1001/jamacardio.2018.0136.

[150] D. Hernando, S. Roca, J. Sancho, Á. Alesanco, and R. Bailón, "Validation of the apple watch for heart rate variability measurements during relax and mental stress in healthy subjects," *Sensors*, vol. 18, no. 8, 2018, doi: 10.3390/s18082619.

[151] S. Shareghi, M. Tavakol, K. Lindborg, and C. Alfaro Vives, "SEEQ mobile cardiac telemetry associated with a high yield of clinically relevant arrhythmias in patients with suspected arrhythmia," *Circulation*, 2016, Accessed: Aug. 18, 2021. [Online]. Available: https://www.ahajournals.org/doi/abs/10.1161/circ.134.suppl_1.16078.

[152] D. Vanegas-Cadavid, Z. Valderrama-Barbosa, and L. Ibatá-Bernal, "Clinical experience in extended cardiac monitoring with the SEEQ^{TM} satellite wireless system," *Rev. Colomb. Cardiol.*, vol. 25, no. 3, pp. e17–e25, 2018, doi: 10.1016/j.rccar.2017.09.008.

[153] J. M. Engel, N. Chakravarthy, G. Nosbush, M. Merkert, M. D. Richard Fogoros, and A. Chavan, "Comparison of arrhythmia prevalence in NUVANT Mobile Cardiac Telemetry System patients in the US and India," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2014, pp. 2730–2733, doi: 10.1109/EMBC.2014.6944187.

[154] J. M. Engel, V. Mehta, R. Fogoros, and A. Chavan, "Study of arrhythmia prevalence in NUVANT mobile cardiac telemetry system patients," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2012, pp. 2440–2443, doi: 10.1109/EMBC.2012.6346457.

[155] R. P. B. Tonino, K. Larimer, O. Eissen, and M. R. Schipperus, "Remote patient monitoring in adults receiving transfusion or infusion for hematological disorders using the VitalPatch and accelerateIQ monitoring system: quantitative feasibility study," *JMIR Hum. Factors*, vol. 6, no. 4, pp. 1–7, 2019, doi: 10.2196/15103.

[156] P. Langley *et al.*, "Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings," *Med. Eng. Phys.*, vol. 34, no. 10, pp. 1441–1447, 2012, doi: 10.1016/j.medengphy.2012.02.002.

[157] R. Kohavi, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” *Int. Jt. Conf. Artif. Intell.*, no. March 2001, 1995.

[158] C. Song, K. Liu, X. Zhang, L. Chen, and X. Xian, “An Obstructive Sleep Apnea Detection Approach Using a Discriminative Hidden Markov Model from ECG Signals,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1532–1542, 2016, doi: 10.1109/TBME.2015.2498199.

[159] O. Faust, “Documenting and predicting topic changes in Computers in Biology and Medicine: A bibliometric keyword analysis from 1990 to 2017,” *Informatics Med. Unlocked*, vol. 11, pp. 15–27, Jan. 2018, doi: 10.1016/j.imu.2018.03.002.

[160] J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61. Elsevier Ltd, pp. 85–117, Jan. 01, 2015, doi: 10.1016/j.neunet.2014.09.003.

[161] Y. Bengio, P. Simard, and P. Frasconi, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” *IEEE Trans. Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994, doi: 10.1109/72.279181.

[162] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[163] M. Schuster and K. K. Paliwal, “Bidirectional Recurrent Neural Networks,” *IEEE Trans. SIGNAL Process.*, vol. 45, no. 11, 1997, Accessed: May 03, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/650093/.

[164] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, 2005, doi: 10.1016/j.neunet.2005.06.042.

[165] F. Chollet, “Keras: Deep learning library for theano and tensorflow,” *https://keras.io/k,7(8)*, 2015. https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=F.+Chollet%2C+et+al.%2C+Keras%3A+Deep+Learning+Library+for+Theano+and+Tensorflow%2C+%282015%29%2Cp.+8+++https%3A%2F%2Fkeras.io%2Fk&btnG=&authuser=1 (accessed May 03, 2021).

[166] A. Gulli and S. Pal, *Deep Learning With Keras*. Packt Publishing Ltd, 2017.

[167] M. Abadi *et al.*, “TensorFlow: A system for large-scale machine learning,” in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016*, 2016, pp. 265–283.

[168] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” 2010, Accessed: May 03, 2021. [Online]. Available: http://www.iro.umontreal.

[169] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” pp. 1–16, 2016.

[170] Y. Gal and Z. Ghahramani, “Bayesian Convolutional Neural Networks with Bernoulli Approximate Variational Inference,” pp. 1–12, 2015, [Online]. Available: http://arxiv.org/abs/1506.02158.

[171] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014, Accessed: May 03, 2021. [Online]. Available: https://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf?utm_campaign=buffer&utm_content=buffer79b43&utm_medium=social&utm_source=twitter.com.

[172] J. F. Mas, “Receiver Operating Characteristic (ROC) Analysis,” 2018, pp. 465–467.

[173] S. S. Chugh *et al.*, “Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study,” *Circulation*, vol. 129, no. 8, pp. 837–847, 2014, doi: 10.1161/CIRCULATIONAHA.113.005119.

[174] J. S. Healey *et al.*, “Subclinical Atrial Fibrillation and the Risk of Stroke,” *N. Engl. J. Med.*, vol. 366, no. 2, pp. 120–129, 2012, doi: 10.1056/nejmoa1105575.

[175] C. Pahl and B. Lee, “Containers and clusters for edge cloud architectures-A technology review,” *Proc. - 2015 Int. Conf. Futur. Internet Things Cloud, FiCloud 2015 2015 Int. Conf. Open Big Data, OBD 2015*, pp. 379–386, 2015, doi: 10.1109/FiCloud.2015.35.

[176] E. Lashgari, D. Liang, and U. Maoz, “Data augmentation for deep-learning-based electroencephalography,” *Journal of Neuroscience Methods*, vol. 346. Elsevier

B.V., p. 108885, Dec. 01, 2020, doi: 10.1016/j.jneumeth.2020.108885.

[177] J. M. Johnson and T. M. Khoshgoftaar, “Survey on deep learning with class imbalance,” *J. Big Data*, vol. 6, no. 1, pp. 1–54, Dec. 2019, doi: 10.1186/s40537-019-0192-5.

[178] J. Zheng, J. Zhang, S. Danioko, H. Yao, H. Guo, and C. Rakovski, “A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients,” *Sci. Data*, vol. 7, no. 1, pp. 1–8, 2020, doi: 10.1038/s41597-020-0386-x.

[179] A. J. Demski and M. L. Soria, “ecg-kit a Matlab Toolbox for Cardiovascular Signal Processing,” *J. Open Res. Softw.*, vol. 4, pp. 2–5, 2016, doi: 10.5334/jors.86.

[180] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, “A Wavelet-Based ECG Delineator Evaluation on Standard Databases,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, 2004, doi: 10.1109/TBME.2003.821031.

[181] T. Fushiki, “Estimation of prediction error by using K-fold cross-validation,” *Stat. Comput.*, vol. 21, no. 2, pp. 137–146, 2011, doi: 10.1007/s11222-009-9153-8.

[182] O. Faust, R. Barika, A. Shenfield, E. J. Ciaccio, and U. R. Acharya, “Accurate detection of sleep apnea with long short-term memory network based on RR interval signals,” *Knowledge-Based Syst.*, vol. 212, p. 106591, 2021, doi: 10.1016/j.knosys.2020.106591.

[183] A. C. Fisher, A. Eleuteri, D. Groves, and C. J. Dewhurst, “The Ornstein-Uhlenbeck third-order Gaussian process (OUGP) applied directly to the un-resampled heart rate variability (HRV) tachogram for detrending and low-pass filtering,” *Med. Biol. Eng. Comput.*, vol. 50, no. 7, pp. 737–742, Jul. 2012, doi: 10.1007/s11517-012-0928-2.

[184] G. D. Clifford, “ECG statistics, noise, artifacts, and missing data,” *Adv. Methods Tools ECG Anal.*, pp. 55–99, 2006.

[185] P. Laguna, G. B. Moody, and R. G. Mark, “Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals,” *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 698–715, 1998, doi: 10.1109/10.678605.

[186] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, “Deep learning for time series classification: a review,” *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019, doi: 10.1007/s10618-019-00619-1.

[187] F. Chollet, “Keras, https://github. com/fchollet/keras,” 2015. .

[188] M. Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” 2016, [Online]. Available: http://arxiv.org/abs/1603.04467.

[189] B. Hanin, “Which neural net architectures give rise to exploding and vanishing gradients?,” *Adv. Neural Inf. Process. Syst.*, vol. 2018-Decem, no. iii, pp. 582–591, 2018.

[190] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.

[191] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[192] H. Fujita and D. Cimr, “Computer Aided detection for fibrillations and flutters using deep convolutional neural network,” *Inf. Sci. (Ny).*, vol. 486, pp. 231–239, 2019, doi: 10.1016/j.ins.2019.02.065.

[193] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, “Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network,” *Inf. Sci. (Ny).*, vol. 405, pp. 81–90, 2017, doi: 10.1016/j.ins.2017.04.012.

[194] N. Henzel, J. Wróbel, and K. Horoba, “Atrial fibrillation episodes detection based on classification of heart rate derived features,” *Proc. 24th Int. Conf. Mix. Des. Integr. Circuits Syst. Mix. 2017*, pp. 571–576, 2017, doi: 10.23919/MIXDES.2017.8005278.

[195] U. Desai, R. J. Martis, U. R. Acharya, C. G. Nayak, G. Seshikala, and R. Shetty K, “Diagnosis of multiclass tachycardia beats using recurrence quantification analysis and ensemble classifiers,” *J. Mech. Med. Biol.*, vol. 16, no. 1, 2016, doi: 10.1142/S0219519416400054.

[196] U. R. Acharya *et al.*, “Automated characterization of arrhythmias using nonlinear

features from tachycardia ECG beats," *2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc.*, pp. 533–538, 2017, doi: 10.1109/SMC.2016.7844294.

[197] H. Ibrahim and I. Owis Mohamed, "Automatic arrhythmia detection using support vector machine based on discrete wavelet transform," *J. Med. Imaging Heal. Informatics*, vol. 6, no. 1, pp. 204–209, 2016, doi: 10.1166/jmihi.2016.1611.

[198] Y. Xia, N. Wulan, K. Wang, and H. Zhang, "Detecting atrial fibrillation by deep convolutional neural networks," *Comput. Biol. Med.*, vol. 93, no. July 2017, pp. 84–92, 2018, doi: 10.1016/j.compbiomed.2017.12.007.

[199] A. Petrenas, V. Marozas, and L. Sörnmo, "Low-complexity detection of atrial fibrillation in continuous long-term monitoring," *Comput. Biol. Med.*, vol. 65, pp. 184–191, 2015, doi: 10.1016/j.compbiomed.2015.01.019.

[200] X. Zhou, H. Ding, B. Ung, E. Pickwell-MacPherson, and Y. Zhang, "Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy," *Biomed. Eng. Online*, vol. 13, no. 1, 2014, doi: 10.1186/1475-925X-13-18.

[201] A. Muthuchudar and S. S. Baboo, "A study of the processes involved in ecg signal analysis," *Int. J. Sci. Res. Publ.*, pp. 1–3, 2013, Accessed: Jun. 22, 2021. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.416.4304&rep=rep1&type=pdf#page=559.

[202] C. Yuan, Y. Yan, L. Zhou, J. Bai, and L. Wang, "Automated atrial fibrillation detection based on deep learning network," *2016 IEEE Int. Conf. Inf. Autom. IEEE ICIA 2016*, no. August, pp. 1159–1164, 2017, doi: 10.1109/ICInfA.2016.7831994.

[203] S. M. Pudukotai Dinakarrao and A. Jantsch, "ADDHard: Arrhythmia detection with digital hardware by learning ECG signal," *Proc. ACM Gt. Lakes Symp. VLSI, GLSVLSI*, pp. 495–498, 2018, doi: 10.1145/3194554.3194647.

[204] M. Salem, S. Taheri, and J. S. Yuan, "ECG Arrhythmia Classification Using Transfer Learning from 2- Dimensional Deep CNN Features," *2018 IEEE Biomed. Circuits Syst. Conf. BioCAS 2018 - Proc.*, 2018, doi: 10.1109/BIOCAS.2018.8584808.

[205] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Futur. Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013, doi: 10.1016/j.future.2013.01.010.

[206] A. M. Hay, “The derivation of global estimates from a confusion matrix,” *Int. J. Remote Sens.*, pp. 1395–1398, 1988, Accessed: Jun. 28, 2021. [Online]. Available: https://www.tandfonline.com/doi/pdf/10.1080/01431168808954945.

[207] D. M. W. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” *Int. J. Mach. Learn. Technol.*, pp. 37–63, 2011, [Online]. Available: http://arxiv.org/abs/2010.16061.

[208] U. R. Acharya, O. Faust, N. Adib Kadri, J. S. Suri, and W. Yu, “Automated identification of normal and diabetes heart rate signals using nonlinear measures,” *Comput. Biol. Med.*, vol. 43, no. 10, pp. 1523–1529, 2013, doi: 10.1016/j.compbiomed.2013.05.024.

[209] U. R. Acharya, S. V. Sree, and J. S. Suri, “Automatic detection of epileptic eeg signals using higher order cumulant features,” *Int. J. Neural Syst.*, vol. 21, no. 5, pp. 403–414, 2011, doi: 10.1142/S0129065711002912.

[210] M. Stickdorn, M. Hormess, A. Lawrence, and J. Schneider, “This is Service Design Doing: Applying Service Design Thinking in the Real World,” *This is Serv. Des. doing*, p. 1030, 2018, [Online]. Available: https://www.thisisservicedesigndoing.com/methods.

[211] S. Paszkiel, “Using BCI in IoT Implementation,” *Stud. Comput. Intell.*, vol. 852, pp. 111–128, 2020, doi: 10.1007/978-3-030-30581-9_13.

[212] J. Trinder *et al.*, “Autonomic activity during human sleep as a function of time and sleep stage,” *J. Sleep Res.*, vol. 10, no. 4, pp. 253–264, 2001, doi: 10.1046/j.1365-2869.2001.00263.x.

[213] B. Yu, L. Xu, and Y. Li, “Bluetooth Low Energy (BLE) based mobile electrocardiogram monitoring system,” *2012 IEEE Int. Conf. Inf. Autom. ICIA 2012*, no. June, pp. 763–767, 2012, doi: 10.1109/ICInfA.2012.6246921.

[214] S. Pasha, “Thingspeak Based Sensing and Monitoring System for IoT with Matlab Analysis,” *Int. J. New Technol. Res.*, vol. 2, no. 6, pp. 19–23, 2016.

[215] O. Faust, W. Yu, and U. Rajendra Acharya, “The role of real-time in biomedical science: A meta-analysis on computational complexity, delay and speedup,” *Comput. Biol. Med.*, vol. 58, pp. 73–84, 2015, doi: 10.1016/j.compbiomed.2014.12.024.

[216] O. S. Lih *et al.*, “Comprehensive electrocardiographic diagnosis based on deep learning,” *Artif. Intell. Med.*, vol. 103, no. November 2019, p. 101789, 2020, doi: 10.1016/j.artmed.2019.101789.

[217] J. Ramshur, “Design, Evaluation, and Application of Heart Rate Variability Analysis Software (HRVAS),” no. April, 2010, doi: 10.13140/RG.2.2.33667.81444.

[218] N. Keidar, Y. Elul, A. Schuster, and Y. Yaniv, “Visualizing and Quantifying Irregular Heart Rate Irregularities to Identify Atrial Fibrillation Events,” *Front. Physiol.*, vol. 12, no. February, 2021, doi: 10.3389/fphys.2021.637680.

[219] J. Wang, “A deep learning approach for atrial fibrillation signals classification based on convolutional and modified Elman neural network,” *Futur. Gener. Comput. Syst.*, vol. 102, pp. 670–679, 2020, doi: 10.1016/j.future.2019.09.012.

[220] O. Faust, B. H. . Sputh, R. Acharya U, and A. R. Allen, “A Pervasive Design Strategy for Distributed Health Care Systems,” *Open Med. Inform. J.*, vol. 2, no. 1, pp. 58–69, 2008, doi: 10.2174/1874431100802010058.

[221] O. Faust, M. Kareem, and N. Lei, “Atrial fibrillation detection service validation tool,” *Softw. Impacts*, p. 100117, 2021, doi: 10.1016/j.simpa.2021.100117.

[222] O. Faust, R. Acharya U, B. H. C. Sputh, and L. C. Min, “Systems engineering principles for the design of biomedical signal processing systems,” *Comput. Methods Programs Biomed.*, vol. 102, no. 3, pp. 267–276, 2011, doi: 10.1016/j.cmpb.2010.05.002.

[223] O. Faust, R. Martis, L. Min, G. L. . Zhong, and W. Yu, “Cardiac arrhythmia classification using electrocardiogram,” *ingentaconnect.com*, no. 3, p. 448, 2013, Accessed: Aug. 02, 2021. [Online]. Available: https://www.ingentaconnect.com/contentone/asp/jmihi/2013/00000003/00000003/art00016.

[224] U. R. Acharya *et al.*, "Application of nonlinear methods to discriminate fractionated electrograms in paroxysmal versus persistent atrial fibrillation," *Comput. Methods Programs Biomed.*, vol. 175, pp. 163–178, 2019, doi: 10.1016/j.cmpb.2019.04.018.

[225] O. Faust, M. Kareem, and N. Lei, "Atrial fibrillation detection service validation tool," *Softw. Impacts*, p. 100117, 2021, doi: 10.1016/j.simpa.2021.100117.

Appendix A- Converis-IRAS Ethics Approval form SHU

Stroke Risk monitoring service

Ethics Review ID: ER12933858

Workflow Status: Approved with Advisory Comments

Type of Ethics Review Template: IRAS - projects requiring NHS or HMPPS ethics

Primary Researcher / Principal Investigator

Murtadha Kareem

(Faculty of Science, Technology and Art)

Converis Project Application:

Q1. Is this project ii) Doctoral research

Director of Studies

Oliver Faust

(Faculty of Science, Technology and Art)

Supervisory Team

Prof. Reza Saatchi

Prof. Marcos Rodrigues

(Communication and Computing Research Centre),(Centre for Automation and Robotics Research)

Q3b. External Investigator Details: STH

Principal Investigator Name and Title: Dr

Ali Ali

Job Title: Consultant Stroke and Geriatrics Employer: Sheffield Teaching Hospitals NHS Trust Directorate: Geriatrics and Stroke Medicine Address: Glossop Road, Sheffield, S10 2JF

Telephone: 0114 2711768 Email: ali.ali@sheffield.ac.uk

Q6. Will the research involve any of the following?

i) Participants under 5 years old: No
ii) Pregnant women: No
iii) 5000 or more participants: No
iv) Research being conducted in an overseas country: No
Q7. If overseas, specify the location:

Q8. Is the research externally funded?: Yes

Funder Name:

Grow MedTech

Q9. Will the research be conducted with partners and subcontractors? Yes

Q9b. If yes, outline how you will ensure that their ethical policies are consistent with university policy: For this project, our partners come from Sheffield Teaching Hospital and Isansys [1]. Details onthese two partners follow below:

Sheffield teaching hospital NHS

Isansys

We will be using the Isansys healthcare platform. This product is a fully certified Class IIa CE-marked and Class II 510(k) cleared medical device. In order to get certification, the company must comply with NHS standards and be consistent with NHS ethical policies. By the rule of inference, the company is alsoconsistent with university policy.

[1] https://www.isansys.com/

Q10. Does the research involve one or more of the following?

i. Patients recruited because of their past or present use of the NHS or Social Care: Yes
ii. Relatives/carers of patients recruited because of their past or present use of the NHS or SocialCare: No
iii. Access to data, organs, or other bodily material of past or present NHS patients: Yes
iv. Foetal material and IVF involving NHS patients: No
v. The recently dead in NHS premises: No
vi. Participants who are unable to provide informed consent due to their incapacity even if the projectis not health related: No
vii. Prisoners or others within the criminal justice system recruited for health-related research: No
viii. Prisoners or others within the criminal justice system recruited for non-health-related research: No
ix. Police, court officials or others within the criminal justice system: No
Is this a research project as opposed to service evaluation or audit?: No

Q11. Category of academic discipline: Physical Sciences and Engineering

Q12. Methodology: Quantitative

P8 - Attachments

Are you uploading any recruitment materials (e.g. posters, letters, etc.)?

Non Applicable

Are you uploading a participant information sheet?

Yes

Are you uploading a participant consent form?

Yes

Are you uploading details of measures to be used (e.g. questionnaires, etc.)?

Yes

Are you uploading an outline interview schedule/focus group schedule?

Non-Applicable

Are you uploading debriefing materials?

Yes

Are you uploading a Risk Assessment Form? : Yes

Are you uploading a Serious Adverse Events Assessment (required for Clinical Trials andInterventions)?

Non-Applicable

Are you uploading a Data Management Plan?

Yes

Are you uploading a draft IRAS application and supporting documents? : Yes

Upload:

Full patient information leaflet.docxParticipant Consent Form (1).docxRisk Assessment Form (1).docx

Stroke risk monitoring service symptom profile questionnaire V1 13.03.20.doc

Data management Plan.docx SRMS Aphasia PoF000099.docx

Full DatasetTrialForm (1).pdf

P9 - Adherence to SHU Policy and Procedures

Primary Researcher / PI Sign-off:

I can confirm that I have read the Sheffield Hallam University Research Ethics Policy andProcedures: true

I can confirm that I agree to abide by its principles and that I have no personal or commercialconflicts of interest relating to this project.: true

Date of PI Sign-off: 22/04/2020

Director of Studies Sign-off:

I confirm that this research will conform to the principles outlined in the Sheffield Hallam UniversityResearch Ethics policy: true

I can confirm that this application is accurate to the best of my knowledge: true

Upload:

Date of submission and supervisor sign-off: 23/04/2020

Director of Studies Sign-off

Oliver Faust

This section to be completed by Lead Reviewer (or FREC if escalated)

Ann Macaskill

P10 - Review

Comments collated by Lead Reviewer (Or FREC if escalated): This a review of an IRAS form. Somemore work is needed on the IRAS form before it can be submitted.

QA6-1 Summary of the study needs to be made more comprehensive, It should outline exactly wha t the study is about and what is required of participants in the study. The language used needs to be easily understood by lay members so some terms may need to be defined. e.g. 'internet of things." Later talk about three cohorts and procedure but a summary is needed here.

Q9. You state here that recruitment will likely continue after the research – that is not ethical – suggestdelete or amend – has this already been approved by IRAS?

QA22. Amend grammar errors and remove word operations – it is misleading.

QA59. You write that the recruitment will take 6 months yet your ethics dates are for 3 months only – you must extend or shorten your study dates but most importantly they must correspond, so the study is for 3 or6 months be clear.

There are two PiS forms – the first one in the upload needs many revisions – the second one is much better

– why are there two? The same form should be used throughout. See PiS comments below:

1. Point 3: Opening statement should be reworded – it reads coercive it must be very clear that taking part inthe study is optional
2. Point 4: Correct grammar and replace 'help us proof' with 'can take part.'
3. Point 5: Correct grammar – voluntary not voluntarily – writing is persuasive rather than objective and informative, simply write the relevant information excluding 'offering the best of both worlds' and other purported inducements.
4. Point 9: There should be a question mark followed by an answer rather than this bald statement.
5. Point 11: 'Your data will help to proof" – should be removed – these data may also disprove the feasibilityof this approach – remove all reference to proving anything throughout and 'lofty goal and advancing science.'
6. Point 12: You may need to do this remotely now make sure you include a means to do this whilstupholding social distancing – this may have to be over the telephone or internet platform.
7. You need contact details on PiS. You need details of person outside of the research team that can be contacted for complaints – usually a named person in the department or research centre. You also need aGDPR statement here for HRA approvals.

5. PIS consent forms need IRAS number on and ID/Participant code box.

6. In the Converis upload you should select academic for all docs

7. Stroke Risk Monitoring Form: The possibilities for discomfort listed here should be added to the PiS. Thestatement ‘did you miss any days altogether’ does not make sense – what does it refer to, make this clear what information you require from the patient here.

8. The Risk Assessment and the entire proposal must address COVID – how will you uphold socialdistancing and conduct the study? You must include information about methods to mitigate againstbreaching social distancing rules.

9. Why are there two PiS forms?

10. Data preservation: Include COVID related caveats here.
Lead Reviewers Comments Following Resubmission: The main issue shave been addressed apart foruploading a consent form. Two forms of the PIS have been uploaded instead. This needs to be addressedbefore s=the IRAS submission

Final Decision to be completed by Lead Reviewer (or FREC if escalated): Approved with advisorycomments

Date of Final Decision: 23/07/2020

P12 - Post Approval Amendments

Amendment 1

Title of Amendment 1: Changing the study title: Validating of an Atrial Fibrillation detection algorithm

Details of Amendment 1: IRAS application form

The title of the study is modified to Validating of Atrial Fibrillation Detection Algorithm instead of Stroke RiskMonitoring Service as requested by NHS Research Ethics Approval

Date of Amendment 1: 11/12/2020

Upload:

FullDatasetTrialForm-6.pdf

In my judgement amendment 1 should be: Amendment Approved

Date of Amendment Outcome 1: 07/05/2021

Amendment 2

Title of Amendment 2: Consent form

Details of Amendment 2: -Minor modification is required to add on the form. This was related how would we inform the patients of the final outcome of the study. For example, sending the final result by letter or email. in addition to that, consenting a patient if he/she agrees to have 2 methods of heart rhythm monitoring applied for 3 days, and that the data from these devices will be downloaded and analysed in an anonymous manner.

-Updating the version control

Date of Amendment 2: 11/12/2020

Upload:

Participant_Consent_Form_11_12_2020_V2.doc

In my judgement amendment 2 should be: Amendment Approved

Date of Amendment Outcome 2: 07/05/2021

Amendment 3

Title of Amendment 3: Participant Information sheet

Details of Amendment 3: - Updating

-The version controls

- Study title

-GDPR

- Ensuring that the language used for identifying the sponsor is belong to Sheffield Hallam University. Forexample, when any word comes with 'we', it means that this related to the sponsor.

Date of Amendment 3: 22/02/2021

Upload:

Participant_Information_Sheet_22_02_2021_V3.docx.docx.doc

In my judgement amendment 3 should be: Amendment Approved

Date of Amendment Outcome 3: 07/05/2021

Appendix B- IRAS -NHS/HRA Ethics approval

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

[redacted]

Dr Oliver Faust
Senior Lecturer
Sheffield Hallam University
Howard St
Sheffield
S1 1WB

03 June 2021

Dear Dr Faust

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Validation of an Atrial Fibrillation detection algorithm**
IRAS project ID:	[redacted]
Protocol number:	
REC reference:	
Sponsor	**Sheffield Hallam University**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The standard conditions document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is [redacted] Please quote this on all correspondence.

Yours sincerely,
Gemma Warren

Approvals Specialist

[redacted]

Copy to: *Dr Keith Fildes*

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
Covering letter on headed paper [REC rebuttal]	1	11 December 2020
Covering letter on headed paper [REC rebuttal]	2	22 February 2021
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Sheffield Hallam University - 2020-21 TWIMC (PI)]	1	01 August 2020
GP/consultant information sheets or letters [GP letter]	2	11 December 2020
Interview schedules or topic guides for participants [Patient interview topic guide]	1	29 October 2020
IRAS Application Form [IRAS_Form_27082020]		27 August 2020
Letter from funder [Letter from the funder]	1	17 August 2020
Letters of invitation to participant [Patient invitation letter]	2	11 December 2020
Organisation Information Document [Organisation information document]	1	17 August 2020
Other [Patient aphasia]	2	11 December 2020
Participant consent form [Consent]	2	11 December 2020
Participant information sheet (PIS) [PIS]	3	22 February 2021
Research protocol or project proposal [Study Protocol]	2	11 December 2020
Response to Request for Further Information [Response to Validation Queries]		17 August 2020
Response to Request for Further Information [Validation Queries 2]		28 August 2020
Response to Request for Further Information [Response to Validation Queries 3]		01 September 2020
Schedule of Events or SoECAT [Schedule events]	1	17 August 2020
Summary CV for Chief Investigator (CI) [CV CI]	1	10 July 2020
Summary CV for student [Student CV]	1	10 July 2020
Validated questionnaire [questionnaire]	1	17 August 2020
Validated questionnaire [questionnaire]	2	29 October 2020
Validated questionnaire [questionnaire]	2	11 December 2020

Appendix C- Letter of Access for STH

Sheffield Teaching Hospitals NHS
NHS Foundation Trust
Research Department

Letter of access for researchers who do not require an honorary research contract

15 March 2021

[redacted]

Dear Murtadha

STH ref: [redacted]
Study title: Validation of an Atrial Fibrillation detection algorithm
Principal Investigator: Ali Ali

Letter of access for research
This letter confirms your right of access to conduct research through Sheffield Teaching Hospitals NHS Foundation Trust for the purpose and on the terms and conditions set out below. This right of access commences on **15 March 2021** and ends on **16 January 2022** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permissionconfirmation from the individual organisation(s) of their agreement to conduct the research.

The information supplied about your role in research at Sheffield Teaching Hospitals NHS Foundation Trust has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

Status
You are considered to be a legal visitor to Sheffield Teaching Hospitals NHS Foundation Trust premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

Reporting Arrangements
While undertaking research through Sheffield Teaching Hospitals NHS Foundation Trust you will remain accountable to your substantive employer/place of study **Sheffield Hallam University** but you are required to follow the reasonable instructions of **Ali Ali** in this NHS organisation or those instructions given on their behalf in relation to the terms of this right of access.

Legal Claims
Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

Policies and Procedures
You must act in accordance with Sheffield Teaching Hospitals NHS Foundation Trust policies and procedures, which are available to you upon request, and the Research Governance Framework.
You are required to co-operate with Sheffield Teaching Hospitals NHS Foundation Trust in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on Sheffield Teaching Hospitals NHS Foundation Trust premises. You must observe the same standards of care and propriety

in dealing with patients, staff, visitors, equipment and premises as is expected of any other contract holder and you must act appropriately, responsibly and professionally at all times.
If you have a physical or mental health condition or disability which may affect your research role and which might require special adjustments to your role, if you have not already done so, you must notify your employer and Sheffield Teaching Hospitals NHS Foundation Trust Occupational Health Service prior to commencing your research role at the Trust.

Confidentiality
You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice and the Data Protection Act 2018. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on Trust premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation does not accept responsibility for damage to or loss of personal property.

Duration and Termination
This NHS organisation may revoke this letter and may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. You must not undertake regulated activity if you are barred from such work. If you are barred from working with adults or children this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity and you MUST stop undertaking any regulated activity immediately.

Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

Indemnity and Liability
Sheffield Teaching Hospitals NHS Foundation Trust will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 2018. Any breach of the Data Protection Act 2018 may result in legal action against you and/or your substantive employer.

Change in Status
If your current role or involvement in research changes, or any of the information provided in your Research Passport changes, you must inform your employer through their normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely

[redacted]

Professor Simon Heller, Director of Research
Sheffield Teaching Hospitals NHS Foundation Trust

cc: Bryony Plumb, Senior Administrator, Sheffield Hallam University