PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
University of Mohamed Khider – BISKRA
Faculty of Exact Sciences, Science of Nature and Life
Computer Science Department

Ordre N°: IA_StartUp 01/M2/2023

Dissertation

Submitted in fulfilment of the requirements for the Master's degree in

Computer science

Option: **Artificial Intelligence (AI)**

E-health Platform for Monitoring Human Life

By:

AFREN Radja

AHMID Salsabil

ZERDOUMI Hajer

Members of the jury:

TIBERMACINE Okba	**M.C.A**	**President**
KAHLOUL Laid	**PROF**	**Supervisor**
MERIZIG Abdelhak	**M.C.B**	**Supervisor**
MOKHTARI Bilal	**M.C.A**	**Supervisor**
BENDAHMANE Asma	**M.A.A**	**Examiner**

Academic year 2022-2023

Acknowledgements

"Saying thank you is more than good manners.

It is good spirituality."

- Alfred Painter -

First and foremost, praises and thanks to my God, **Allah**, for everything.

I would like to extend my heartfelt acknowledgment to my beloved **Mother**, she is the radiant moon that illuminates my nights. Next, I want to express my sincere thanks to my dear **Father**, who has always been a guiding light in my life.

I would also like to express my profound gratitude to my dear siblings and extended family members, **Soufian, Selma, Zahra, Wafa, and Djouhaina,** as well as my sweet chicks **Akram, Sedjoud, and Ratil**. In particular, I want to extend my special thanks to my dear **Wafa,** the house doctor, who has been my rock and confidante throughout this incredible journey.

Furthermore, I wish to express my deepest appreciation to my supervisor, **Professor KAHLOUL Laid**, for his unwavering belief in my abilities and for providing me with the invaluable opportunity to undertake this remarkable project titled **"Deep Learning Approach for the Retinal Optical Coherence Tomography (OCT) Images Classification"**. His encouragement has been a constant source of motivation, driving me to persist with unwavering determination. I am also grateful to **Dr.MERIZIG Abdelhak** and **Dr.MOKHTARI Bilal**, Their invaluable assistance and guidance have greatly contributed to conducting the research.

Finally, I want to extend my sincere appreciation to my dear friends **Aicha, Dalila, Hajer, Salsabil, Saliha, Sara, Yasmine, Nada, Amani, Nomidia, Rayan, Radhia, and Donia**.

AFREN Radja.

Acknowledgements

Praise be to God first and foremost

First of all, I thank **Allah**, the Almighty, who gave me the strength, the patience, and the will to accomplish this modest work. This thesis would not have been accomplished without the support and guidance of several individuals to whom I would like to express my gratitude in this acknowledgment.

I would like to express my sincere gratitude to **Dr. Mokhtari** for his valuable guidance and mentorship throughout my thesis on **emotion recognition using vital signals**. His expertise and encouragement have been instrumental in shaping my research work.

I would also like to extend my heartfelt thanks to **Pr. Kahloul** and **Dr. Merizig** for their unwavering support and help. Their insightful comments and constructive feedback have greatly contributed to the success of my research work.

The most eminent thanks go to my dearest parents who have always been there for me, for the sacrifices made towards me, for their support, and for all the efforts made for my education and training. I would also like to extend my gratitude to my brothers **Sliman** and **Houdhaifa**, and my sisters **Ahlam** and **Wafa** for their continuous support and encouragement.

Finally, I cannot end without thanking my closest friends **Sara Kerboub**, **Radja Afren**, **Selsabil Barkat**,**Kater El Nada Mohamdi**, **Rayan Ben Brahim**, **Salsabil Ahmid**, and all my family for their unwavering support and encouragement throughout my academic journey.

ZERDOUMI Hajer.

Acknowledgements

I express my sincere gratitude to **Allah** Almighty, who has granted me the ability to complete my thesis on **"Prediction of heart disease using the Internet of Medical Things (IoMT)."** I would like to acknowledge and appreciate those who stayed up for me and sacrificed for my education. They have provided me with everything I needed to continue my studies and have continuously encouraged me until I reached this point. To **my dear mother and father**, thank you very much. This graduation and success are dedicated to both of you and **my siblings**, whose support has been invaluable in overcoming numerous obstacles and difficulties.

I am deeply indebted to my supervisor, **Dr. Merizig Abdelhak**, for their unwavering support and guidance, which greatly contributed to the quality of this work. Their patience, encouragement, valuable advice, and, most importantly, their unwavering belief in my abilities played a crucial role in my success.

I would also like to extend my heartfelt to **Pr. Kahloul Laid** and **Dr. Mokhtari Bilal** for their valuable advice and support throughout my research journey.

Additionally, I wish to extend my heartfelt appreciation to the esteemed members comprising the jury. I am deeply grateful to the jury president **Dr. Tibermacine Okba**, and the examiner **Dr. Bendahmane Asma** for undertaking the task of assessing my work and graciously devoting their time, and knowledge throughout the evaluation phase.

Lastly, I want to convey my thanks to all my teachers, friends, and colleagues in the computer science department. Their support, insights, and camaraderie have played an instrumental role in shaping my academic experience.

AHMID Salsabil.

Abstract

This project presents an AI-powered healthcare platform that utilizes advanced technologies to revolutionize healthcare. The platform offers services such as retinal damage detection, mental health monitoring, and heart disease prediction. The retinal damage classification service achieves a 95.39% accuracy using DenseNet, a convolutional neural network (CNN), to analyze OCT scans and categorize retinal diseases (choroidal neovascularization, diabetic macular edema, multiple drusen).

Additionally, the platform includes an emotion recognition service with a 99.89% accuracy. It monitors patients emotional states using electrocardiogram (ECG) signals and the Internet of Medical Things (IoMT). By extracting features like heart rate variability (HRV) and utilizing a Random Forest classifier.

The platform uses deep learning models to accurately predict heart diseases using patient information. By leveraging IoMT capabilities for data extraction, the platform incorporates a deep neural network (DNN) achieving 91% accuracy and a sequence-to-sequence (Seq2Seq) model with a mean squared error (MSE) of 0.1 and root mean squared error (RMSE) of 0.3 for predicting sequence-based ECG data. Also a 98% accuracy rate in ECG data classification. The platform also promptly notifies attending doctors of optimal patient care in critical situations.

Our services provide healthcare professionals with precise information for informed decision-making and targeted treatments in ophthalmology, mental health issues, and cardiac diseases. Moreover, the three applications we have developed using these models have demonstrated superior accuracy, confirming their performance and effectiveness compared to alternative approaches.

Keywords: Healthcare, Internet of Medical Things (IoMT), Artificial Intelligence (AI), retinal damage detection, emotion recognition, heart diseases prediction.

Résumé

Ce projet présente une plateforme de soins de santé alimentée par l'intelligence artificielle qui utilise des technologies avancées pour révolutionner les soins de santé. La plateforme propose des services tels que la détection des lésions rétiniennes, la surveillance de la santé mentale et la prédiction des maladies cardiaques. Le service de classification des lésions rétiniennes atteint une précision de 95% en utilisant DenseNet, un réseau neuronal convolutif (CNN), pour analyser les scanner OCT et catégoriser les maladies rétiniennes (néovascularisation choroïdienne, œdème maculaire diabétique, multiples drusen).

De plus, la plateforme comprend également un service de reconnaissance des émotions avec une précision de 99.89%. Il surveille les états émotionnels des patients en utilisant des signaux d'électrocardiogramme (ECG) et l'Internet des objets médicaux (IoMT). En extrayant des caractéristiques telles que la variabilité du rythme cardiaque (HRV) et en utilisant un classifieur Random Forest.

La plate-forme utilise des modèles d'apprentissage en profondeur pour prédire avec précision les maladies cardiaques à l'aide des informations sur les patients. En tirant parti des capacités de l'IoMT pour l'extraction de données, la plate-forme intègre un réseau de neurones profonds (DNN) atteignant une précision de 91% et un modèle séquence à séquence (Seq2Seq) avec une erreur quadratique moyenne (MSE) de 0,1 et erreur quadratique moyenne racine (RMSE) de 0,3 pour prédire les données ECG basées sur la séquence. Également un taux de précision de 98% dans la classification des données ECG. La plateforme informe également rapidement les médecins traitants des soins optimaux aux patients dans des situations critiques.

Notre service fournit aux professionnels de la santé des informations précises pour une prise de décision éclairée et des traitements ciblés en ophtalmologie, en problèmes de santé mentale et dans les maladies cardiaques. De plus, les trois applications que nous avons développées avec ces modèles ont démontré une précision supérieure, confirmant ainsi leur performance et leur efficacité par rapport aux approches alternatives.

Mots clés: Santé, Internet des objets médicaux, intelligence artificielle (IA), détection des dommages rétiniens, reconnaissance des émotions, prédiction des maladies cardiaques.

Contents

Acknowledgements i
Abstract iv
Résumé v
List of Figures xiv
List of Tables xvi
List of Abbreviations xvii

1 General introduction **1**
1.1 Introduction 1
1.2 Organisation of the dissertation 3

2 State of the art **5**
2.1 Introduction 5
2.2 Health care definition 6
2.3 Health care types 6
2.4 Traditional health care problems 7
2.4.1 Diagnosis problems 7
2.4.2 Patient monitoring 8
2.4.3 Surveillance and monitoring problems 8
2.5 Applications in healthcare 8
2.6 Types of used data 9
2.6.1 Electrocardiogram 9
2.6.2 Optical Coherence Tomography 11
2.7 Internet of Medical Things 13

2.8 Internet of Medical Things architecture 13
2.9 Prediction of heart disease 15
2.9.1 Normal structure and function of the heart anatomy 15
2.9.2 Blood vessels 17
2.9.3 The causes of heart disease 17
2.9.4 Types of heart disease 18
2.10 Retinal damage 20
2.10.1 The retina 20
2.10.2 Age-related Macular Degeneration 21
2.10.3 Choroidal Neovascularization 22
2.10.4 Diabetic Macular Edema 23
2.10.5 Multiple drusen 24
2.11 Emotion recognition 25
2.11.1 Defining emotion 25
2.11.2 Emotion recognition in human-computer interaction 25
2.11.3 Applications of emotion recognition systems using Electrocardiogram . . . 27
2.11.4 Emotion models 28
2.12 Related work 31
2.12.1 Synthesis comparative study 33
2.13 Conclusion 36

3 Machine Learning and Deep Learning 39

3.1 Introduction 39
3.2 Machine learning 39
3.2.1 Random Forest 40
3.3 Deep Learning 40
3.3.1 Artificial Neural Network 41
3.3.2 Convolution Neural Networks 41
3.3.3 The basics of Convolutional Neural Network 42
3.3.4 Prominent Convolutional Neural Network Architectures 44

3.3.5 Recurrent Neural Network 48
3.3.6 Long Short-Term Memory 48
3.3.7 Bidirectional Long Short-Term Memory 48
3.3.8 Sequence to Sequence Models 49
3.3.9 Encoder-Decoder 49
3.3.10 Transfer learning 51
3.4 Conclusion 51

4 Contribution and system design 53

4.1 Introduction 53
4.2 Analysis of the requirements 53
4.3 Proposed architecture 55
4.4 The proposed approach for Internet of Medical Things 56
4.5 The proposed approach for prediction of heart disease process 59
4.5.1 Data set description 59
4.5.2 The proposed approach for risk assessment of heart diseases 61
4.5.3 The proposed approach for the prediction of Electrocardiogram 65
4.5.4 The proposed approach for Electrocardiogram classification 67
4.6 The proposed approach for the retinal Optical Coherence Tomography images classification 69
4.6.1 Data set description 70
4.6.2 Preprocessing phase 70
4.6.3 Data splitting 72
4.6.4 Training phase of our deep learning model 73
4.6.5 Prediction phase 74
4.7 The proposed approach for emotion recognition using vital signals 75
4.7.1 Data set description 77
4.7.2 Augmantation of the DREAMER dataset 77
4.7.3 Labels transformation 78
4.7.4 Features extraction 80

4.7.5 Features selection . . . 80
4.7.6 Classification . . . 84
4.8 Conclusion . . . 85

5 Implementation and results **87**
5.1 Introduction . . . 87
5.2 Devlopement tools and used platforms . . . 87
5.2.1 Used platforms . . . 88
5.2.2 Used IoT board and sensors . . . 92
5.3 System interfaces and examples . . . 95
5.3.1 Detecting retinal damage application . . . 95
5.3.2 Prediction of heart disease application . . . 103
5.3.3 Emotion recognition application . . . 116
5.4 Obtained results and discussion . . . 119
5.4.1 Prediction of heart disease . . . 119
5.4.2 Retinal damage classification . . . 139
5.4.3 Emotion recognition using vital signals . . . 147
5.5 Challenges encountered . . . 152
5.6 Conclusion . . . 153

6 Conclusion and perspectives **154**

References **156**

List of Figures

2.1 The ECG Waveform in a Healthy and Normal Heart [1]. 10
2.2 Ocular Coherence Tomography Imaging [2]. 12
2.3 Category of HIoT application [3]. 13
2.4 IoT Architecture [4]. 15
2.5 Conduction system of the heart 16
2.6 Arteries, capillaries, and veins. 17
2.7 ECG Waves and Their Relation to Heart Nodes [5]. 19
2.8 the components of the eye [6]. 20
2.9 The difference between a normal eye and an eye with AMD [7]. 22
2.10 Stages of Choroidal Neovascularization (CNV) [8]. 23
2.11 The Difference between a normal eye and an eye with Diabetic Macular Edema [9]. 24
2.12 Understanding the differences between a healthy retina, retina with Wet AMD, and retina with Drusen [6]. 24
2.13 Plutchik's Color Wheel: Depicting the Eight Primary Emotions [10]. 29
2.14 Valence-Arousal model for emotion representation [11]. 30
2.15 Valence-Arousal-Dominance model for emotion representation [10]. 30

3.1 Deep Learning Architecture: Exploring the Structure of Neural Networks [12]. . . . 41
3.2 Composition of Convolutional Neural Networks Architecture [13]. 42
3.3 Example of convolutional layer slides [14]. 43
3.4 Demonstration of Max-Pooling [15]. 44
3.5 DenseNet architecture [16]. 45

3.6 A 5-layer Dense Block with a Growth Rate of 4, featuring Full Connectivity within Layers [17]. . . . 45
3.7 Network architectures of VGG16 and VGG19 [11]. . . . 47
3.8 InceptionV3 architecture [18]. . . . 47
3.9 LSTM vs RNN [19]. . . . 48
3.10 Internal architecture of encoder [20]. . . . 50
3.11 Internal architecture of decoder [20]. . . . 50

4.1 An overview of the general e-health platform Architecture. . . . 56
4.2 The proposed Wiring diagram of ECG measurement device. . . . 57
4.3 The wiring diagram of heart rate measurement device [4]. . . . 57
4.4 The proposed process approach for prediction of Heart Disease. . . . 59
4.5 The plan of the proposed approach. . . . 69
4.6 Analysis of Data Sample Distribution in the dataset. . . . 70
4.7 CSV File Structure: Image Paths and Labels Representation. . . . 71
4.8 Analysis of Data Sample Distribution in the Train Set. . . . 72
4.9 The Proposed Architecture for Retinal Damage Classification. . . . 73
4.10 Architecture of an ECG-based emotion recognition system. . . . 76
4.11 Mean location of film clips in two-dimensional affective space as rated by the participants in DREAMER [21]. . . . 79
4.12 Heart Rate Variability[22]. . . . 80
4.13 Visual representation: Random forest applied to extracted features. . . . 85

5.1 Python. . . . 88
5.2 TensorFlow. . . . 88
5.3 Keras. . . . 88
5.4 Scikit-learn. . . . 88
5.5 NumPy. . . . 89
5.6 Pandas. . . . 89
5.7 Matplotlib. . . . 89
5.8 Neurokit. . . . 89

5.9 Kaggle. 90
5.10 Android studio. 90
5.11 Arduino IDE. 90
5.12 Raspberry Pi. 91
5.13 Firebase. 91
5.14 Flutter. 91
5.15 Kotlin. 91
5.16 ONNX. 92
5.17 ESP8266-based development board. 93
5.18 AD8232 ECG sensor. 93
5.19 The Raspberry Pi 3 Model B+ [23]. 93
5.20 Connvertisseur analogique numérique ADS1115. 94
5.21 Pluse sensor. 94
5.22 The logo and first splash screen. 95
5.23 Onboarding screens. 96
5.24 Multi-provider login. 97
5.25 Phone OTP (One-Time Password) verification. 97
5.26 Forgot password procedure. 98
5.27 Signup procedure. 98
5.28 Home page interfaces. 99
5.29 Retina services interfaces. 100
5.30 Medical student teaching tool. 101
5.31 User profile interfaces. 102
5.32 Cardiovascular specialist opinion about the application. 103
5.33 Start Interfaces. 104
5.34 Login Interface. 105
5.35 Verification code interface: SMS Access and Auto-Retrieval 106
5.36 Complete registration interface. 107
5.37 The home screen with the options. 108
5.38 The conversations screen. 109

5.39 Create and manage new groups and select group members. 110
5.40 The monitoring heart rate screen. 111
5.41 Service selection. 112
5.42 Instant doctor notification. 113
5.43 The doctor's input screen. 114
5.44 The ECG monitoring screen. 115
5.45 Splash and user authentication screen. 116
5.46 Home Interface and ECG Visualization screen. 117
5.47 Emotion recognition service. 118
5.48 Original structure. 120
5.49 The modified dataset structure. 120
5.50 Confusion matrix. 123
5.51 Receiver operating characteristic. 125
5.52 The ECG5000 dataset. 126
5.53 The Seq2seq architecture. 128
5.54 Results 129
5.55 Input and Output sequence. 129
5.56 The balanced dataset. 130
5.57 The Structure of the dataset. 131
5.58 The ROC curve of our model with balanced data and without balanced data. 138
5.59 The training and validation accuracy and loss curves for the custom CNN model. . 142
5.60 Confusion matrix obtained using Custom CNN model. 142
5.61 The training and validation accuracy and loss curves for the custom VGG16 model. 143
5.62 The training and validation accuracy and loss curves for the custom VGG19 model. 144
5.63 The training and validation accuracy and loss curves for the custom InceptionV3 model. 144
5.64 Obtained results for custom DenseNet model. 145
5.65 Accuracy comparison between different classifiers without data augmentation (emotion expressed in terms of valence)in DREAMER dataset. 148

5.66 Accuracy comparison between different classifiers with data augmentation in DREAMER dataset. 149
5.67 Accuracy of the combined dataset using different classifiers. 150
5.68 Confusion Matrix for Emotion Classification on Combined Dataset. 151

List of Tables

2.1 Summary of recent studies on classifying retinal damage. 34
2.2 Related work on prediction of heart disease. 35
2.3 Related work on emotion recognition. 36

4.1 Average Heart Rates by Age Group and Gender [4]. 58
4.2 Description of the Time domain features selected. 81
4.3 Description of the frequency domain features selected. 82
4.4 Description of the nonlinear domain features selected. 83

5.1 Comparison between development board. 92
5.2 Dataset Information 121
5.3 Model Hyperparameters 121
5.4 Loss and accuracy values 122
5.5 Performance Metrics 123
5.6 Dataset Information 127
5.7 Model Hyperparameters. 127
5.8 Comparison between different classifiers without a balanced dataset. 132
5.9 Comparison between different classifiers with a balanced dataset. 135
5.10 Comparison with Related work on prediction of heart disease. 139
5.11 Results of Proposed Architectures for Retinal Damage Classification. 140
5.12 Classification Report for DenseNet Model. 146
5.13 Performance Metrics for DenseNet Model. 146
5.14 Table of Model Comparison. 146

5.15 Comparison with Existing Literature using the DREAMER dataset. 152

List of Abbreviations

Abbreviation	Meaning
AHP	Analytical Hierarchy Process
ANN	Artificial Neural Network
ANS	Autonomic Nervous System
BPM	Beat Per Minute
BVP	Blood Volume Pressure
CAD	Coronary Artery Disease
CNS	Central Nervous System
CNNs	Convolutional Neural Networks
CNV	Choroidal Neovascularization V
CVD	Cardiovascular Disease
DEM	Discrete Emotional Model
DE	Differential Evolution
DME	Diabetic Macular Edema
DSS	Decision Support Systems
DT	Decision Trees
EEG	Electroencephalogram
ECG	Electrocardiogram
EMG	Electromyogram
EHR	Electronic Health Records
FFT	Fast Fourier Transformation
FN	False Negatives
FP	False Positives
FNN	Feedforward Neural Network
GSR	Galvanic Skin Response
HCI	Human Computer Interaction
HF	High Frequency
HIoT	Healthcare Internet of Things

HIS	Health Information Exchange
HR	Heart Rate
HRFLM	Hbrid Random Forest with a Linear Model
HRV	Heart Rate Variability
IBI	Interbeat Interval
KMC	Kasturba Medical College
KNN	K-Nearest Neighbors
LA	Left Arm
LF	Low Frequency
LSTM	Long Short-Term Memory
LR	Logistic Regression
MBN	Market Business News
MAE	Mean Absolute Error
MLP	Multilayer Perceptron
MSE	Mean Squared Error
MI	Myocardial Infarction
NN	Neural Network
NPV	Negative Predictive Value
OCT	Optical Coherence Tomography
ONNX	Open Neural Network Exchange
OTP	One Time Password
PPG	PhotoPlethysmoGram
PSD	Power Spectral Density
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
RSP	Respiration
RL	Right Leg
RA	Right Arm
RF	Random Forest
RMSE	Root Mean Squared Error

SCR	Skin Conductive Resistance
Seq2Seq	Sequence to Sequence
SVC	Support Vector Classifier
TL	Transfer Learning
TN	True Negatives
TP	True Positives
VEGF	Vascular Endothelial Growth Factor
VAD	Valence Arousal Dominance
WHO	World Health Organization
XGB	XGBoost

Chapter 1

General introduction

1.1 Introduction

Healthcare plays a vital role in preserving and enhancing human life by providing a wide range of medical services aimed at promoting well-being and preventing diseases. However, according to the study reported in [24] the healthcare industry faces numerous challenges that hinder the delivery of high-quality care. These challenges include increasing demands on healthcare systems, resource limitations, complex diagnoses, and the need for timely and accurate decision-making. Heart disease, as reported by the World Health Organization [25], remains the leading cause of global mortality, resulting in approximately 17.9 million deaths annually. The prevalence of diabetes has significantly risen, reaching 422 million affected individuals according to the study conducted in [26]. According to Wong et al. [27], age-related macular degeneration (AMD) contributes to approximately 8.7% of global blindness cases. Additionally, Varma et al. [28] estimated a prevalence of around 2% for drusen, which is a characteristic sign of AMD, in the general population.

To overcome these obstacles and enhance healthcare services, researchers and professionals are developing technological advancements and intelligent solutions. Poongodi et al.[29] propose a platform for patient monitoring that aims to reduce the time of intervention in critical situations. On the other hand, Young et al.[30] conducted a systematic review and thematic synthesis of qualitative research to explore the experiences and reported impact of participation in venue-based exercise interventions for people with stroke in the UK. The mentioned works focus on specific interventions, but they are not implemented in Algeria and lack AI integration.

They also do not address the needs of individuals with poor emotional states, which can affect their physical health. Algeria requires comprehensive healthcare solutions that cover both physical and emotional well-being, highlighting the need for AI-based technologies to improve healthcare delivery.

This dissertation introduces an innovative healthcare platform that leverages the potential of artificial intelligence (AI) [31], and advanced technologies. The platform offers a comprehensive solution to address the specific challenges encountered in the healthcare industry. The primary objective is to empower healthcare professionals with tools that facilitate informed decision-making, leading to improved patient outcomes and an elevated healthcare experience. By promoting efficiency, effectiveness, and patient-centric care, this transformative platform aims to shape the future of healthcare.

The proposed healthcare platform incorporates several intelligent services that leverage AI and machine learning techniques. One of these services focuses on retinal damage classification using Optical Coherence Tomography (OCT) scanner images. The platform can accurately analyze and classify different types of retinal diseases by applying deep learning algorithms. This capability empowers healthcare professionals to make informed decisions and provide targeted treatments, ultimately improving patient care in the field of ophthalmology.

In addition to retinal damage classification, the platform integrates an emotion recognition service that utilizes the Internet of Medical Things (IoMT) to monitor patients' ECG signals. Through advanced machine learning algorithms, the platform can interpret these signals and classify patients' emotional states. This feature has significant implications for mental health diagnosis and treatment, as it enables healthcare providers to gain insights into patients' emotional well-being and tailor interventions accordingly.

Moreover, the healthcare platform incorporates deep learning models for predicting heart diseases. By leveraging the power of AI, the platform can analyze various factors, such as medical history, ECG data, and other relevant patient information, to identify potential risks and predict the likelihood of cardiovascular conditions. This predictive capability enables early intervention and personalized treatment plans, ultimately leading to better patient outcomes and improved management of heart diseases.

By combining retinal damage classification, emotion recognition, and heart disease prediction, the healthcare platform offers a comprehensive suite of services that address critical aspects of patient care.

1.2 Organisation of the dissertation

This dissertation consists of an introduction, four chapters, and a general conclusion, and it is organized as follows:

- Chapter 1 provides an overview of the current state of the art in healthcare, highlighting key concepts and addressing existing issues in the field.
- Chapter 2 explores machine learning and deep learning algorithms. This chapter serves as a technical background for the subsequent chapters.
- Chapter 3 presents the architectural design of the platform and discusses the models employed for the various healthcare services, providing detailed insights into the platform's structure, functionalities, and integration of intelligent services.
- Chapter 4 focuses on the practical implementation and showcases the results obtained. It describes the platform's infrastructure, tools, and interfaces, demonstrating the effectiveness and potential impact of the healthcare platform through examples.
- Finally, the general conclusion offers a summary of the dissertation's key findings and contributions, provides a perspective on future work, and discusses potential advancements in healthcare monitoring and decision support systems.

“I believe, if you zoom out into the future, and you look back, and you ask the question, ‘What was Apple’s greatest contribution to mankind?’ it will be about health.”

Tim Cook

Chapter 2

State of the art

2.1 Introduction

Due to the significant development that has taken place with the advent of the fourth industrial revolution, artificial intelligence (IA) [31], and the Internet of Things (IOT), scientists have chosen to incorporate these technologies into various fields, including healthcare [32]. Artificial intelligence has progressively transformed medical practice, thanks to recent advancements in digital data acquisition, machine learning, and computing infrastructure. AI applications have expanded into areas that were previously believed to be solely within the realm of human experts. Similarly, the Internet of Things [4] has emerged as a crucial technology in healthcare systems, enhancing efficiency, cost-effectiveness, and comfort for individuals. It has also facilitated monitoring solutions for patients within and outside hospital settings.

In this chapter, we will have addressed the challenges encountered in traditional healthcare, patient monitoring, and diagnosis. Furthermore, we will have defined the concept of the Internet of Medical Things. Additionally, we will have discussed three vital healthcare systems in this chapter: emotion recognition for mental health, heart disease prediction for physical health, and the classification of retinal damage. Finally, we will have outlined the motivations behind the startup and the solutions it offers.

2.2 Health care definition

Healthcare or Health care has many definitions. According to the medical dictionary [33], health care is the services provided to individuals or societies by health services agents or professions to promote health, preserve, monitor, or restore.

The Definition of MBN (Market business news) [34] : "Healthcare refers to the efforts that medical professionals make to restore our physical and mental well-being. The term also includes the 3 Background provision of services to maintain emotional well-being".

2.3 Health care types

Depending on the health problem, healthcare is divided into 4 different types.

- **Primary healthcare:** According to authors in [35], Primary healthcare includes at least: education concerning prevailing health problems and the methods of preventing and controlling them; promotion of food supply and proper nutrition; an adequate supply of safe water and basic sanitation; maternal and child health care, including family planning; immunization against the major infectious diseases; prevention and control of locally endemic diseases; appropriate treatment of common diseases and injuries; and provision of essential drugs.

- **Secondary healthcare:** This means that you will be cared for by someone who has more specific experience with your condition. For example, medical conditions that need secondary care services are cancer treatment, medical care for pneumonia and other acute and sudden infections, and care for broken bones. As illustrated by [36].

- **Tertiary healthcare:** In agreement with [37], tertiary care has become a common feature in certain specialties for rare conditions, or where the diagnostic or treatment facilities are scarce or require scarce combinations of resources, or which remain essentially the subject of research. These facilities are commonly found in medical schools and teaching hospitals. According to [35] it includes specialist cancer management, neurosurgery, cardiac surgery, transplant services, plastic surgery, treatment for severe burns, advanced neonatology services, palliative, and other complex medical and surgical interventions.

- **Quaternary healthcare:** Is defined by [35] as an extension of tertiary care and is highly specialized and not widely accessed, usually only offered in a very limited number of national or international centers. Experimental medicine and some uncommon types of diagnostic or surgical procedures are considered quaternary healthcare.

2.4 Traditional health care problems

The use of traditional healthcare systems produces many problems, which affect healthcare services. Here we will present some of these problems and explain how they affect the quality of care:

2.4.1 Diagnosis problems

Diagnostic problems are the difficulties that health care providers may face when trying to determine the general condition of the patient in order to reach the appropriate treatment in a timely manner as shown in [38]. Here are some common diagnostic problems:

- **Diagnostic uncertainty:** No matter how much data we collect, how many observations we make, or how many tests we run, we will never be able to diagnose a patient with absolute certainty. A diagnosis is a conjecture about the type of illness a patient is experiencing, one that is derived from observations by drawing conclusions. As the inferential process progresses, the accumulation of data that either supports a given diagnosis or refute competing hypotheses increases our confidence as [clinicians] in that diagnosis. Our goal is to lessen diagnostic uncertainty so that we can make the best possible therapeutic decisions, not to reach certainty in accordance with [39].

- **Time:** The factor of time is crucial to the diagnostic process [40]. Most diseases develop gradually over time, and the patient's symptoms may not appear for some time after the disease first manifests. It may also take some time before a patient's symptoms are identified as belonging to a particular diagnosis.

- **Population trends:** As proven by [41] [42], The diagnostic process is becoming significantly more complex due to population trends like the aging of the population, and clinicians must take into account comorbidity, polypharmacy, associated medication side effects, as well as disease and medication interactions. For instance, fatigue and confusion rather than typical symptoms like chest pain or radiating arm pain may be present in acute Myocardial Infarction (MI).

2.4.2 Patient monitoring

Patient monitoring [43] is a crucial part of healthcare because it provides you with the knowledge you need to more accurately recognize, foresee, and treat patients' needs throughout the continuum of care. Constant patient data monitoring, including blood pressure, respiration rate, heart rate and ECG, and many more parameters[44]. the World Health Organization (WHO) reports that more than 40% of WHO member states report having less than 10 physicians per 10,000 people, and more than 26% report having fewer than 3 in relation to the world population as shown in [4].

2.4.3 Surveillance and monitoring problems

As indicated by [45], Surveillance has been defined as an ongoing and systematic undertaking that involves gathering, analyzing, interpreting, and sharing descriptive information with the purpose of monitoring health issues. In order to assess the state of the public's health, track trends, identify public health priorities, and gauge the success of public health initiatives, surveillance, and monitoring activities are a cornerstone of public health practice [46]. The challenges of surveillance and monitoring refer to the challenges of compiling and analyzing information on the prevalence of diseases. This is due to the fact that conventional devices frequently provide inaccurate information, which, as we previously stated in [39], results in uncertainty in diagnosis and delays in response.

2.5 Applications in healthcare

The medical community is buzzing about digital healthcare technology. According to [47], healthcare technology has the potential to transform the way we approach healthcare, and it's already changing how patients are treated. Healthcare solutions have a wide range of applications across various fields, including:

- **WebMD:** Gives quick access to healthcare information, a symptom checker, medical reminders, and a directory of providers in a certain area. Additionally, users may learn about specific prescriptions, and the effects of medications, and even explore therapies and diagnoses As confirmed by [48].

- **HealthTap:** The application offers you answers to millions and millions of questions from real doctors spanning various topics. Within 24 hours, a doctor will answer. It also makes it easier to check your lab results and manage your medications during planned live consultations with doctors, dentists, and psychiatrists as shown in [48].

- **Pocket Pharmacist:** is an application that offers guidance on how to cure various diseases depending on their symptoms. Additionally, the app provides patients with information on medical diseases, their causes, and much more like it indicated in [48].

2.6 Types of used data

In our study, we utilized several tools, which are defined as follows:

2.6.1 Electrocardiogram

ECG is a technique that records and measures the electrical activity of the heart during various phases and perspectives, depending on the situation and configuration [49]. The ECG graph consists of a series of deflections or waves: the P wave, the Q wave, the R wave, S wave, and the T wave [50], and often the U wave [51] as shown in Figure 5.44. Each wave within the ECG provides valuable information that can be utilized to comprehend the condition of the heart [52].

- P-wave: Represents the electrical depolarization of the atria before contraction.

- Q-wave: Indicates the initial depolarization of the ventricular septum.

- R-wave: Reflects the depolarization of the ventricular myocardium, particularly the main mass of the ventricles.
- S-wave: Represents the depolarization of the posterior basal region of the ventricles.
- T-wave: Represents the rapid repolarization of the ventricles, preparing them for the next cardiac cycle.

The R peak-to-peak interval distance is commonly utilized to calculate the Inter Beat Interval (IBI) for Heart Rate (HR) detection [50]. In order to extract Heart Rate Variability (HRV) characteristics from ECG signals, it is crucial to perform QRS detection to accurately sort the RR intervals [53].

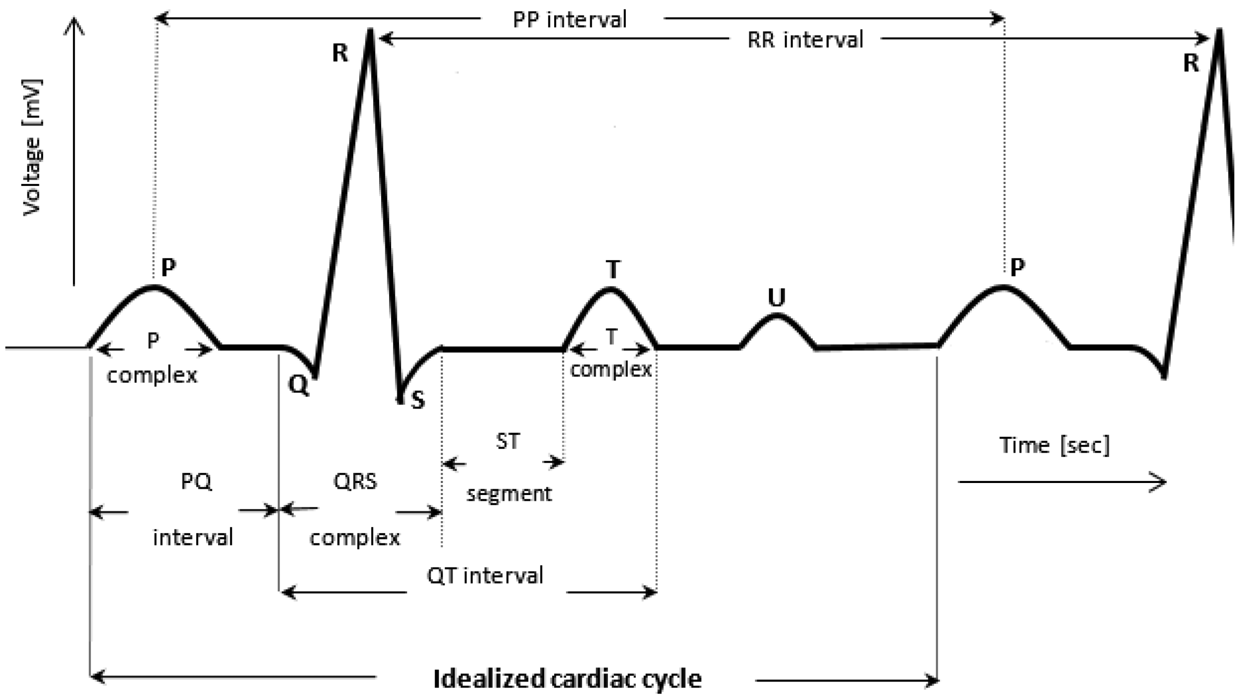


Figure 2.1: The ECG Waveform in a Healthy and Normal Heart [1].

In the study referenced in [54], it was found that an ECG has multiple applications, including:

- Heart rate tracking.
- Identifying symptoms such as shortness of dizziness, fainting, breath, palpitations, and angina pectoris.

- Assessing the effectiveness of medications and implanted mechanical devices in the heart.
- Monitoring heart conditions in individuals with a history of heart disease.
- Evaluating the extent of myocardial hypertrophy.

Heartbeat: According to [55] heart attack is not always indicated by a Heartbeat. But, if you experience chest pains or breathing difficulties or if it's a new symptom, it might be an early sign of a heart attack.

Heart rate: According to the authors in [56] heart rate is The frequency of heartbeats per minute in accordance with [57]. A normal heart rate is between 60 and 100 beats per minute. An increased risk of a heart attack may be associated with a resting heart rate greater than 76 beats per minute, As confirmed by [55].

2.6.2 Optical Coherence Tomography

According to the authors in [58], Optical Coherence Tomography (OCT) has emerged as a groundbreaking advancement in the field of optical imaging, offering a unique and innovative approach to capturing highly detailed images of biological tissues. As evidenced by [59], This imaging technique utilizes light to generate 2D and 3D images with a remarkable resolution down to the micrometer scale. With its wide range of applications in medical imaging and research, OCT has become an invaluable tool as mentioned in [60]. It employs a non-invasive, high-resolution optical imaging technology that relies on the interference between the signal from the object being examined and a local reference signal [59].

One of the remarkable features of OCT mentioned in [60], is its ability to produce real-time cross-sectional images, providing a two-dimensional representation of the object being scanned in both the lateral and axial coordinates. In the context of ocular health, OCT is particularly advanced, serving as an advanced eye scan suitable for individuals of all ages. Similar to ultrasound imaging, OCT employs light waves instead of sound waves to visualize the various layers comprising the back of the eye. Furthermore, certain OCT systems even enable simultaneous capture of a digital photograph of the eye's surface, allowing for a comprehensive analysis and cross-referencing of any areas of concern [2].

Figure 2.2 visually depicts the process and results of Ocular Coherence Tomography Imaging, showcasing the remarkable capabilities and applications of this advanced imaging modality.

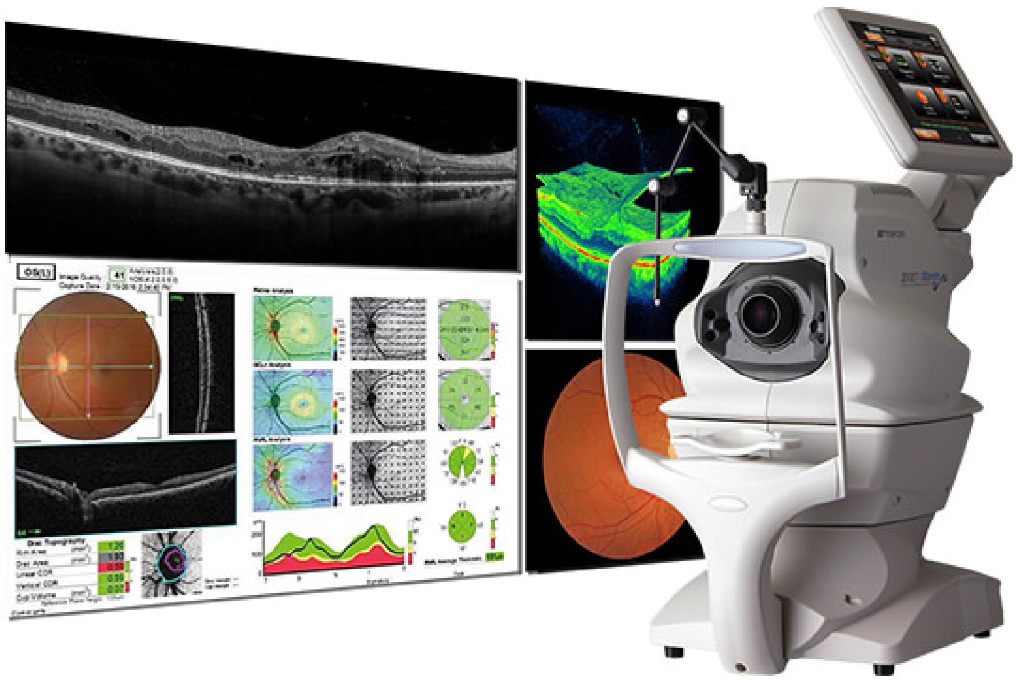


Figure 2.2: Ocular Coherence Tomography Imaging [2].

Benefits of OCT

The main benefits of OCT are:

- Live images below the surface at near microscopic resolution.
- Immediate and direct imaging of tissue morphogenesis.
- No ionizing radiation.
- There is no sample or subject preparation and no communication.

2.7 Internet of Medical Things

The term "Internet of Medical Things," or IoMT, refers to the expanding network of internet-connected medical equipment and software programs that can gather, transmit, and analyze healthcare data. IoMT encompasses a wide range of gadgets, including medical equipment used in hospitals and clinics as well as wearable fitness trackers and health monitors (see figure 2.3). IoMT is often referred to as healthcare IoT in accordance with [61]. By giving healthcare practitioners access to real-time data, enabling remote patient monitoring, enhancing patient outcomes, and lowering healthcare costs, the usage of IoMT has the potential to completely transform the way healthcare is delivered. IoMT devices may gather information on a range of health parameters, including blood pressure, blood sugar levels, and heart rate, and send this information to healthcare professionals for evaluation and therapy.

Figure 2.3: Category of HIoT application [3].

2.8 Internet of Medical Things architecture

An emphasis on medical devices and healthcare applications distinguishes the Internet of Medical Things from the larger Internet of Things (IoT), which shares a similar architecture [62]. The architecture typically includes the following layers (see figure 2.4):

1. Perception Layer: This layer consists of sensors, wearables, and other devices that collect health-related data. These devices can include blood pressure monitors, heart rate monitors, glucose monitors, and other medical sensors. The data collected by these devices is transmitted to the next layer of the architecture, According to authors in [63].

2. Network Layer: This layer consists of the communication infrastructure that connects the sensing layer to the cloud. This layer can include Bluetooth, Wi-Fi, cellular, or other wireless communication technologies. The network layer is responsible for transmitting the collected data to the cloud for further processing, As confirmed by [64].

3. Middleware Layer: processes and analyzes the collected data using databases, data analytics tools, machine learning algorithms, and other technologies that can provide insights into a patient's health, as shown in [63].

4. Application Layer: This layer consists of software applications that allow healthcare providers and patients to access and analyze health-related data. These applications can include electronic health records (EHRs), patient portals, and other healthcare-related applications, As indicated by [65].

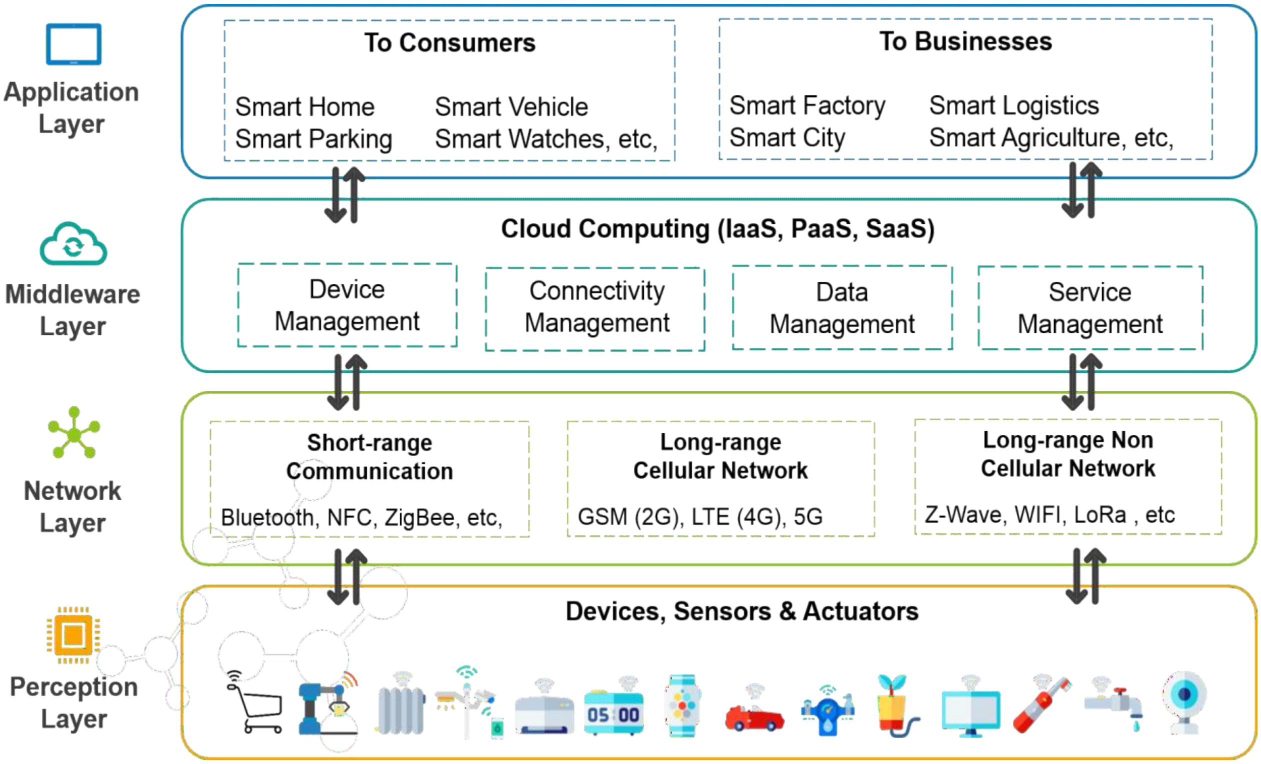


Figure 2.4: IoT Architecture [4].

2.9 Prediction of heart disease

Today, everywhere in the world, There are several persons whose health might be harmed by inadequate health care monitoring, including elderly patients with heart disease, While 17.9 million fatalities per year are reported, cardiovascular illnesses constitute the biggest cause of mortality worldwide, corresponding to 32% of all fatalities worldwide. The majority of Cardiovascular disease (CVD) fatalities occur in low- and middle-income nations According to WHO [25]. As a result, it is essential to identify cardiac illnesses as soon as possible or to predict if necessary additional treatment with counseling and medicine can start.

2.9.1 Normal structure and function of the heart anatomy

The heart is the mechanism that drives the body's blood to circulate. Your heart contracts each time it beats in order to pump blood. The right atrium, right ventricle, left atrium, and left ventricle are the names of the atria and ventricles that make up each of the two parts of the heart As depicted in [66] [67], As shown in Figure 2.5.

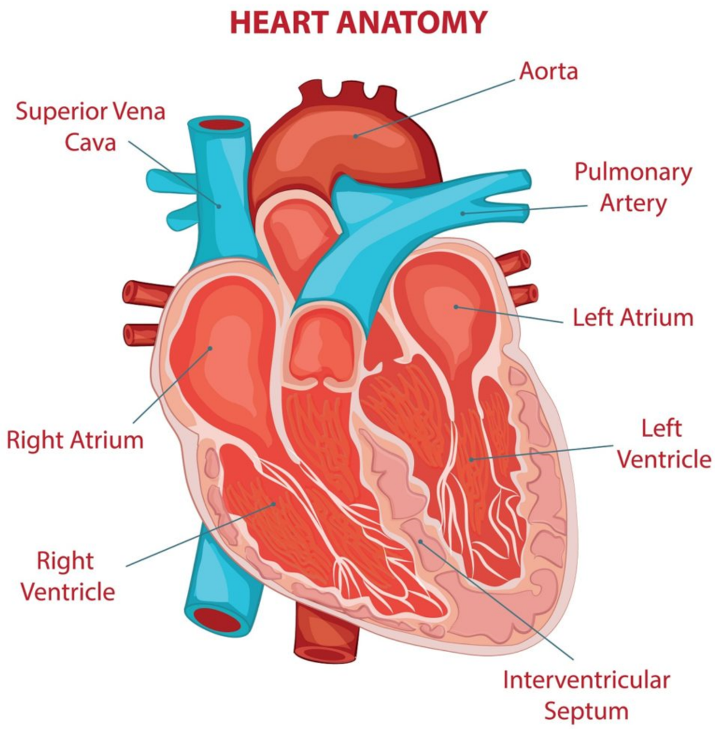


Figure 2.5: Conduction system of the heart [68].

In human anatomy and According to [67] the heart follow some step to pump blood:

1. The heart follows a specific pattern of blood flow. First Deoxygenated blood enters the heart through the vena cava and flows into the right atrium.

2. Next, It then moves to the right ventricle and is pumped to the lungs to pick up oxygen.

3. Oxygenated blood returns to the heart through the pulmonary veins, enters the left atrium, and then moves to the left ventricle.

4. Finally, it is pumped out through the aorta to supply the rest of the body.

The heart and the vessels through which blood circulates in the body make up the cardiovascular system.

2.9.2 Blood vessels

The arteries, veins, and capillaries are the three primary kinds of blood vessels that carry blood throughout the body According to authors in [67][69].

1. Arterioles are large blood channels that carry blood from the heart.

2. Veins are large blood channels that carry blood back to the heart.

3. The capillaries are the small blood vessels that connect the arteries and veins and through which the exchange of oxygenated and deoxygenated blood takes place, As shown in Figure 2.6.

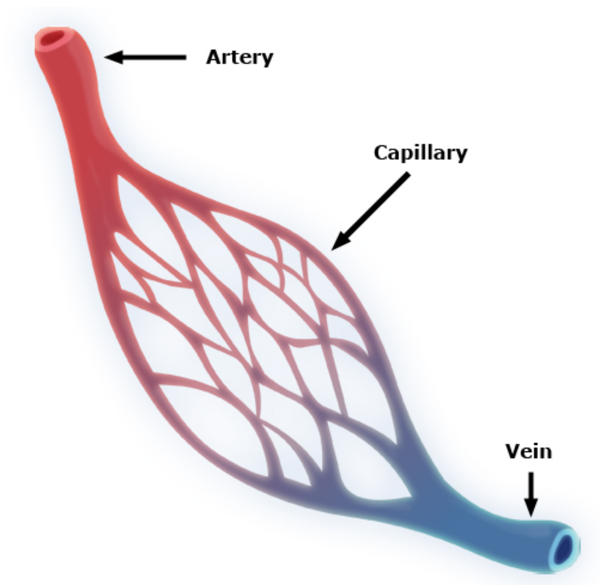


Figure 2.6: Arteries, capillaries, and veins [66].

.

2.9.3 The causes of heart disease

The primary risk factors for heart disease include high blood pressure, high blood cholesterol, age, excessive alcohol intake, physical inactivity, and diabetes. These factors contribute to the development of heart disease and compromise overall cardiovascular health, As confirmed by [70].

2.9.4 Types of heart disease

Heart Disease comes in a variety of forms, of which we have listed a few.

Coronary Artery Disease (CAD): Is the leading cause of death worldwide and the most common type of heart disease. also known as ischemic heart disease. Atherosclerosis happens when plaque builds up in the arteries that carry blood to the heart. This accumulation is known as atherosclerosis. As a result, the heart gets less oxygen and nutrients from the blood supply. Heart failure and arrhythmias are possible due to weak heart muscle, meaning the heart cannot pump blood the way it should. A heart attack may also result from artery plaque rupture due to obstructions that limit blood flow. The most prevalent sign of CAD is angina or discomfort and pain in the chest, According to [70].

Myocardial infarction: There are several names for myocardial infarction, including heart attack and coronary thrombosis. The blood flow to the heart is interrupted during an infarction, and this may lead to damage or destruction of part of it, In agreement with [71].

Cardiac Arrest: When a person's heart stops circulating blood throughout their body and they cease breathing regularly, it is known as cardiac arrest. Outside of a hospital, 90% of persons who experience a cardiac arrest pass die, frequently within minutes, As indicated by [71]. However, half of the cardiac arrests happen to people who did not know they had a heart problem[66].

Arrhythmia: Is a term used to describe an irregular heartbeat caused by malfunctioning electrical impulses that regulate the heart's rhythm. This condition can lead to irregular or excessively fast heartbeats (see figure 5.44), and in some cases, it can have severe consequences or even be fatal, As confirmed by [72] [73].

There are several types of arrhythmias, As defined by [73]:

- Tachycardia: This refers to a rapid heartbeat, where the heart beats faster than the normal resting rate.
- Bradycardia: In contrast, bradycardia is characterized by a slow heartbeat, where the heart beats slower than the normal resting rate.

- Premature contractions: These are abnormal early heartbeats that can occur before the regular heartbeat.

- Atrial fibrillation: This is a specific type of arrhythmia characterized by an irregular and often rapid. A person could have a sensation of a racing or fluttering heart, In agreement with [72].

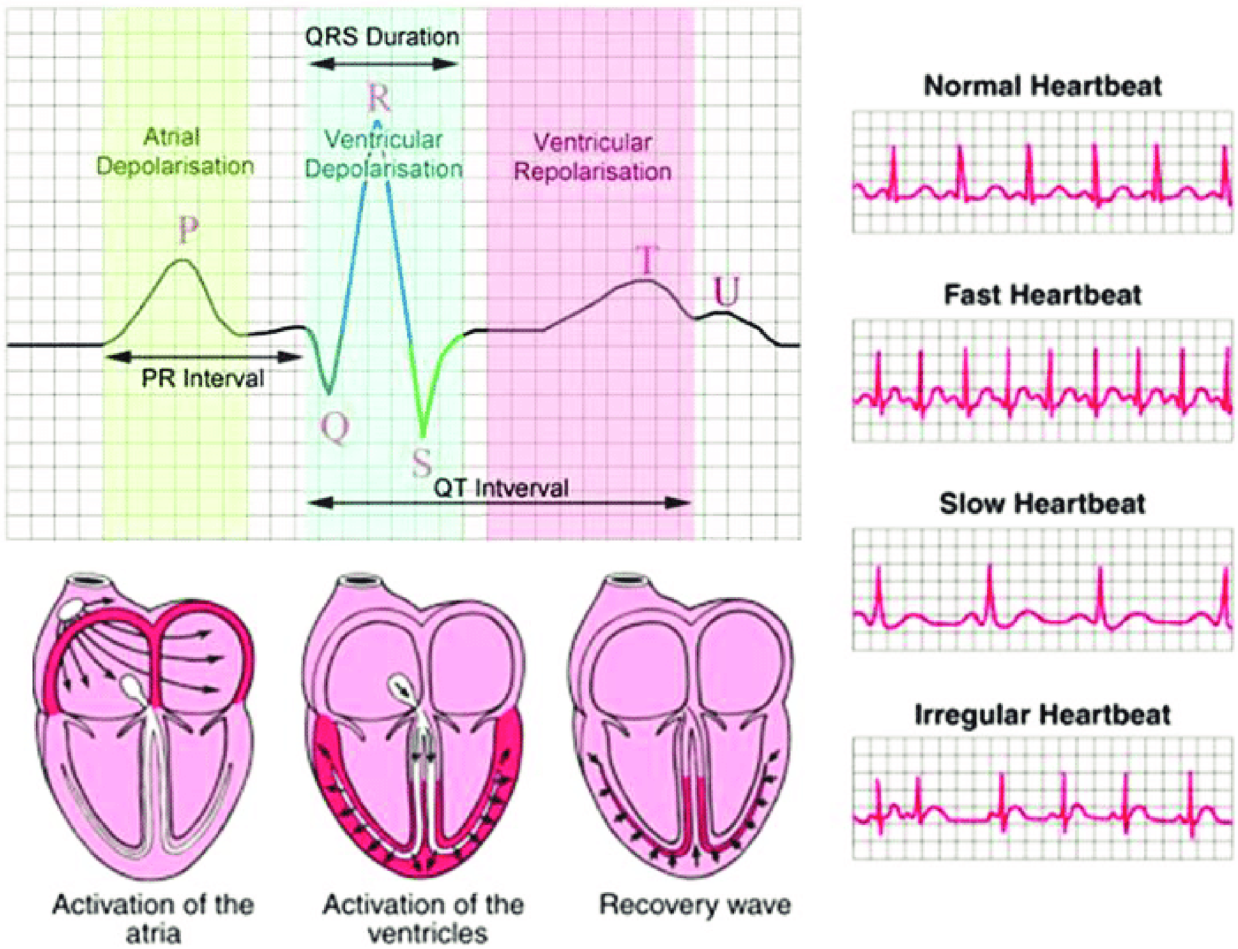


Figure 2.7: ECG Waves and Their Relation to Heart Nodes [5].

Heart failure: That can be brought on by issues with the pumping or relaxing mechanisms. and is often called congestive heart failure due to the fluid accumulation in the lungs, liver, legs, and feet, As confirmed by [66]. Untreated coronary artery disease, excessive blood pressure, arrhythmias, and other disorders can lead to heart failure. heart failure can be fatal, however, getting treatment for heart-related disorders quickly can help avoid problems, as shown in [73].

2.10 Retinal damage

The ability to see is highly dependent on the complex structure and proper functioning of the retina. Any type of stress or injury inflicted upon the retina can have severe consequences, causing disturbances in its architecture and leading to vision impairment, loss of visual function, and potentially even total blindness As confirmed by [74]. The delicate equilibrium necessary for clear and accurate vision can be easily disrupted, underscoring the vulnerability of this essential sensory organ. In this section, we will delve into the definition of the retina and explore several diseases that pose significant risks to this crucial ocular structure.

2.10.1 The retina

The retina [13][75], situated at the posterior of the eye, is a highly specialized tissue layer. Its primary function within the visual system is to capture and process incoming light, transforming it into electrical signals that can be conveyed to the brain. This intricate layer consists of different types of cells, according to [76] it including photoreceptor cells, neurons, and supportive cells, all working together as sensory receptors to facilitate our perception and comprehension of visual stimuli. With its complex structure and intricate neural networks, the retina enables the formation of intricate visual images, playing a vital role in our capacity to see and comprehend the surrounding world. Figure 2.8 shows understanding the components of the eye: an illustrated guide.

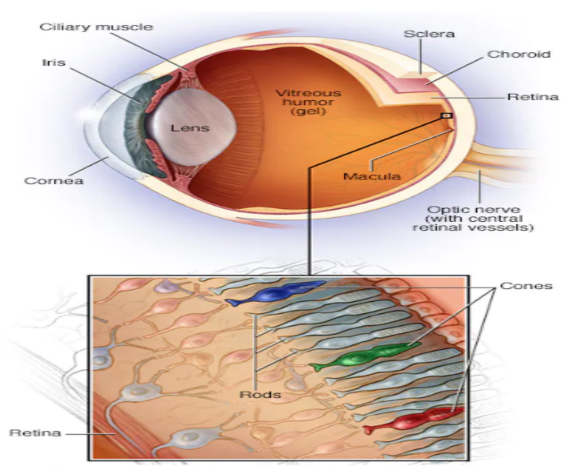


Figure 2.8: the components of the eye [6].

Function of the retina [76]

- To absorb photons of light.
- Translate light into a biochemical message.
- Translate biochemical messages into an electrical impulse.
- Transmit electrical impulses to the brain via ganglion cells[1].

Symptoms of problems with retinal [77]

Many retinal diseases share some common signs and symptoms. These may include:

- Blurry or distorted vision.
- Loss of peripheral vision.
- Double vision (diplopia).
- Light sensitivity.
- Your vision is getting noticeably worse.

2.10.2 Age-related Macular Degeneration

Age-related macular degeneration (AMD) is a prevalent and irreversible condition that leads to significant vision loss, often resulting in legal blindness, especially among the elderly population, typically individuals aged 60 and above [78].

Progressive vision-threatening ocular disease, affecting central region of the retina the macula (see Figure 2.9) . Due to the degenerative nature of AMD, which first affects the retinal pigment epithelial cells before progressing to the photoreceptors, central vision can be disturbed or lost in part, leading to legal blindness [7].

AMD is classified into two forms: wet and dry. These forms are distinguished by the presence or absence of blood vessels that have abnormally invaded the retina. The majority of AMD

[1] Ganglion cells:bridging neurons that connect the retinal input to the visual processing centres within the central nervous system.

cases fall under the atrophic dry AMD subtype, while the neovascular wet AMD subtype is characterized by the presence of choroidal neovascularization (CNV), where new blood vessels grow beneath the retina [7] [79].

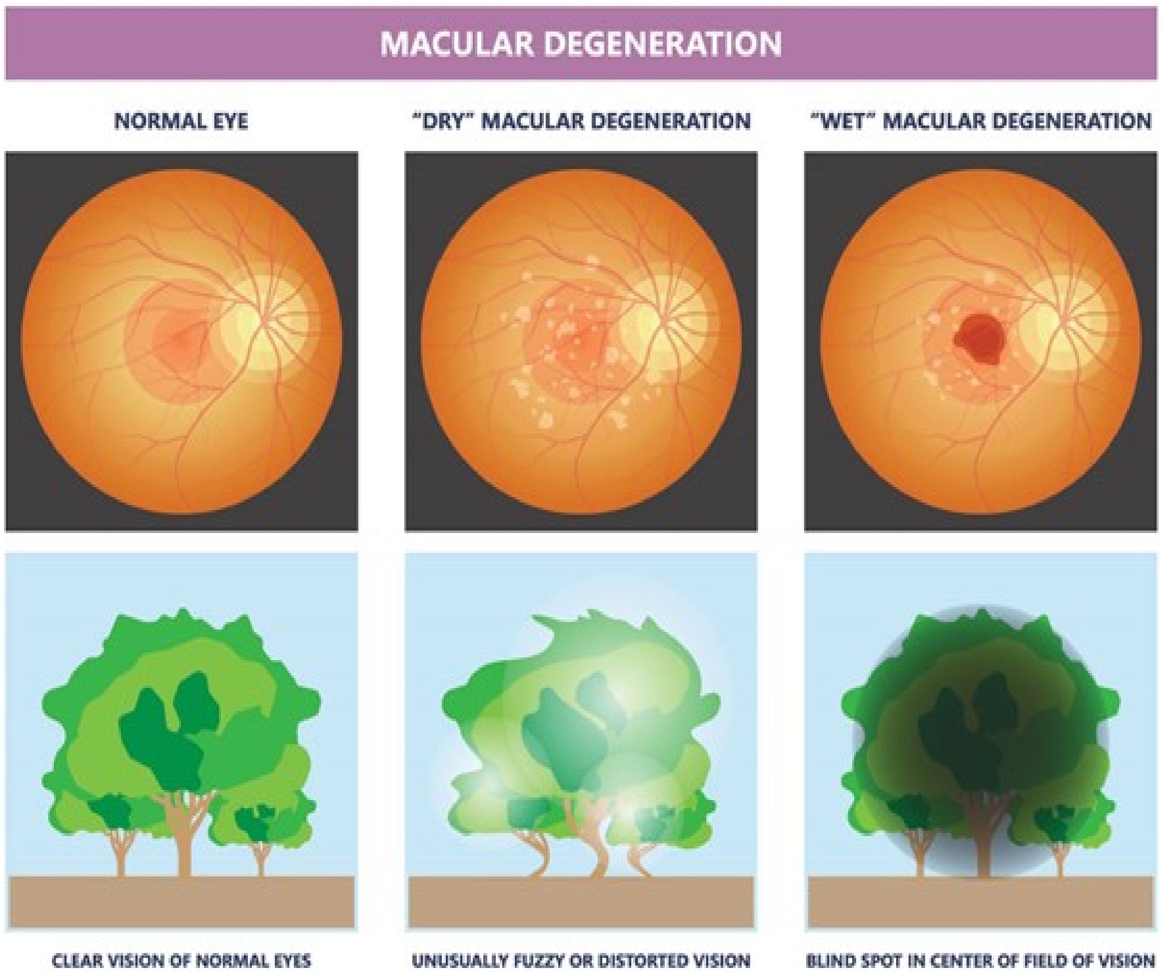


Figure 2.9: The difference between a normal eye and an eye with AMD [7].

2.10.3 Choroidal Neovascularization

"New blood vessels" are what is meant by choroidal neovascularization. The choroid, a layer beneath the retina that contains blood vessels, is where these abnormal new blood vessels originate. Vascular Endothelial Growth Factor (VEGF), which the retinas of people with AMD overproduce(see Figure 2.9), causes new blood vessels to sprout from the choroid and then grow into the retina. It is possible for blood fluid, and even red blood cells, to enter the retina through new arteries since they leak in contrast to normal vessels (see Figure 2.10). Due to the formation of a "blister" in the retina, which is ordinarily flat, this fluid can rapidly distort vision. This fluid can harm the retina over days to months, killing photoreceptors [8].

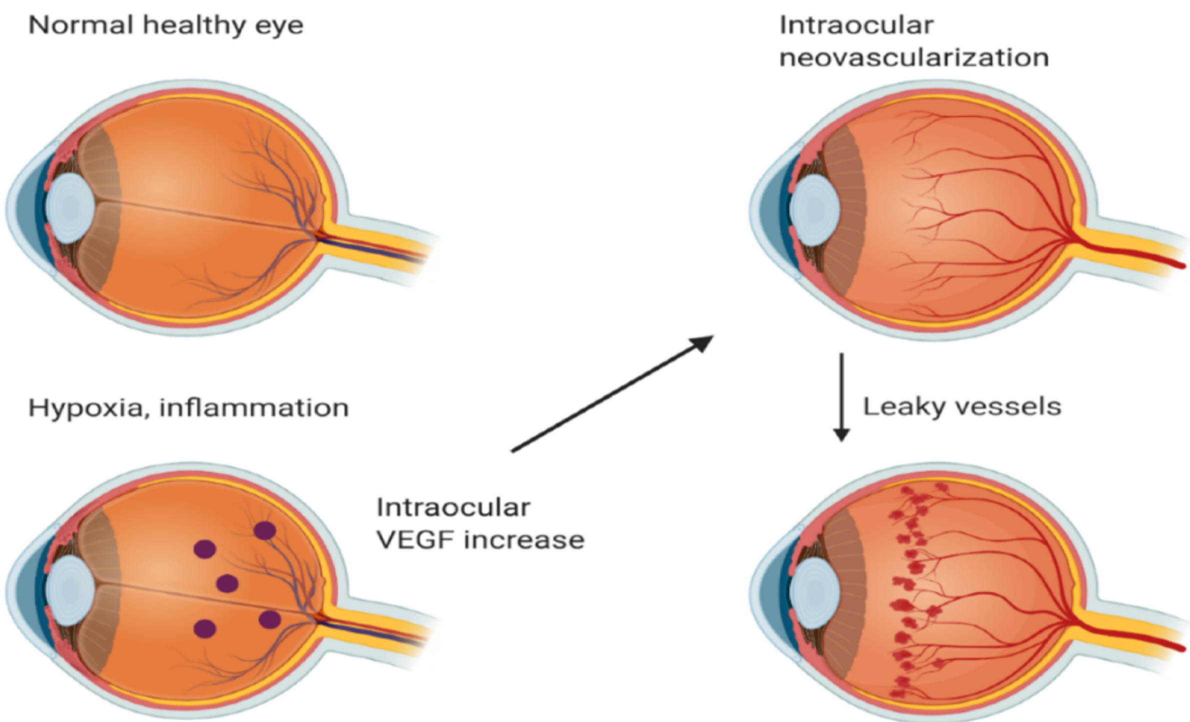


Figure 2.10: Stages of Choroidal Neovascularization (CNV) [8].

2.10.4 Diabetic Macular Edema

High blood sugar levels in diabetics can harm or totally block the small blood vessels in the retina. The blood arteries develop tiny swellings known as micro-aneurysms, which allow fluid to flow into the retina. The macula, a region of the retina, may expand as a result of the fluid (see Figure 2.11). Blindness or vision issues may result from it. Eye pain or redness, blurring or spotty vision, steadily decreasing eyesight, and sudden vision loss are all signs of diabetic macular edema [13][76][80]. Diabetic macular edema symptoms are Eye pain or redness, blurred or patchy vision, gradually worsening vision, and sudden vision loss as it mentioned in [13].

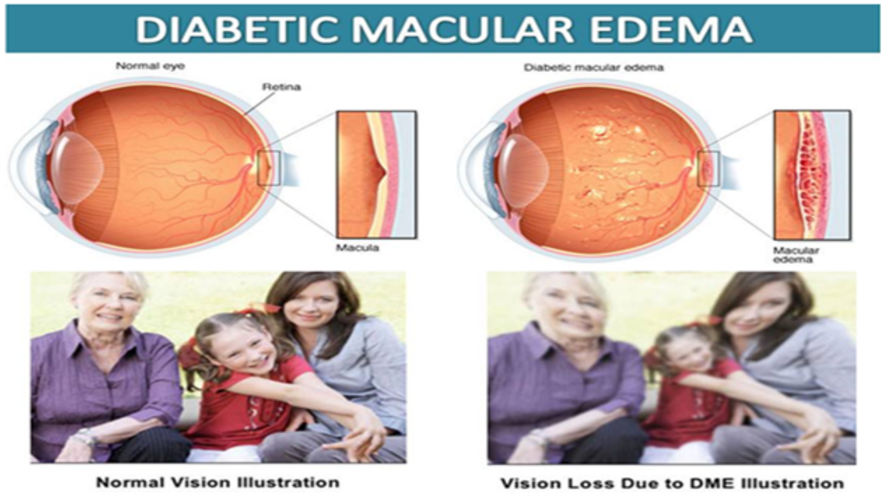


Figure 2.11: The Difference between a normal eye and an eye with Diabetic Macular Edema [9].

2.10.5 Multiple drusen

Drusen, characterized by yellow deposits located beneath the retina, consist of lipids and proteins. It is important to note that drusen alone do not cause age-related macular degeneration (AMD) (refer to Figure 2.12). However, the presence of drusen increases the likelihood of developing AMD, which is responsible for vision loss [81]. Drusen are commonly associated with aging and are typically observed in individuals aged 60 and above. Although most people with drusen do not experience symptoms, the development of drusen signifies an increased risk of developing wet AMD [82].

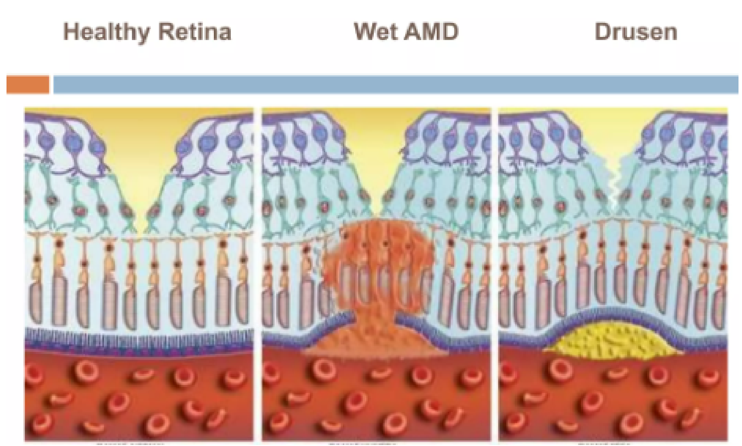


Figure 2.12: Understanding the differences between a healthy retina, retina with Wet AMD, and retina with Drusen [6].

2.11 Emotion recognition

Emotion recognition has emerged as a significant area of research in the healthcare field. It involves the detection and analysis of human emotions, allowing healthcare professionals to gain valuable insights into patients' emotional well-being.

In the upcoming subsection, we will provide a comprehensive definition of emotion.

2.11.1 Defining emotion

According to Moors et al. [83], there is no currently established scientific definition for emotion due to the varied understanding of emotions from the many theories and perspectives proposed. In their study, Mauss et al. [84] described emotions as reactions to events or stimuli, with a short duration, and involving coordinated responses in language, behavior, physiology, and neural activity.

In the next subsection, we will delve into the topic of emotion recognition in human-computer interaction.

2.11.2 Emotion recognition in human-computer interaction

In the study conducted by Tivatansakul et al. [85], it was found that the recognition and awareness of emotions are crucial in human-computer and human-robot interaction. The authors emphasize the importance of incorporating emotion recognition in HCI to facilitate natural and meaningful interactions, as it allows computers and their applications to better understand and respond to human emotions. Below are the main categories of techniques proposed for detecting and classifying user emotions.

Speech emotion recognition: Language serves as a communication medium between individuals, and according to Dhuheir et al. [86], researchers have employed machines to interact with people and extract emotions. However, achieving natural human-machine interaction requires significant effort. In the study by Ververidis et al. [87], emotional speech recognition is proposed as a method to automatically recognize emotional or physical states through the human voice. The authors highlight the importance of analyzing the paralinguistic aspect of speech in identifying emotional states.

Emotion recognition from facial expression: Analyzing a user's facial expressions to extract and validate emotional cues plays a crucial role in enhancing the level of interaction in human-machine communication systems. According to Tivatansakul et al., [85], this approach detects human emotions by analyzing facial muscle movements, including those of the eyes, mouth, eyebrows, and facial texture. Various studies have employed computer vision systems to automatically analyze and detect changes in facial movements using visual information.

Behavioral cues: In a study on human behavior [88], it was found that interactions and actions exhibited by individuals provide valuable insights into their emotional states and dynamics. The authors emphasized the importance of observing individuals' actions and discerning their emotions as they engage in various activities, as this can greatly assist psychologists in providing targeted and effective treatment.

Emotion recognition using biological signals: Research interesting in the affective computing sector [89] has been generalized to physiological modalities through the accelerated development of technological solutions in healthcare. In the study in [90] it is mentioned that the main physiological signals commonly used to assess human emotions are Blood Volume Pressure (BVP), Skin Conductive Resistance (SCR), Electromyogram (EMG), Galvanic Skin Response(GSR), Electrocardiogram (ECG), Electroencephalogram (EEG), temperature, respiration (RSP). EEG, ECG, RSP, GSR, and EMG change in certain ways when a person faces certain situations. Physiological signals are the body's responses to the central nervous system (CNS) and autonomic nervous systems (ANS) according to Connon's theory [91]. One of the main advantages of the latter method is that the CNS and ANS are activated almost involuntarily and therefore not easily controlled.

Multimodal emotion recognition: A study on multimodal emotion recognition [92], aims to replicate the predictive abilities of humans by combining biological and behavioral characteristics. Unlike unimodal emotion recognition, which focuses on a single data source, multimodal emotion recognition involves processing multiple data sources simultaneously. The study highlights the need for robust approaches in order to address the variations in predictive accuracies observed among different multimodal methods.

In the next section, we will focus on the application of ECG in emotion recognition and its relevance across various fields.

2.11.3 Applications of emotion recognition systems using Electrocardiogram

Emotion recognition systems have many potential applications, including healthcare, marketing, entertainment, e-learning, security, and human monitoring. According to [93], there were three main areas of application that used ECG as a primary data source for emotion recognition. In the upcoming subsection, we will explore the various applications of emotion recognition systems using ECG.

Monitoring: Monitor human emotions and assess behavior during specific tasks and reactions in crisis situations. For example, in [94] real-time ECG data processing can be used during car racing, emotions and psychological situations commonly affect driver behavior and reactions, which can lead to accidents. According to a study conducted on emotion recognition by physiological signals [93] it was found that this approach can support clinical diagnosis when the patient's expressiveness is limited or absent as in Parkinson's disease, Huntington's disease, cortical lesions, and stroke. So the goal is therefore to assess the patient's emotional state.

Medicine: According to a study on emotion recognition [95], utilizing ECG signals can be particularly useful in working with autistic patients and individuals with intellectual disabilities. The ECG signals, which reflect the activity of ANS, provide insights into a person's genuine emotional state. Additionally, the authors in another study [96] demonstrate that real-time emotion detection from ECG signals is beneficial and effective in detecting and analyzing negative emotions, enabling immediate assistance.

Marketing: Emotion recognition can be used for marketing, the system proposed in [97] can be used for website optimization, and a system capable suggests taking a break when emotions are detected by webcam and ECG sensors are classified as negative.

Emotion recognition can be achieved through the utilization of various modalities, and these modalities will be elaborated on in the following sections.

2.11.4 Emotion models

Experts from different backgrounds [83][84] have tried to uncover the universal definition of emotion, however, the exact definition is not widely recognized by psychologists, and none of them reached an agreement to establish a single emotional model.Psychologists commonly adopt two prevalent approaches to model emotions, namely discrete categories and emotional dimensions, as highlighted in a relevant study [89].

Discrete Emotional Model: In the study by Song et al. [98], the researchers define DEM as an approach that classifies emotions into standard terms. For example, Robert Plutchik [99], proposed eight main emotions: surprise, fear, disgust, sadness, anger,anticipation, joy, and trust, and arranged them in a colored wheel, as shown in Figure 2.13. Emotional discrete models use words to reflect their feelings instead of quantitative analysis. In the study conducted by Shu et al. [10], it was observed that emotions such as mixed feelings are challenging to precisely describe in words, making their analysis difficult.

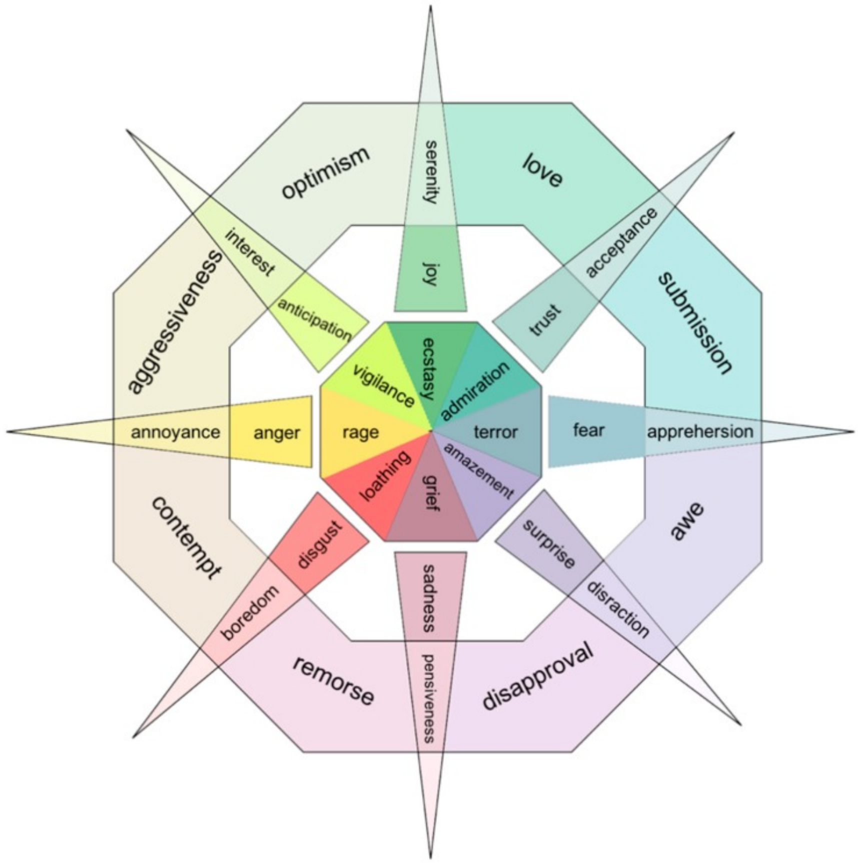


Figure 2.13: Plutchik's Color Wheel: Depicting the Eight Primary Emotions [10].

Affective Dimensional Model: When psychology researchers conducted in-depth research, as mentioned in a study by Shu et al.[10], they discovered connections between various emotions such as hate and love, pleasure and displeasure, and like and dislike. These emotions encompass specific states with varying intensities, with happiness being described as either very happy or somewhat happy. This research emphasizes the complexity and wide range of human emotions, which can be organized and understood through fundamental dimensions.

Lang [100] proposed that human emotions can be represented on a two dimensional plane based on valence and arousal, as depicted in figure 2.14.

Mehrabian [101] extends the emotion model from 2D to 3D (see Figure 2.15). This extended model includes an additional dimension called dominance. The components of Valence Arousal Dominance (VAD) model are summarized as follows:

- **Valence:** Refers to the positivity or negativity of an emotion.

- **Arousal:** It measures the degree of excitement, stimulation, or energy experienced during an emotional state.
- **Dominance:** Reflects the perceived level of control or power within an emotional experience.

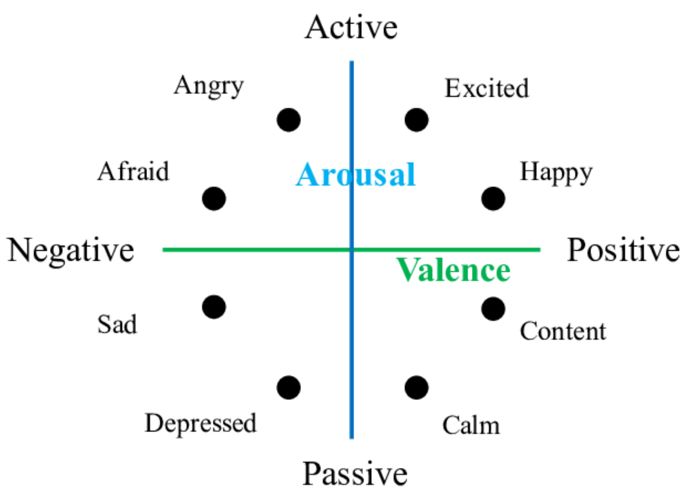


Figure 2.14: Valence-Arousal model for emotion representation [11].

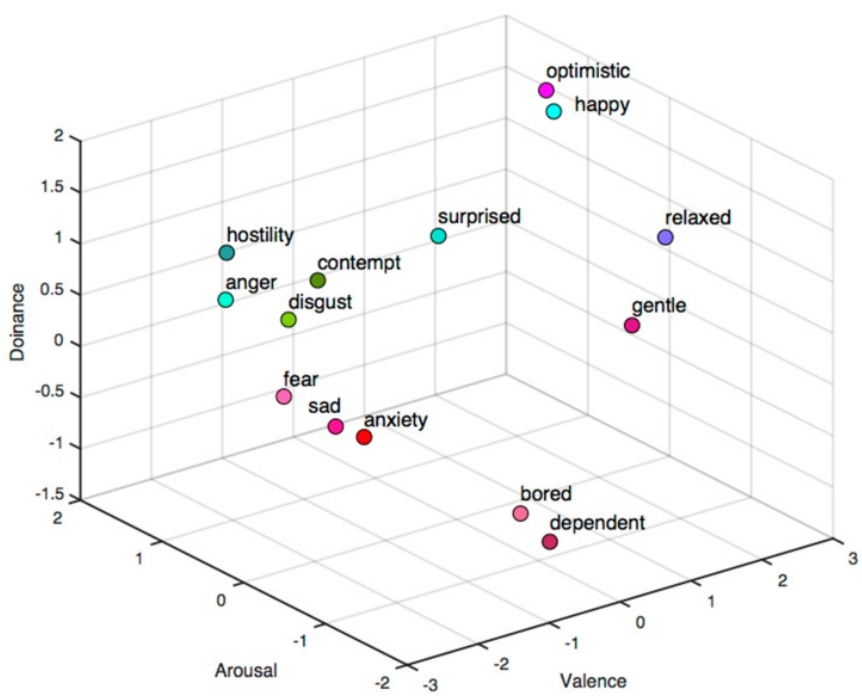


Figure 2.15: Valence-Arousal-Dominance model for emotion representation [10].

Binary Emotional Model: This model proposed by [102], categorizes emotions into two broad categories, positive and negative. This model suggests that all emotions can be classified as either positive or negative, depending on their subjective valence. Positive emotions are those that create a pleasant feeling or experience, such as happiness, joy, and love. Negative emotions, on the other hand, are those that create an unpleasant feeling or experience, such as sadness, anger, and fear .as shown in figure 2.14.

In addition, it has been mentioned in the study in [89] reducing the emotional model to just two classes can result in a more streamlined and efficient emotion recognition system. This simplification can lead to higher accuracy in both the training and testing of the models. Additionally, the studies [103][104] have further supported these findings, emphasizing the potential benefits of simplifying the emotional model into two classes for improving the efficiency and accuracy of emotion recognition systems.

2.12 Related work

In the domains of retinal damage classification, heart disease detection, and emotion recognition, a multitude of deep learning-based and machine-learning models have been developed to automate these processes. This section focuses on showcasing three notable examples in each of these fields. These examples aim to highlight specific problems, the challenges they present, the proposed solutions, and the inherent limitations associated with each approach.

- **Abdelhafid Errabih et al (2022) [105]**. The objective of this study was to automate the diagnosis of retinal diseases by utilizing OCT images. To achieve this, the researchers employed deep learning techniques and utilized pre-trained neural network models, specifically VGG16 and InceptionV3. It is important to note that the dataset used in the study was relatively small, with 1000 images for each class. Transfer learning and fine-tuning approaches were implemented to enhance classification performance. The findings indicated that the VGG16 model outperformed InceptionV3, achieving an impressive accuracy of 0.93. However, there is still potential for further enhancements, such as incorporating additional retinal disorders and developing a comprehensive application that provides ophthalmologists with advanced capabilities for analyzing OCT images.

- **Jen Hong Tan et al (2018) [106]**. This study addresses the issue of AMD in the elderly and emphasizes the importance of early detection to prevent vision impairment. To tackle this problem, the researchers developed a sophisticated fourteen-layer deep CNN model. The model achieved impressive accuracy rates of 91.17% and 95.45% through evaluations using blindfold and ten-fold cross-validation strategies, respectively. In addition to its high accuracy. However, practical implementation in clinics may encounter challenges related to hardware and software compatibility, as well as integration with existing systems.

- **Philippe M Burlina et al (2017) [107]**. The focus of this study is on AMD in the elderly and the significance of early detection for preventing vision impairment. To address this issue, the researchers devised a sophisticated fourteen-layer deep CNN model. Notably, the model demonstrated remarkable accuracy rates of 91.17% and 95.45% in a blindfold and ten-fold cross-validation evaluations, respectively. The model's high accuracy is a notable achievement. However, the practical implementation of the model in clinics may encounter challenges pertaining to hardware and software compatibility, as well as integration with existing systems.

- In [108], a full model comprised of a modified differential evolution (DE) approach, a fuzzy analytical hierarchy process (AHP), and a feedforward neural network was created by Vivekanandan and Sriman. (FNN). The modified differential evolution approach was used to choose the most crucial traits. In order to forecast heart illness, a condensed set of characteristics was input into an optimized model for a fuzzy AHP with FNN. The simulation findings that are being provided are based on the 83 percent accurate modified differential evolution technique.

- In [109] heart disease may now be predicted more accurately thanks to a method developed by Mohan et al. that combines a random forest (RF) and linear model (LM) to find key traits. The method makes the most of RF and LM. The model is put into practice utilizing numerous feature combinations and well-known classification techniques. The prediction accuracy was reported by the authors to be 88.7.

- A deep convolutional neural network (CNN) algorithm was created by Mohammad et al.

[110] In the suggested study, the algorithm aimed to accurately classify heartbeats, specifically five different arrhythmias. The researchers analyzed patient data using a deep CNN approach. the suggested method demonstrated average accuracies of 93.4% for arrhythmia classification and 95.9% for MI classification.

- The study in [111] demonstrated a framework for emotion recognition using multimodal physiological signals from the respiratory belt (RB), PPG, and fingertip temperature (FTT). It is shown that decision-level fusion from multiple classifiers improved the accuracy rate of emotion recognition both for arousal and valence dimensions.

- This paper [21] presents DREAMER, a multi-modal database consisting of EEG and ECG signals recorded from 23 participants where each participant rated his emotional response along the scales of valence, arousal, and dominance.

- This article [112] presents a robust physiological model for the recognition of human emotions, called a Deep Physiological Affect Network. the model is based on a convolutional long short-term memory (ConvLSTM) network and a new temporal margin-based loss function. this paper uses EEG signals to recognize calm, anger, and sadness.

2.12.1 Synthesis comparative study

This section aims to provide a comparative analysis of different related work mentioned in the previous subsection. The criteria for comparison include the authors, publication year, purpose of the work, methodology, dataset utilized, and accuracy achieved. By considering these factors, we can gain a comprehensive understanding of the similarities and differences among the studies and evaluate their respective contributions to the field.

- The table below (see Table 2.1) represents earlier work conducted on classifying retinal damage.

Table 2.1: Summary of recent studies on classifying retinal damage.

Authors	**Abdelhafid Errabih et al** [105]	**Jen Hong Tan et al** [106]	**Philippe M Burlina et al** [107]
Year	2022	2018	2017
Purpose of the work	Identifying Retinal Diseases on OCT Image (CNV, DRUSEN, and DME)	Automatically diagnose AMD at an early stage	Automatically detect AMD
Method	Combination of deep learning and transfer learning techniques, utilizing the VGG16 and InceptionV3 neural network models	A fourteen-layer deep CNN model was developed and evaluated using blindfold and ten-fold cross-validation strategies	Developing deep convolutional neural networks specifically trained for automated grading of AMD from fundus images
Used dataset	Public dataset kermany 2018	The Ophthalmology Department of Kasturba Medical College (KMC) dataset	The National Institutes of Health AREDS dataset
Number of classes	4	4	2
Accuracy	93.12%	91.4%	88.4%

- The table 2.2 represents earlier work on the prediction of heart disease:

Table 2.2: Related work on prediction of heart disease.

Authors	Vivekanandan,T., and [108]	Mohan,S.et al. [109]	Mohammad et al [110]
Methodology	Prediction of heart disease	Prediction of cardio-vascular disease	Classification of heart-beats
Approach	Modifed DE with Fuzzy AHP FFNN	Hybrid random forest with a linear model	Deep convolutional neural networks
Use of IoT Device	No	No	No
Accuracy	83%	88.7%	93.4%
prediction of sequence ECG	No	No	No
classification of sequence ECG	No	No	Yes

- The table 2.3 represents the related work on emotion recognition:

Table 2.3: Related work on emotion recognition.

Authors	Ayata et al [111]	katsigiannis et al. [21]	kim et al [112]
Paper title	Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems	DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices	Deep Physiological Affect Network for the Recognition of Human Emotions
Used dataset	DEAP	DREAMER	DEAP
Bio-signals	RB, PPG, and FTT	ECG, and EEG	EEG
Emotion recognized	Arousal and valence	Arousal,valence, and dominance	Calm, angry, sad
Accuracy	73.08%	62,32%	89,96%

2.13 Conclusion

Several technologies can reduce the overall costs of preventing or managing chronic disease, these include devices that constantly monitor health indicators, devices that automatically administer treatments, or devices that track health data in real time when a patient self-administers treatment. Due to the increased access to high-speed internet and smartphones, many patients have begun to use these technologies to manage various health needs, and now that artificial intelligence (AI) has appeared in healthcare, it has reshaped the way we diagnose, treat, and monitor patients and has dramatically improved outcomes by producing more accurate diagnoses and enabling more personalized treatments. Also, The Internet of Things has changed the world, providing powerful applications and systems in various fields and contributing to facilitating people's lives and accelerating economic growth, including industry, transportation, and health care, which provides a safe, effective, and easily deployed health monitoring system that can guarantee high-quality healthcare services with a fraction of the cost.

To understand the context of the work, the chapter has provided a summary of the challenges in traditional healthcare and patient monitoring, serving as a foundation for comprehending the capabilities offered by healthcare systems based on AI and the Internet of Medical Things. In this regard, three different systems affecting both physical and mental healthcare have been discussed. The upcoming chapter will offer a comprehensive exploration of the fundamental concepts and algorithms in machine learning and deep learning.

“Without intelligence there is no value.”

Kiva Allgood

Chapter 3

Machine Learning and Deep Learning

3.1 Introduction

In healthcare organizations, the implementation of AI applications [32] has become prevalent for tasks like diagnosis, treatment recommendations, patient engagement, and administrative activities. AI has demonstrated the capability to outperform humans in certain healthcare tasks.

In the previous chapter, we have provided an overview of the current state of the art in the field. The objective of this research has been to develop a system that harnesses the power of artificial intelligence algorithms, specifically deep learning for classifying retinal diseases, and Machine Learning for emotion recognition and predicting heart disease.

In this chapter, we will delve into the algorithms and techniques employed to develop and achieve the desired functionality of this system. Through careful selection and implementation of various approaches, we have created robust and efficient models that are well-suited for their respective tasks.

3.2 Machine learning

Machine learning as defined in [113], is a field of study within AI and computer science that centers on utilizing algorithms and data to replicate human learning processes, with the aim of gradually enhancing its precision. By using statistical techniques, algorithms are trained to classify or predict outcomes, generating insights that can inform decision-making processes within applications and ideally boost critical growth metrics.

3.2.1 Random Forest

RF is a machine learning algorithm that leverages an ensemble of decision trees to make predictions. According to [114], it combines the predictions from multiple decision trees to achieve improved accuracy and robustness in the final prediction. The algorithm works by building a collection of decision trees, each trained on a random subset of the training data and considering only a subset of features at each node. This randomness helps to reduce overfitting and introduce diversity among the trees. The final prediction is made by aggregating the individual predictions from each tree, either through majority voting for classification or averaging for regression, as mentioned in [115].

3.3 Deep Learning

Deep learning is a subset of machine learning [116]. The primary characteristic of deep learning is the utilization of neural networks that comprise numerous hidden layers (see Figure 3.1). Therefore, the term "deep" in this context pertains to the substantial number of layers, exceeding two hidden layers [76]. These techniques are more efficient than machine learning due to their deep structure and ability to autonomously learn the significant features from the dataset and produce an output [117] [118]. Deep neural networks are employed to mimic the learning mechanism of the human brain and extract relevant features from vast amounts of data, including but not limited to sound, text, and images, as it mentioned in [119].

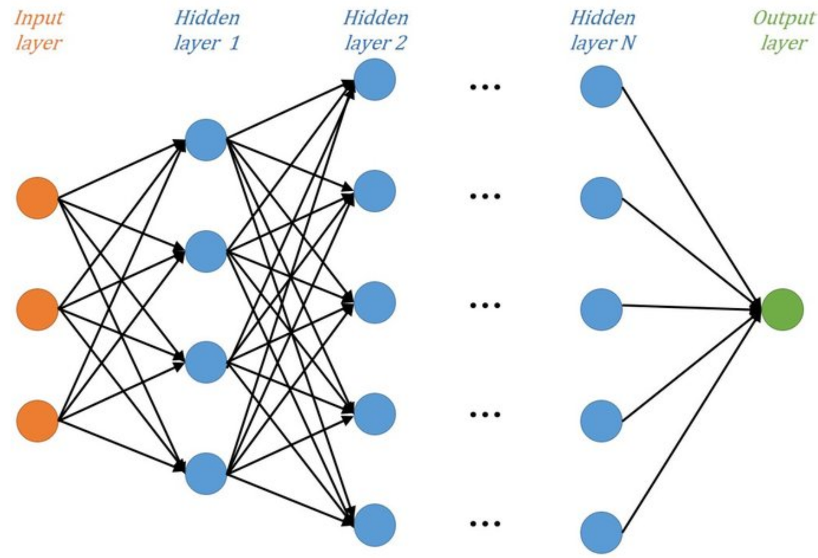


Figure 3.1: Deep Learning Architecture: Exploring the Structure of Neural Networks [12].

3.3.1 Artificial Neural Network

ANN consists of an input layer, one or more hidden layers, and an output layer. The input layer is in charge of gathering the problem-specific input characteristics. Between the input and output layers are the "hidden layers," which are the intermediary layer. Because their neurons do not immediately interact with the input or output. The network's last layer, the output layer, offers the network's output or prediction [120] [121].

3.3.2 Convolution Neural Networks

CNNs, or Convolutional Neural Networks, are a specialized type of deep learning neural network that has been developed specifically for the processing and analysis of visual data like images and videos. The primary goal of CNNs is to replicate the capabilities of human vision. These networks are comprised of three main types of computational layers: convolutional, pooling, and fully connected like it mentioned in [122]. The architecture of CNNs is structured into two main blocks, as depicted in Figure 3.2, and also incorporates additional layer types that are tailored to specific tasks [13].

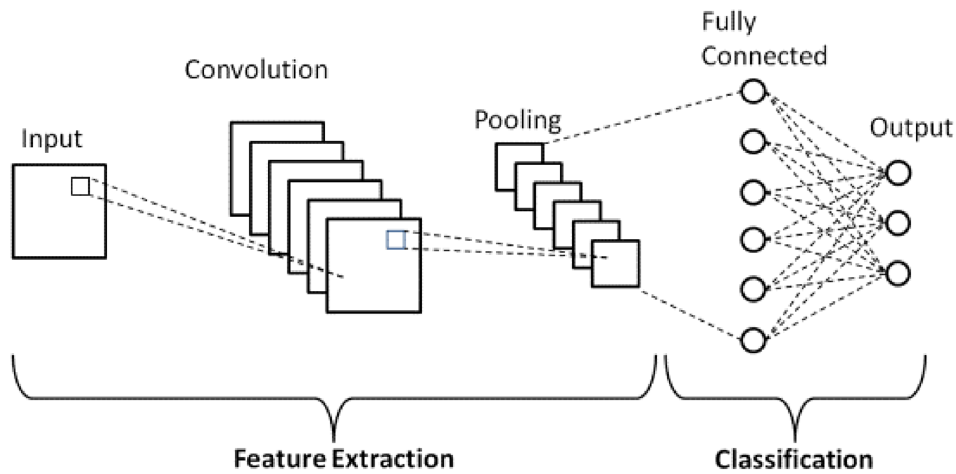


Figure 3.2: Composition of Convolutional Neural Networks Architecture [13].

3.3.3 The basics of Convolutional Neural Network

A CNN comprises three distinct types of computational layers: convolutional, pooling, and fully connected as defined by [123].

Convolutional layer: In [13][105], the primary objective of CNN is to detect specific features in the input images. This is achieved by applying convolution to the input picture using a filter or kernel, which serves as the fundamental operator of linear image processing. The convolution process involves sliding a window representing the desired feature over the image and calculating the convolution product between the feature and each portion of the scanned image. In computer terms, an image is represented as a matrix of pixels with varying values. Figure 3.3 illustrates an example of the sliding process in a convolutional layer.

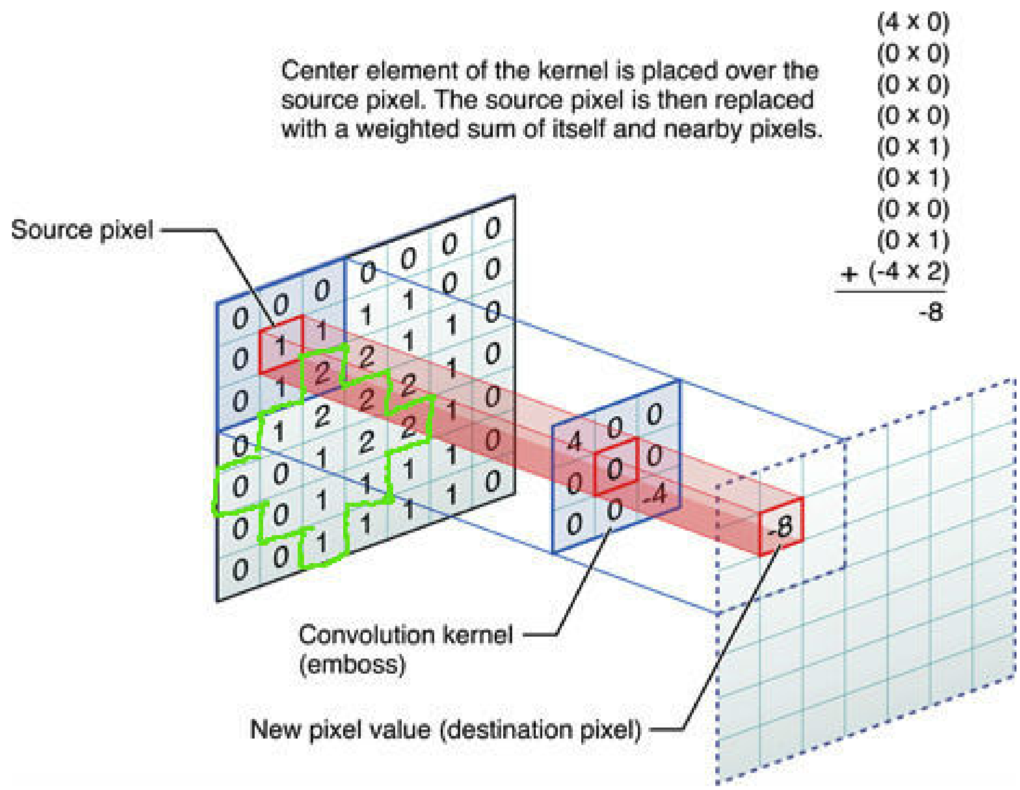


Figure 3.3: Example of convolutional layer slides [14].

Pooling Layer: The main idea of pooling is down-sampling in order to reduce the complexity for further layers [15]. We examine three types of pooling layers: min-pooling, average pooling, and max-pooling.

In this example for the pink part on the left of the map features in Figure 3.4, the important feature is 6, and we do the same to the rest parts of the map features.

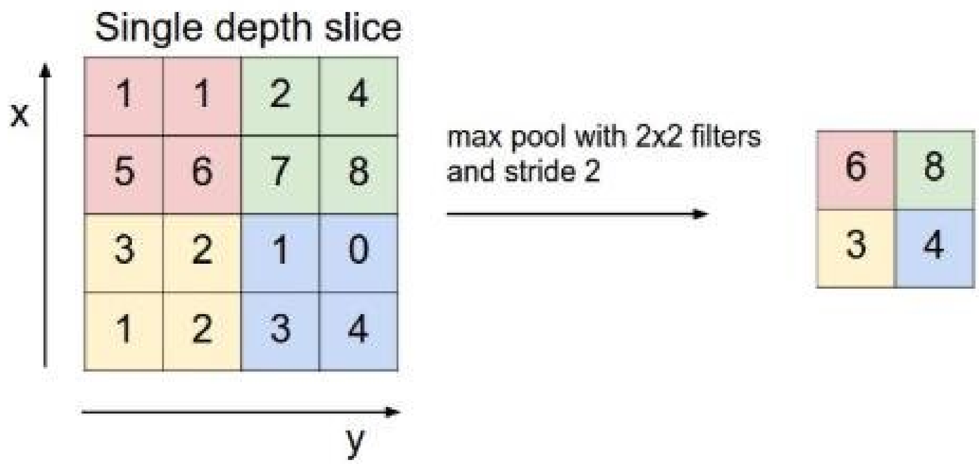


Figure 3.4: Demonstration of Max-Pooling [15].

Fully Connected Layer: The authors in [124] suggest that the fully connected layer is an essential component of neural networks, especially in tasks such as image recognition and classification. It follows the classic neural network architecture, where every input from one layer is connected to all activation units of the next layer. Each input unit has distinct weights for each output unit, enabling the conversion of inputs into a vector for further analysis. The fully connected output layer generates predicted labels for each class by evaluating the weighted inputs. In the context of image classification, the last fully-connected layer calculates scores for each class using the extracted features obtained from earlier stages of the network.

3.3.4 Prominent Convolutional Neural Network Architectures

We will specifically concentrate on four popular architectures, namely DenseNet, VGG16, VGG19, and InceptionV3, as they are the chosen models for our work.

- **DenseNet:** DenseNet [16] comprises a feature layer (convolutional layer) that captures low-level features from images, multiple dense blocks, and transition layers connecting adjacent dense blocks (as shown in Figure 3.5). According to [17], it establishes connections from one layer to all the subsequent layers in a feed-forward manner (as depicted in Figure 3.6).

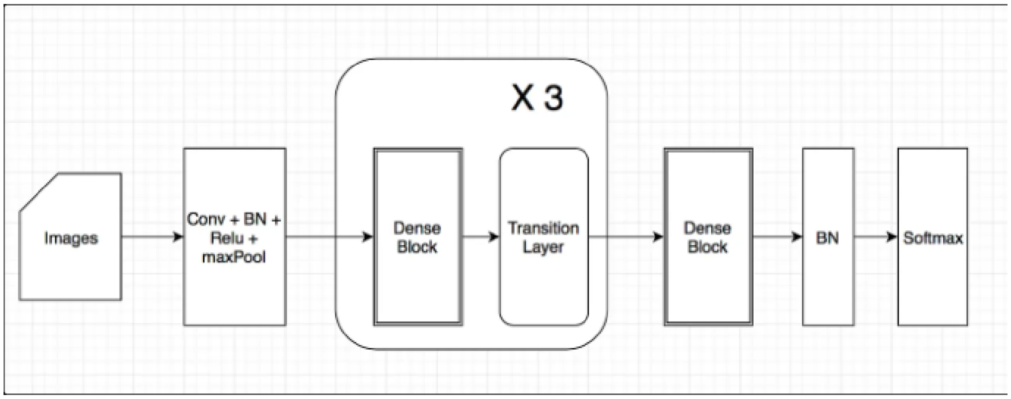


Figure 3.5: DenseNet architecture [16].

Dense block: These blocks consist of a series of densely connected layers, where each layer is directly connected to all preceding layers within the same block as illustrated in Figure 3.6. This dense connectivity pattern allows for the efficient flow of information and facilitates feature reuse throughout the network. By promoting direct connections between layers, dense blocks enable the network to capture and leverage both local and global features effectively as mentioned in [16][17][125].

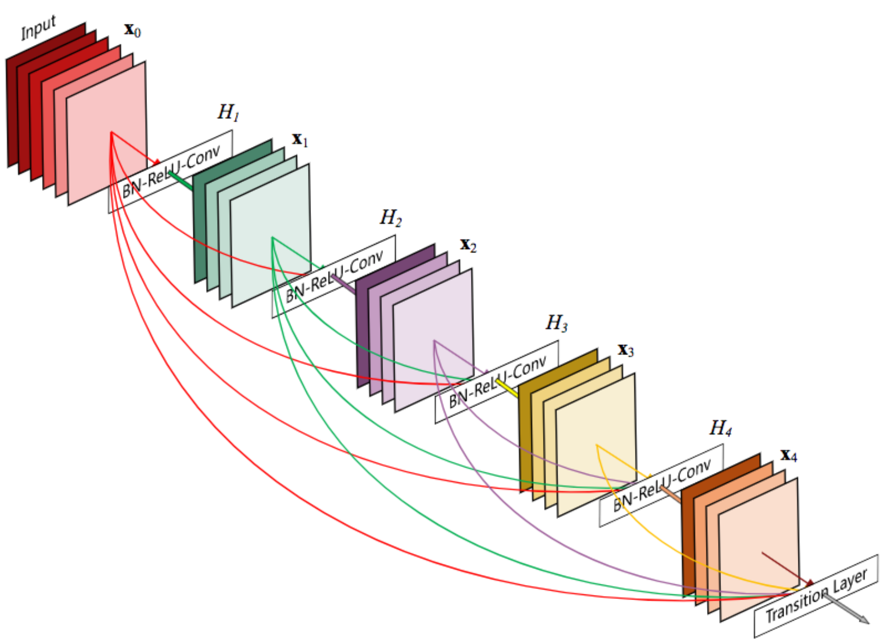


Figure 3.6: A 5-layer Dense Block with a Growth Rate of 4, featuring Full Connectivity within Layers [17].

Transition layer: The transition layers [125] in DenseNet architecture serve as connectors between adjacent dense blocks, facilitating the flow of information. These layers are responsible for reducing the spatial dimensions of the feature maps, while preserving the channel information. According to [16], Each transition layer consists of a batch normalization layer, a 1×1 convolutional layer, and a 2×2 average pooling layer . This combination enables the transition layer to decrease the width and height of the feature maps while maintaining the same number of channels. Overall, these transition layers act as bridges, ensuring smooth connectivity and feature reuse between dense blocks.

- **VGG16:** Also known as VGGNet, is a convolutional neural network (CNN) model with 16 layers. It was one of the top-performing models in the ILSVRC-2014 competition. Despite its success, the simplicity of the VGGNet16 architecture is its main appeal. The model was proposed by K. Simonyan and A. Zisserman from Oxford University and presented in a paper titled "Very Deep Convolutional Networks for Large-Scale Image Recognition" like it mentioned in [126]. The VGG16 architecture includes key features of convolutional neural networks, making it a popular choice in computer vision tasks.

- **VGG19:** According to [11][127], the model's names "19" and "16" denote the number of weight layers, or convolutional layers, in the network. The VGG19 model shares the same concept as the VGG16, with the exception that it includes 19 layers, as opposed to 16. As such, VGG19 has three additional convolutional layers compared to VGG16. The network architectures of VGG16 and VGG19 are illustrated in Figure 3.7, with VGG16 shown in the top figure and VGG19 shown in the bottom figure.

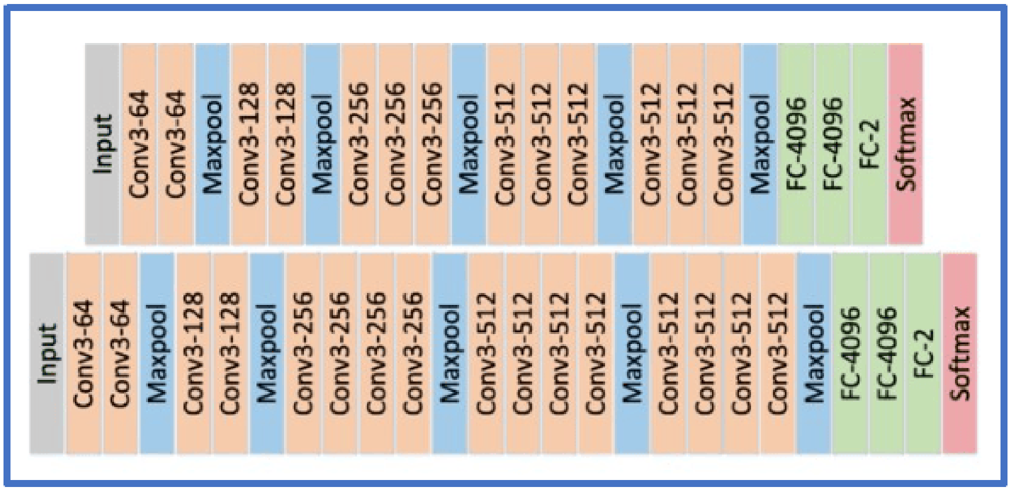


Figure 3.7: Network architectures of VGG16 and VGG19 [11].

- **InceptionV3:** Essentially, it is a 27-layer deep Convolutional Neural Network (CNN).

  "(Inception Layer) is a combination of all those layers (namely, 1×1 Convolutional layer, 3×3 Convolutional layer, 5×5 Convolutional layer)" with their output filter banks concatenated into a single output vector forming the input of the next stage [128]. For more detailed information, refer to Figure 3.8, which illustrates the architecture of InceptionV3.

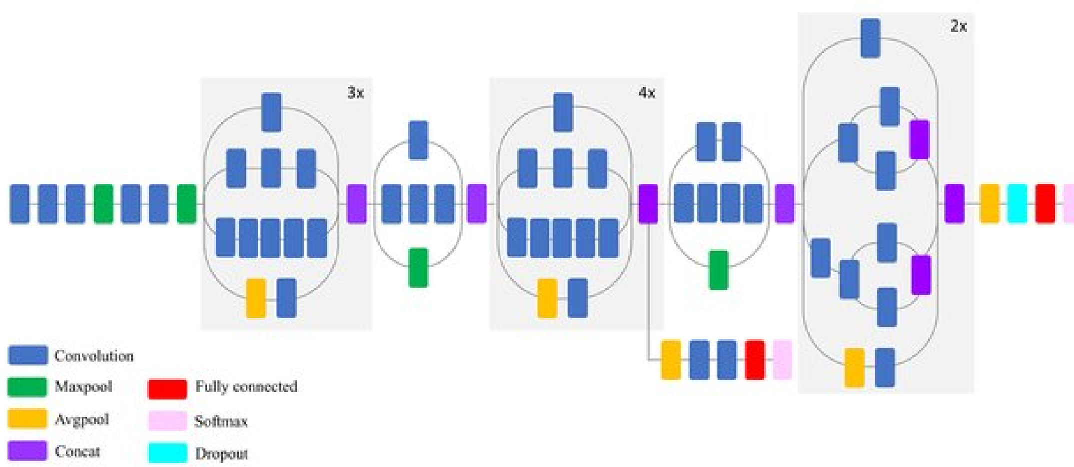


Figure 3.8: InceptionV3 architecture [18].

3.3.5 Recurrent Neural Network

This particular kind of artificial neural network makes use of time series data or sequential data.RNNs have feedback loops built in to help them recall information, but they are ineffective at learning long-term dependencies. Simply put, RNNs can only process events that occurred recently. Neural networks with long-term memory cells are needed to process distant temporal events or long-term dependencies, as shown in [129] the next figure 3.9 explains the differences between RNN and LSTM.

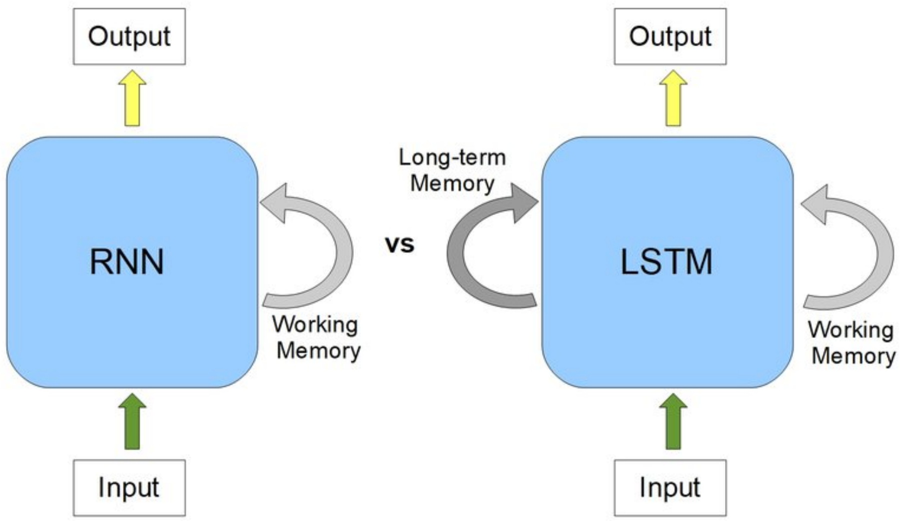


Figure 3.9: LSTM vs RNN [19].

3.3.6 Long Short-Term Memory

LSTM stands for Long Short-Term Memory and is a type of RNN that can handle and analyze sequential data, including audio, video, text, and other types, According to [129].

3.3.7 Bidirectional Long Short-Term Memory

BiLSTM is a powerful recurrent neural network used primarily in natural language processing. Unlike standard LSTMs, BiLSTMs process input in both forward and backward directions, leveraging information from both sides. This enables them to effectively model sequential dependencies between words and phrases in a given sequence. In essence, BiLSTMs include an additional LSTM layer that reverses the information flow, allowing the input sequence to be processed in a backward manner within that layer, As defined by [130].

3.3.8 Sequence to Sequence Models

Sequence To Sequence, or Seq2Seq, is a model used in applications requiring sequence prediction, in which both, the input and the output are a sequence. such as language modeling and machine translation [131].

3.3.9 Encoder-Decoder

The encoder-decoder architecture is widely recognized as a robust and extensively used machine learning framework for accomplishing sequence-to-sequence tasks like machine translation, text summarization, and question answering. Its effectiveness and prevalence make it a formidable choice in the field of artificial intelligence, as shown in [132].

Encoder: The internal state vectors, or context vectors, which are known as the hidden state and cell state vectors in the case of LSTM, are created by the encoder after it has received the input sequence and summarized the data. We only keep the internal states after discarding the encoder's outputs. In order to aid the decoder in making precise predictions, this context vector strives to include all relevant information for all input items [20], As shown in Figure 3.10.

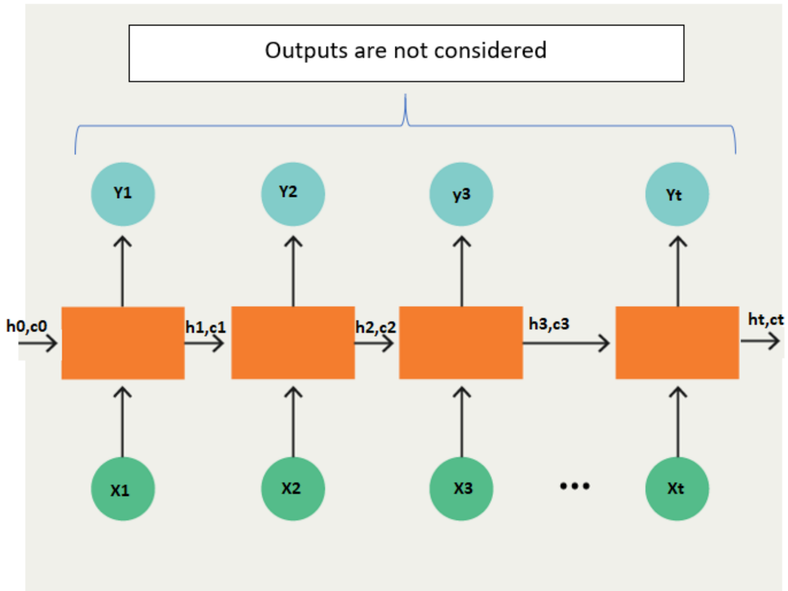


Figure 3.10: Internal architecture of encoder [20].

Decoder: An LSTM is used in the decoder, and its beginning states are set to match those of the encoder's LSTM's end states. Specifically, the first cell of the decoder network receives the context vector from the encoder's final cell as an input. The decoder begins creating the output sequence using these starting states, and these outputs are also taken into account for subsequent outputs [20], As shown in Figure 3.11.

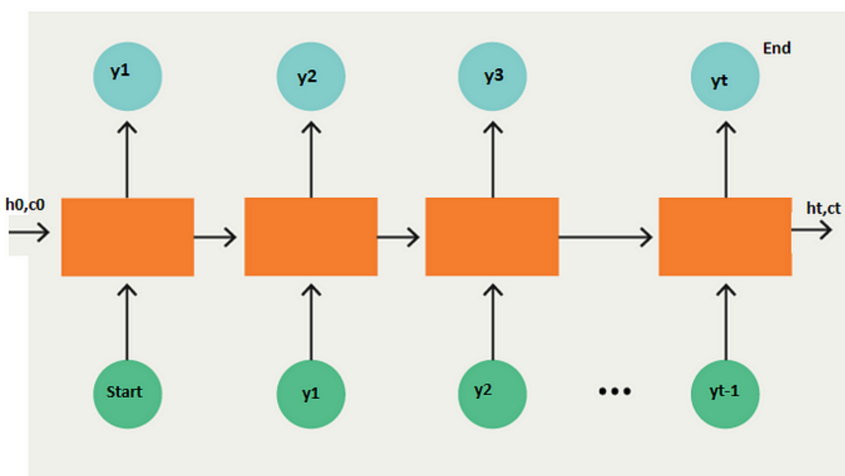


Figure 3.11: Internal architecture of decoder [20].

3.3.10 Transfer learning

TL is a powerful technique in deep learning that involves leveraging knowledge from a related task to enhance learning in a new task [133]. Acoording to [13][134][135], It is commonly used by utilizing pre-trained neural network models, which are initially trained on different tasks like classifying natural images. The weights of these pre-trained models can be used as a starting point for training on a new task, providing a significant advantage in terms of time and computational resources. In the study conducted by [105], the authors highlight the significance of transfer learning is favored in the field due to its ability to overcome the challenges of building and training new networks from scratch. By utilizing pre-trained model weights, the process of training the new model becomes simpler and faster.

3.4 Conclusion

In conclusion, this chapter has provided key definitions of algorithms in machine and deep learning. The next chapter will delve into the detailed explanation of our proposed architecture for the healthcare platform. We will highlight our specific contributions to addressing various problematic areas within this work.

“The Internet of Things has the potential to change the world, just as the Internet did. Maybe even more so.”

Kevin Ashton

Chapter 4

Contribution and system design

4.1 Introduction

In this chapter, our focus has shifted towards the development of models for each addressed problem: the classification of retina damage, the prediction of heart diseases, and emotion recognition using vital signals. For each problem, we have presented our proposed models, which have included the architectural design and underlying algorithms. Moreover, we have discussed the datasets used in our research, emphasizing their relevance and unique characteristics. The preprocessing steps have been thoroughly explained, as they are vital in preparing the data for model training and evaluation. Additionally, we have delved into the feature extraction and selection process, highlighting the specific features chosen to capture relevant information from the data. By combining these steps, we have ensured that our data is suitable for feeding into the proposed models, ultimately leading to accurate predictions.

4.2 Analysis of the requirements

By giving importance to promoting health, preventing illnesses, and intervening early, individuals can experience notable improvements in their overall health and well-being as mentioned in [136]. CVDs, which claim an estimated 17.9 million lives annually, are the leading cause of mortality worldwide, according to the WHO [25]. For the past 20 years, heart disease has continued to be the top cause of mortality worldwide. However, it is currently responsible for more deaths than ever. Since 2000, there have been more than 2 million extra heart disease deaths, bringing the total to about 9 million in 2019. Presently, 16% of all fatalities are caused by heart disease as mentioned in [137]. This highlights the pressing need for an effective system to predict and manage heart conditions. Additionally, the prevalence of diabetes has significantly risen, with the number of affected individuals reaching 422 million in 2014 from 108 million in 1980 as shown in [26]. In [27] authors explain that diabetes poses risks such as DME and blindness, emphasizing the importance of early detection and prompt treatment to prevent vision loss. AMD contributes to approximately 8.7% of global blindness cases, with a 2% estimated prevalence of drusen, a characteristic sign of AMD, in the general population as shown in [28]. Consequently, according to [138] there is a clear need for improved methods in detecting and managing these conditions, specifically DME, AMD, and drusen. OCT has emerged as a widely used diagnostic tool in ophthalmology for these conditions. However, the increasing use of OCT screening has led to a higher workload for ophthalmologists in interpreting the images, highlighting the necessity for an automated diagnostic screening system to alleviate their burden and enhance efficiency. In the field of emotion recognition, as demonstrated by [139] current approaches relying on behavioral cues face limitations due to inherent uncertainty. Therefore, it is crucial to explore alternative perspectives in emotion recognition beyond behavior-based methods. Remarkably, recent findings [140], Indicate that the ECG signal, primarily utilized for heart disease detection, can also provide valuable insights into emotional states.

We have developed a comprehensive healthcare platform that offers three essential services. The first service focuses on detecting retinal damage by analyzing OCT images. To achieve higher classification accuracy, we employ a deep learning model trained on a dataset specifically curated for this purpose. The second service revolves around emotion recognition, leveraging ECG signals to identify and classify various emotional states. We extract relevant features from the ECG signals to enhance the accuracy of our machine-learning model. Lastly, the platform utilizes deep learning models to accurately classify individuals as healthy or affected by heart disease based on ECG signal analysis. By leveraging IoMT technology, we enable continuous monitoring of heart rates and ECG signals for real-time tracking of patients' cardiovascular well-being. This integration of deep and machine learning technologies revolutionizes the diagnosis and monitoring of heart diseases, providing precise and efficient healthcare services. In critical situations, our platform promptly notifies healthcare professionals of potential risks through an integrated notification system, facilitating immediate intervention and necessary medical assistance. This approach enhances patient safety, improves outcomes, and saves lives by leveraging real-time data analysis and communication capabilities.

4.3 Proposed architecture

In the proposed E-health care system, several models are utilized to improve diagnosis speed and save time (see Figure 4.1). Each model focuses on a specific aspect of health monitoring and diagnosis. Here is a description of each model:

1. **Emotion Recognition**: In this model, machine learning techniques, specifically the Random Forest algorithm, are employed to detect and classify the emotional state of the patient as positive or negative. The IoT device (ECG sensor) captures the relevant information, which is then processed by the Random Forest model. Based on the extracted features, the model classifies the emotional state and triggers a notification to the patient's relatives in case of a negative emotional state.

2. **Heart Diseases**: DNNs integrated with the IoMT are utilized to assess the risk of developing heart disease and monitor patients efficiently. By leveraging IoMT devices, such as heart rate sensors or ECG monitors. Furthermore, a seq2seq model is employed to predict future ECG patterns based on the patient's historical data, enabling proactive monitoring. Additionally, a classification model using DNNs swiftly identifies abnormal ECG signals in real-time, When an abnormality is detected, the application promptly notifies the doctor, allowing for swift intervention to ensure the patient's well-being.

3. **Retinal Damage**: The DenseNet architecture is employed for the classification of retinal damage in this model. Doctors use an application to capture or insert retinal images obtained from an OCT scanner. The images are processed by the DenseNet model, which extracts features and classifies the presence and severity of retinal damage. This assists doctors in assessing the condition and making informed decisions.

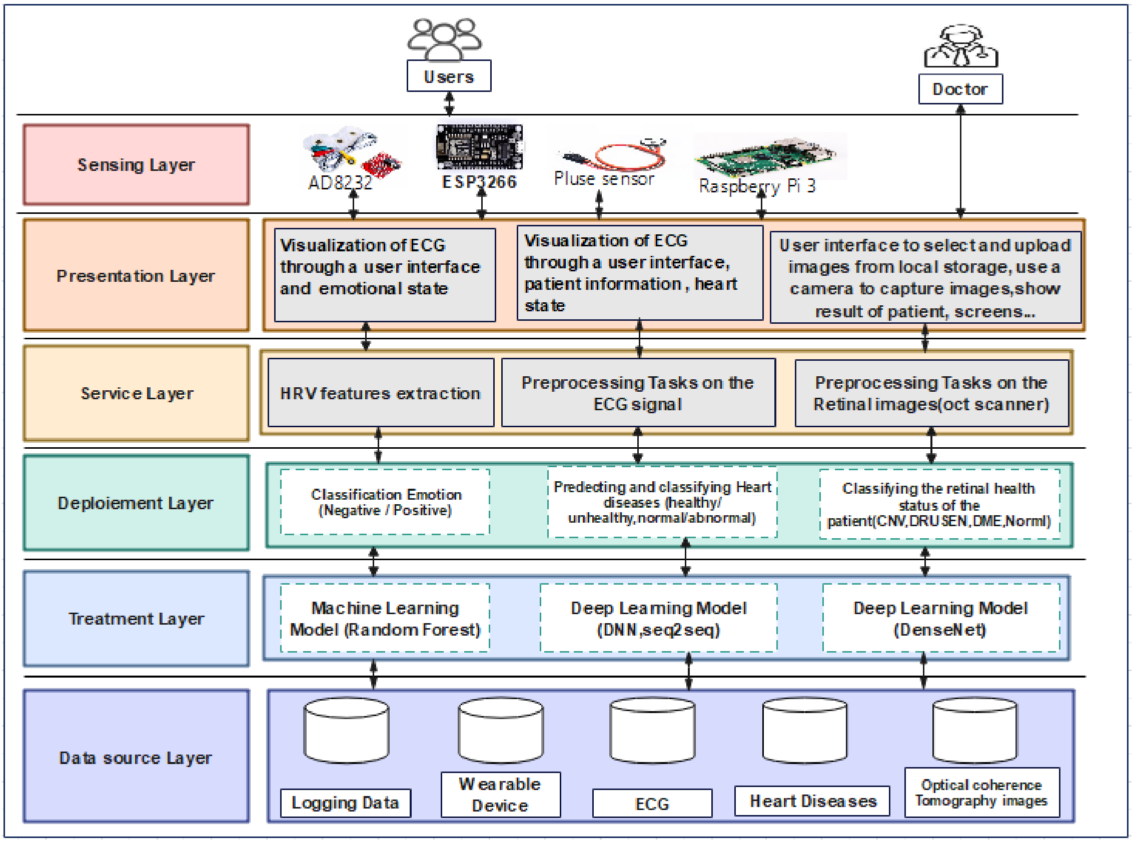


Figure 4.1: An overview of the general e-health platform Architecture.

4.4 The proposed approach for Internet of Medical Things

Here's a proposed wiring diagram for connecting the AD8232 ECG sensor to an ESP8266-based development board for ECG measurement (see figure 4.2). To measure ECG signals accurately, it is essential to affix the AD8232 ECG sensor securely to the human body. The AD8232 ECG sensor requires three electrodes for proper ECG measurement: RA (Right Arm), LA (Left Arm), and RL (Right Leg). We should place the RA electrode near the right collarbone or the right side of the chest, the LA electrode near the left collarbone or the left side of the chest, and the RL electrode on the lower right side of the abdomen or lower right torso. We can use adhesive pads or medical tape to secure the electrodes in place, ensuring good contact with the skin for accurate readings. We should connect the lead wires from the AD8232 ECG sensor to the corresponding electrodes on the human body, allowing the transmission of ECG signals to the sensor for processing.

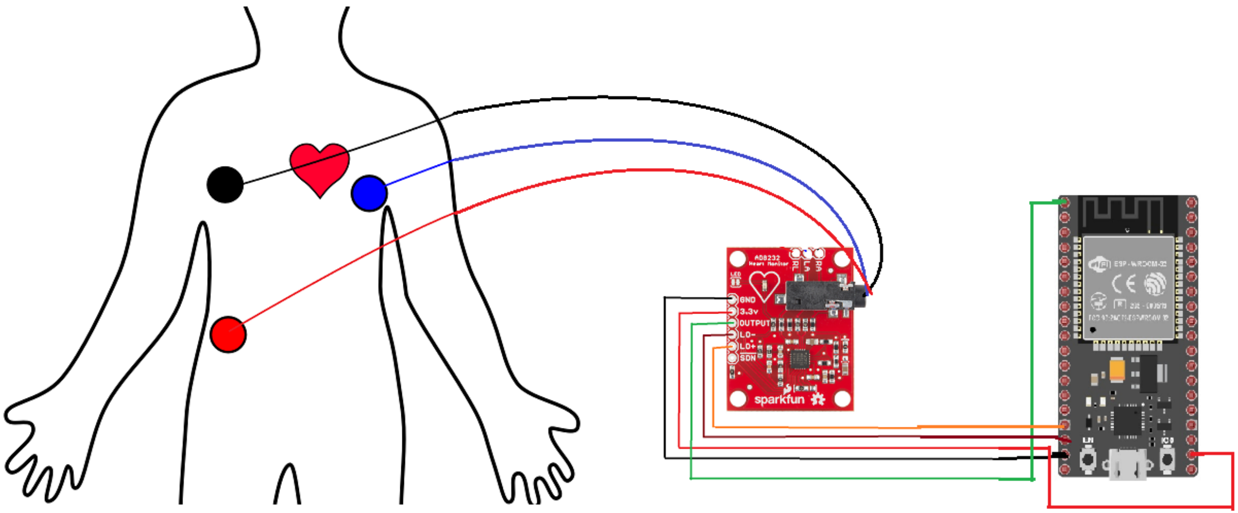


Figure 4.2: The proposed Wiring diagram of ECG measurement device.

The proposed approach for IoMT for monitoring heartbeat involves utilizing a Raspberry Pi 3B+ along with a pulse sensor to monitor and analyze heart rate data. The Raspberry Pi serves as the central device for collecting data from the pulse sensor. The pulse sensor measures the heartbeat and sends the information to the Raspberry Pi for processing (see figure 4.3).

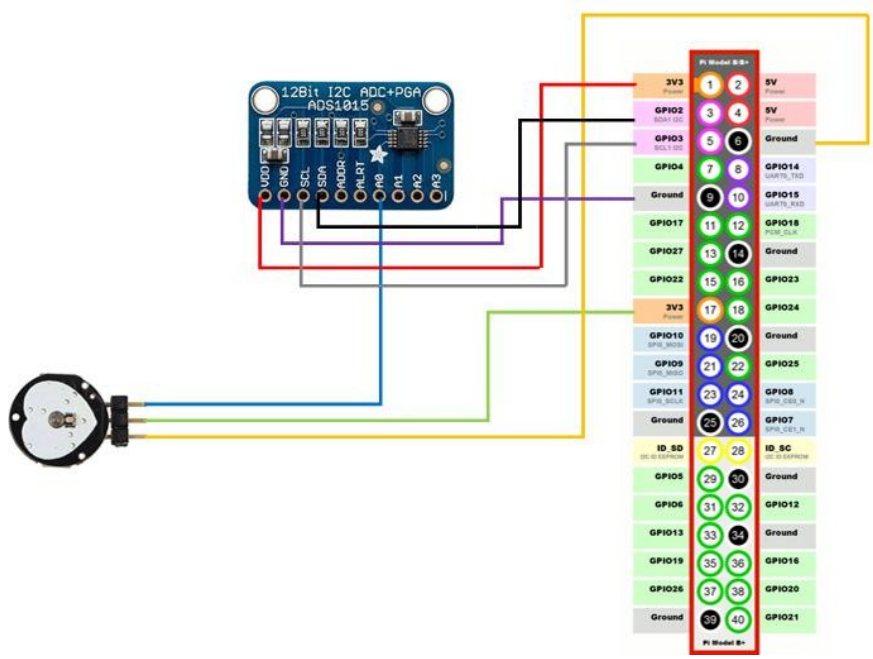


Figure 4.3: The wiring diagram of heart rate measurement device [4].

To establish connectivity and enable real-time data sharing, the Raspberry Pi is integrated

Table 4.1: Average Heart Rates by Age Group and Gender [4].

Age Group	Heart Rate (BPM)	
	Male	Female
4-5	87-104	84-100
6-8	79-94	76-92
9-11	76-91	70-86
12-15	70-87	70-87
16-19	69-85	66-83
20-39	66-82	61-78
40-59	64-79	61-77
60-79	64-78	60-75

with Firebase. The heart rate data collected by the Raspberry Pi is sent to Firebase, where it is stored in a database. A mobile app connects to Firebase to display real-time heart rate data, and to ensure a comprehensive understanding of the heart rate data. The app compares the data to the average ranges from the table to determine if the heart rate is normal or abnormal shown in Table 4.1. The average heart rate ranges provided in the table serve as a reference for comparison. The app compares the current heart rate with the appropriate range based on the user's age group and gender. If the heart rate is outside the normal range, the app alerts the user about the abnormality. This approach provides a convenient and efficient system for monitoring and managing heart rate data.

Our algorithms are based on resting heart rate data from the Centers for Disease Control and Prevention in the United States since the heart rate varies from person to person depending on characteristics including age, sex, and activity.

This approach enables real-time heart rate monitoring and analysis, providing users with immediate insights into their heart health. By combining IoT technology, a Raspberry Pi, a pulse sensor, Firebase, and a mobile app, the proposed solution offers a practical and user-friendly system for monitoring and managing heart rate data.

In the following sections, we present each system used in our platform.

4.5 The proposed approach for prediction of heart disease process

Our goal is to protect individuals from the risks of heart disease and to reduce the number of deaths associated with it. In this part, we will explain our systematic approach to achieving this noble goal. Our system will undertake a series of meticulously planned phases. Initially, a distinction will be made between individuals who display normal heart health, those who are at high risk of developing heart disease, or those who are already diagnosed as patients. Then, it will make accurate predictions based on the ECG readings. Finally, it will accurately classify each ECG as normal or abnormal as shown in the figure 4.4. During these phases, we will meticulously implement a series of steps, which include the pre-processing phase, the training phase, and finally, the prediction phase.

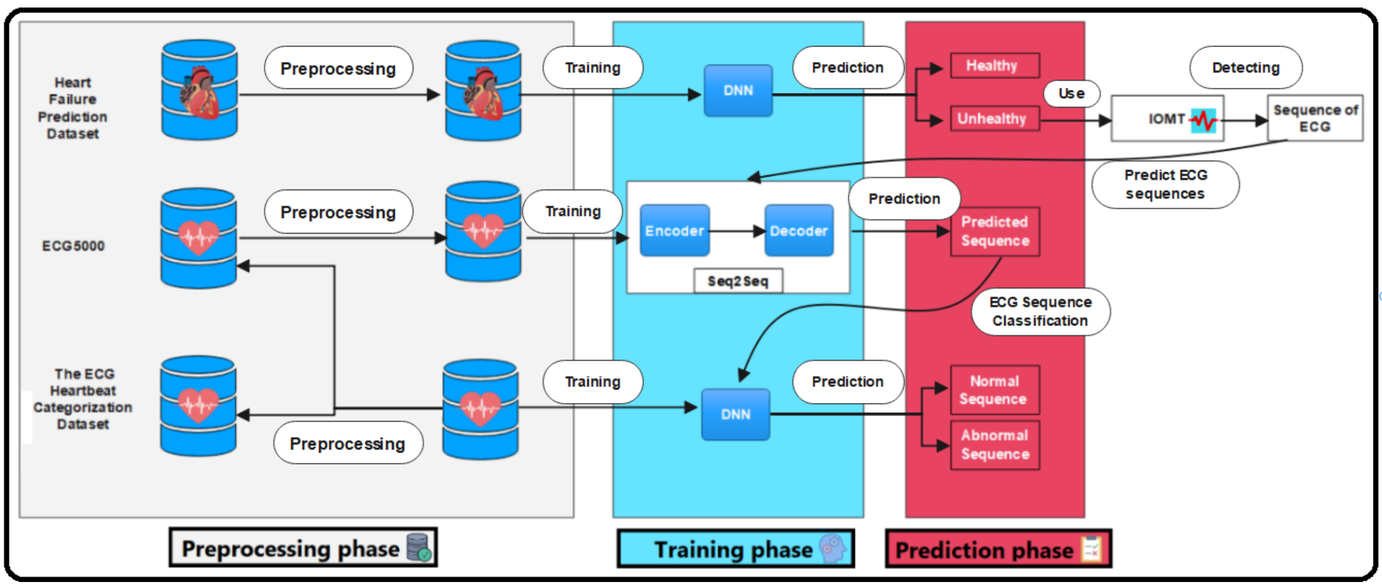


Figure 4.4: The proposed process approach for prediction of Heart Disease.

4.5.1 Data set description

To accomplish our objective of predicting heart disease, our study incorporates three distinct datasets.

Heart failure prediction dataset: Is a compilation of five different heart disease datasets that were previously available independently. These datasets were combined to create a larger dataset for research purposes. The dataset includes 11 common features and is considered the largest heart disease dataset available for research. The five datasets used for its curation are:

- **Cleveland**: This dataset consists of 303 observations.
- **Hungarian**: This dataset consists of 294 observations.
- **Switzerland**: This dataset consists of 123 observations.
- **Long Beach VA**: This dataset consists of 200 observations.
- **Stalog (Heart) Data Set**: This dataset consists of 270 observations.

The total number of observations in the combined dataset is 1190. However, there were 272 duplicated observations, resulting in a final dataset of 918 unique observations. The creators of the dataset are:

- Hungarian Institute of Cardiology, Budapest: Andras Janosi, M.D.
- University Hospital, Zurich, Switzerland: William Steinbrunn, M.D.
- University Hospital, Basel, Switzerland: Matthias Pfisterer, M.D.
- V.A. Medical Center, Long Beach and Cleveland Clinic Foundation: Robert Detrano, M.D., Ph.D.

This dataset serves as a valuable resource for heart disease research and analysis [141].

Attribute information:

- **Age**: The age of the individual [years].
- **Sex**: The gender of the individual.
- **Chest Pain Type**: Describes the type of chest pain experienced by the individual [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic].
- **Resting Blood Pressure**: The resting blood pressure of the individual (in mm Hg).

- **Cholesterol**: The cholesterol levels of the individual (in mg/dL).
- **Fasting Blood Sugar**: Indicates whether the fasting blood sugar level is above or below 120 mg/dL [1: if FastingBS > 120 mg/dl, 0: otherwise].
- **Resting Electrocardiographic Results**: Provides information on the results of resting electrocardiography [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes' criteria].
- **Maximum Heart Rate Achieved**: The highest heart rate achieved during exercise[Numeric value between 60 and 202].
- **Exercise-Induced Angina**: Indicates whether the individual experienced angina during exercise [Y: Yes, N: No].
- **ST Depression**: The ST depression induced by exercise relative to rest.
- **Slope**: The slope of the peak exercise ST segment.
- **HeartDisease**: The presence or absence of a heart Disease [1: heart disease, 0: Normal].

ECG5000: The "ECG5000" dataset is a collection of 5000 ECG signals, each with 140 data points and one column for classification. The original dataset for "ECG5000" is a 20-hour-long ECG downloaded from Physionet. The data was pre-processed in two steps: (1) extract each heartbeat, (2) make each heartbeat equal in length using interpolation [113].

The ECG Heartbeat Categorization Dataset: Is a combination of two well-known datasets in the field of heartbeat classification: the MIT-BIH Arrhythmia Dataset and The PTB Diagnostic ECG Database. These datasets contain a large number of heartbeat signals that are commonly used for training deep neural networks and other machine-learning models [142].

4.5.2 The proposed approach for risk assessment of heart diseases

The proposed approach to heart disease risk assessment includes several stages: pretreatment, training, and prediction. Here we will explain each stage for predictions for heart disease risk assessment.

1. **Preprocessing phase**

   In general, the preprocessing phase involves cleaning and pre-processing the collected data to ensure its quality and suitability for analysis. This step includes various tasks such as normalizing or standardizing numerical features and encoding categorical variables. However, our dataset "Heart Failure Prediction Dataset" [141], is already clean, and we don't need to replace any missing values. So, in this case, the preprocessing phase would focus on tasks such as normalizing numerical features and encoding categorical variables to prepare the data.

   So we define input and output variables based on the problem at hand. In the case of predicting a heart problem, the input variables would include features such as age, gender, blood pressure, and cholesterol levels... while the output variable would be whether or not the patient had a heart problem. We apply one hot cipher to categorical columns. This process involves creating binary variables for each unique category in a categorical column. Each binary variable indicates the presence or absence of a particular class in the data point. This transformation allows the model to effectively understand and process categorical information.

   We split the dataset into training and testing sets. The training set is used to train the machine learning model, while the testing set is used to evaluate the model's performance on unseen data. The usual practice is to allocate a larger portion, such as 80% of the data, for training (80% of the training data, 20% for validation), and the remaining 20% for testing. We use z-score normalization to standardize the input variables. This step helps to bring all variables to a similar scale and prevent any single variable from dominating the training process.

   The formula for calculating a z-score is given by:

$$z = \frac{x - \mu}{\sigma} [56]$$

where:

$$-\mu \text{ is the population mean.} \tag{4.1}$$

$$-\sigma \text{ is the population standard deviation.}$$

2. **Training phase**

   During the training phase, the model learns basic patterns and relationships between traits and the presence or absence of heart disease. we create a deep neural network (DNN) model, We fit the model to the training data. This process involves iterating through the training data multiple times and adjusting the weights and biases of the model using optimization algorithms Adam. The goal is to minimize the difference between the predicted outputs and the true outputs.

3. **Prediction phase**

   We assess the performance of the model by employing multiple metrics, which include the confusion matrix, accuracy, precision, Negative Predictive Value (NPV), specificity, sensitivity, F1 score, and Receiver Operating Characteristic [143] [144].

   The confusion matrix provides valuable information regarding the quantities of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics are crucial for understanding the model's performance in binary classification tasks. True positives represent the number of correctly predicted positive instances, true negatives represent the number of correctly predicted negative instances, false positives indicate the number of negative instances incorrectly predicted as positive, and false negatives represent the number of positive instances incorrectly predicted as negative.

   Accuracy: Accuracy is a metric that measures the overall correctness of the model's predictions. It is calculated using the equation 4.2.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.2}$$

Precision: also known as the positive predictive value, measures the proportion of correctly predicted positive instances among all instances predicted as positive. It quantifies the model's ability to avoid false positives:

$$Precision = \frac{TP}{TP+FP} \tag{4.3}$$

NPV: measures the proportion of correctly predicted negative instances among all instances predicted as negative. It quantifies the model's ability to avoid false negatives:

$$NPV = \frac{TN}{TN+FN} \tag{4.4}$$

Specificity: also known as the true negative rate, measures the proportion of correctly predicted negative instances among all actual negative instances. It quantifies the model's ability to identify true negatives:

$$Specificity = \frac{TN}{TN+FP} \tag{4.5}$$

Sensitivity: also known as recall or true positive rate, measures the proportion of correctly predicted positive instances among all actual positive instances. It quantifies the model's ability to identify true positives:

$$Sensitivity = \frac{TP}{TP+FN} \tag{4.6}$$

F1 Score: The F1 score is a comprehensive metric that takes into account both precision and recall, providing a balanced measure of the model's performance. It is the harmonic mean of precision and recall, considering both false positives and false negatives:

$$F1\ Score = \frac{2\ x\ (Precision\ x\ Sensitivity)}{Precision + Sensitivity} \tag{4.7}$$

Receiver Operating Characteristic (ROC) is a graphical plot that illustrates the performance of a classification model at various classification thresholds. It displays the tradeoff between sensitivity and specificity as the threshold for classifying positive instances changes. The area under the ROC curve (AUC) is often used as a summary measure of the model's performance, with a higher AUC indicating better discrimination power.

These metrics, including precision, NPV, specificity, sensitivity, and the ROC curve, enhance our understanding of the model's performance, enabling us to evaluate its effectiveness in correctly classifying positive and negative instances.

4.5.3 The proposed approach for the prediction of Electrocardiogram

We will now describe the proposed approach for future ECG prediction using the pre-treatment phase, the training phase, and the prediction phase. We used the ECG50000 dataset, which was originally intended for classification, but we modified it for sequence prediction, and seq2seq trained on it.

1. **Preprocessing phase**

   We have dropped the classification column from the dataset, as it is unsuitable for sequence prediction. We verified that all ECG sequences are the same length. If not, we may perform data pre-processing, such as padding or truncating sequences to a fixed length. We divided the pre-processed data set into training and test sets. We allocated 20% of the data for testing and used the remaining 80% for training. We took care to maintain sequence order during splitting to maintain temporal relationships.

2. **Training phase**

   In this phase, we define the length of the input sequence and the length of the output sequence. Then we divide the training set into input sequences and target sequences. We implemented the Seq2Seq model architecture, which consists of an encoder and a decoder. The encoder processed the input sequence and encoded the information into a fixed-length context vector. The decoder takes the context vector as input and generates the expected output sequence. We compiled a Seq2Seq model with an appropriate loss function (such as a mean squared error) and an Adam optimizer. Finally, we superimposed the model on the training data and determined the number of epochs and batch size for training.

3. **Prediction phase**

   In this phase, we implemented the Seq2Seq model architecture, which consists of an encoder and a decoder. The encoder processed the input sequence and encoded the information into a fixed-length context vector. The decoder takes the context vector as input and generates the expected output sequence. We compiled a Seq2Seq model with an appropriate loss function such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and mean absolute error (MAE). These loss functions measure the discrepancy between the predicted output and the target output. MSE calculates the average squared difference, while RMSE calculates the square root of the average squared difference and an Adam optimizer. Finally, we superimposed the model on the training data and determined the number of epochs and batch size for training.

   Mean Squared Error: The MSE is calculated by taking the average of the squared differences between the predicted values and the actual values [145].

$$
\begin{aligned}
&MSE = \frac{1}{n}\sum_{i=1}^{n}(y_{\text{pred},i} - y_{\text{actual},i})^2 \\
&\text{where:} \\
&\quad - MSE \text{ represents the Mean Squared Error.} \\
&\quad - n \text{ is the total number of data points.} \\
&\quad - y_{\text{pred},i} \text{ is the predicted value for the } i\text{-th data point.} \\
&\quad - y_{\text{actual},i} \text{ is the actual value for the } i\text{-th data point.}
\end{aligned}
\tag{4.8}
$$

   Root Mean Squared Error: The RMSE is the square root of the MSE, providing a measure of the average magnitude of the errors in the original unit of the target variable [145].

$$
\begin{aligned}
&RMSE = \sqrt{MSE} \\
&\text{where:} \\
&\quad - RMSE \text{ represents the Root Mean Squared Error.} \\
&\quad - MSE \text{ is the Mean Squared Error calculated earlier.}
\end{aligned}
\tag{4.9}
$$

   Mean Absolute Error: The MAE is calculated by taking the average of the absolute differences between the predicted values and the actual values [145].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_{\text{pred},i} - y_{\text{actual},i}|$$

where:

$$\begin{aligned} &- MAE \text{ represents the Mean Absolute Error.} \\ &- n \text{ is the total number of data points.} \\ &- y_{\text{pred},i} \text{ is the predicted value for the } i\text{-th data point.} \\ &- y_{\text{actual},i} \text{ is the actual value for the } i\text{-th data point.} \end{aligned} \tag{4.10}$$

By calculating these metrics, you can assess the performance of the Seq2Seq model in terms of how well it predicts the next sequence based on the given input sequence and the quality of the generated sequences.

4.5.4 The proposed approach for Electrocardiogram classification

In this part, we will explain the proposed approach to classify the ECG pre-treatment phase, training phase, and prediction phase. The aim of this approach is to accurately classify heart rhythms as normal or abnormal, enabling effective detection and analysis of potential cardiac abnormalities.

1. **Preprocessing phase**

   In the preprocessing phase, we aim to prepare the data for training our ECG classification model. We will work with two datasets: the ECG Heartbeat Categorization Dataset and ECG5000. Our goal is to unify the labels from both datasets into two classes: normal and abnormal. Additionally, we will remove any duplicate data to maintain data integrity. Once the preprocessing is complete, we will proceed to the training phase.

2. **Training phase**

   During the training phase, we will train our ECG classification model using the preprocessed data. First, we will split the data into an 80% training set and a 20% test set, ensuring that the data is representative of different classes. Next, we create a deep neural network (DNN) model, we will fit the model using the training data, allowing it to learn patterns and relationships within the ECG signals.

3. **Prediction phase**

   In the prediction phase, we will utilize the trained model to predict the classes of new, unseen data. We will feed the ECG signals into the model and obtain predictions for each heartbeat. To evaluate the performance of our model, we will employ various metrics. These include accuracy, which measures the overall correctness of the predictions, and the confusion matrix, which provides insights into the classification results for each class. Additionally, we will calculate other important metrics such as recall score, precision score, ROC, F-Beta, and F1 score. These metrics help assess the model's performance with respect to true positives, false positives, and false negatives.

   F-Beta score : is a metric commonly used in binary classification tasks to evaluate the model's performance, taking into account both precision and recall. It is an extension of the F1 score, which balances precision and recall equally, but allows you to assign different weights to precision and recall based on the value of the beta parameter.

$$F_\beta = (1+\beta^2)\cdot\frac{\text{precision}\cdot\text{recall}}{(\beta^2\cdot\text{precision})+\text{recall}}$$

where:

$$\begin{aligned}&-F_\beta \text{ represents the F-beta score.}\\&-\beta \text{ is the weighting factor.}\\&-\text{precision is the precision value.}\\&-\text{recall is the recall value.}\end{aligned} \tag{4.11}$$

4.6 The proposed approach for the retinal Optical Coherence Tomography images classification

To develop our retinal damage classification system, we utilize OCT datasets as the input data. Our approach consists of two main steps: preprocessing and classification. A crucial aspect of our approach is the utilization of transfer learning, a technique in deep learning where a model trained on one task is applied as a starting point for a different task. In our case, we employ a pre-trained model that was initially trained on a large dataset of natural images and adapt it to detect retinal diseases.

Before passing the data to the deep learning model, it is essential to load and preprocess the dataset. The dataset is quite large, with a size of approximately 5.8 gigabytes, containing around 84,495 images. Once the preprocessing phase is completed, we split the resulting dataset and proceed to the second step, which involves training the data using a custom DenseNet architecture specifically designed for retinal disease classification.

In the following sections, we provide a detailed description of the dataset used in this work, followed by an explanation of the preprocessing pipeline. We then delve into the specifics of the architecture employed. Finally, we discuss the prediction process of our model. Figure 4.5 illustrates the organizational structure of the proposed approach.

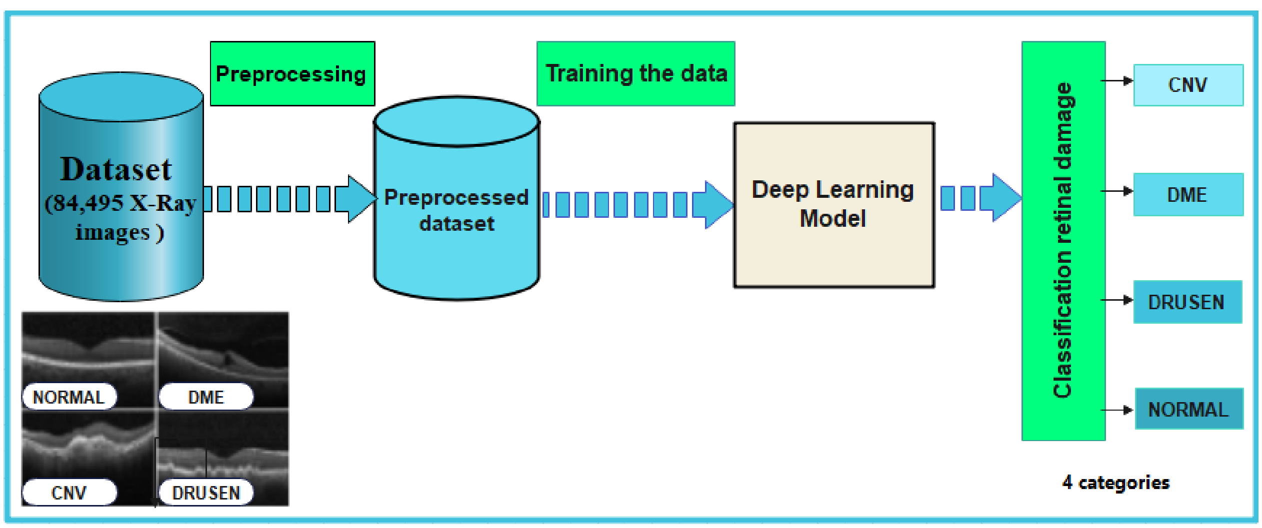
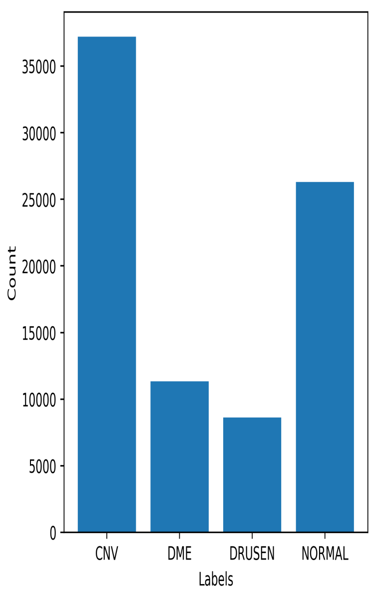


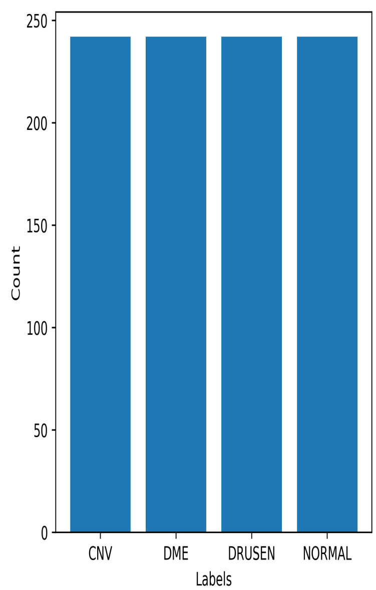
Figure 4.5: The plan of the proposed approach.

4.6.1 Data set description

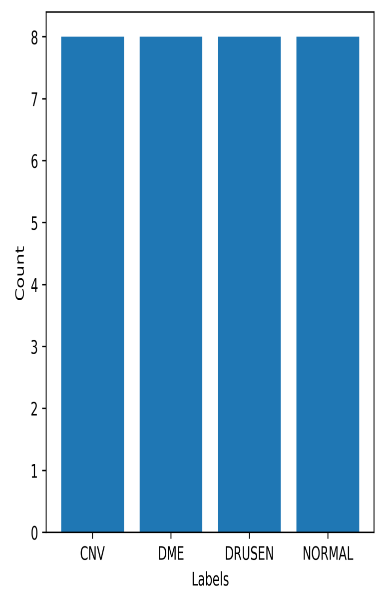
The dataset utilized in this method consists of OCT images acquired from Spectralis OCT, developed by Heidelberg Engineering in Germany. These images were collected from retrospective cohorts of adult patients across various institutions [146]. The dataset comprises a total of 84,495 X-Ray images in JPEG format, which are categorized into four distinct groups: NORMAL, CNV, DME, and DRUSEN. To ensure organized management, the dataset is structured into three main folders: train, test, and validation (see Figure 4.6).



(a) Class Distribution Visualization - Distribution of Train Data.



(b) Class Distribution Visualization - Distribution of Test Data.



(c) Class Distribution Visualization - Distribution of Validation Data.

Figure 4.6: Analysis of Data Sample Distribution in the dataset.

4.6.2 Preprocessing phase

The preprocessing phase, a crucial step in data science, is performed prior to training and can be time-consuming. It encompasses several important steps, including dataset loading, creation of image paths and labels, dataset balancing, image resizing, and conversion of data into arrays. These preprocessing steps are necessary to prepare the data for further processing and training of the proposed DenseNet model.

- **Creating a CSV file:** To create a CSV file, we will store the image paths along with their corresponding labels. This involves iterating over the filenames for each class and adding tuples containing the image path and label (0 for CNV, 1 for DME, 2 for DRUSEN, and 3 for NORMAL) to a list As illustrated in Figure 4.7.

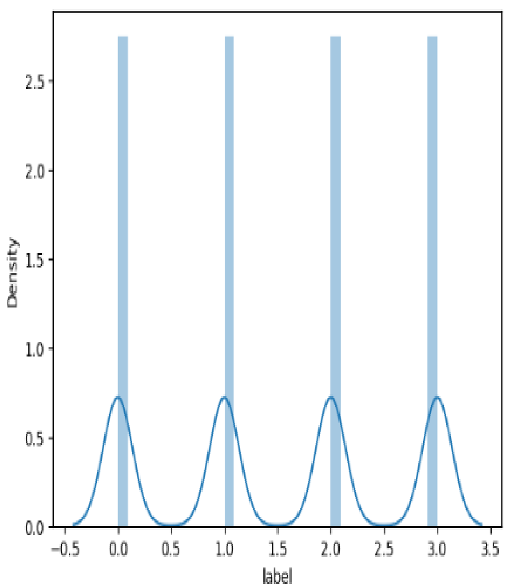
  The main objective of converting the grayscale images from the dataset into a CSV file format is to effectively manage memory requirements and facilitate efficient data handling during subsequent processing steps.

	image_path	label
0	/kaggle/input/kermany2018/oct2017/OCT2017 /tra...	0
1	/kaggle/input/kermany2018/oct2017/OCT2017 /tra...	0
2	/kaggle/input/kermany2018/oct2017/OCT2017 /tra...	0
3	/kaggle/input/kermany2018/oct2017/OCT2017 /tra...	0
4	/kaggle/input/kermany2018/oct2017/OCT2017 /tra...	0

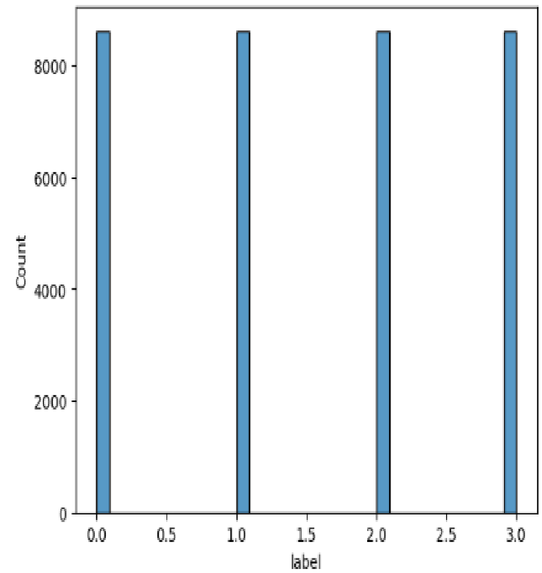
Figure 4.7: CSV File Structure: Image Paths and Labels Representation.

- **Dataset balancing:** Is a crucial step performed to address class imbalance, which occurs when the number of samples in different classes is significantly uneven like it shown in Figure 4.6. To ensure that each class is represented by an equal number of images. It begins by determining the minimum count of images per label in the dataset. Then, the data is grouped based on the 'label' column, allowing for separate handling of each class (refer to Figure 4.8).

  The purpose of dataset balancing is to enable the model to learn equally from all classes, thereby preventing bias towards the majority class. By providing an equal representation of each class, the model becomes more capable of generalizing and achieving higher accuracy when predicting all classes. Balancing the data plays a vital role in improving the overall performance and fairness of the model in classification tasks.



(a) Class Distribution Visualization - Density of Train Data.



(b) Class Distribution Visualization - Balanced Distribution of Train Data.

Figure 4.8: Analysis of Data Sample Distribution in the Train Set.

- **Resize Images:** Resizing the images is a common preprocessing step in deep learning to ensure uniformity, improve computational efficiency, and enhance the model's ability to extract meaningful features and generalize to new data.

  During this phase, We will resize every image in the balanced Data to a standardized size of 64x64 pixels.

  Using large-sized images directly in a CNN architecture requires a correspondingly large input shape size. This results in an increased number of parameters in the CNN model, which in turn leads to longer training times.

4.6.3 Data splitting

In the realm of deep learning, the commonly employed approach for dividing a dataset is to create separate sets for training, validation, and testing. In our study, we adopted a similar strategy where our dataset was partitioned into three subsets. The training set consisted of 24,124 images per class, the validation set contained 10,340 images, and the remaining images were allocated to the test set.

4.6.4 Training phase of our deep learning model

Once the preprocessing and generalization of the training data have been completed, the next crucial step is to train our deep learning architecture. In this phase, we will delve into a comprehensive explanation of the proposed architecture.

For the purpose of retinal damage classification, we have introduced a modified version of the DenseNet model in our proposed architecture. Our architecture comprises multiple dense blocks, transition blocks, and pooling layers. These components work together to enhance feature extraction and information flow within the network.

The model architecture begins with an initial convolutional layer, which is then followed by a dense block, as depicted by the blue cards in the Figure 4.9. Afterwards, a transition block, represented by the yellow cards, is inserted, and this alternating pattern of dense and transition blocks continues. To reduce the spatial dimensions of the feature maps to a vector, a global average pooling layer is employed. Finally, a fully connected layer with softmax activation is added for multi-class classification.

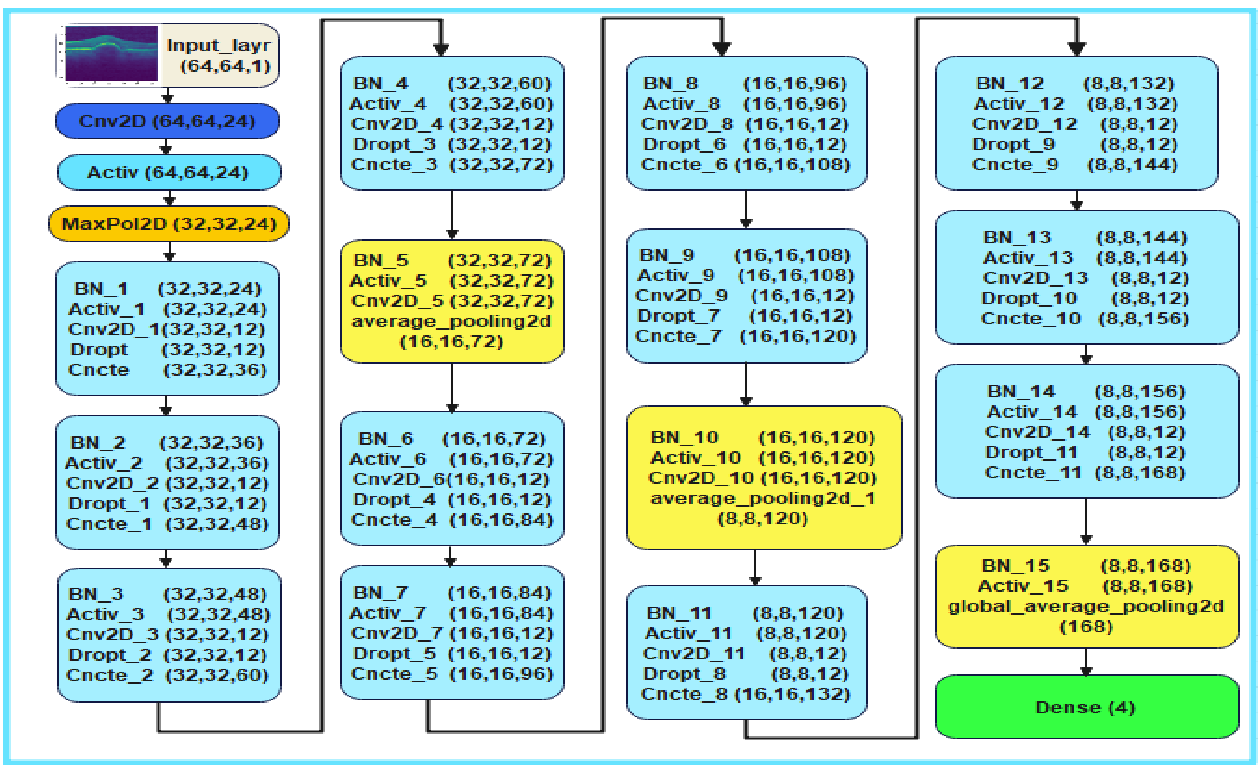


Figure 4.9: The Proposed Architecture for Retinal Damage Classification.

The dense blocks are responsible for feature extraction and consist of multiple convolutional layers with a fixed growth rate. Each layer within a dense block receives inputs from all preceding layers, encouraging feature reuse and information flow. Batch normalization and ReLU activation are applied after each convolutional layer to improve training stability and introduce non-linearity.

The transition blocks are used to reduce the dimensionality of feature maps and control the number of parameters in the network. They include a convolutional layer followed by average pooling, which reduces the spatial dimensions while preserving the number of channels as illustrated in the Figure 4.9.

4.6.5 Prediction phase

Moving on to the prediction step, we load the model and the stored weights from the training stage. Then, we perform a prediction operation to evaluate whether the model has been trained effectively and is capable of performing the task accurately.

To evaluate the performance of the model, we employ different evaluation metrics, including TP, TN, FP, and FN. These metrics are utilized in calculating various performance measures, such as accuracy, precision, and F1 score. The equations for accuracy, precision, and F1 score were derived in the previous section as shown in equations 4.2, 4.3, and 4.7.

4.7 The proposed approach for emotion recognition using vital signals

This work specifically emphasizes the utilization of machine learning techniques in the development of emotion recognition systems. The process of developing such systems involves several distinct steps including data collecting, preprocessing, feature extraction, feature selection, classification, and validation.

The ECG analysis process in machine learning-based approaches heavily relies on a well-annotated dataset. However, creating a high-quality annotated ECG dataset remains a significant challenge. It should be noted that models trained on small datasets struggle to generalize the collected data to validation and test sets, leading to imprecise study results. Consequently, such models often encounter the issue of overfitting, where in they fit well to the training data but fail to generalize effectively to new and unseen data. In this study, three different datasets, namely DREAMER, SWELL, and WESAD, were utilized for emotion recognition.

To enhance the diversity of samples used for training the model, a data augmentation process is applied to the DREAMER dataset. This is particularly important because DREAMER is a multi-modal database specifically designed for analyzing emotions. By augmenting the dataset, we can increase the variety of samples, which ultimately improves the model's ability to learn and yield more accurate results.

Since these datasets have varying label representations, a label transformation process was performed to unify them under a negative/positive model.

In the preprocessing step, it is highlighted that ECG signals exhibit higher voltage amplitudes and are less prone to interferences. This characteristic implies that additional processing may not be necessary for these signals, as mentioned in the study of the DREAMER dataset [21].

The next step in the emotion recognition process involves the extraction of relevant features. Specifically, in this context, HRV features are computed. HRV features provide valuable insights into the variations in the timing between consecutive heartbeats, which can be indicative of an individual's emotional state. By extracting HRV features, we capture essential information from the ECG signals that can be utilized in subsequent stages of emotion recognition.

After features extraction, the next step is feature selection. This involves selecting commonly used HRV features from different domains such as time, frequency, and nonlinear domains. These features provide valuable insights into the characteristics of heart rate variability.

After the feature selection step, the next stage in the emotion recognition model development is applying classifiers to train the model. One commonly used classifier is the Random Forest classifier, which has shown promising results in various applications, including emotion recognition. The performance of the Random Forest classifier in emotion recognition can be observed in several studies, such as those conducted on [104], and [147]. The architecture of emotion recognition models is presented in Figure 4.10.

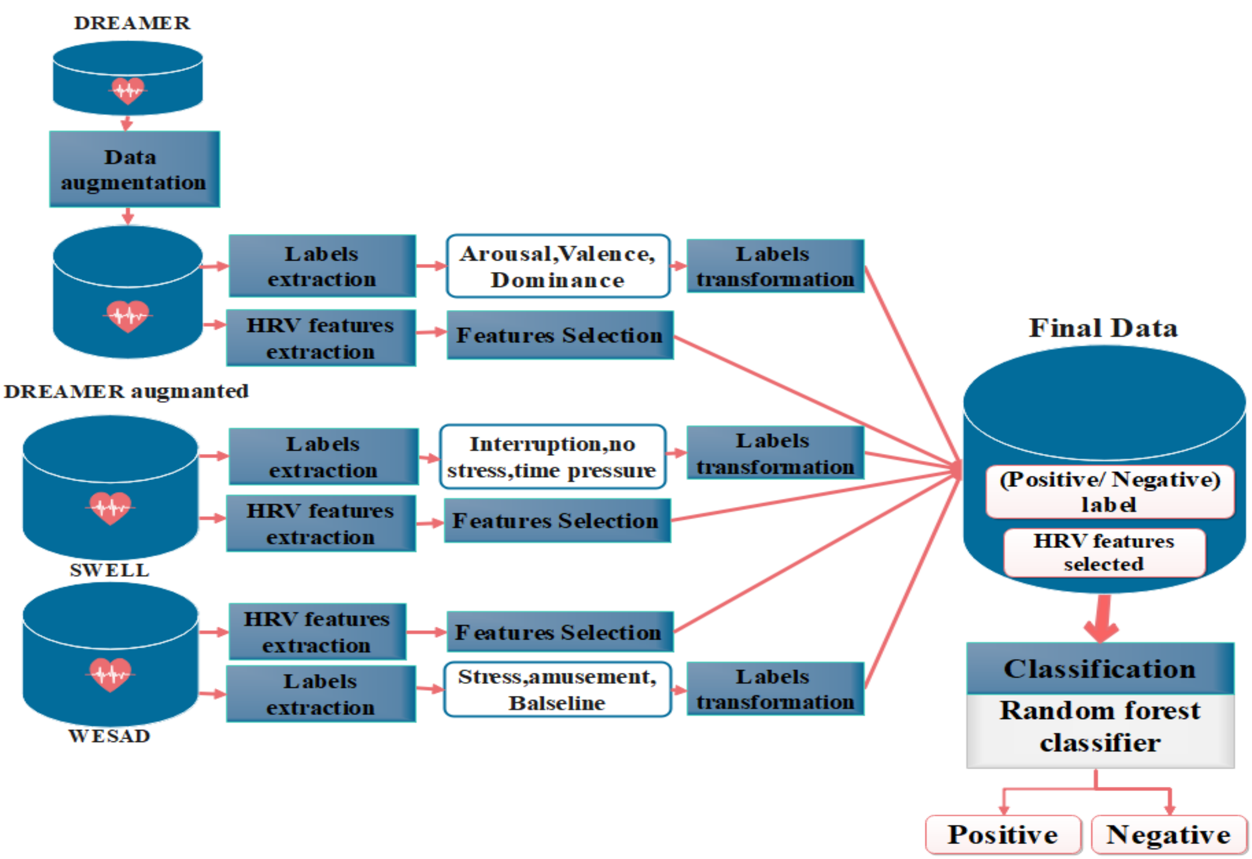


Figure 4.10: Architecture of an ECG-based emotion recognition system.

In the next subsections, we present a detailed description of each step.

4.7.1 Data set description

This study incorporates three distinct datasets, namely DREAMER, SWELL, and WESAD, to investigate emotion recognition.

– **DREAMER** [21]: The dataset consists of 18 video clips watched by 23 participants. The focus of the dataset is on collecting EEG and ECG data. For ECG measurements, a wireless, portable device from Shimmer was used, with a sampling rate of 256 Hz. The dataset also includes self-annotations of valence, arousal, and dominance (ADM) provided by the participants.

– **SWELL [103]:** Is a multimodal dataset focused on stress research. It involves 25 participants who were involved in stress-inducing tasks. The dataset includes physiological signals, specifically ECG and SC (skin conductance) signals. The ECG data was captured using a Mobi device from TMSi, with a sampling rate of 2048 Hz. The participants' self-assessment of stress levels was based on two models: ADM and categorization into three stress levels: time pressure, interruption, and no stress.

– **WESAD [104]:** It comprises data collected from 15 subjects who were involved in a variety of tasks. The dataset includes multiple biosignals such as ECG, BVP, EDA, EMG, RSP, and temperature. ECG signals were recorded using a RespiBAN Professional device at a sampling rate of 700 Hz. The subjects self-annotated their emotions, categorizing them as amusement, neutral, or stress.

In the upcoming subsection, we will introduce and explain the process applied to the DREAMER dataset to increase the number of samples.

4.7.2 Augmantation of the DREAMER dataset

The dataset utilized in this study consists of 414 ECG samples and is categorized as a small ECG dataset. It specifically belongs to the ECG DREAMER dataset, which serves as a multimodal database for analyzing emotions. However, due to the limited size of the dataset, the precision of the study results may be affected. To mitigate this issue, data augmentation techniques are employed to combat overfitting in models. Additionally, it is important to note that this study lacks diversity in the type of data, as it focuses on two specific datasets: one comprising emotions such as no stress, interruption, and time pressure in SWELL, and the other consisting of stress, baseline, and amusement in WESAD.

In this study, the baseline records from the DREAMER dataset were chosen as positive emotion references, while the stimuli records were also extracted for analysis. To increase the dataset size from the initial 414 samples, each of the 414 ECG signals was divided into 1-minute segments. Consequently, the dataset expanded to include a total of 1311 samples.

In the next subsection, we will introduce an explanation of the label transformation process that will be employed to unify the labels of the datasets into a binary model, comprising negative and positive categories.

4.7.3 Labels transformation

When utilizing multiple datasets for emotion recognition, such as DREAMER, SWELL, and WESAD, it is common to encounter different labeling schemes. In order to unify the labels and facilitate the training of the emotion recognition model, a negative/positive model can be employed.

In this approach, the original labels are transformed into a binary format, where one class represents the negative emotional state and the other class represents the positive emotional state. This transformation allows for consistency across datasets and simplifies the classification task.

By adopting a negative/positive model, the emotion recognition model can effectively capture and differentiate between negative and positive emotional states, regardless of the original labels used in each dataset. This approach provides a standardized framework for training and evaluating the model's performance across multiple datasets.

The label adjustments in the datasets were made as follows:

– **DREAMER:** In the DREAMER dataset, participants rated their emotional responses in terms of valence, arousal, and dominance using a scale ranging from 1 to 5 (see figure 4.11). Within this database, a positive/negative model was employed, where a valence value of less than 3 indicated a negative emotional state, while any other value indicated a positive emotional state. The determination of the emotional state was based on the valence value provided by the participants.

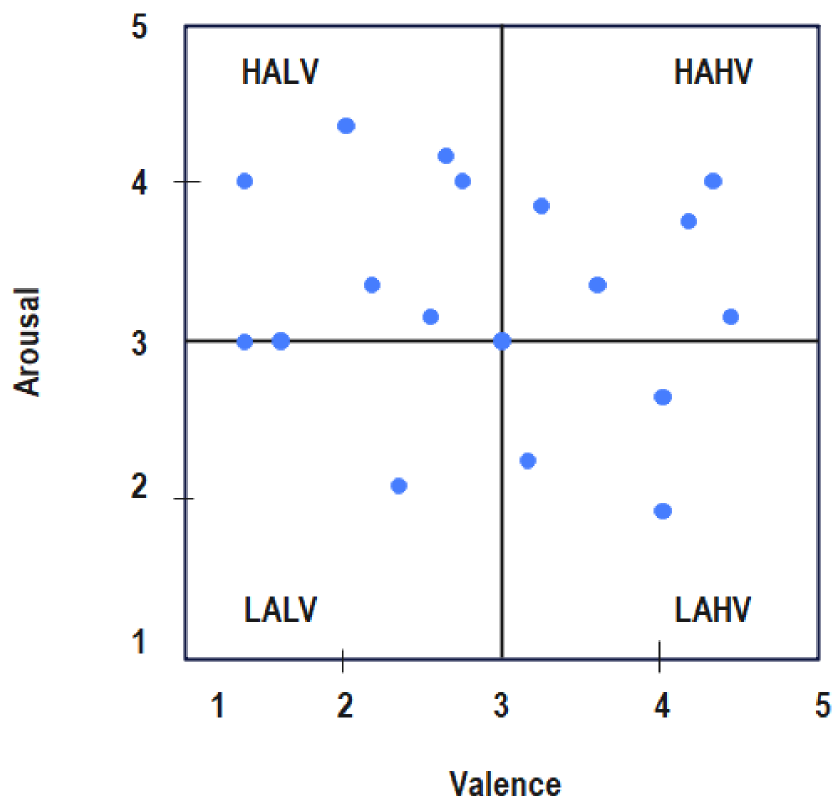


Figure 4.11: Mean location of film clips in two-dimensional affective space as rated by the participants in DREAMER [21].

– **SWELL:** The condition 'no stress' was labeled as positive, while conditions 'interruption' or 'time pressure' were labeled as negative.

– **WESAD:** The condition 'stress' was labeled as negative, while conditions 'baseline' or 'amusement' were labeled as positive.

In the next subsection, we will elucidate the subsequent step in the emotion recognition process, which entails the extraction of pertinent Heart Rate Variability (HRV) features.

4.7.4 Features extraction

ECG signals, specifically HRV signals, have been widely recognized in the literature as playing a critical role in the field of emotion assessment research. In fact, HRV is determined by the time intervals between consecutive heartbeats, commonly referred to as RR intervals as shown in figure 4.12. It has been widely acknowledged in the literature that HRV plays a crucial role in research on emotion assessment, as it provides valuable insights into the physiological changes associated with emotional states, as noted in the study by Malik et al. [148]. In the upcoming subsection, we will provide a detailed explanation of the selected HRV features.

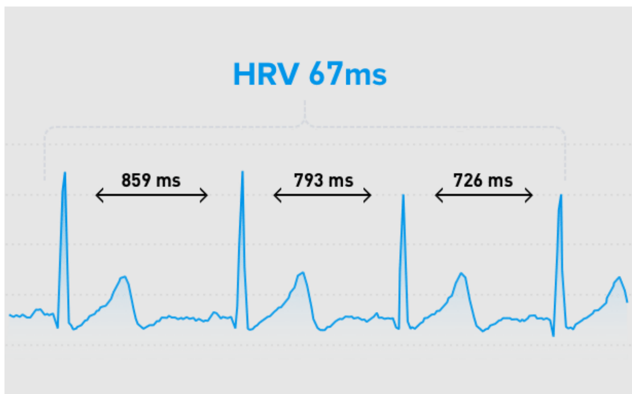


Figure 4.12: Heart Rate Variability[22].

4.7.5 Features selection

Feature selection has been widely acknowledged as a valuable technique for enhancing the performance of emotion recognition systems. By selecting relevant features, both the training and testing accuracy of the system can be improved. Additionally, the computational time required for classification is significantly reduced. These advantages contribute to the overall effectiveness and efficiency of emotion recognition systems as mentioned in the study [89].

In the field of HRV feature analysis, there are three main domains: the time domain, frequency domain, and time-frequency domain. Each domain offers a unique perspective on HRV analysis. Here is a detailed explanation of each domain based on the study in [89]:

- **Time domain:** This domain focuses on measuring the level of variability in the IBI. The measurements are typically expressed as the natural logarithm of the original units or in the original units themselves, which helps to achieve a more normally distributed representation.

  In our study, we employed temporal domain features including MEAN_RR, MEDIAN_RR, SDSD, SDRR, RMSSD, and pNN50. The equations 4.12, 4.13 , 4.14, and 4.15 calculates respectively MEAN RR, SDRR, RMSSD, and pNN50. The description of MEDIAN_RR, and SDSD are presented in the table 4.4. These features capture different aspects of HRV and provide valuable insights into the temporal characteristics of the data.

Table 4.2: Description of the Time domain features selected.

HRV index	**Short description**	**Equation/Reference**
MEAN RR	Mean value over all RR intervals of a measurement in [ms]. RRi is the ith RR interval and N is the number of all RR intervals of the measurement.	$$\bar{RR} = \frac{1}{N}\sum_{i=1}^{N} RR_{i+1} \quad (4.12)$$ see.[149]
MEDIAN RR	Median of all RR intervals	see.[147]
SDRR	Standard deviation of all interval	$$\sqrt{\frac{\sum_{i=1}^{N}(RR_i - \bar{RR})^2}{N-1}} \quad (4.13)$$ see.[149]

RMSSD	Square root of the mean of the sum of the squares of the difference between adjacent RR intervals	$$\sqrt{\frac{\sum_{i=1}^{N-1}(RR_{i+1}+RR_i)^2}{N-1}} \tag{4.14}$$ see.[147]
SDSD	Standard deviation of all interval of differences between adjacent RR intervals	see.[147]
PNN50	percentage of adjacent RR intervals differing by more than 50 ms	$$\frac{\sum_{i=1}^{N}(\lvert RR_i+RR_{i+1}\rvert > 50ms)}{N-1} \tag{4.15}$$ see.[147]

- **Frequency domain:** Analysis involves the evaluation of power distribution at different frequencies through the use of fast Fourier transformation (FFT). By transforming the amplitude of FFT, a power spectral density (PSD) is obtained, which represents the power at each frequency component. In the frequency domain, we utilize the following features: LF (low-frequency power), HF (high-frequency power), and LF/HF ratio, the description of these parameters are presented in the table 4.4. These features provide insights into the distribution of power across different frequency bands, allowing us to assess the autonomic nervous system's activity and balance.

Table 4.3: Description of the frequency domain features selected.

HRV index	**Short description**	**Reference**
LF	Low (0.04Hz - 0.15Hz) frequency band of the HRV power spectrum.	see.[148]
HF	High (0.15Hz - 0.4Hz) frequency band of the HRV power spectrum.	see.[148]
LF/HF	Ratio of LF to HF.	see.[148]

– **Nonlinear domain:** Analysis of HRV involves assessing the unpredictability and complexity of the HRV time series. This is achieved through the extraction of features from Poincaré geometric plots, which provide a detailed pattern detection mechanism using scatter plots. These features capture the nonlinearity and intricacy of the HRV dynamics, offering valuable insights into the underlying complexity of the autonomic nervous system.

  In the nonlinear analysis section, we utilize two important features: SD1 and SD2. Their descriptions can be found in the table 4.3. Equation 4.16 is employed to calculate SD1, while equation 4.17 is used to calculate SD2. These features contribute to a comprehensive understanding of the nonlinear dynamics of HRV in emotion recognition.

Table 4.4: Description of the nonlinear domain features selected.

HRV index	**Short description**	**Equation/Reference**
SD1	Poincaré plot descriptor of the short-term HRV.	$$\sqrt{variance\frac{(RR_i - RR_{i+1})}{\sqrt{2}}} \tag{4.16}$$ see.[147]
SD2	Poincaré plot descriptor of the long-term HRV.	$$\sqrt{variance\frac{(RR_i + RR_{i+1})}{\sqrt{2}}} \tag{4.17}$$ see.[147]

In the upcoming subsection, we will explain the final step in the implementation of the emotion recognition model, which involves the classification of emotions using machine learning techniques.

4.7.6 Classification

In the field of machine learning, there are three main categories of learning methods: supervised learning, unsupervised learning, and hybrid learning. In this study, we focused on supervised learning, specifically using emotion labels based on a positive/negative model. For classification purposes, we employed the random forest classifier.

The random forest algorithm operates in several key steps:

1. **Bootstrap Sampling:** Randomly selecting samples from the training dataset with replacement allows us to create multiple subsets, called bootstrapped samples. This sampling technique introduces randomness and diversity into each subset, making each tree in the forest unique.

2. **Tree Creation:** For each bootstrapped sample, create a decision tree. At each node of the tree, randomly select a subset of features (approximately the square root of 11 features in this case) to consider for splitting.

3. **Splitting Nodes:** Split each internal node of the tree using the selected features and a splitting criterion (Gini impurity). The split is chosen to maximize the separation of classes in the resulting child nodes.

4. **Recursive Process:** The process of splitting nodes continues recursively until a stopping criterion is met. This criterion could be reaching the maximum depth of the tree or having too few samples to split further.

5. **Prediction:** The prediction of the individual trees is aggregated, typically using majority voting, to produce the final prediction. This ensemble approach enhances the model's performance, leveraging the diversity and randomness of the trees while reducing the risk of overfitting.

Figure 4.13 provides a visual representation of the application of the Random Forest algorithm to extract HRV features. It showcases how the algorithm processes and analyzes these features to make accurate classifications.

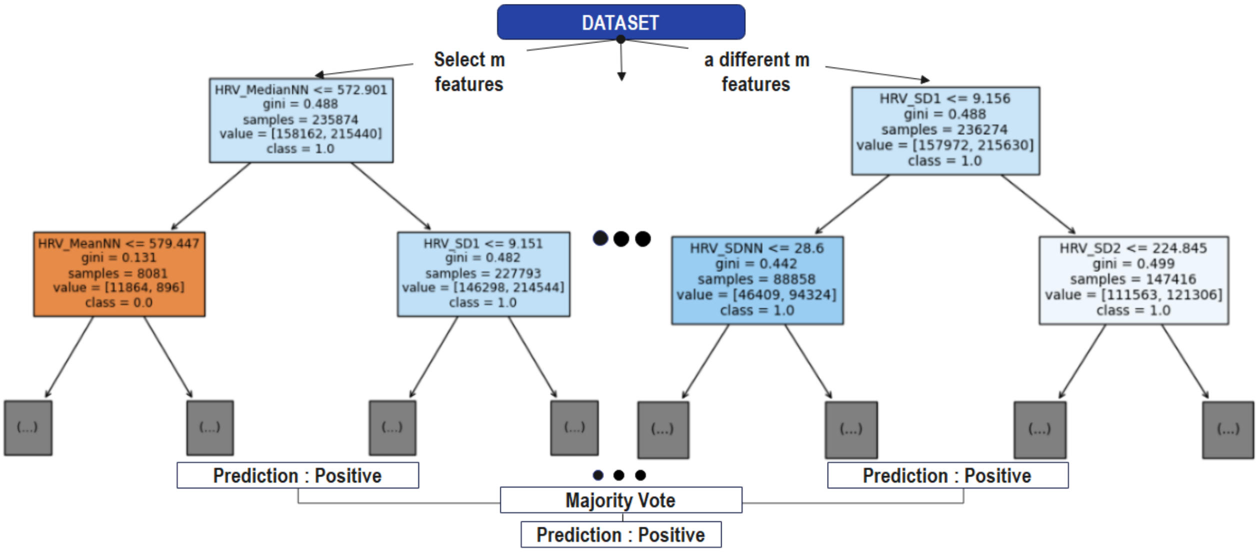


Figure 4.13: Visual representation: Random forest applied to extracted features.

The accuracy of the model demonstrated promising results, indicating its effectiveness in emotion recognition tasks.

4.8 Conclusion

In this chapter, we have provided an overview of the different problems addressed in this thesis and have presented the architectural framework for each problem. We have also described the datasets utilized in our research. Furthermore, we have outlined the preprocessing pipeline employed. Additionally, we have discussed the feature extraction and selection methods applied to extract relevant information from the data. Subsequently, we have delved into the techniques and algorithms utilized for the analysis and modeling of the data. In the next chapter, we will provide an overview of the tools, libraries, and frameworks utilized throughout our system.

“AI will not replace doctors, but instead will augment them, enabling physicians to practice better medicine with greater accuracy and increased efficiency.”

Benjamin Bell

Chapter 5

Implementation and results

5.1 Introduction

In the previous chapter, we showcased our models designs for the development of a healthcare platform. In this chapter, our focus shifts towards the implementation phase. We aim to introduce the tools and platforms that facilitated the execution of our code, enabling us to bring these designs to life and construct a remarkable platform aligned with our envisioned model designs.

Furthermore, this chapter will delve into the outcomes we achieved through an engaging research journey and the exploration of various application interfaces. We will present and analyze the results we obtained, highlighting the insights gained and lessons learned throughout the process.

By addressing these aspects, we aim to provide a comprehensive overview of the implementation process, the supportive tools employed, and the tangible outcomes obtained, ultimately showcasing the evolution of our healthcare platform.

5.2 Devlopement tools and used platforms

Development platforms and tools are crucial components in the software development process as they furnish developers with indispensable resources and frameworks to construct, test, and launch applications. Here, we present an overview of the tools that have facilitated the creation of our platform.

5.2.1 Used platforms

- **Python**:

Is a general-purpose programming language that is known for its simplicity and readability. Python is often used for data science, machine learning, deep learning, and web development [150].

Figure 5.1: Python.

- **TensorFlow**:

Is an open-source framework developed by Google that provides a Python-friendly library for numerical computation. It enables users to perform machine learning, deep learning, and other statistical and predictive analytics tasks efficiently [151].

Figure 5.2: TensorFlow.

- **Keras**:

Is an open-source software library that provides a Python interface for building artificial neural networks. It acts as a user-friendly interface for the TensorFlow library, enabling developers to leverage its powerful capabilities. Keras is specifically designed for deep learning, emphasizing fast experimentation and efficient research [152].

Figure 5.3: Keras.

- **Scikit-learn**:

This is an open-source machine-learning library for Python that provides a user-friendly interface and a comprehensive set of tools and algorithms for performing various machine-learning tasks [153].

Figure 5.4: Scikit-learn.

- **NumPy**:

This is a Python library that enhances the programming language by introducing support for large and multi-dimensional arrays and matrices. It also offers a wide range of high-level mathematical functions specifically designed to operate on these arrays [154].

Figure 5.5: NumPy.

- **Pandas**:

Is a highly efficient and versatile open-source tool for data analysis and manipulation. It is designed to provide a fast, powerful, and user-friendly experience for working with data in Python [155].

Figure 5.6: Pandas.

- **Matplotlib**:

Is a powerful Python library integrated with NumPy that enables the creation of various types of plots and visualizations. It serves as a resourceful tool for handling and analyzing large-scale numerical data [156].

Figure 5.7: Matplotlib.

- **NeuroKit2**:

Is a user-friendly package that simplifies biosignal processing for researchers and clinicians who have limited programming or biomedical signal processing knowledge. By utilizing only two lines of code, users can effortlessly analyze physiological data and utilize advanced processing techniques [157].

Figure 5.8: Neurokit.

- **Kaggle**:

Is a leading global data science community that offers a wealth of powerful tools and resources to support individuals in their data science pursuits. With its vast platform, Kaggle provides a collaborative space for data enthusiasts, scientists, and machine learning practitioners to connect, share knowledge, and engage in competitive data challenges.
Kaggle offers users unrestricted access to GPUs in our notebooks, allowing them to utilize GPU resources for up to 30 hours per week at no cost [158].

Figure 5.9: Kaggle.

- **Android studio**:

It serves as the primary integrated development environment (IDE) for building Android applications. It offers a wide array of tools and features that empower developers to create, test, and debug their Android apps efficiently [159].

Figure 5.10: Android studio.

- **Arduino IDE**:

Is a user-friendly software application that simplifies programming and development for development boards. It offers tools for writing, compiling, and uploading code onto the boards, along with an intuitive workflow and extensive library support. This IDE empowers users of all skill levels to effortlessly create projects and prototypes using Arduino boards [160].

Figure 5.11: Arduino IDE.

- **Raspberry Pi OS**:

Derived from the Debian operating system, is an open-source and highly recommended operating system specifically designed for Raspberry Pi devices. With its extensive community support and compatibility with the Raspberry Pi hardware, it offers a stable and versatile platform for various projects and applications [161].

Figure 5.12: Raspberry Pi.

- **Firebase**:

Is a comprehensive platform developed by Google that provides developers with a backend infrastructure and a suite of services for building and managing web and mobile applications. It simplifies common app development tasks, including authentication, real-time database management, cloud storage, and hosting [162].

Figure 5.13: Firebase.

- **Flutter**:

Is an open-source UI toolkit developed by Google that enables developers to create visually appealing and high-performance applications for mobile, web, and desktop platforms using a single code-base [163].

Figure 5.14: Flutter.

- **Kotlin**:

A contemporary programming language, that enables cross-platform execution on JVM and JavaScript. It boasts concise syntax and seamless integration with Java, making it a favored choice for Android app development. With its emphasis on readability and comprehensive toolset, Kotlin streamlines the development of robust and high-performing applications [164].

Figure 5.15: Kotlin.

- **ONNX (Open Neural Network Exchange)**:

Is an open-source framework that provides a standardized format for representing and exchanging AI models. It supports both deep learning and traditional machine learning models, allowing interoperability between different frameworks and platforms. This enables easier model development, deployment, and collaboration in the field of artificial intelligence [165].

Figure 5.16: ONNX.

5.2.2 Used IoT board and sensors

Our study focuses on developing a device that is both cost-effective and compact, comprising a development board and an ECG sensor. This device aims to detect and log ECG readings from patients, with the recorded data stored in Firebase. To achieve this, we require a development board equipped with a Wi-Fi module for seamless wireless connectivity and analog ports to facilitate communication with the ECG sensor. In pursuit of this objective, we conducted a comparison of various development boards, and the results are presented in table 5.1.

Table 5.1: Comparison between development board.

Features	**ESP8266**	**Arduino UNO**	**Raspberry Pi 3 B+**
Size	Small (25.4 x 50.8 mm)	Medium (53.3 x 68.6 mm)	Relatively large (85 x 56 x 17 mm)
Wi-Fi Module	Built-in	Not built-in, requires additional module or shield	Built-in
Analog Ports	Limited (1 ADC pin)	6 analog input pins	No built-in analog inputs, external ADCs can be used
Cost	Low	Moderate	Higher

In the context of our study, where we require a device with a low cost and small size, as well as a Wi-Fi module and analog port, the ESP8266 development board appears to be a suitable choice. Below, we will provide definitions and descriptions of the development board and sensor that were used in our project.

- **ESP8266-based development board** [166]:

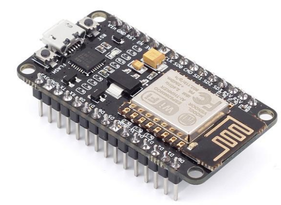


Figure 5.17: ESP8266-based development board.

(see figure 5.17) refers to a hardware board that utilizes the ESP8266 Wi-Fi module as its core component. These development boards are designed to provide an easy and convenient platform for prototyping and creating projects based on the ESP8266 module.

- **The AD8232 ECG sensor** [167]:

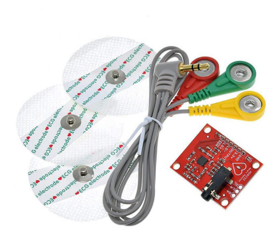


Figure 5.18: AD8232 ECG sensor.

(see figure 5.18)is a specialized electronic device used for measuring and detecting ECG signals. It is designed to capture the electrical activity of the heart and convert it into readable data. The AD8232 sensor offers a compact and easy-to-use solution for monitoring heart rate and analyzing ECG waveforms in various applications such as healthcare, fitness tracking, and biomedical research.

- **The Raspberry Pi 3 Model B+** [23]:

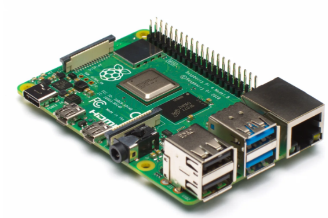


Figure 5.19: The Raspberry Pi 3 Model B+ [23].

(see figure 5.19) is the last version of the third-generation single-board computer, featuring various improvements over its predecessors. It includes a 1.4GHz 64-bit quad-core processor, dual-band wireless LAN, Bluetooth 4.2/BLE for wireless connectivity, faster Ethernet for improved networking capabilities, and Power-over-Ethernet (PoE) support with an additional PoE HAT (Hardware Attached on Top) accessory.

- **ADS1115**:

(see figure 5.20) is a 4-channel analog-to-digital converter known for its low power consumption and high precision with a 16-bit resolution. Unlike the Raspberry Pi, which is a digital-only computer lacking built-in hardware for reading analog inputs, we have incorporated the ADS1115 into the setup. This device acts as an intermediary between the Raspberry Pi 3 Model B+ and the sensor, enabling the conversion of the sensor's analog output signal into a highly accurate digital signal.

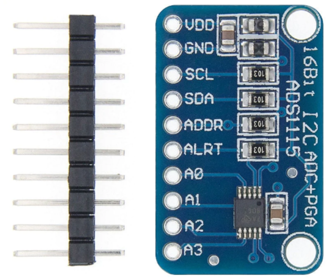


Figure 5.20: Connvertisseur analogique numérique ADS1115.

- **Pluse sensor** [168]:

(see figure 5.21) is a specialized device designed to detect and monitor the heart rate of an individual. It works by detecting the volume changes in a blood vessel, which occur with each heartbeat. Typically, the sensor is clipped onto an earlobe or a fingertip, where it can accurately measure the pulsations in the blood vessels. The sensor captures this data and provides real-time information about the heart rate.

Figure 5.21: Pluse sensor.

5.3 System interfaces and examples

In this section, we are providing an overview of each app screen on our platform, highlighting its key features and functionalities.

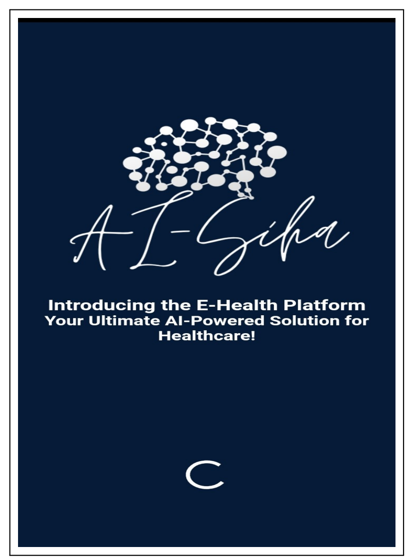
5.3.1 Detecting retinal damage application

In this portion, we will present the application that will ultimately incorporate the three services mentioned earlier. Currently, the application is equipped with a solitary service that is specifically designed for the detection of retinal damage.

- **Splash screens:**

  The left side features an image of the logo refsymbolizing the connection between health and artificial intelligence, representing the core theme of the content. On the right side, you will find the initial interface providing further insights into our chosen topic (see Figure 5.22).

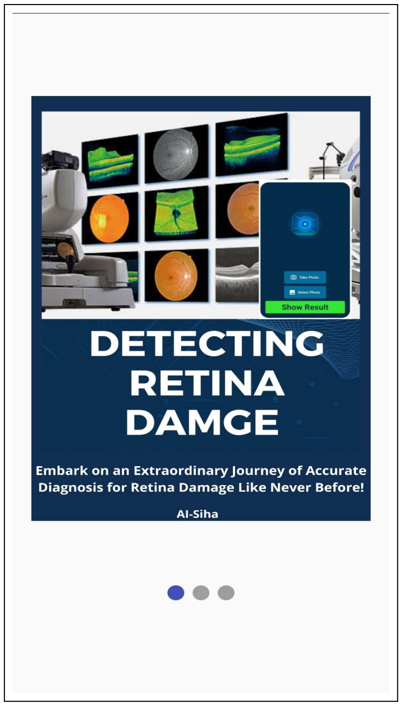
(a) The Logo of our Application.



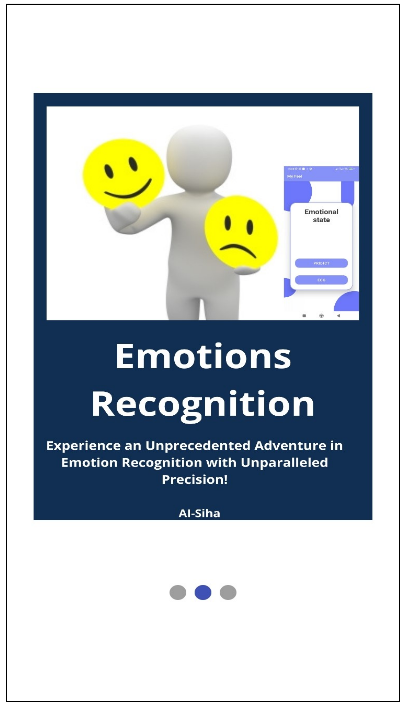
(b) Splash screen: A glimpse into the general context of our application.

Figure 5.22: The logo and first splash screen.

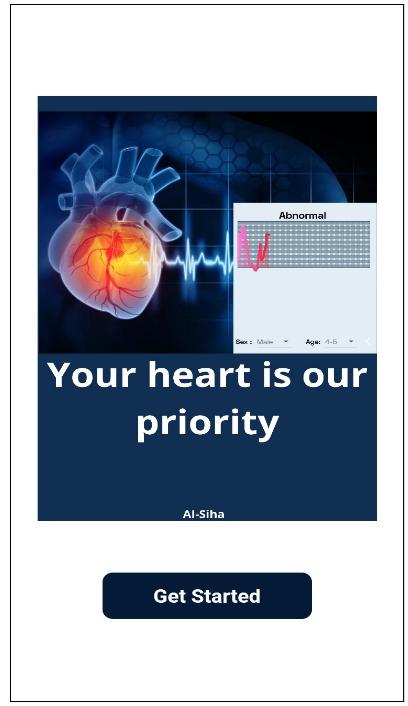
Next, we will discuss the services that our application will offer in the future, as well as the corresponding interfaces illustrated in the provided Figure 5.23.



(a) Splash screen for detecting retinal damage service.



(b) Splash screen for emotion recognition service.

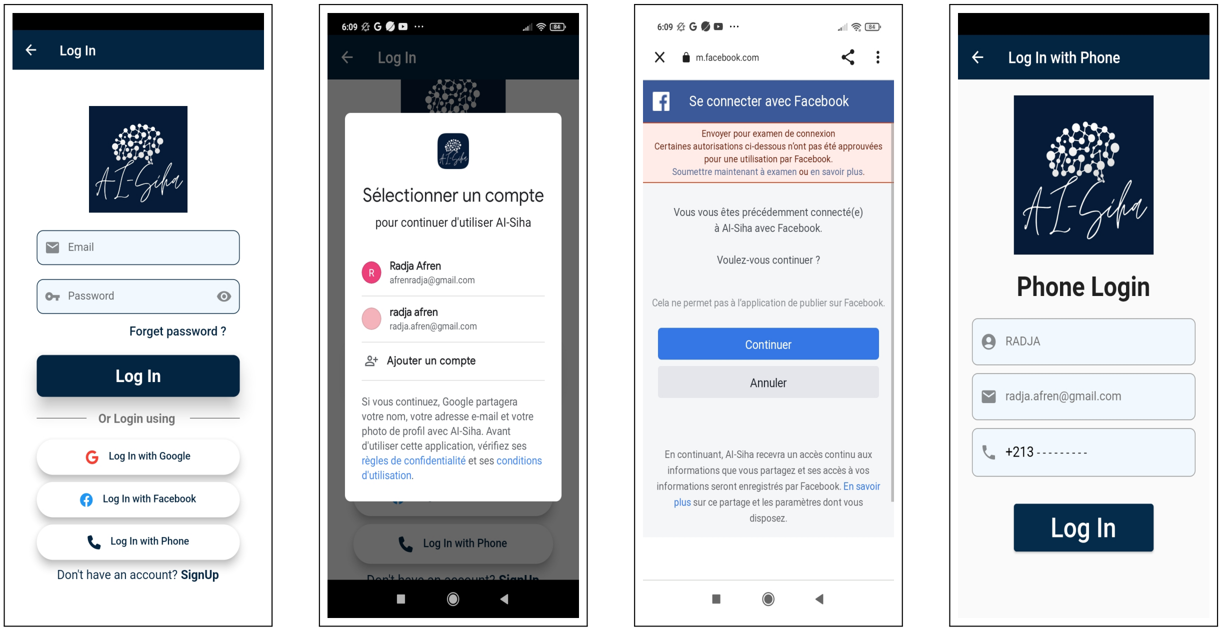


(c) Splash screen for heart deases service.

Figure 5.23: Onboarding screens.

- **Login and registration interfaces:** Upon navigating through the initial screens, the core services of the application become apparent. Following this, users can access the registration interface (refer to Figure 5.24a), enabling them to register using Gmail (refer to Figure 5.24b), Facebook (refer to Figure 5.24c), or their phone number. Verification is carried out through a notification containing a secret number (refer to Figure 5.25).

  In case of password loss, users can retrieve it (refer to Figure5.26). Additionally, after agreeing to the terms of use, users are presented with the choice to create an account. In the page of terms and condition ther's page of provider where the users gain access to the service provider's page where they can leave comments and provide feedback about the application as depicted in the accompanying Figure 5.27.



(a) The Interface of the Login Screen.

(b) Login with gmail provider.

(c) Login with Facebook provider.

(d) Login with phone provider.

Figure 5.24: Multi-provider login.

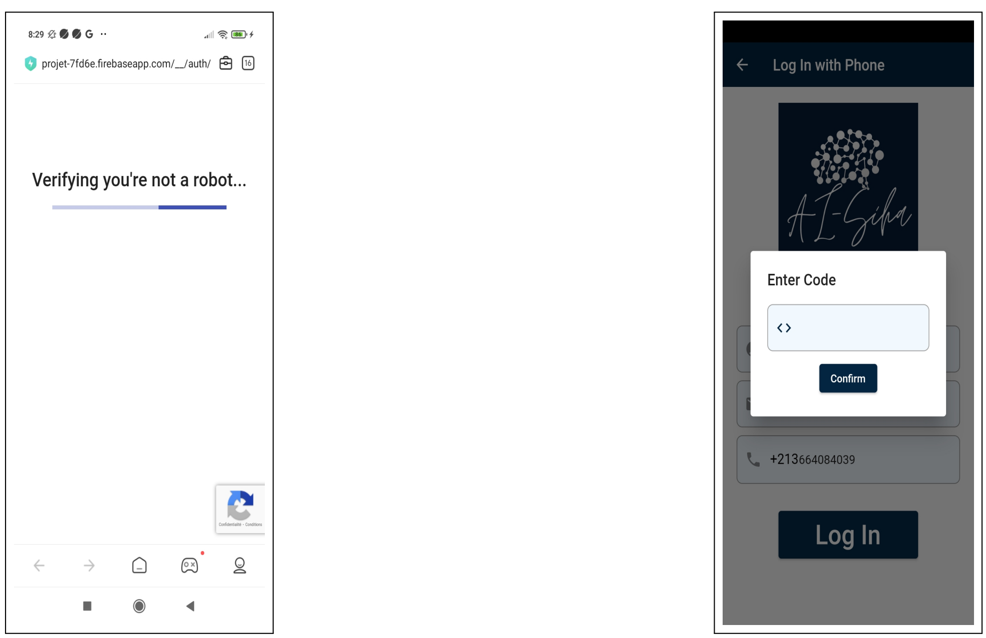
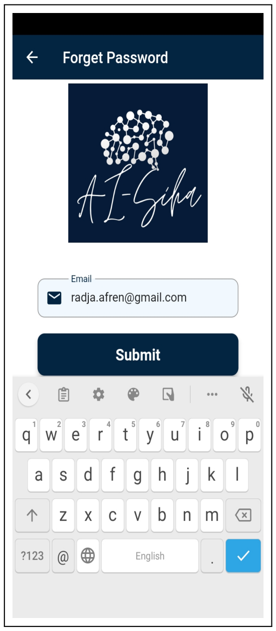
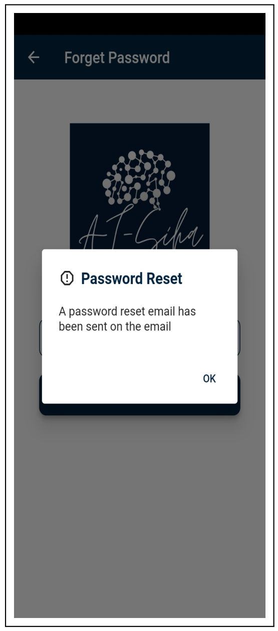


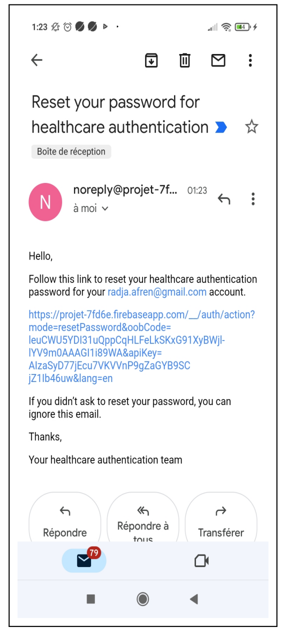
Figure 5.25: Phone OTP (One-Time Password) verification.



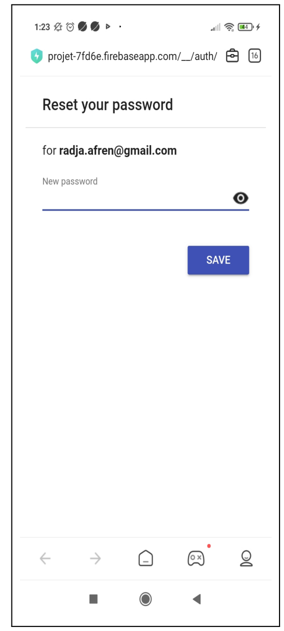
(a) Forgot Password Interface.



(b) Email for Password Reset.



(c) Service Provider Email Notification.



(d) Change Password Interface.

Figure 5.26: Forgot password procedure.

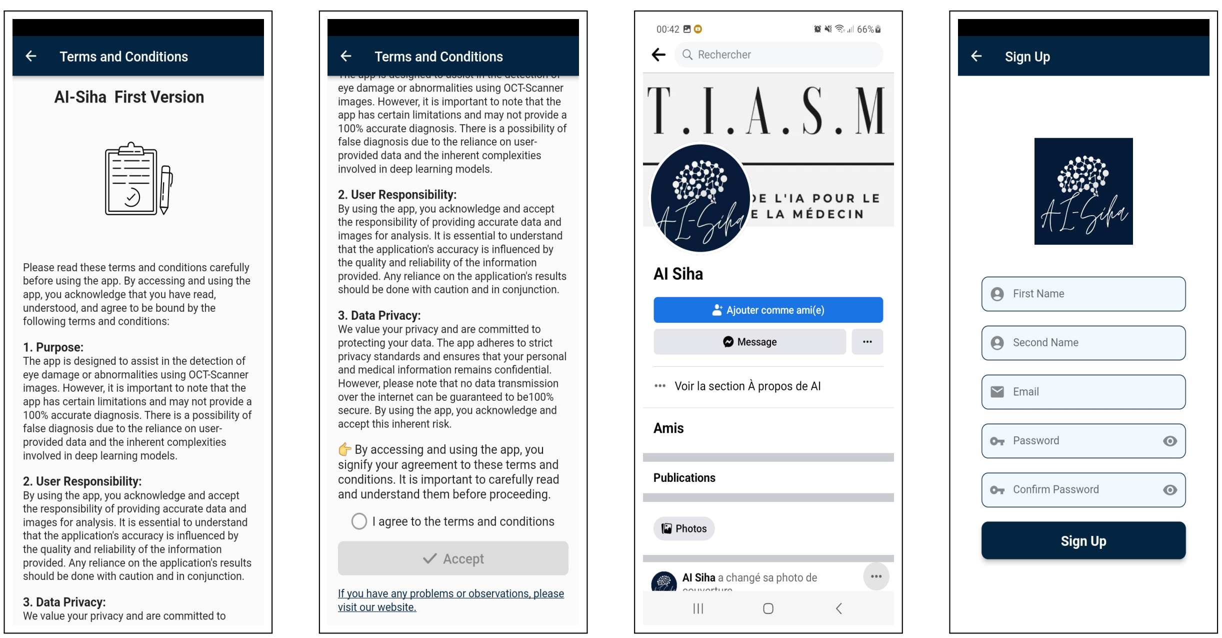


Figure 5.27: Signup procedure.

- **Home page:** The home page features mobile interfaces that showcase the available services within the application (see Figure 5.28). Clicking on each interface opens a dedicated interface for utilizing the respective service. Currently, the only functioning service is the detection of retinal diseases.

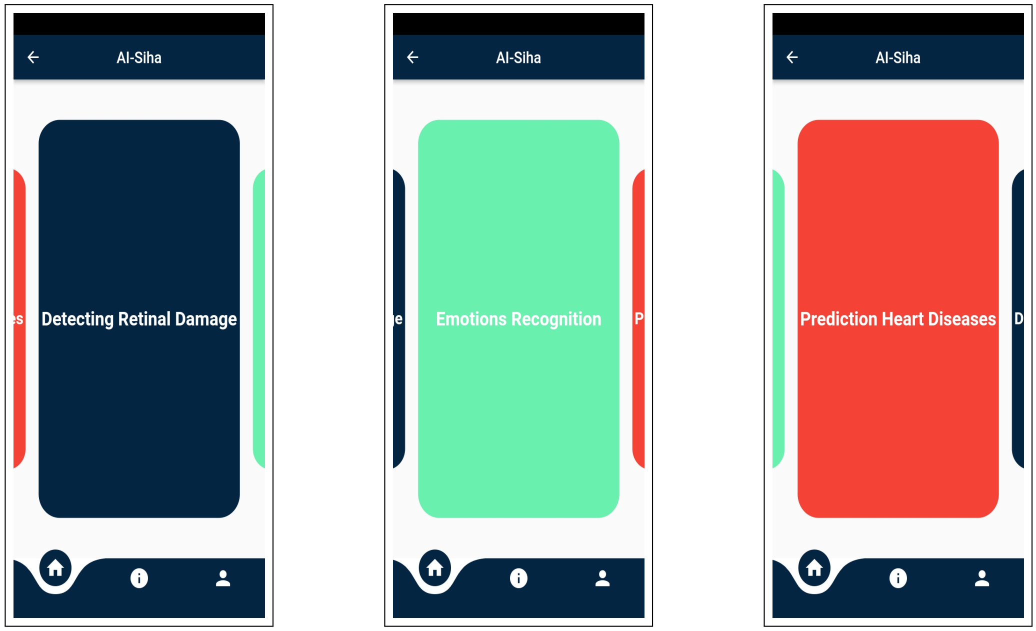
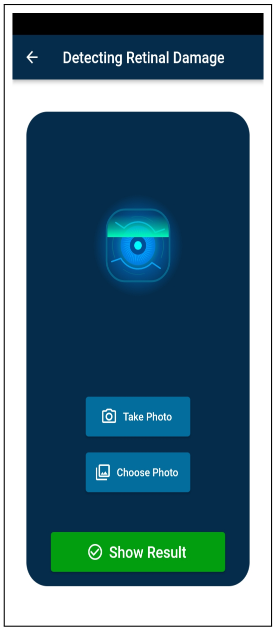


Figure 5.28: Home page interfaces.

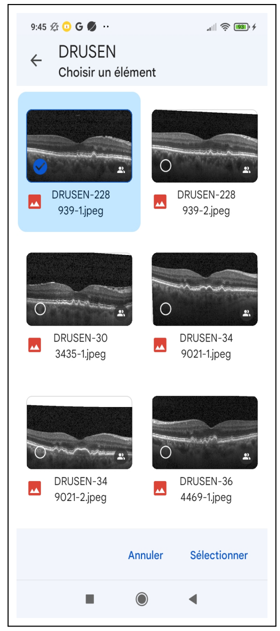
- **Retina services:**

  Within this section, we can utilize the service for detecting retinal damage. This can be done by selecting an image from the phone's gallery or using the phone's camera (see Figure 5.29a). Subsequently, we have the option to edit the image by cropping or rotating it to obtain the desired results, as illustrated in the accompanying Figure 5.29d and in Figure 5.29h. Finally, the diagnosis results are announced (see Figure 5.29f).

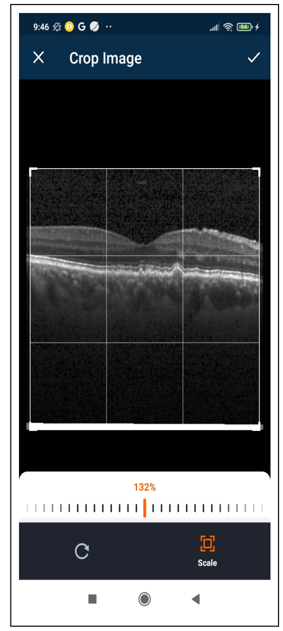


(a) Retina services page.

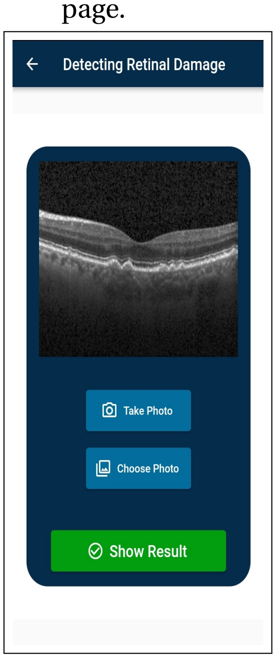
(b) No image selected interface



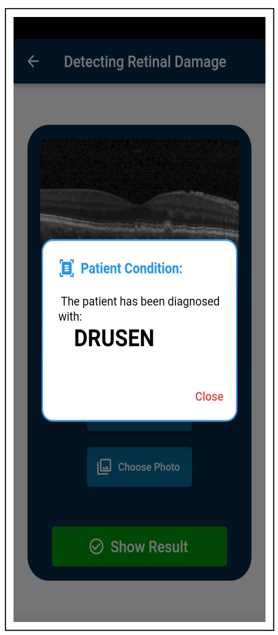
(c) Image selection interface.



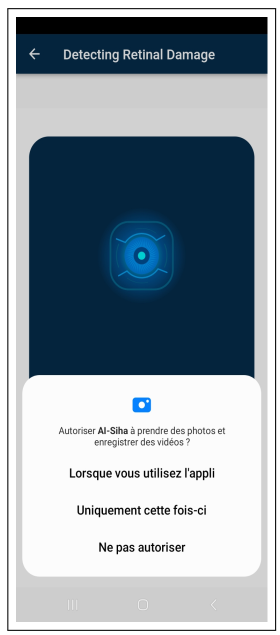
(d) Selected image cropping interface.



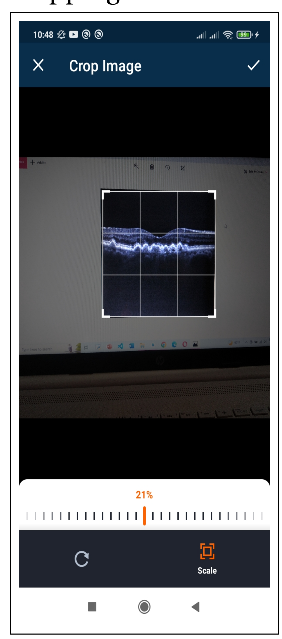
(e) Selected image page.



(f) Result display interface.



(g) Camera authorization interface.



(h) Image cropping interface.

Figure 5.29: Retina services interfaces.

- **Medical student teaching tool:** In this section, we will introduce an educational tool designed to facilitate learning for medical students through this application. It provides concise descriptions of diseases, including their symptoms, signs, and suggested treatment options (see Figure 5.30).

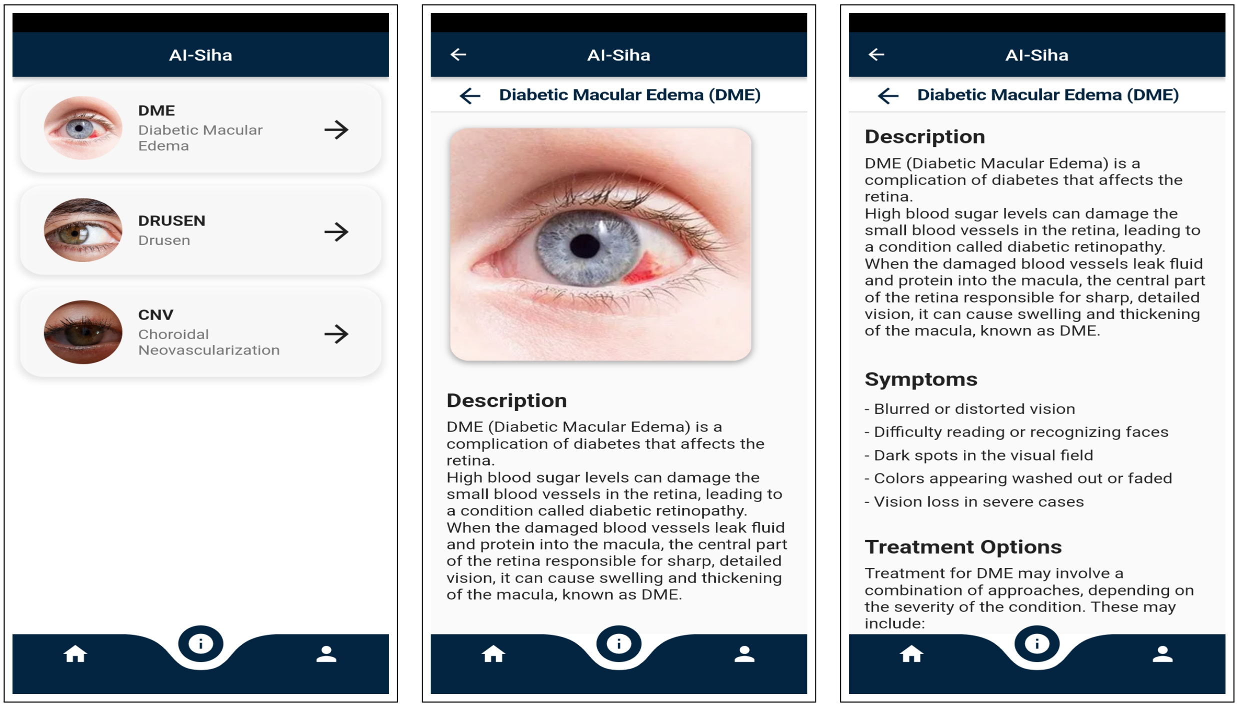


Figure 5.30: Medical student teaching tool.

- **User profile interface:** Towards the end, there is a dedicated section for displaying user personal information and providing the option to log out from the application (see Figure 5.31).

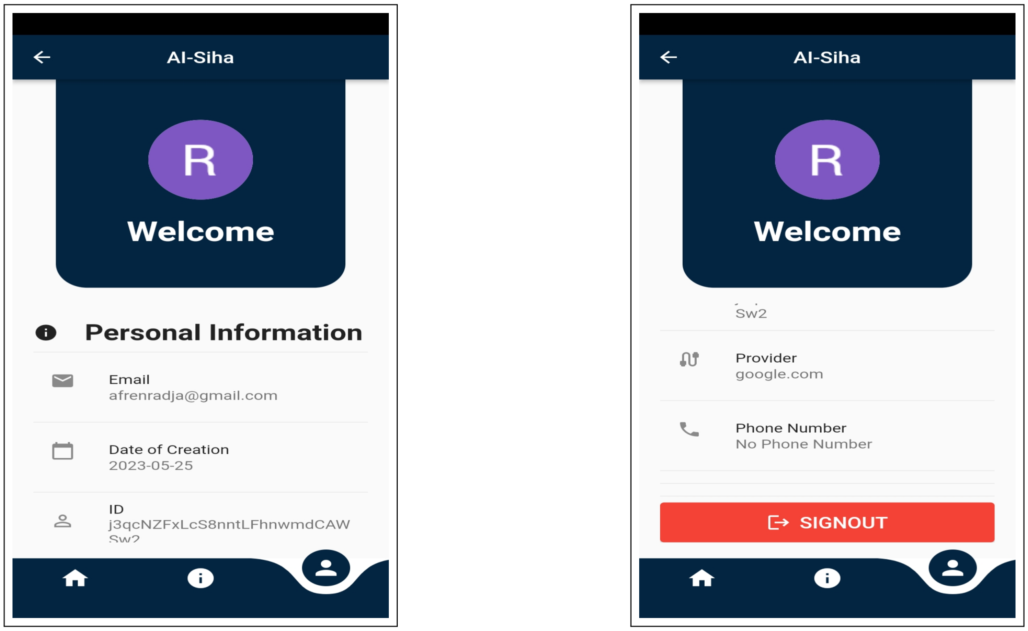


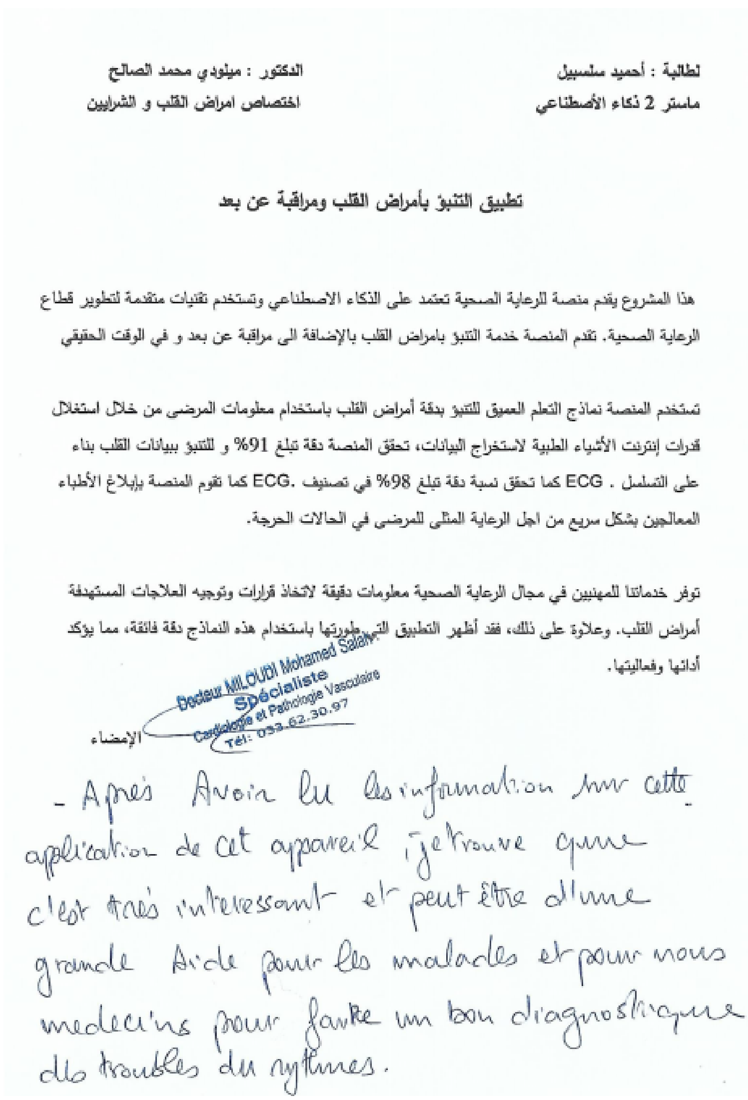
Figure 5.31: User profile interfaces.

5.3.2 Prediction of heart disease application

This section will present our application that predicts heart diseases and monitors heartbeats or ECG readings.

The main objective of the application is to identify individuals with heart disease and monitor their heart health using IoMT (Internet of Medical Things) devices such as ECG monitors or heart rate sensors. In the case of any danger, the application swiftly notifies the doctor to save the patient.

Figure 5.32 is the opinion of Dr. Salah Miloudi Mohamed Cardiovascular specialist about the application.



لطالبة : أحميد سلسبيل
ماستر 2 ذكاء الأصطناعي

الدكتور : ميلودي محمد الصالح
اختصاص امراض القلب و الشرايين

تطبيق التنبؤ بأمراض القلب ومراقبة عن بعد

هذا المشروع يقدم منصة للرعاية الصحية تعتمد على الذكاء الاصطناعي وتستخدم تقنيات متقدمة لتطوير قطاع الرعاية الصحية. تقدم المنصة خدمة التنبؤ بامراض القلب بالإضافة الى مراقبة عن بعد و في الوقت الحقيقي

تستخدم المنصة نماذج التعلم العميق للتنبؤ بدقة أمراض القلب باستخدام معلومات المرضى من خلال استغلال قدرات إنترنت الأشياء الطبية لاستخراج البيانات، تحقق المنصة دقة تبلغ 91% و للتنبؤ ببيانات القلب بناء على التسلسل . ECG كما تحقق نسبة دقة تبلغ 98% في تصنيف .ECG كما تقوم المنصة بإبلاغ الأطباء المعالجين بشكل سريع من اجل الرعاية المثلى للمرضى في الحالات الحرجة.

توفر خدماتنا للمهنيين في مجال الرعاية الصحية معلومات دقيقة لاتخاذ قرارات وتوجيه العلاجات المستهدفة أمراض القلب. وعلاوة على ذلك، فقد أظهر التطبيق التي طورتها باستخدام هذه النماذج دقة فائقة، مما يؤكد أدائها وفعاليتها.

Docteur MILOUDI Mohamed Salah
Spécialiste
Cardiologie et Pathologie Vasculaire
Tél: 032.62.30.97

الإمضاء

- Après Avoir lu les information sur cette application de cet appareil je trouve que c'est très interessant et peut être d'une grande Aide pour les malades et pour nous medecins pour faire un bon diagnostique des troubles du rythmes.

Figure 5.32: Cardiovascular specialist opinion about the application.

(a) Splash interface.

(b) Onboarding interface.

Figure 5.33: Start Interfaces.

Figure 5.33a illustrates the initial screen that is briefly displayed upon launching the mobile application. Figure 5.33b represents the Onboarding screen, which appears after the initial screen when the mobile application is launched. The Onboarding screen provides information to new users, allowing them to familiarize themselves with the app's features.

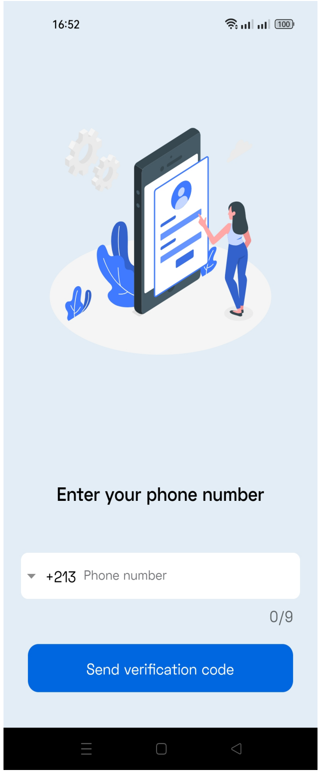
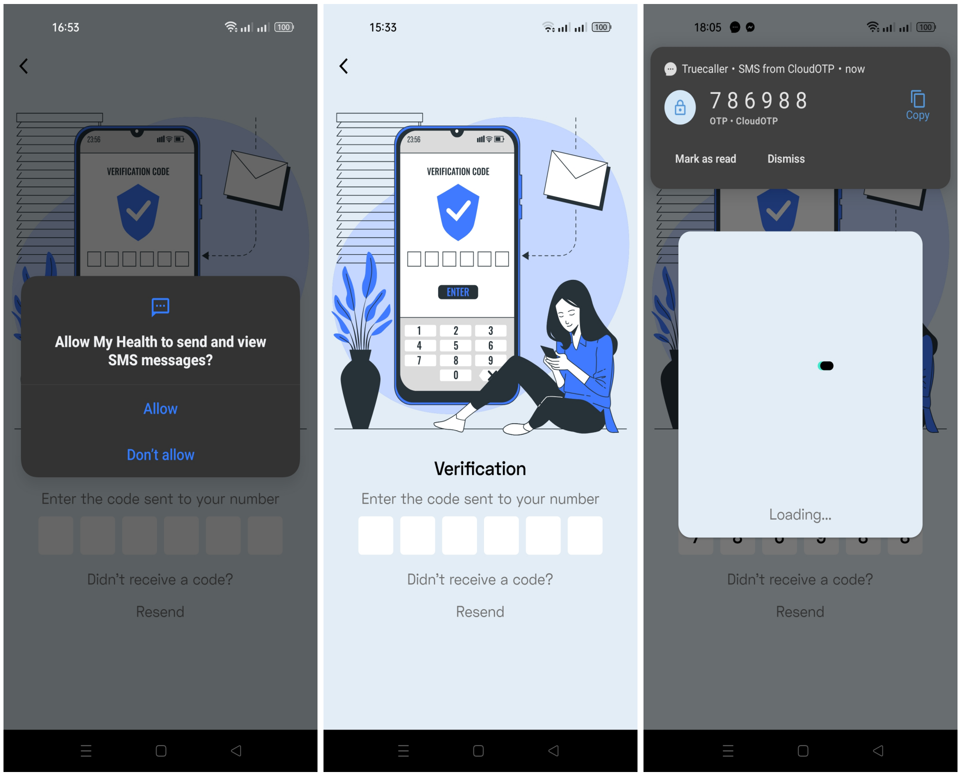


Figure 5.34: Login Interface.

Figure 5.34 demonstrates the login interface, where users are required to enter their phone number to receive an OTP (One-Time Password) for authentication. Users will be prompted to input their phone number, and upon verification, an OTP will be sent to their registered number. The figure may also depict the option for users to select their country or region during the login process, enabling them to choose their appropriate location for authentication purposes. the figure includes a button labeled "Send Verification Code" in the login interface. This button allows users to initiate the process of receiving the OTP for authentication. When users click on this button, a verification code will be sent to their registered phone number, enabling them to proceed with the login process.



(a) SMS Message Access Permission.

(b) Verification code interface.

(c) Automated verification code retrieval.

Figure 5.35: Verification code interface: SMS Access and Auto-Retrieval

Figure 5.35 depicts the screen for entering the verification code of the OTP (One-Time Password). In this screen, the user is presented with options to either allow or deny the mobile application's permission to send and view SMS messages s shown in figure 5.51b. If the user chooses to allow the app access, the application will automatically retrieve the verification code from the received SMS, simplifying the login process. However, if the user denies permission, they will need to manually input the verification code received.

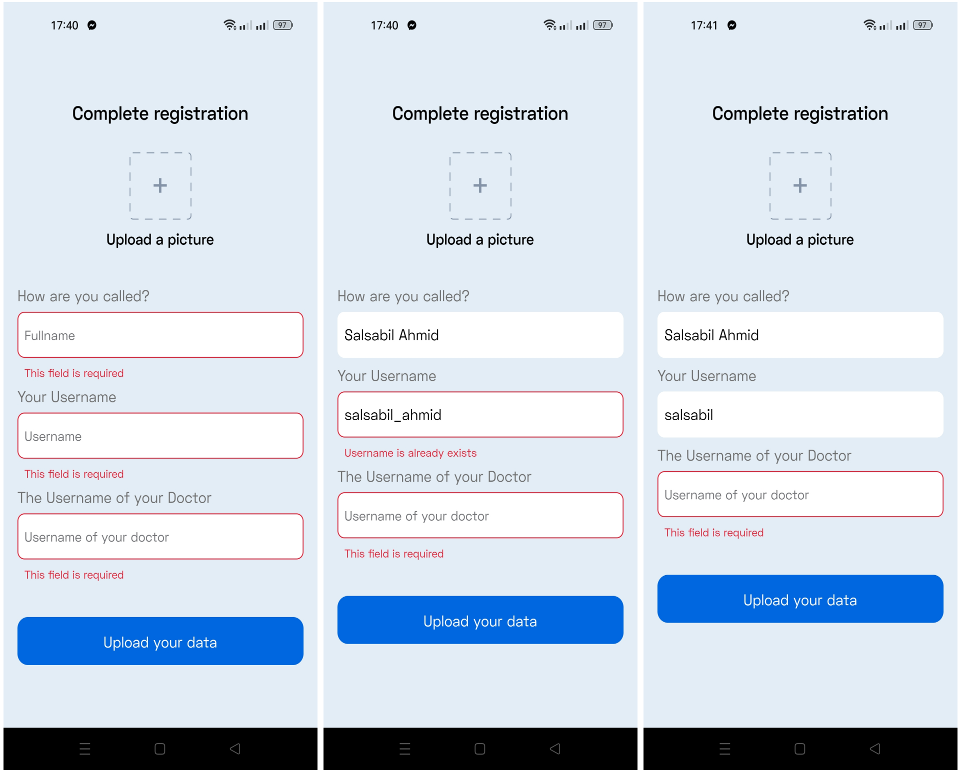


Figure 5.36: Complete registration interface.

Figure 5.50 showcases the complete registration screen, which includes the following fields:

1. **Name:** The user can enter their full name in this field.

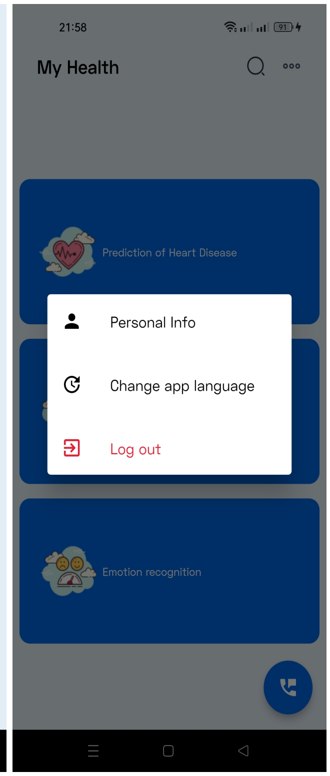
2. **Username:** The user can input their desired username in this field.

3. **Doctor's Username:** This field is specifically for healthcare professionals, allowing them to enter their unique username.

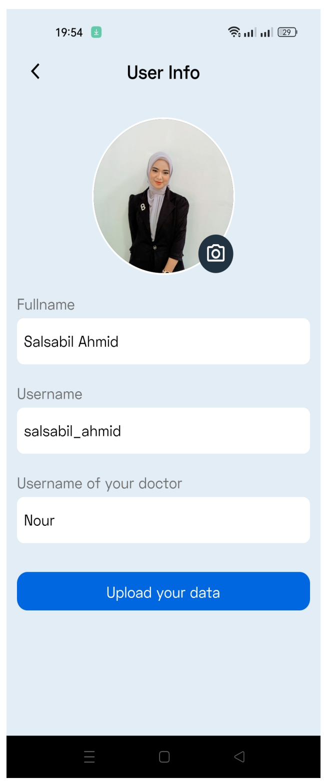
If any of the fields are left blank, or if the entered username is incorrect, already taken, or doesn't meet the required criteria, the screen will display an error message indicating the specific issue encountered. This feedback on the screen helps users identify and correct the errors in their registration details.

(a) Home interface.

(b) Home interface.



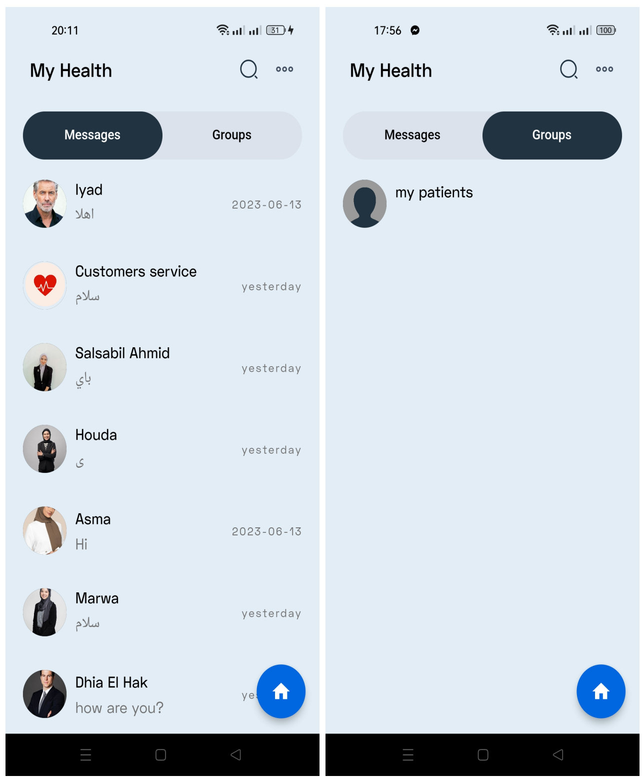
(c) The options.



(d) The profile.

Figure 5.37: The home screen with the options.

The home screen (see figure 5.37) of the application features three buttons for accessing different services or features. Additionally, users have the ability to search for other users using their usernames. Furthermore, they can navigate to view and modify their personal information and profile as shown in figure 5.51a. The home screen also provides options for changing the application's language, allowing users to switch between Arabic and English. Lastly, users can log out from their accounts using the respective option. In addition to the mentioned features, the home screen also provides a button for accessing the chat screen.



(a) Individual. (b) Group.

Figure 5.38: The conversations screen.

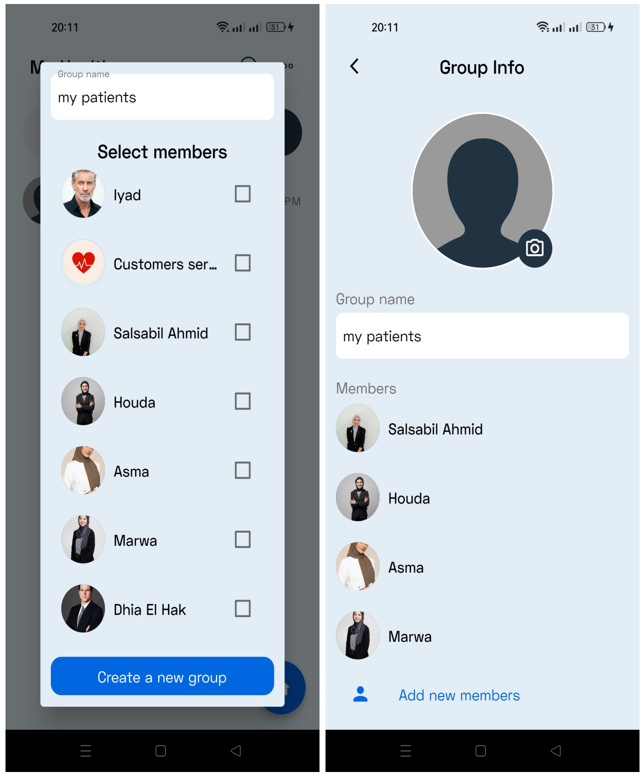


Figure 5.39: Create and manage new groups and select group members.

The chat screen in the mobile application allows users to engage in both individual and group conversations. Users can view and send messages in individual chats, as well as participate in group chats 5.38. They can create new groups, select members for group conversations, and manage group settings such as adding or removing members and updating group information 5.39. The chat screen serves as a central hub for users and doctors to communicate, providing them with the flexibility to stay connected with others and efficiently manage.

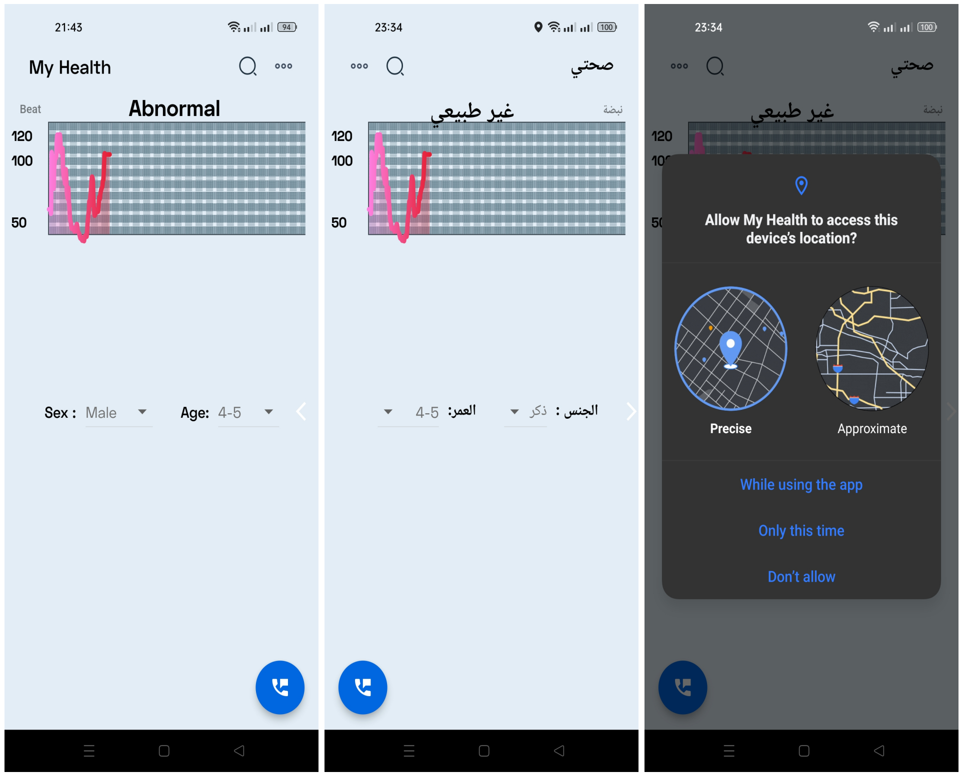


Figure 5.40: The monitoring heart rate screen.

The monitoring heart rate screen 5.40 in the mobile application allows users to view a real-time chart of their heart rate. They can select their gender and set the age interval for accurate analysis. If an abnormal heart rate is detected, the app automatically retrieves the user's location and sends a message to their doctor, providing all the necessary information for immediate assistance. This feature combines real-time monitoring, customizable settings, and automated communication to enhance cardiovascular health monitoring and enable prompt medical intervention when needed.

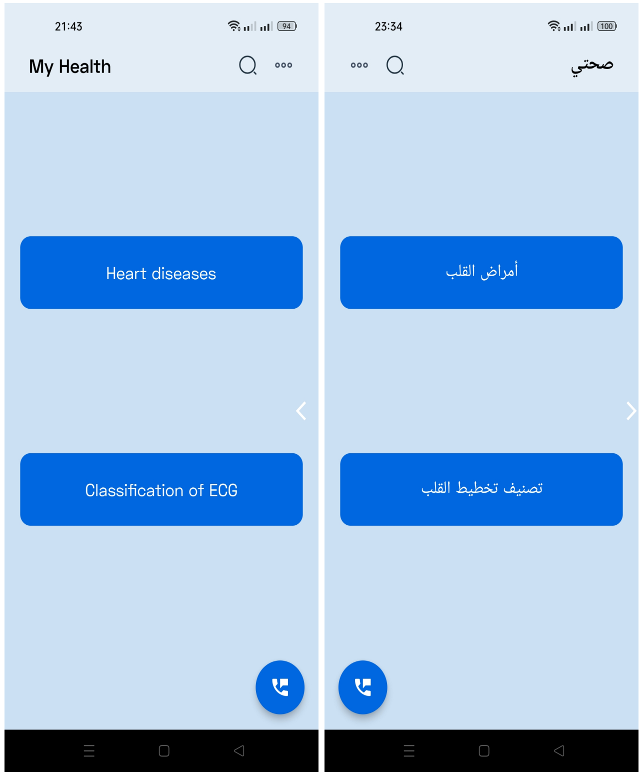


Figure 5.41: Service selection.

In this screen 5.41, users are presented with a choice between two services: heart disease information and ECG classification. By providing these distinct options, the screen accommodates diverse user needs and enables them to select the service that best suits their requirements and interests, whether it be accessing information about heart diseases or utilizing ECG classification tools.

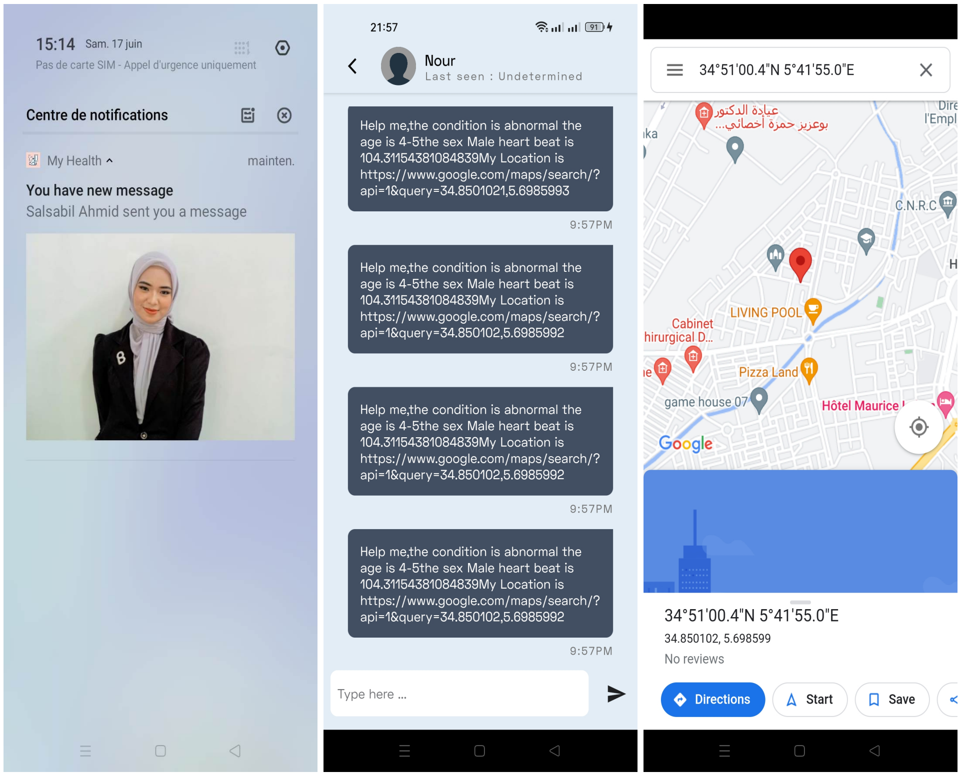


Figure 5.42: Instant doctor notification.

The application sends a real-time notification to the doctor when an abnormal heart rate or ECG classification is detected. The notification includes a message containing crucial information such as the patient's heart rate, age, sex, and location. When the doctor opens the message, the application displays a map with the patient's location and may provide directions for a quick response, as shown in figures 5.42. This feature aims to facilitate prompt medical intervention and potentially save lives by ensuring doctors have immediate access to relevant patient information and can reach them efficiently.

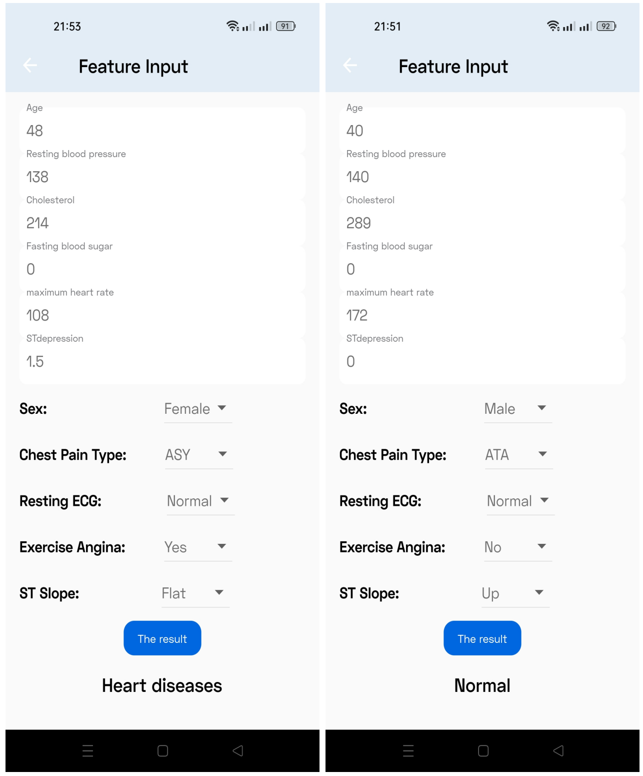


Figure 5.43: The doctor's input screen.

The doctor's input screen 5.43 allows them to enter essential information about the patient's medical condition related to heart disease. This includes the patient's age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, resting electrocardiogram results, maximum heart rate, exercise-induced angina, ST depression induced by exercise, and ST slope. These data points assist the doctor in assessing the patient's cardiac health, making accurate diagnoses, and determining appropriate treatment plans.

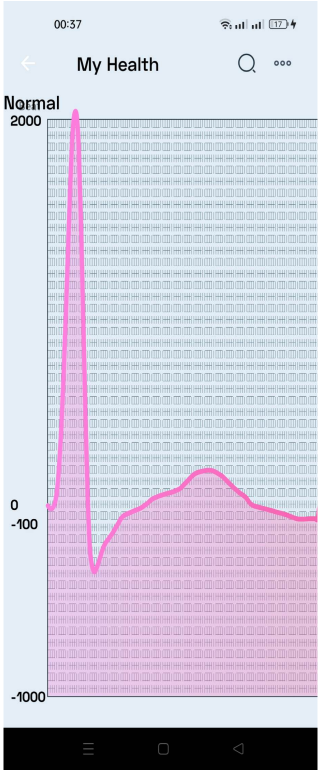
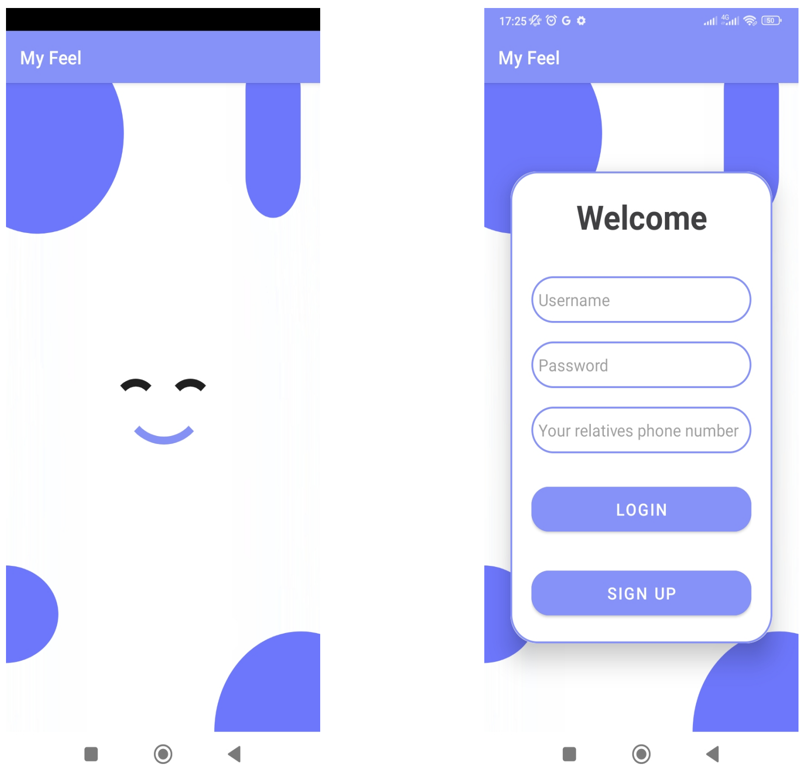


Figure 5.44: The ECG monitoring screen.

The ECG monitoring screen 5.44 in the mobile app enables users to observe a live chart of their ECG readings. This function allows users to track their heart's electrical activity in real time, providing valuable information about their cardiovascular health. When an irregular ECG pattern is detected, the app automatically retrieves the user's location and sends a notification to their assigned doctor, supplying all the relevant details for immediate assistance. By integrating real-time monitoring, customizable options, and automated communication, this screen enhances the monitoring of cardiovascular health and ensures timely medical intervention when necessary.

5.3.3 Emotion recognition application

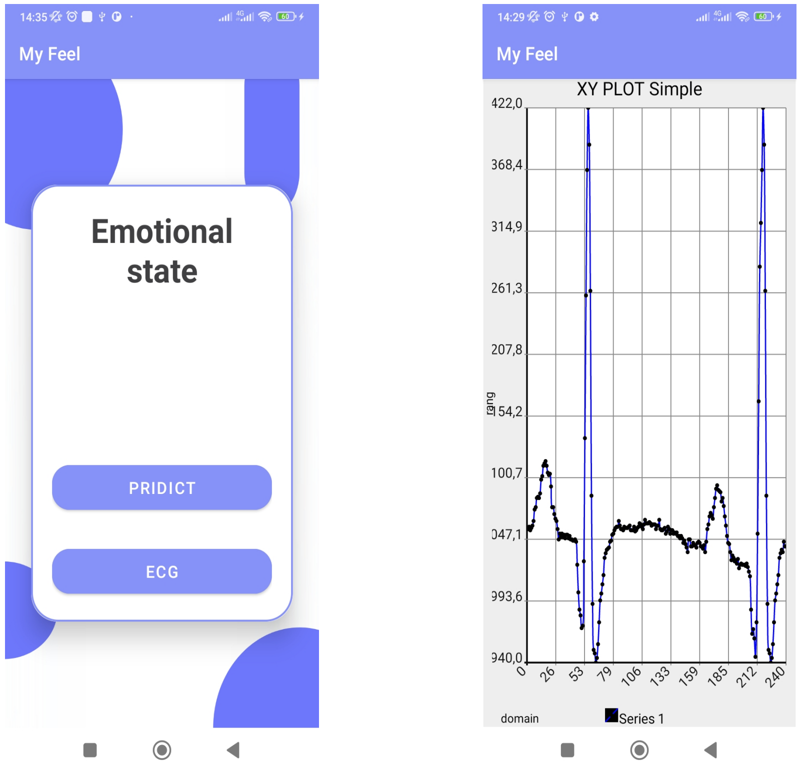
For the emotion recognition application, the user journey begins with a captivating splash screen (see figure 5.45a) that sets the mood and captures the user's attention. After the splash screen, users are seamlessly directed to the login and sign-up screen as shown in figure 5.45b. This ensures a smooth transition into the application, where users can either log in with their existing credentials or create a new account to access the full range of features and functionalities offered by the application.



(a) Splash screen. (b) Authentication screen.

Figure 5.45: Splash and user authentication screen.

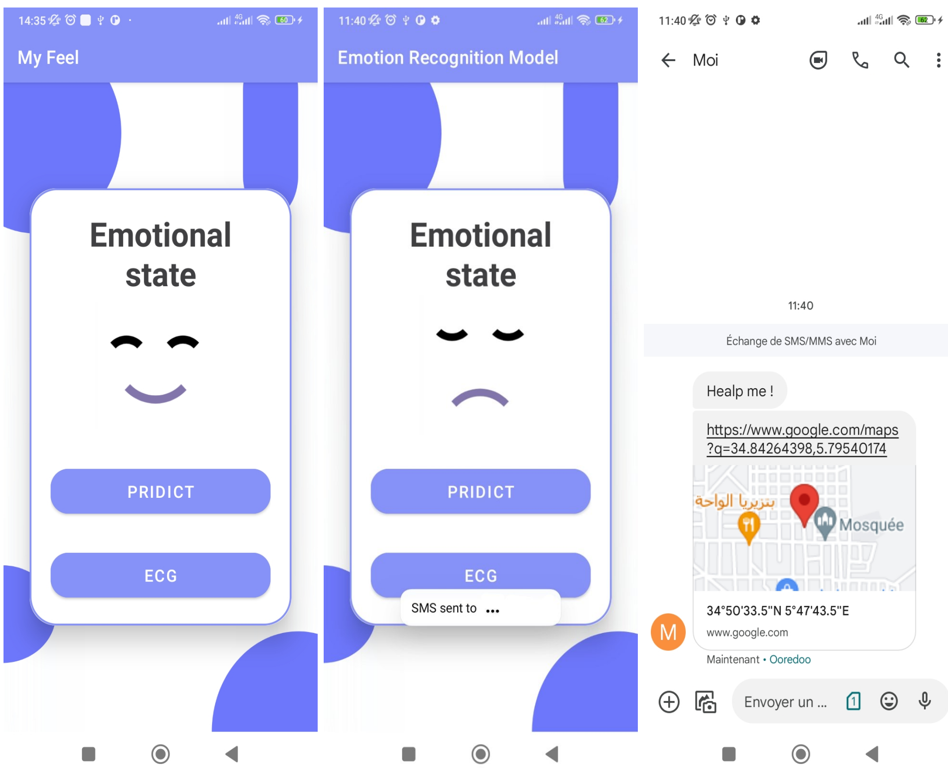
After login or sign up, users are directed to an interface for monitoring their emotional state, as depicted in figure 5.46a. The interface features two buttons: one for predicting the emotional state and the other for displaying the user's ECG data from Firebase, as illustrated in figure 5.46b.



(a) Home interface. (b) The ECG visualization.

Figure 5.46: Home Interface and ECG Visualization screen.

When the user clicks on the predict button, if the emotional state is positive, a smiley face is shown (see figure 5.47a). However, if the emotional state is negative, the interface displays a sad face and automatically sends an SMS to the user's designated relatives, requesting assistance (see figure 5.47b). The SMS includes a Google Maps link pinpointing the user's current location, facilitating easy navigation, and providing immediate support as shown in the figure 5.47c.



(a) Positive state screen. (b) Negative state screen. (c) SMS notification.

Figure 5.47: Emotion recognition service.

5.4 Obtained results and discussion

In this section, we will discuss the different results obtained from each model used in our project.

5.4.1 Prediction of heart disease

In this part, we will discuss the different results obtained from each model used in the project focused on the protection of heart disease. First, we will examine the Heart Disease Detection model, which aims to detect and differentiate between healthy individuals and those with heart disease. This model provides valuable insights for diagnosis.

Next, we present our findings on predicting the next sequence of ECG based on the past sequence of ECG. By utilizing seq2seq, we have developed models capable of accurately forecasting the upcoming ECG patterns. This predictive capability holds the potential for early detection of heart abnormalities and allows for timely medical intervention.

Finally, we will delve into the classification of ECG sequences into normal and abnormal patterns. By leveraging machine learning algorithms, we have created a classification model that effectively distinguishes between these two categories. This classification aids in identifying abnormal ECG sequences, enabling healthcare professionals to intervene promptly and provide appropriate care.

Heart disease detection

In binary classification tasks for heart disease detection, a common approach is to label one class as "healthy" and the other class as "unhealthy" or "diseased." Assigning the label "1" to the unhealthy class and "0" to the healthy class is standard practice.

Figure 5.48 represents the database or dataset used for training the model, where each row corresponds to a patient and each column represents a feature or attribute. The last column contains the labels, with "1" indicating an unhealthy patient and "0" indicating a healthy patient.

	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
0	40	M	ATA	140	289	0	Normal	172	N	0.0	Up	0
1	49	F	NAP	160	180	0	Normal	156	N	1.0	Flat	1
2	37	M	ATA	130	283	0	ST	98	N	0.0	Up	0
3	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
4	54	M	NAP	150	195	0	Normal	122	N	0.0	Up	0
...	...	...	...	...	...	...	...	...	...	...	...	...
913	45	M	TA	110	264	0	Normal	132	N	1.2	Flat	1
914	68	M	ASY	144	193	1	Normal	141	N	3.4	Flat	1
915	57	M	ASY	130	131	0	Normal	115	Y	1.2	Flat	1
916	57	F	ATA	130	236	0	LVH	174	N	0.0	Flat	1
917	38	M	NAP	138	175	0	Normal	173	N	0.0	Up	0

918 rows × 12 columns

Figure 5.48: Original structure.

We modified the original structured dataset for heart disease analysis by renaming categories and applying one-hot encoding. The categorical values were updated to ensure consistency and improve interpretability. Additionally, we standardized the numerical features using the formula z-score normalization, which helped in scaling the data and ensuring equal contribution from all features as shown in the figure 5.49.

	Age	RestingBP	Cholesterol	FastingBS	MaxHR	Oldpeak	HeartDisease	Sex_0	Sex_1	ChestPainType_0	...	ChestPainType_2	ChestPainType_3	RestingECG_0	RestingECG_1	RestingECG_2	ExerciseAngina_0	ExerciseAngina_1	ST_Slope_0
0	40	140	289	0	172	0.0	0	0	1	0	...	0	0	1	0	0	1	0	0
1	49	160	180	0	156	1.0	1	1	0	0	...	1	0	1	0	0	1	0	0
2	37	130	283	0	98	0.0	0	0	1	0	...	0	0	0	1	0	1	0	0
3	48	138	214	0	108	1.5	1	1	0	0	...	0	1	1	0	0	0	1	0
4	54	150	195	0	122	0.0	0	0	1	0	...	1	0	1	0	0	1	0	0

5 rows × 21 columns

Figure 5.49: The modified dataset structure.

To evaluate the performance of our model accurately, we split the dataset into three subsets: a training subset, a validation subset, and a testing subset as shown in Table 5.6. This division allowed us to train the model on the training subset, optimize hyperparameters as shown in Table 5.7, and finally assess the model's performance on the testing subset. to create a reliable and effective predictive model for heart disease.

Table 5.2: Dataset Information

Dataset	Initial Shape
Initial Dataset	(918, 12)
Modified Dataset	(918, 21)
Training Set	(734, 20)
Testing Set	(184, 20)
Training Set (After Split)	(587, 20)
Validation Set	(147, 20)

Table 5.3: Model Hyperparameters

Hyperparameter	Value
Model Architecture	Sequential
Hidden Layers	4
Hidden Units	128, 64, 32, 32
Activation Function	ReLU
Regularization	L2 (0.01)
Output Units	1
Output Activation	Sigmoid
Loss Function	Binary Crossentropy
Optimizer	Adam

By training the model on this labeled dataset, it learns to identify patterns and relationships between the input features and the corresponding labels, enabling it to classify new, unseen patients as either healthy or unhealthy based on their feature values.

Results and discussion

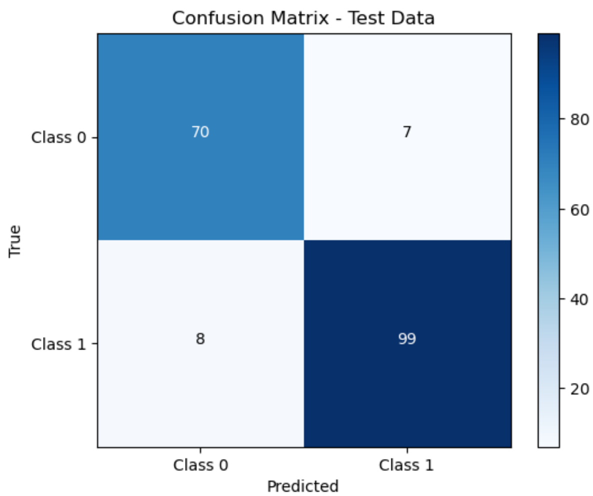
The metrics employed for evaluation encompass training accuracy, validation accuracy, training loss, and validation loss. To effectively showcase the model's performance, we provide the following:

Table 5.4: Loss and accuracy values

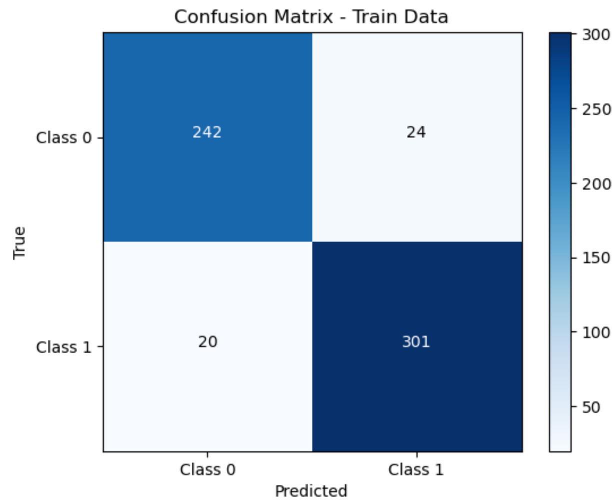
	Loss	Accuracy
Validation	0.367909312	0.911564648
Test	0.413065344	0.918478250
Train	0.330613493	0.925042569

Confusion matrix

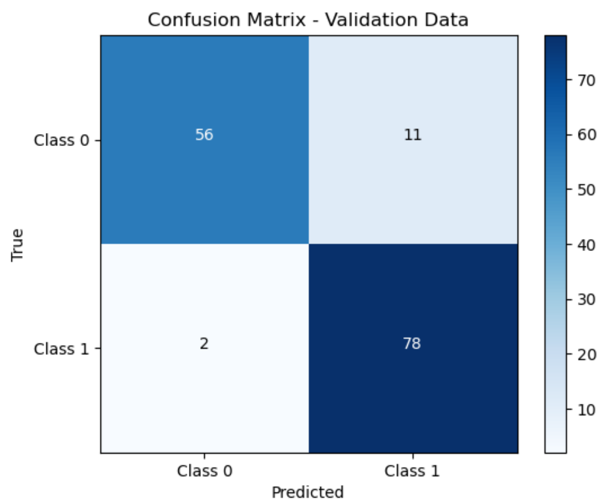
We plot the Confusion Matrix to see how well our model performs in each dataset (train, test, and validation). So we got the results reported in Figure 5.50 :



(a) Confusion matrix for the test dataset.



(b) Confusion matrix for the training dataset.



(c) Confusion matrix for the validation dataset.

Figure 5.50: Confusion matrix.

The confusion matrix typically consists of four values: TP, TN, FP, and FN. With these values, we can calculate performance metrics such as accuracy, precision, recall, and F1 score, which provide insights into how well the model is performing as shown in Table 5.5.

Table 5.5: Performance Metrics

Data	Precision	NPV	Accuracy	Specificity	Sensitivity	F1 Score
Test	0.933962	0.897436	0.918478	0.909091	0.925234	0.929577
Validation	0.876404	0.965517	0.911565	0.835821	0.975	0.923077
Train	0.926154	0.923664	0.925043	0.909774	0.937695	0.931888

analyze the given results for the training, test, and validation data sets:

- F1 Score: It is the harmonic mean of precision and sensitivity (or recall). It provides a balance between these two metrics. Higher values indicate better performance. The F1 scores for all three datasets are relatively high, indicating good overall performance.

- Precision: It measures the proportion of correctly predicted positive instances out of the total predicted positive instances. Higher precision indicates fewer false positives. The precision values for all three datasets are reasonably high, suggesting that the models have a low rate of false positives.

- NPV: It measures the proportion of correctly predicted negative instances out of the total predicted negative instances. Higher NPV indicates fewer false negatives. The NPV values for the training and test datasets are relatively high, while the validation dataset has a significantly higher NPV. This suggests that the model performs better at correctly predicting negative instances in the validation data.

- Accuracy: It measures the overall correctness of the predictions, considering both true positives and true negatives. Higher accuracy indicates better overall performance. The accuracy values for all three datasets are relatively high, indicating good overall prediction accuracy.

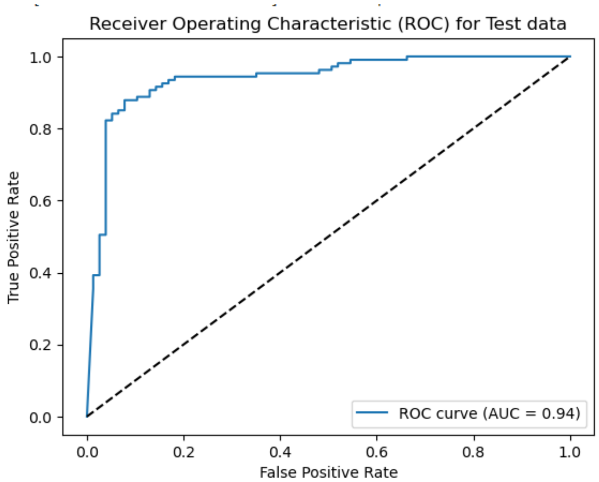
- Specificity: It measures the proportion of correctly predicted negative instances out of the total actual negative instances. Higher specificity indicates a lower false positive rate. The specificity values for all three datasets are relatively high, suggesting that the models have a low rate of false positives.

- Sensitivity: It measures the proportion of correctly predicted positive instances out of the total actual positive instances. Higher sensitivity indicates a lower false negative rate. The sensitivity values for all three datasets are relatively high, indicating that the models have a low rate of false negatives.

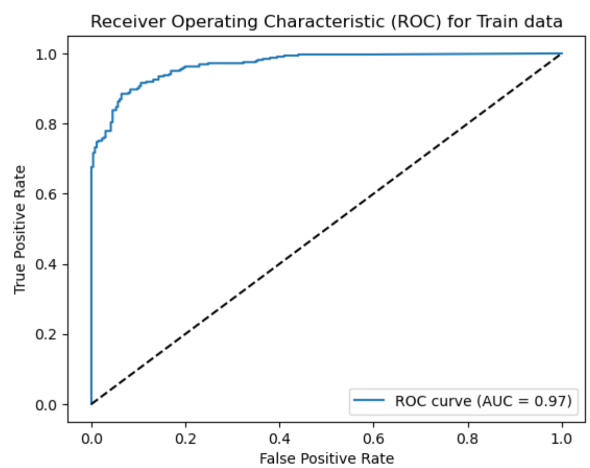
ROC curves

The ROC curves are commonly used to evaluate the performance of binary classification models. The area under the ROC curve (AUC) is a metric that quantifies the overall performance of the model, with values ranging from 0 to 1. A higher AUC indicates better discrimination ability and a more accurate model. As we see in Figure 5.51a, Figure 5.51b, and Figure 5.53 the AUC values for the ROC curves of the different datasets are as follows:

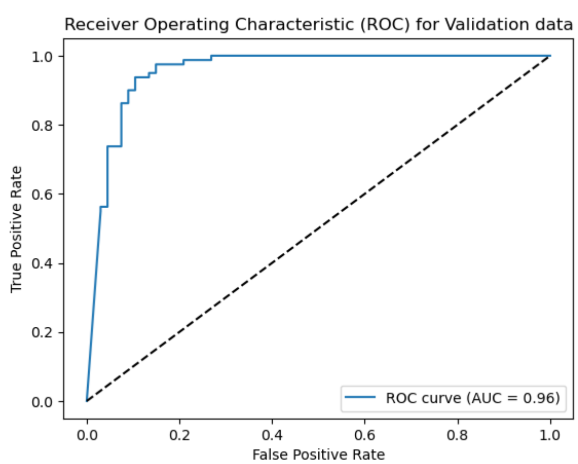
- Test Dataset ROC curve AUC: 0.94
- Train Dataset ROC curve AUC: 0.97
- Validation Dataset ROC curve AUC: 0.96



(a) ROC for the test dataset.



(b) ROC for the training dataset.



(c) ROC for the validation dataset.

Figure 5.51: Receiver operating characteristic.

We observe that these AUC values suggest that our model performs well on all three datasets. The highest AUC of 0.97 is observed in the training dataset, indicating that our model achieves excellent discrimination ability on the data it was trained on. The AUC values of 0.94 and 0.96 for the test and validation datasets, respectively, suggest that our model generalizes well to unseen data and can effectively classify instances from these datasets.

Prediction of Sequence Electrocardiogram

In our attempt to predict sequences of ECG based on the ECG5000 dataset as shown in Figure5.52, we initially tried to utilize the Datasat ECG Heartbeat Categorization Dataset. However, we encountered challenges as the dataset contained a significant number of zero values, and various algorithms failed to yield satisfactory results. As an alternative, we turned to the ECG5000 dataset, which consists of 5000 rows and 141 columns. We found that excluding the last column, which was intended for classification purposes, was not required for the sequence prediction task.

	0	1	2	3	4	5	6	7	8	9	...	131	132	133	134	135	136	137	138	139	140
0	1.0	-0.112522	-2.827204	-3.773897	-4.349751	-4.376041	-3.474986	-2.181408	-1.818286	-1.250522	...	0.160348	0.792168	0.933541	0.796958	0.578621	0.257740	0.228077	0.123431	0.925286	0.193137
1	1.0	-1.100878	-3.996840	-4.285843	-4.506579	-4.022377	-3.234368	-1.566126	-0.992258	-0.754680	...	0.560327	0.538356	0.656881	0.787490	0.724046	0.555784	0.476333	0.773820	1.119621	-1.436250
2	1.0	-0.567088	-2.593450	-3.874230	-4.584095	-4.187449	-3.151462	-1.742940	-1.490659	-1.183580	...	1.284825	0.886073	0.531452	0.311377	-0.021919	-0.713683	-0.532197	0.321097	0.904227	-0.421797
3	1.0	0.490473	-1.914407	-3.616364	-4.318823	-4.268016	-3.881110	-2.993280	-1.671131	-1.333884	...	0.491173	0.350816	0.499111	0.600345	0.842069	0.952074	0.990133	1.086798	1.403011	-0.383564
4	1.0	0.800232	-0.874252	-2.384761	-3.973292	-4.338224	-3.802422	-2.534510	-1.783423	-1.594450	...	0.966606	1.148884	0.958434	1.059025	1.371682	1.277392	0.960304	0.971020	1.614392	1.421456
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4995	4.0	-1.122969	-2.252925	-2.867628	-3.358605	-3.167849	-2.638360	-1.664162	-0.935655	-0.866953	...	0.205543	-0.472419	-1.310147	-2.029521	-3.221294	-4.176790	-4.009720	-2.874136	-2.008369	-1.808334
4996	2.0	-0.547705	-1.889545	-2.839779	-3.457912	-3.929149	-3.966026	-3.492560	-2.695270	-1.849691	...	1.218185	1.258419	1.907530	2.280888	1.895242	1.437702	1.193433	1.261335	1.150449	0.804932
4997	2.0	-1.351779	-2.209006	-2.520225	-3.061475	-3.065141	-3.030739	-2.622720	-2.044092	-1.295874	...	-0.896575	-1.512234	-2.076075	-2.586042	-3.322799	-3.627311	-3.437038	-2.260023	-1.577823	-0.684531
4998	2.0	-1.124432	-1.905039	-2.192707	-2.904320	-2.900722	-2.761252	-2.569705	-2.043893	-1.490538	...	-2.495989	-2.821782	-3.268355	-3.634981	-3.168765	-2.245878	-1.262260	-0.443307	-0.559769	0.108568
4999	2.0	0.728813	0.192597	-0.733884	-1.779456	-2.345908	-2.977565	-3.380053	-3.417164	-3.030925	...	1.219499	1.267275	1.678989	2.483389	2.569073	2.122891	1.753963	1.538975	1.713781	1.309382

5000 rows × 141 columns

Figure 5.52: The ECG5000 dataset.

In order to accurately assess the performance of our seq2seq model in predicting ECG sequences, we partitioned the ECG5000 dataset into two distinct subsets: a training subset and a testing subset, as illustrated in Table 5.6. This partitioning enabled us to train the model on the training subset, fine-tune the hyperparameters (as documented in Table 5.7), and ultimately evaluate the model's performance on the testing subset. Our primary objective was to develop a dependable and efficient predictive model.

Table 5.6: Dataset Information

Dataset	**sequences**
Training Set	4000
Testing Set	1000
Total	5000

Table 5.7: Model Hyperparameters.

Hyperparameter	**Value**
Encoder Bidirectional LSTM Units	64
Decoder LSTM Units	128
Optimizer	Adam
Loss Function	MSE
Metrics	MSE, MAE, RMSE
Number of Epochs	100
Batch Size	32

The model architecture : is a Seq2Seq encoder-decoder model. It consists of an encoder network that processes the input sequence and a decoder network that generates the output sequence.

The encoder part of the model takes an input sequence of size 50 with one feature per time step. This input is passed through a bidirectional LSTM layer with 128 units. The bidirectional LSTM layer produces outputs for both the forward and backward sequences, resulting in an output shape of (None, 50, 128). The final hidden state of the forward and backward sequences, as well as the final cell state, are concatenated, resulting in a shape of (None, 64) for both.

The decoder part of the model takes an input sequence of size 90 with one feature per time step. This input is concatenated with the previously concatenated hidden and cell states from the encoder, resulting in shapes of (None, 128) for both concatenations. The concatenated input is then passed through an LSTM layer with 128 units, resulting in an output shape of (None, 90, 128). Finally, a dense layer with one unit is applied to predict the output sequence, resulting in a shape of (None, 90, 1).

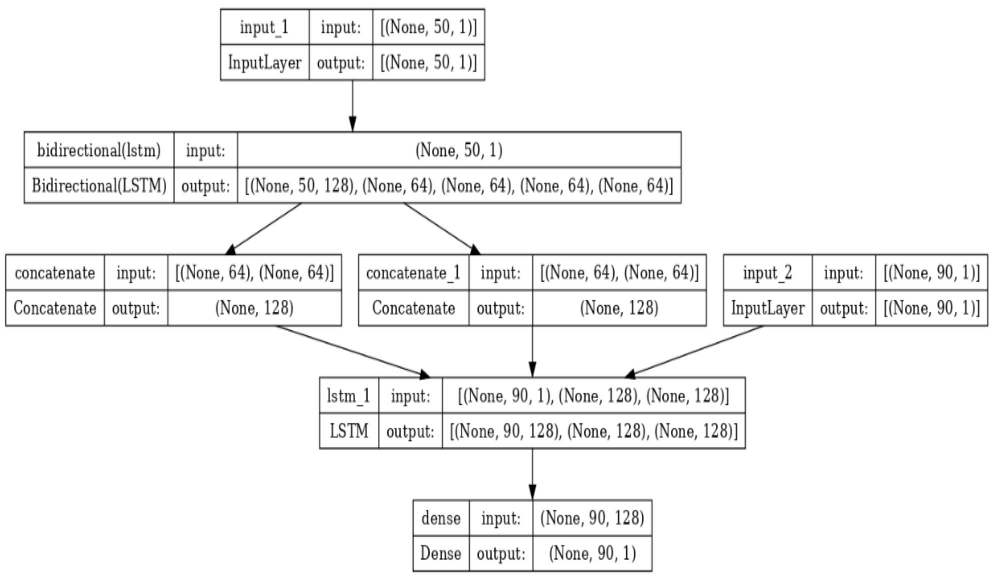
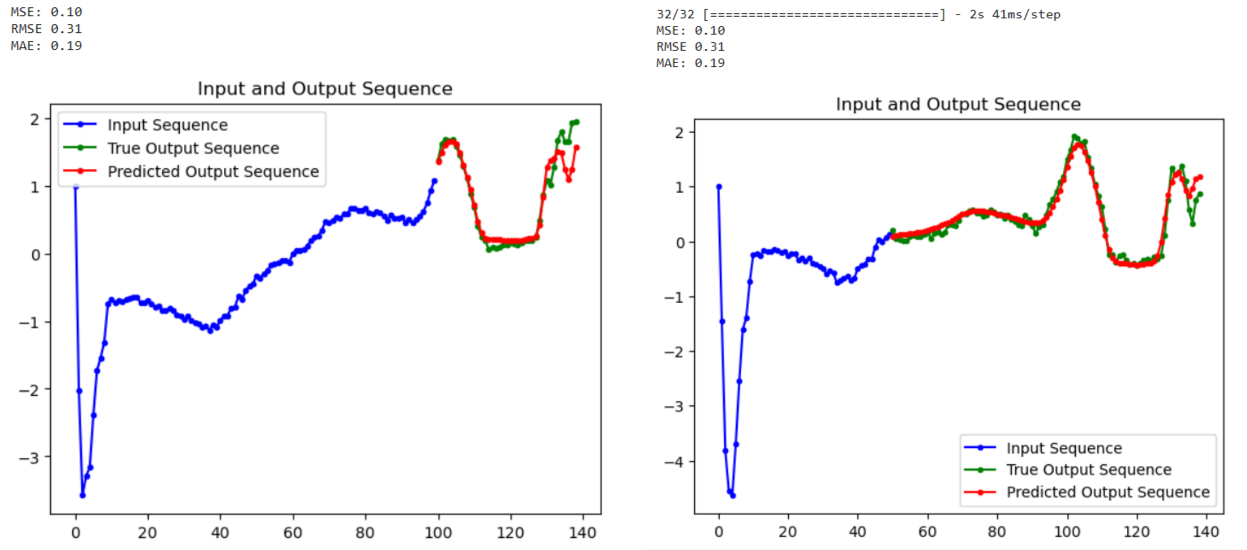


Figure 5.53: The Seq2seq architecture.

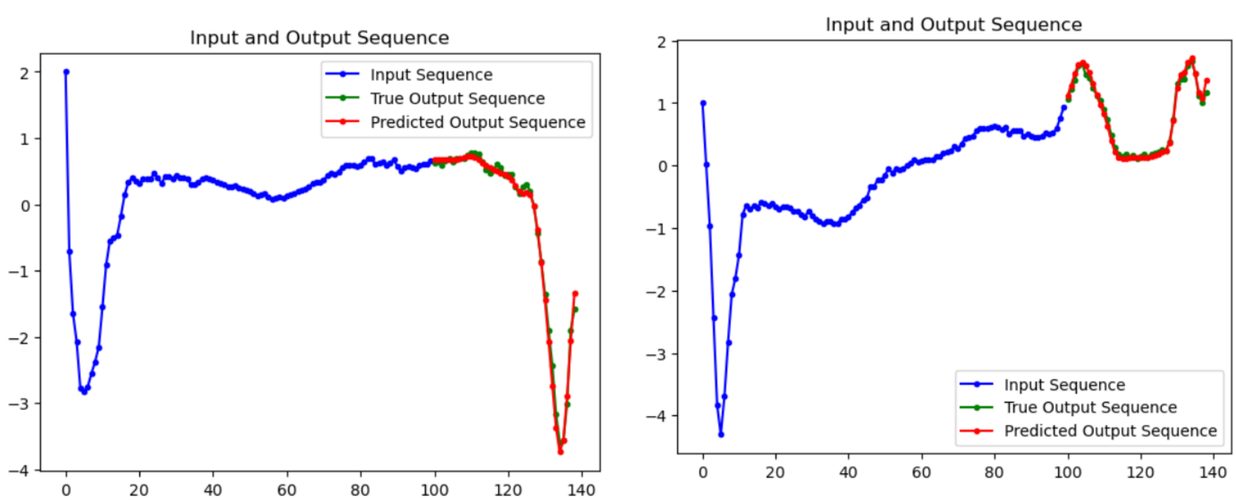
Based on the analysis of our results, we observed that when using a sequence length of 100 as input and 40 as output, our model achieved an MSE of 0.10, an RMSE of 0.31, and an MAE of 0.19, as depicted in Figure 5.58a. Surprisingly, we obtained similar results when modifying the input and output sequence lengths to 50 and 90, respectively, as shown in Figure 5.58b. In the case of MSE, a value of 0.10 indicates that, on average, the squared difference between the predicted values and the actual values is 0.10. and RMSE value of 0.31, indicates the square root of the average squared difference between the predicted values and the actual values.MAE value of 0.19, represents the average absolute difference between the predicted values and the actual values.



(a) (b)

Figure 5.54: Results

Upon closer examination in Figure 5.55, where the blue line represents the input sequence and the green line represents the true output sequence, we observe a striking similarity between the green (true output) and red (predicted output) lines. This observation signifies that our model was successful in accurately predicting the output sequence.



(a) (b)

Figure 5.55: Input and Output sequence.

These findings highlight the robustness of our model's performance, as it was able to achieve comparable results despite variations in the input and output sequence lengths. Overall, these results indicate the effectiveness of our approach in accurately predicting ECG sequences and emphasizing the reliability of our model.

Classification of Sequence Electrocardiogram

In our pursuit of classifying ECG data sequences, we employed the ECG5000 dataset (as depicted in Figure 5.52). Additionally, we meticulously grouped and balanced the datasets (as illustrated in Figure 5.56) to ensure a fair representation of different categories. Specifically, we organized a total of 62,876 ECG sequences, each comprising 140 values. This resulted in 31,438 sequences categorized as normal and an equal number of 31,438 sequences categorized as abnormal (as depicted in Figure 5.57).

	output	1	2	3	4	5	6	7	8	9
0	0.0	1.000000	0.515464	0.508591	0.498282	0.491409	0.470790	0.463918	0.408935	0.278351
1	1.0	0.958194	0.804348	0.162207	0.030100	0.225753	0.227425	0.219064	0.217391	0.217391
2	1.0	0.987421	0.865828	0.672956	0.494759	0.236897	0.067086	0.014675	0.014675	0.023061
3	0.0	1.000000	0.639701	0.367427	0.032431	0.000000	0.053101	0.154312	0.180684	0.197078
4	1.0	1.000000	0.938650	0.579755	0.266871	0.159509	0.125767	0.113497	0.076687	0.101227
...	...	...	...	...	...	...	...	...	...	...
62871	1.0	1.000000	0.635379	0.025271	0.043321	0.072202	0.086643	0.083032	0.097473	0.093863
62872	1.0	1.000000	0.778603	0.369985	0.078752	0.008915	0.098068	0.148588	0.141159	0.126300
62873	0.0	1.000000	0.779352	0.673077	0.341093	0.222672	0.342105	0.308704	0.300607	0.230769
62874	0.0	0.961353	0.886473	0.142512	0.000000	0.200483	0.200483	0.222222	0.210145	0.205314
62875	1.0	0.913201	0.777577	0.506329	0.195298	0.018083	0.025316	0.075949	0.099458	0.088608

62876 rows × 141 columns

Figure 5.56: The balanced dataset.

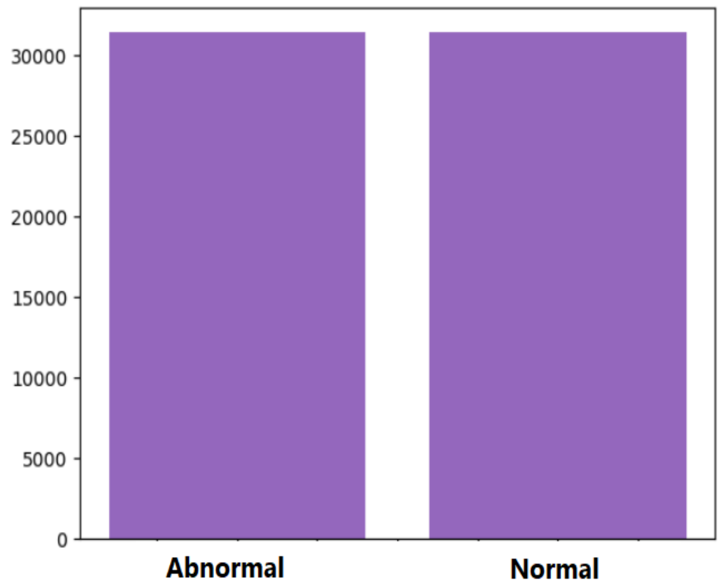


Figure 5.57: The Structure of the dataset.

To achieve this, we extensively evaluated various machine learning algorithms, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Trees (DT), Random Forest (RF), XGBoost (XGB), and implemented a DNN. Additionally, we also explored the applicability of Naive Bayes and Support Vector Machines (SVC) in this context. By comparing the performance of these algorithms, as depicted in Table 5.8 and in Table 5.9. we aim to develop an effective method for accurately classifying ECG sequences based on their characteristics.

The final column "Testing value" in Table 5.8 and in Table 5.9 represents the test value, which combines the values of recall, accuracy, F-1 score, and F-Beta score. These metrics are grouped together in one column because their values are close and similar.

Table 5.8: Comparison between different classifiers without a balanced dataset.

Model	Training Accuracy	Testing Accuracy	Testing Score
LR	89.39	89.27	89.27
KNN	98.21	97.41	97.41
DT	100.00	95.01	95.01
RF	100.00	97.47	97.47
XGB	99.80	97.84	97.84
Naive Bayes	81.08	80.93	80.93
SVC	92.65	92.63	92.63
DNN	99.84	98.51	98.51

Based on these results, the DNN model performs the best in terms of testing accuracy and F1 score. It achieves an accuracy of 98.51%, indicating excellent performance in classifying the testing data. Additionally, it has the highest F1 score of 98.51%, which indicates a good balance between precision and recall.

Following the DNN model, the RF and XGB models also exhibit strong performance, with testing accuracies of 97.47% and 97.84%, respectively. These models have high F1 scores as well.

The other models, such as LR, KNN, DT, Naive Bayes, and SVC, also achieve reasonably good accuracies, but their performance is slightly lower compared to the top-performing models.

Overall, the DNN model appears to be the best choice for this classification task, followed by the RF and XGB models.

It's important to note that accuracy alone is not always the best metric to evaluate a model's performance, especially when dealing with imbalanced datasets. Therefore, considering other metrics like recall, precision, and F1 score can provide a more comprehensive understanding of a model's strengths and weaknesses.

Confusion Matrices:

LR:

$$\begin{bmatrix} 667 & 1834 \\ 141 & 15763 \end{bmatrix}$$

Starting with LR, we see that it achieved a relatively high TN count of 15,763, indicating it correctly predicted the negative class instances. However, the FP count is quite high at 1,834, suggesting a significant number of false alarms or incorrect positive predictions. The TP count of 667 indicates correctly predicted positive instances, but the FN count of 141 suggests some instances were missed.

KNN:

$$\begin{bmatrix} 2123 & 378 \\ 99 & 15805 \end{bmatrix}$$

Moving on to KNN, we observe a higher TP count of 2,123, indicating more accurately predicted positive instances. The TN count of 15,805 is also high, indicating accurate negative predictions. However, the FP count of 378 and FN count of 99 suggest some misclassifications. KNN demonstrates reasonably good performance.

DT:

$$\begin{bmatrix} 2023 & 478 \\ 440 & 15464 \end{bmatrix}$$

The DT model shows a similar TP count as LR but with a lower TN count. The FP count of 478 indicates a higher number of false positive predictions compared to LR. Additionally, the FN count of 440 suggests some instances were incorrectly classified as negative.

RF:

$$\begin{bmatrix} 2066 & 435 \\ 30 & 15874 \end{bmatrix}$$

In the case of RF, we observe a higher TP count of 2,066, indicating a strong ability to correctly predict positive instances. The TN count of 15,874 is quite high, suggesting accurate negative predictions. Moreover, the low FP count of 435 and extremely low FN count of 30 indicate fewer misclassifications.

XGB:

$$\begin{bmatrix} 2195 & 306 \\ 92 & 15812 \end{bmatrix}$$

Next, the XGB model exhibits a higher TP count of 2,195, indicating accurate positive predictions. The TN count of 15,812 is also high, suggesting accurate negative predictions. The FP count of 306 and FN count of 92 are relatively low, indicating fewer misclassifications.

Naive Bayes:

$$\begin{bmatrix} 652 & 1849 \\ 1661 & 14243 \end{bmatrix}$$

Naive Bayes, on the other hand, shows a low TP count of 652 and a high FN count of 1,661, indicating a substantial number of misclassifications for positive instances. The TN count of 14,243 indicates accurate negative predictions. However, the high FP count of 1,849 suggests a significant number of false positive predictions.

SVC:

$$\begin{bmatrix} 1201 & 1300 \\ 56 & 15848 \end{bmatrix}$$

Moving to the SVC, we observe a lower TP count of 1,201 and a relatively high FN count of 56, indicating some misclassifications of positive instances. The TN count of 15,848 suggests accurate negative predictions. However, the FP count of 1,300 is relatively high, indicating a higher number of false positive predictions.

DNN:

$$\begin{bmatrix} 2333 & 168 \\ 107 & 15797 \end{bmatrix}$$

Finally, the DNN model shows a higher TP count of 2,333, indicating accurate positive predictions. The TN count of 15,797 is also high, suggesting accurate negative predictions. Moreover, the low FP count of 168 and FN count of 107 indicate a lower number of misclassifications. DNN, being a deep learning model, demonstrates strong performance due to its ability to capture complex patterns and relationships in the data.

In conclusion, when considering the accuracy, precision, and recall, models like Random Forest, XGBoost, and Deep Neural Network perform well with high TP and TN counts and lower FP and FN counts. Logistic Regression, K-Nearest Neighbors, Decision Tree, and Support Vector Classifier show relatively mixed performance. Naive Bayes, however, displays suboptimal results.

Table 5.9: Comparison between different classifiers with a balanced dataset.

Model	**Training Accuracy (%)**	**Testing Accuracy (%)**	**Testing Score (%)**
LR	73.85	74.02	74.02
KNN	96.24	94.39	94.39
DT	100.0	90.22	90.22
RF	100.0	95.49	95.49
XGB	98.76	94.90	94.90
Naive Bayes	61.87	62.28	62.28
SVC	88.87	88.41	88.41
DNN	99.08	96.11	96.11

Based on these metrics, the DNN model achieves the highest training accuracy of 99.08% and testing accuracy of 96.12%. It also has high recall, precision, and F1 scores, all of which are at 96.12%. These values indicate that the DNN model performs very well on the given classification task, showing a high level of accuracy in both training and testing data.

Among the other models, the Decision Tree and Random Forest models also exhibit high training accuracy, but their testing accuracies are slightly lower. The K-Nearest Neighbors and XGBoost models also perform well with high accuracy and balanced recall, precision, and F1 scores.

On the other hand, the Logistic Regression, Naive Bayes, and Support Vector Classifier models have lower accuracies and performance scores, indicating that they might not be as effective for this particular classification task.

Confusion Matrices:

LR:

$$\begin{bmatrix} 4305 & 1997 \\ 1270 & 5004 \end{bmatrix}$$

The LR model achieved a TP count of 4,305, indicating accurate positive predictions. The TN count of 5,004 suggests accurate negative predictions. However, the FP count of 1,997 indicates a relatively high number of false positive predictions. The FN count of 1,270 suggests some instances were incorrectly classified as negative.

KNN:

$$\begin{bmatrix} 5833 & 469 \\ 237 & 6037 \end{bmatrix}$$

The KNN model achieved a TP count of 5,833, indicating accurate positive predictions. The TN count of 6,037 suggests accurate negative predictions. The FP count of 469 indicates a relatively low number of false positive predictions. The FN count of 237 suggests some instances were incorrectly classified as negative.

DT:

$$\begin{bmatrix} 5659 & 643 \\ 587 & 5687 \end{bmatrix}$$

DT achieved a TP count of 5,659, indicating accurate positive predictions. The TN count of 5,687 suggests accurate negative predictions. However, the FP count of 643 indicates a higher number of false positive predictions compared to LR and KNN. The FN count of 587 suggests some instances were incorrectly classified as negative.

RF:

$$\begin{bmatrix} 5941 & 361 \\ 206 & 6068 \end{bmatrix}$$

The RF model achieved a TP count of 5,941, indicating accurate positive predictions. The TN count of 6,068 suggests accurate negative predictions. The FP count of 361 indicates a relatively low number of false positive predictions. The FN count of 206 suggests some instances were incorrectly classified as negative.

XGB:

$$\begin{bmatrix} 5917 & 385 \\ 256 & 6018 \end{bmatrix}$$

The XGBoost (Extreme Gradient Boosting) model achieved a TP count of 5,917, indicating accurate positive predictions. The TN count of 6,018 suggests accurate negative predictions. The FP count of 385 indicates a relatively low number of false positive predictions. The FN count of 256 suggests some instances were incorrectly classified as negative.

Naive Bayes:

$$\begin{bmatrix} 2618 & 3684 \\ 1060 & 5214 \end{bmatrix}$$

Naive Bayes achieved a TP count of 2,618, indicating accurate positive predictions. The TN count of 5,214 suggests accurate negative predictions. However, the FP count of 3,684 indicates a higher number of false positive predictions. The FN count of 1,060 suggests some instances were incorrectly classified as negative.

SVC:

$$\begin{bmatrix} 5271 & 1031 \\ 427 & 5847 \end{bmatrix}$$

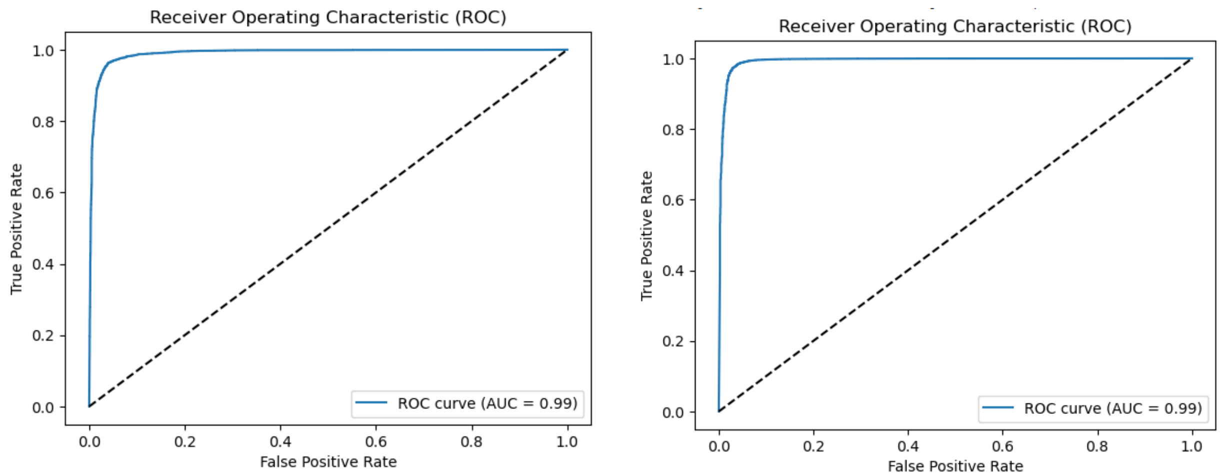
The SVC achieved a TP count of 5,271, indicating accurate positive predictions. The TN count of 5,847 suggests accurate negative predictions. However, the FP count of 1,031 indicates a higher number of false positive predictions. The FN count of 427 suggests some instances were incorrectly classified as negative.

DNN:

$$\begin{bmatrix} 6023 & 266 \\ 222 & 6065 \end{bmatrix}$$

The DNN model achieved a TP count of 6,023, indicating accurate positive predictions. The TN count of 6,065 suggests accurate negative predictions. The FP count of 266 indicates a relatively low number of false positive predictions. The FN count of 222 suggests some instances were incorrectly classified as negative.

In conclusion, models like K-Nearest Neighbors, Random Forest, XGBoost, and Deep Neural Network exhibit good overall performance with higher TP and TN counts and lower FP and FN counts. Logistic Regression, Decision Tree, Naive Bayes, and Support Vector Classifier show relatively mixed performance.



(a) with balanced data

(b) without balanced data.

Figure 5.58: The ROC curve of our model with balanced data and without balanced data.

Our model's ROC curve as shown in Figure 5.58, which represents the performance of our classifier, demonstrates exceptional results with an AUC of 0.99. This high AUC value indicates that our model has a remarkable ability to discriminate between positive and negative instances, regardless of the data balance. Whether the data is balanced or unbalanced, our model consistently achieves a high true positive rate while maintaining a low false positive rate. This robustness suggests that our model's performance is not significantly affected by the distribution of positive and negative instances in the dataset. Overall, our model's consistent and reliable performance with an AUC of 0.99 underscores its effectiveness in accurately predicting and classifying instances.

In summary, this section highlights the outcomes achieved through different models in our heart disease protection project. We explore the Heart Disease Detection model, the prediction of the next ECG sequence based on past data, and the classification of ECG sequences into normal and abnormal patterns. These results contribute to the advancement of cardiovascular health and enhance patient care by facilitating early detection and intervention.

Comparison with related Work

we thoroughly examine the analysis of findings derived from various articles predicting heart disease. The results are presented and discussed in detail in Table 5.10

Table 5.10: Comparison with Related work on prediction of heart disease.

Authors	Vivekanandan,T., and [108]	Mohan,S.et al. [109]	Mohammad et al [110]	Our work
Approach	Modifed DE with Fuzzy AHP FFNN	Hybrid random forest with a linear model	Deep convolutional neural networks	DNN and Seq2seq
Use of IoT Device	No	No	No	Yes
Evaluation	83% (Accuracy)	88.7% (Accuracy)	93.4% (Accuracy)	91%, 98% (Accuracy), 0.1 (MSE) 0.3 (RMSE) and 0.1 (MAE)
Identify individuals with heart disease	Yes	Yes	No	Yes
prediction of sequence ECG	No	No	No	Yes
classification of sequence ECG	No	No	Yes	Yes

5.4.2 Retinal damage classification

The model utilized in the project to detect retinal damage underwent several architectural variations, allowing us to compare different approaches found in the existing literature and enhance our modeling process. The results of these diverse architectures are elaborated in the following table 5.11.

Table 5.11: Results of Proposed Architectures for Retinal Damage Classification.

Architecture	**Train Accuracy**	**Validation Accuracy**
Custom CNN Model	92.56%	96.88%
Custom VGG16 Model	98.02%	100%
Custom VGG19 Model	98.00%	100%
Custom InceptionV3 Model	97.14%	100%
Custom DenseNet Model	95.39%	92.01%

Initially, we developed a custom CNN model from scratch, but it did not produce satisfactory results. The model achieved a train accuracy of 92.56% and a validation accuracy of 96.88%. To improve the performance, we explored transfer learning by utilizing pre-trained architectures such as VGG16, VGG19, InceptionV3, and DenseNet.

When applying VGG16, VGG19, and InceptionV3 architectures pre-trained on ImageNet, we observed an increase in accuracy. The train accuracies achieved were 98%, 98.02%, and 97.14% respectively, with a perfect validation accuracy of 100%. However, there was a concern that these models might be overfitting the training data since the validation accuracy exceeded the train accuracy. Overfitting occurs when the model becomes too focused on the specific details of the training data, hindering its ability to generalize to new data.

To address this issue, we specifically utilized the DenseNet architecture. Experimental results demonstrated that the proposed DenseNet model outperformed the previous architectures. It achieved a train accuracy of 95.39%, and a validation accuracy of 92.01%.

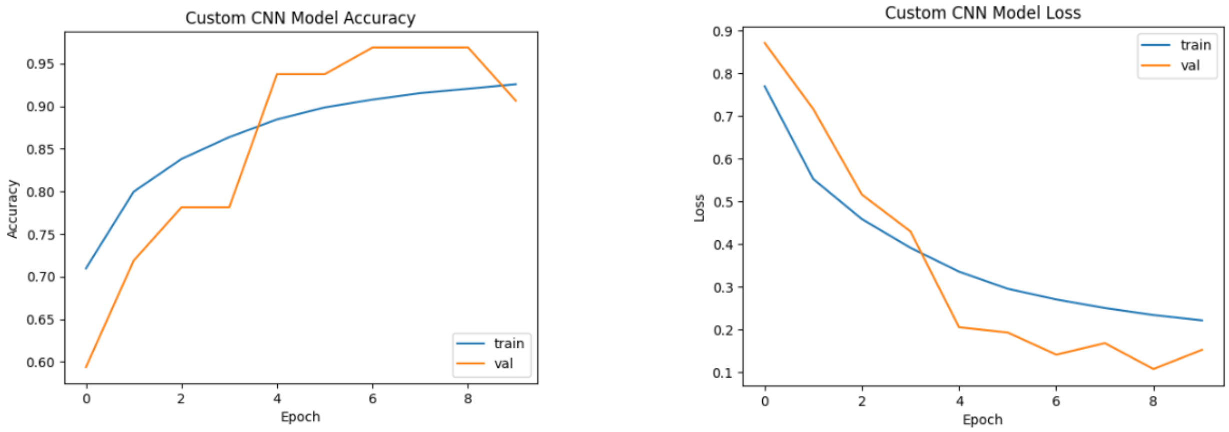
In our pursuit of achieving our desired goal, we employed various architectures (see Table 5.11) and techniques, each playing a crucial role in our approach. Below, we outline the key structures utilized, highlight their significant components, and present the outcomes achieved.

Custom CNN model: In our initial endeavor to classify OCT images, our approach involved creating and training a model from scratch. To facilitate this process, we splited the dataset into three distinct sets: 83,484 images were allocated for training, 968 images for validation, and 32 images for testing. Additionally, we resized all the images to a standardized dimension of 224x224x3.

The architecture of the model comprised of four convolutional layers, which were followed by max pooling and dropout layers with a dropout rate of 0.25. These layers were designed to extract relevant features from the input images and mitigate the risk of overfitting. Subsequently, the model included a flatten layer to convert the output of the convolutional layers into a 1D feature vector. This was followed by a fully connected layer, another dropout regularization layer, and an output layer with softmax activation. The softmax activation was utilized for multiclass classification, allowing the model to assign probabilities to each class.

The initial results of the model were disappointing, as it exhibited a higher misclassification rate during the prediction phase, despite being trained on a large dataset. The training and validation accuracy rates were approximately 92.56% and 96.88% respectively, while the training and validation loss rates were around 22.11% and 10.71% respectively (refer to figure 5.59). These findings suggest that the model struggled to accurately classify images, despite achieving relatively high accuracy rates during training and validation.

Upon closer examination, it became evident that the model made a significant number of errors. These errors were primarily due to the similarities between certain classes. Specifically, the model misclassified CNV as NORMAL 66 times, misclassified DME as NORMAL 69 times, and misclassified DRUSEN as CNV 82 times. These observations highlight the challenges posed by the similarities between these classes (refer to figure 5.60).



(a) The accuracy curve.

(b) The loss curve.

Figure 5.59: The training and validation accuracy and loss curves for the custom CNN model.

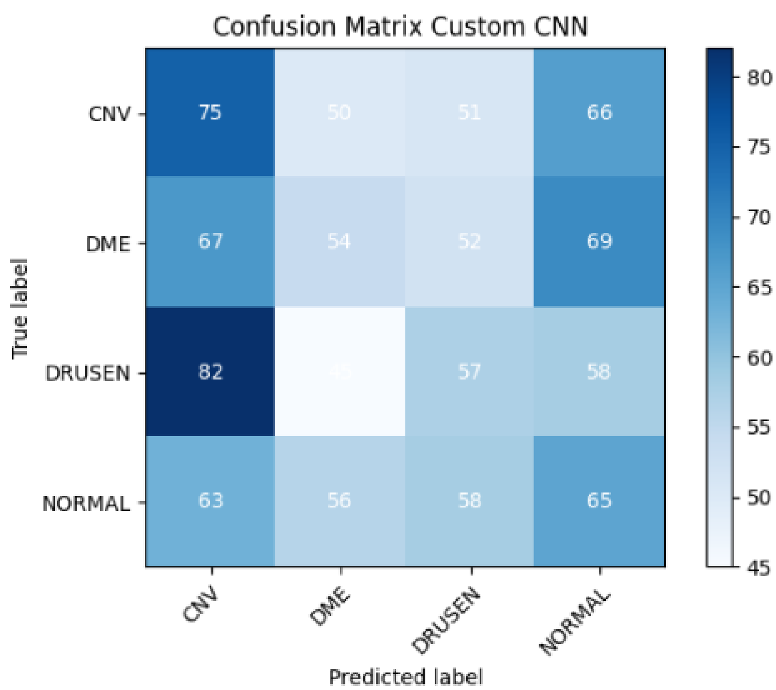
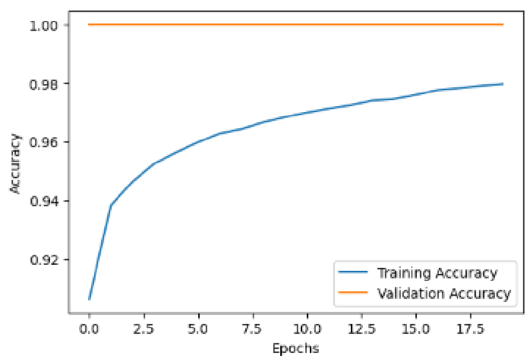


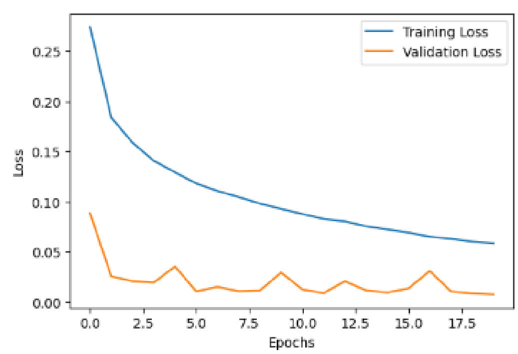
Figure 5.60: Confusion matrix obtained using Custom CNN model.

Looking at the model from a different perspective, we observed that the learning process was time-consuming, indicated by a large number of trainable parameters (6,553,028) and a relatively high error rate. Consequently, we embarked on exploring alternative methods to improve the learning process. One promising approach we considered was harnessing the power of transfer learning from top pre-trained models.

Custom VGG16 model: We implemented a VGG16 CNN architecture with a customized classifier for our model. The classifier included a dropout layer with a rate of 0.2, which helped accelerate the training process by reducing the network's sensitivity to the order of training data. Additionally, we standardized the input images by resizing them to dimensions of 150x150x3, ensuring uniformity in the data. We maintained the same data splitting configuration as the previous architecture. Moreover, we performed two rounds of training: the first to obtain the best weights and the second to fine-tune the model's learning rate. Furthermore, our classifier consisted of a fully connected dense layer with four neurons, activated using the softmax function. By making these modifications, we successfully addressed the issue of training time and achieved a validation accuracy of 100% after training the model. The current model exhibits signs of overfitting the training data, indicating that it has learned the specific details of the training data too extensively. The results are illustrated in the Figure 5.61 .



(a) The accuracy curve.

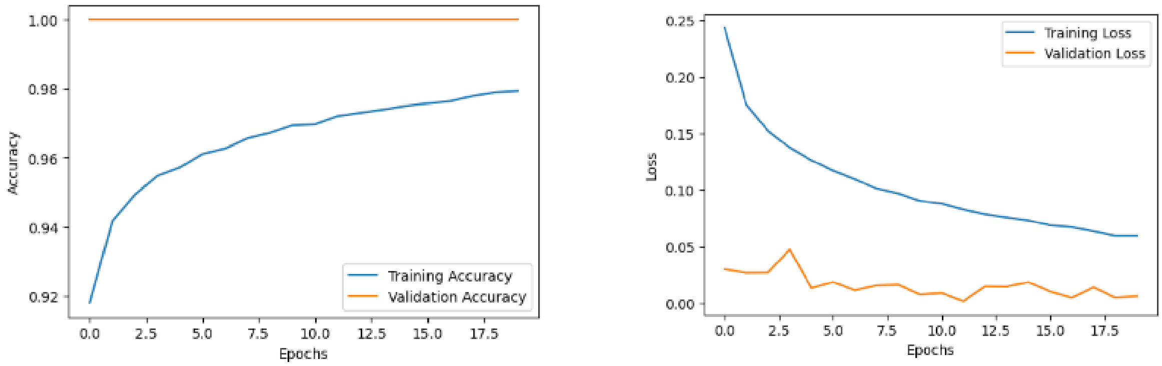


(b) The loss curve.

Figure 5.61: The training and validation accuracy and loss curves for the custom VGG16 model.

As a consequence, the model struggles to generalize well when exposed to new, unseen data. To address this issue, we will explore the utilization of alternative architectures in our model. By experimenting with different architectures, we aim to find a better balance between capturing intricate details in the training data and achieving good generalization performance on unseen data.

Custom VGG19 Model: The enhanced model, featuring three extra convolution layers compared to the previous architecture while keeping the same conditions, shows comparable performance to the VGG16 model. However, it does not effectively tackle the problem of overfitting. Despite the improvements in architecture, the model continues to demonstrate overfitting tendencies, suggesting that solely increasing the number of convolution layers does not address the core issue. The results of the model are illustrated in Figure (5.62).

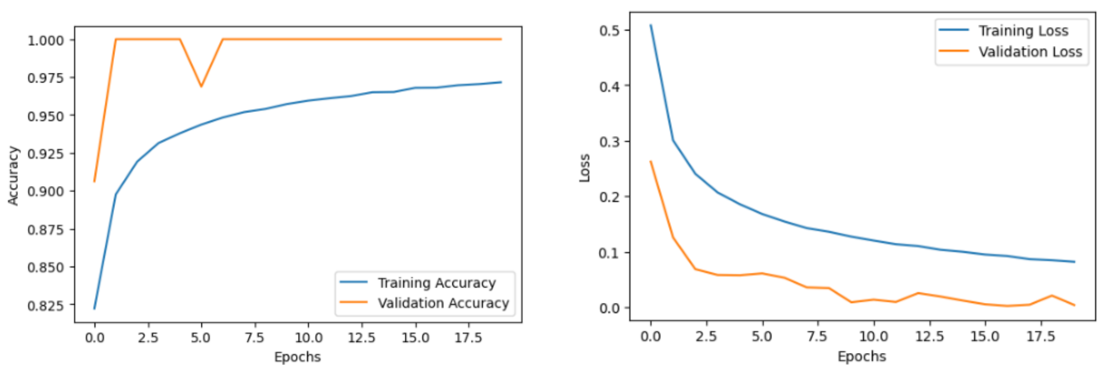


(a) The accuracy curve.

(b) The loss curve.

Figure 5.62: The training and validation accuracy and loss curves for the custom VGG19 model.

Custom InceptionV3 model: The model, incorporating additional types of convolution layers compared to the previous architectures (VGG16 and VGG19), while maintaining the same conditions, did not yield improved performance and failed to effectively address the overfitting problem. The results of this model can be seen in Figure 5.63.



(a) The accuracy curve.

(b) The loss curve.

Figure 5.63: The training and validation accuracy and loss curves for the custom InceptionV3 model.

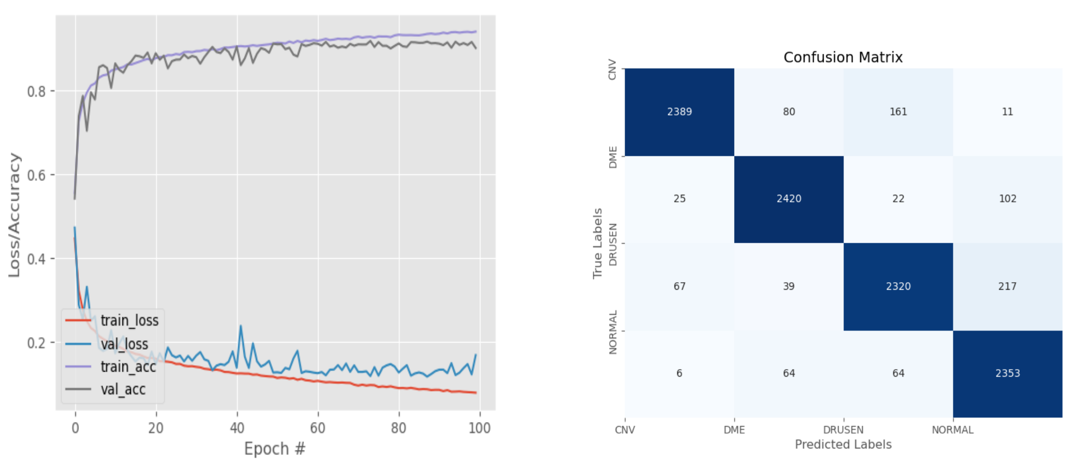
To improve the reliability and performance of our model, we implemented several modifications to the data preprocessing strategy. Initially, the results were suboptimal, leading us to reformat the data. This involved expanding the validation dataset and reorganizing the spliting of images across files to achieve unbiased predictions.

Additionally, we acknowledged the influence of image type preservation on accuracy. Since transfer learning architectures typically operate on 3-channel images, while our images were grayscale.

Custom DenseNet model:

The proposed model, utilizing DenseNet architecture, has proven to be effective in improving the quality of grayscale images with an input dimension of 64x64x1. This model outperformed all previous models, achieving a remarkable 95.39% training accuracy and 92.01% validation accuracy, along with a low training loss of 7.01% and a validation loss of 13%. These exceptional results were obtained through various techniques, including data balancing and splitting. Specifically, the validation dataset was augmented from 968 to 10,340 images, while 24,124 images were allocated for training, leaving the remaining data for testing purposes.

By implementing these new techniques, we not only achieved significant accuracy improvements but also successfully addressed the issue of overfitting. The details of the preprocessing methods utilized can be found in the previous chapter, starting from page 68. For a visual representation of the results, refer to Figure 5.64.



(a) The accuracy and loss curves.

(b) Confusion Matrix.

Figure 5.64: Obtained results for custom DenseNet model.

In order to further evaluate the performance of our model, we computed the precision, F1-score, and recall. The results of these metrics are presented in Table 5.12 and Table 5.13.

Table 5.12: Classification Report for DenseNet Model.

Label	Precision	Recall	F1-Score	Support
CNV	0.96	0.90	0.93	2641
DME	0.93	0.94	0.94	2569
DRUSEN	0.90	0.88	0.89	2643
NORMAL	0.88	0.95	0.91	2487
Accuracy			0.92	10340
Macro Avg	0.92	0.92	0.92	10340
Weighted Avg	0.92	0.92	0.92	10340

Table 5.13: Performance Metrics for DenseNet Model.

Precision	0.916207267200758
Recall	0.9115755398468184
F1-Score	0.9114603119078072

Comparison of our work with previous work:

This section is dedicated to conducting a comprehensive comparison between our results and those achieved in prior studies utilizing the same dataset, albeit with different data organization. The objective is to assess the performance of our model in comparison to other deep learning-based approaches discussed in relevant literature. Overall, our approach exhibits promising results and performs favorably when compared to similar work (refer to Table 5.14).

Table 5.14: Table of Model Comparison.

Work	Method	Train Accuracy	Validation Accuracy	Precision	Recall	F1-Score
[169]	ResNet50	0.88	0.96	0.91	0.89	0.89
[170]	InceptionV3	0.92	0.88	-	-	-
[169]	DenseNet169	0.89	0.90	0.90	0.88	0.88
Our Work	DenseNet (our work)	0.95	0.92	0.91	0.91	0.91

5.4.3 Emotion recognition using vital signals

In our work, our main objective was to improve the accuracy of our emotion recognition system by enhancing the dataset size. To achieve this, we adopted a two-fold approach. Firstly, we focused on data augmentation techniques to increase the number of samples in the DREAMER dataset. Additionally, we combined the augmented DREAMER dataset with two other datasets, namely SWELL and WESAD.

Below, we will discuss the results obtained from our efforts to enhance the dataset size and improve the accuracy of our emotion recognition system.

DREAMER Dataset augmentation: To assess the effectiveness of the proposed ECG data augmentation method in emotion detection using the DREAMER dataset, we conducted comparative experiments. These experiments involved different classifiers and compared their performance with and without data augmentation. Our primary focus was on valence, as it serves as the scale for detecting positive and negative emotions, as mentioned in Chapter 4 of the study.

In our comparative analysis, we will employ various widely-used classifiers that have demonstrated strong performance in emotion recognition systems. The initial classifier we will utilize is the SVM, which has been a popular choice in this field [171]. Another high-performing classifier is the KNN, which has been shown to outperform SVM in specific studies [172][173]. Additionally, we will consider the RF classifier, known for its promising outcomes in previous research [174]. Furthermore, we will explore the effectiveness of neural network-based deep learning classifiers, specifically using the multilayer perceptron (MLP) architecture.

Figure 5.65 presents the accuracy comparison of valence dimension between SVM, RF, KNN, and NN. Without data augmentation, the classifiers achieved the following classification accuracies on the DREAMER dataset:

- SVM classifier: 56.80%
- RF classifier: 48.80%
- KNN classifier: 48%
- NN classifier: 52%

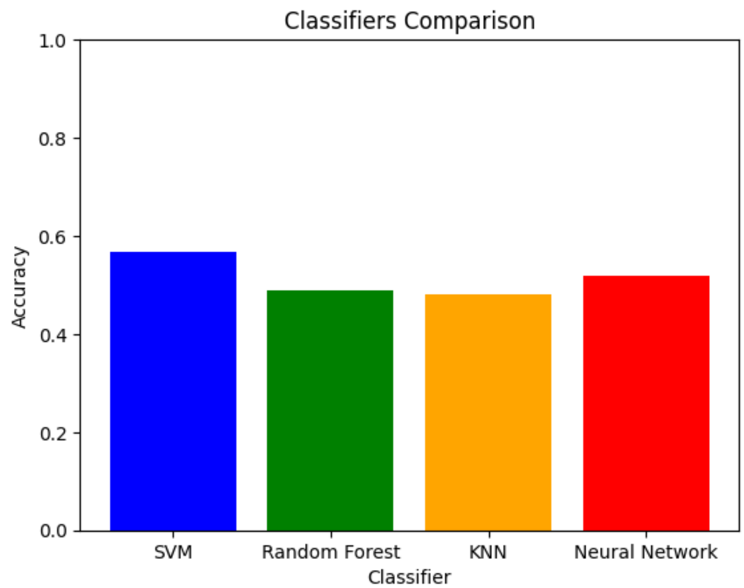


Figure 5.65: Accuracy comparison between different classifiers without data augmentation (emotion expressed in terms of valence)in DREAMER dataset.

With data augmentation, the results showed improved classification accuracies for all classifiers. Figure 5.66 specifically demonstrated that the SVM classifier achieved the highest accuracy of 61.93% in detecting the valence dimension. The NN classifier also performed well with an accuracy of 63.32%. In comparison, the RF classifier achieved an accuracy of 56.35% and the KNN classifier achieved an accuracy of 59.64%.

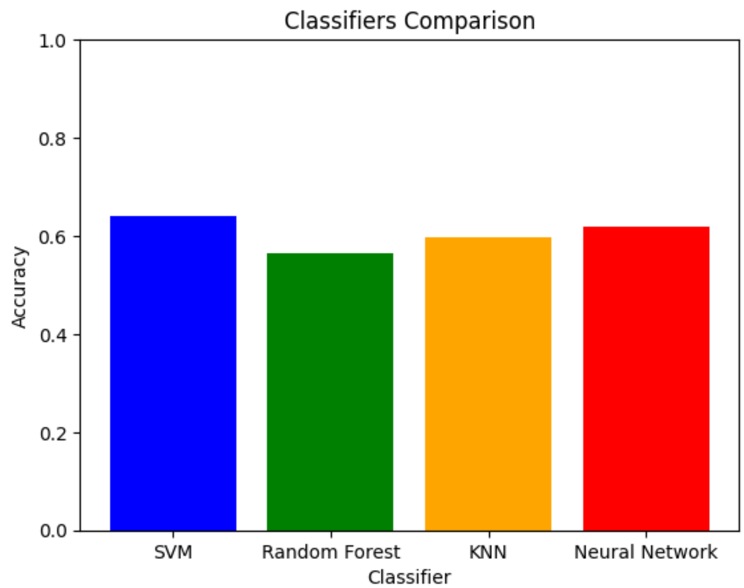


Figure 5.66: Accuracy comparison between different classifiers with data augmentation in DREAMER dataset.

It is worth noting that the accuracies obtained with data augmentation were higher than those without augmentation.

In addition, we will discuss the results achieved by augmenting the DREAMER dataset and merging it with two additional datasets: SWELL and WESAD.

Combining DREAMER with SWELL and WESAD: We understand that the problem of a small dataset in emotion recognition presents significant challenges. When we have limited data, there is a higher risk of overfitting, where our model becomes too specific to the training examples and struggles to generalize effectively to new, unseen data. Furthermore, the unavailability of the dataset to everyone restricts our ability to gather more samples, exacerbating the issue.

However, we can address these limitations by combining the datasets that are available to us. This approach offers several potential benefits in terms of increasing the number of samples and improving accuracy in emotion recognition.

Figure 5.67 illustrates the accuracy results obtained from the combined dataset using various classifiers.

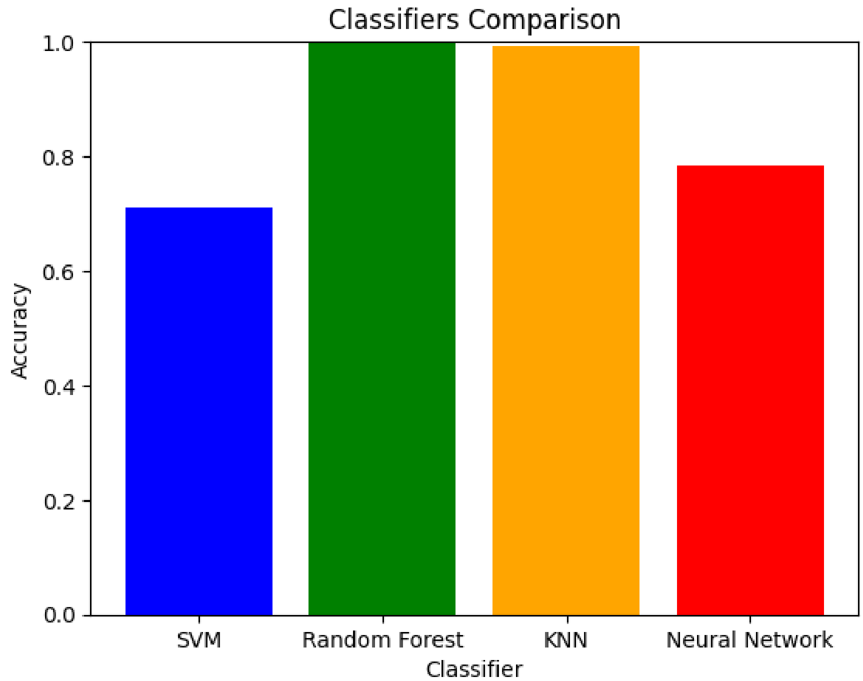


Figure 5.67: Accuracy of the combined dataset using different classifiers.

In the accuracy comparison of different classifiers on the combined dataset, we obtained the following results:

- SVM: 70.96%
- RF: 99.89%
- KNN: 99.30%
- NN: 78.38%

Figure 5.68 presents the confusion matrix, providing a visual representation of the classification results for different emotion categories. In this specific case, the confusion matrix is based on a binary classification task with two classes: 0 representing negative emotions and 1 representing positive emotions. The matrix highlights the accuracy and effectiveness of the emotion classification system in correctly identifying and distinguishing between these two categories within the combined dataset.

The confusion matrix shows that out of the total 160,116 samples, the model accurately predicted 92,372 instances of positive emotions (class 1) and 67,565 instances of negative emotions (class 0). However, there were 104 instances of negative emotions misclassified as positive (false positives) and 75 instances of positive emotions misclassified as negative (false negative). These misclassifications indicate areas where the model can be further improved to enhance its accuracy in predicting emotions.

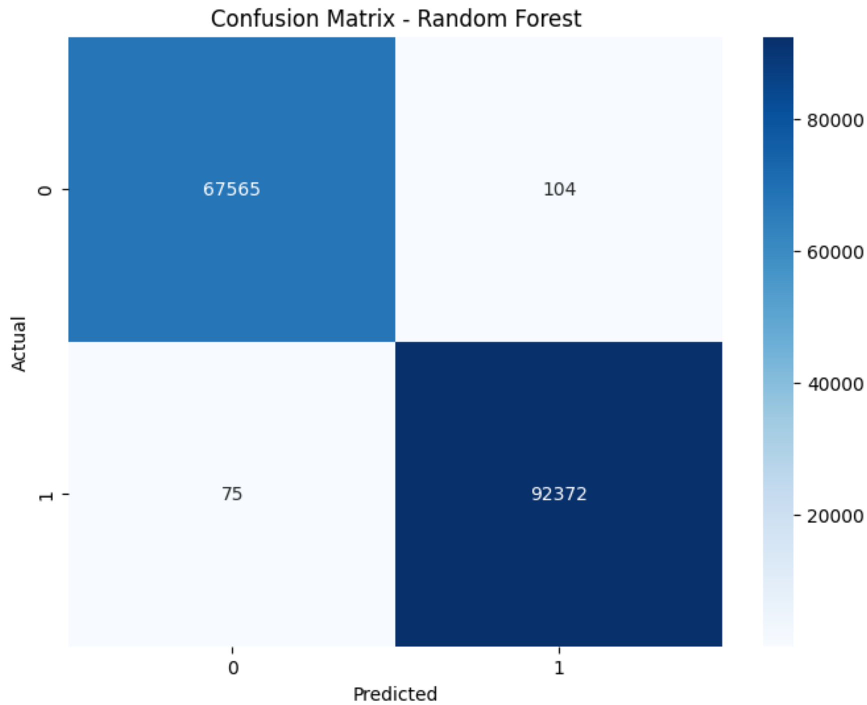


Figure 5.68: Confusion Matrix for Emotion Classification on Combined Dataset.

Comparison with related Work

In table 5.15, we delve into the analysis of results obtained from a range of articles that have explored emotion recognition using the DREAMER dataset. These articles have utilized a variety of methodologies and techniques in order to address the task of accurately identifying emotions from the dataset.

- In [21], the researchers present the DREAMER database and utilize an SVM classifier to classify emotions based on valence, arousal, and dominance levels.
- In [175], the study introduces a self-supervised deep multi-task learning framework for emotion recognition based on electrocardiogram (ECG) data. The framework involves two stages of learning: first, learning ECG representations through a signal transformation

recognition network, and second, transferring the learned weights to an emotion recognition network for classifying emotions.

- In [176], the study proposes enriching the ECG dataset with representative samples and using a CNN classifier to detect human emotions based on valence, arousal, and dominance levels.

Refernces	**Method**	**Accuracy**
[21]	SVM	0.624
[175]	Self-Supervised CNN	0.859
[176]	data augmentation CNN	0.951
Our work	Random forest	0.998

Table 5.15: Comparison with Existing Literature using the DREAMER dataset.

5.5 Challenges encountered

Developing a healthcare platform is a complex undertaking that involves various challenges and obstacles. In this section, we will discuss the challenges encountered during the creation of our platform and the strategies employed to overcome them.

- According to [176], successful ECG analysis using machine learning approaches relies on access to a well-annotated dataset with comprehensive information. However, obtaining ECG data is challenging due to limited availability. Despite our efforts to request data access, only the DREAMER dataset granted us permission to use their data. To enhance the diversity of emotions in our analysis and improve the accuracy of our eHealth platform, we employed a data augmentation technique on the DREAMER dataset. Additionally, we combined the DREAMER dataset, which encompasses nine distinct emotions, with the SWELL and WESAD datasets, which primarily focus on stress detection. By integrating these datasets, we expanded the range of emotions available for training our model and achieved improved results in our analysis.

- We encountered challenges in saving machine learning models like Random Forest as a TFLite file for deployment in our eHealth platform developed with Flutter. To overcome this limitation, we explored alternative options and decided to use ONNX (Open Neural Network Exchange) to save our model. ONNX provides a flexible format for exporting and importing models across different frameworks. With ONNX, we were able to save our model and proceed with developing the Android application using Kotlin, which has good compatibility with ONNX. This allowed us to integrate machine learning classifiers effectively within our eHealth platform.

5.6 Conclusion

In this chapter, we have provided an overview of the libraries, frameworks, tools, sensors, and development boards that have been essential in our domain of study. We have discussed their functionalities and importance in carrying out our project.

Furthermore, we have presented the obtained results from our experiments and have conducted various comparisons to evaluate the performance of different approaches and classifiers. We have discussed the outcomes of each specific problematic and have analyzed the results to gain insights into the effectiveness of our methods.

Chapter 6

Conclusion and perspectives

In the context of a healthcare platform that utilizes deep learning and machine learning for classification tasks such as retinal damage, emotion recognition, prediction of heart disease, detection of heart disease, and classification of sequence ECG based on IoT technology, this work serves as a foundation for future advancements.

Firstly, the prediction and detection of heart disease are critical for proactive healthcare. By leveraging machine learning algorithms and analyzing various risk factors and physiological indicators integrating real-time monitoring capabilities. Future work can focus on improving the accuracy and reliability of heart disease prediction models and exploring new risk factors.

The prediction and classification of sequence ECG based on IoT technology is another notable feature of this platform. By analyzing ECG signals in real-time, the platform can accurately detect abnormal cardiac rhythms, allowing for timely intervention and potentially saving lives. Future work can concentrate on refining the ECG classification algorithms and exploring novel features for improved detection accuracy.

Emotion recognition is another valuable service offered by the platform. By analyzing physiological signals like ECG, the platform can assess an individual's emotional state, p Future work can focus on refining the emotion recognition algorithms, integrating additional sensors to capture more comprehensive data, and expanding the range of emotional states that can be recognized.

Furthermore, the classification of retinal damage plays a crucial role in the early detection and treatment of eye-related disorders. By leveraging advanced image processing techniques, the platform can accurately identify and classify retinal damage, enabling timely interventions and preventing further complications. This work sets the foundation for further enhancements in this area, such as improving the accuracy of retinal damage classification algorithms and expanding the platform's capabilities to cover other ocular conditions.

From a technical engineering perspective, this work lays the groundwork for the development of IoT-based healthcare platforms. It paves the way for future advancements in the field, including the development of IoT devices and sensors for detecting various health conditions and states. Additionally, the creation of a comprehensive and diverse dataset is essential for training and improving the platform's algorithms. Future work can focus on expanding the dataset, ensuring its representativeness, and making it available to the research community for further innovation and collaboration.

In conclusion, the use of deep learning and machine learning for classification tasks in a healthcare platform holds great promise. By leveraging these technologies, the platform can improve the accuracy and efficiency of tasks such as retinal damage classification, emotion recognition, heart disease prediction and detection, and sequence ECG analysis. Future work should focus on refining models, incorporating new techniques, developing IoT devices, and creating comprehensive datasets to drive advancements in the field and provide valuable healthcare services in a unified platform.

References

[1] Andrzej Szczepański and Khalid Saeed. A Mobile Device System for Early Warning of ECG Anomalies. *Sensors*, 14(6):11031–11044, 6 2014.

[2] Optical Coherence Tomography | Medical Eye Centre. Accessed May 30, 2023 https://medicaleyecentre.com/services/optical-coherence-tomography-oct/.

[3] Bikash K. Pradhan, Saugat Bhattacharyya, and Kunal Pal. Iot-based applications in healthcare devices. *Journal of Healthcare Engineering*, 2021, 2021.

[4] Maroua AHMID. *An Agent-Based Approach for the Internet of Things*. PhD thesis, Université de mohamed kheider biskra, 2021.

[5] Formalizing electrocardiogram (ECG) signal behavior in event-B - Scientific. Accessed May 30, 2023 https://www.researchgate.net/figure/ECG-Waves-and-Their-Relation-to-Heart-Nodes_fig1_282920335, 2023.

[6] Retinal diseases - Symptoms and causes - Mayo Clinic. Accessed May 30, 2023 https://www.mayoclinic.org/diseases-conditions/retinal-diseases/symptoms-causes/syc-20355825, 1 2022.

[7] Jerzy Z Nowak and Acta Pol Pharm. Amd–the retinal disease with an unprecised etiopathogenesis: in search of effective therapeutics. *Aging*, 4(6), 2014.

[8] What is Choroidal Neovascularization (CNV)? | BrightFocus Foundation. Accessed May 30, 2023 https://www.brightfocus.org/macular/article/what-choroidal-neovascularization, 7 2021.

[9] Macular Edema Vision. Accessed May 30, 2023 https://animalia-life.club/qa/pictures/macular-edema-vision.

[10] Lin Shu, Jinyan Xie, Mingyue Yang, Ming Zhang, Zhenqi Li, Dan Liao, Xiangmin Xu, and Xinyi Yang. A Review of Emotion Recognition Using Physiological Signals. *Sensors*, 18(7):2074, 6 2018.

[11] Wei Wang, Jinge Tian, Chengwen Zhang, Yanhong Luo, Xin Wang, and Ji Li. An improved deep learning approach and its applications on colonic polyp images detection. *BMC Medical Imaging*, 20(1), 7 2020.

[12] Neural network deep learning. https://www.lexrosere.top/products.aspx?cname=neural+network+deep+learning&cid=95.

[13] Riad Sebti, Siham Zroug, Laid Kahloul, and Saber Benharzallah. A deep learning approach for the diabetic retinopathy detection. In *Innovations in Smart Cities Applications Volume 5: The Proceedings of the 6th International Conference on Smart City Applications*, pages 459–469. Springer, 2022.

[14] Any use of non-rectangular-shaped kernels in convolutional neural networks? Especially when analyzing game boards. https://stats.stackexchange.com/questions/235032/any-use-of-non-rectangular-shaped-kernels-in-convolutional-neural-networks-espe, March 5 2023.

[15] Understanding of a convolutional neural network. Accessed May 30, 2023 https://ieeexplore.ieee.org/document/8308186,, 8 2017.

[16] Hao Gao. The Efficiency of Densenet - Hao Gao - Medium. 6 2018.

[17] Kang Atul. Densely Connected Convolutional Networks – DenseNet, 12 2018.

[18] Masoud Mahdianpari, Bahram Salehi, Mohammad Ahangarzadeh Rezaee, Fariba Mohammadimanesh, and Yun Zhang. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. *Remote Sensing*, 10(7):1119, 7 2018.

[19] Adib Habbou. Qursquo;est ce qursquo;un réseau lstm ? *La revue IA*, 9 2022.

[20] A simple introduction to sequence to sequence models. Accessed May 15, 2023 https://www.analyticsvidhya.com/blog/2020/08/a-simple-introduction-to-sequence-to-sequence-models/, November 2020.

[21] Stamos Katsigiannis and Naeem Ramzan. DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices. *IEEE Journal of Biomedical and Health Informatics*, 22(1):98–107, 1 2018.

[22] WHOOP. What is Heart Rate Variability (HRV)? [The Ultimate HRV Guide], 3 2023.

[23] Raspberry Pi Ltd. Buy a Raspberry Pi 3 Model B+ – Raspberry Pi. Accessed May 15, 2023 https://www.raspberrypi.com/products/raspberry-pi-3-model-b-plus/.

[24] Don Eliseo III Lucero-Prisno, MBN Kouwenhoven, Yusuff Adebayo Adebisi, Adriana Viola Miranda, Dawa Gyeltshen, Mohamed Hoosen Suleman, Isabel Kazanga Chiumia, Mat Lowe, Thinley Dorji, Junjie Huang, et al. Top ten public health challenges to track in 2022. *Public Health Challenges*, 1(3):e21, 2022.

[25] World Health Organization: WHO. Cardiovascular diseases. Accessed May 30, 2023 https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1, 6 2019.

[26] Margaret Chan. Global report on diabetes. *World Health Organization*, 58(12):1–88, 2016.

[27] Wan Ling Wong, Xinyi Su, Xiang Li, Chui Ming G Cheung, Ronald Klein, Ching-Yu Cheng, and Tien Yin Wong. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *The Lancet Global Health*, 2(2):e106–e116, 2014.

[28] Rohit Varma, Neil M. Bressler, Quan V. Doan, Michelle Gleeson, Mark D. Danese, Julie K. Bower, Elizabeth Selvin, Chantal M. Dolan, J. T. Fine, Shoshana Colman, and Adam Turpcu. Prevalence of and Risk Factors for Diabetic Macular Edema in the United States. *JAMA Ophthalmology*, 132(11):1334, 11 2014.

[29] M Poongodi, Ashutosh Sharma, Mounir Hamdi, Ma Maode, and Naveen Chilamkurti. Smart healthcare in smart cities: wireless patient monitoring system using iot. *The Journal of Supercomputing*, pages 1–26, 2021.

[30] Rachel E Young, David Broom, Karen Sage, Kay Crossland, and Christine Smith. Experiences of venue based exercise interventions for people with stroke in the uk: A systematic review and thematic synthesis of qualitative research. *Physiotherapy*, 110:5–14, 2021.

[31] Mei Chen and Michel Decary. Artificial intelligence in healthcare: An essential guide for health leaders. In *Healthcare management forum*, volume 33, pages 10–18. SAGE Publications Sage CA: Los Angeles, CA, 2020.

[32] Thomas H. Davenport and Ravi Kalakota. The potential for artificial intelligence in healthcare. *Future healthcare journal*, 6(2):94–98, 6 2019.

[33] Medical dictionary the free dictionary. Healthcare care definition. https://www.thefreedictionary.com/healthcare, 2023.

[34] Christian Nordqvist. Health care – definition and meaning. Accessed May 30, 2023 https://marketbusinessnews.com/financial-glossary/health-care/., 12 2021.

[35] USSR) International Conference on Primary Health Care (1978 : Alma-Ata. Primary health care : report of the International Conference on Primary Health Care, Alma-Ata, USSR, 6-12 September 1978 / jointly sponsored by the World Health Organization and the United Nations Children's Fund. Accessed May 30, 2023 https://apps.who.int/iris/handle/10665/39228, 6 2012.

[36] Fiona M Burns, Anne M Johnson, James Nazroo, Jonathan Ainsworth, Jane Anderson, Ade Fakoya, Ibidun Fakoya, Andy Hughes, Eva Jungmann, S Tariq Sadiq, et al. Missed opportunities for earlier hiv diagnosis within primary and secondary healthcare settings in the uk. *Aids*, 22(1):115–122, 2008.

[37] What is Tertiary health care - Meaning and definition - Pallipedia. Accessed May 30, 2023 https://pallipedia.org/tertiary-health-care/.

[38] Erin P. Balogh. The Diagnostic Process. Accessed May 30, 2023 https://www.ncbi.nlm.nih.gov/books/NBK338593/, 12 2015.

[39] Jerome P. Kassirer. Our Stubborn Quest for Diagnostic Certainty. *The New England Journal of Medicine*, 320(22):1489–1491, 6 1989.

[40] Laura Zwaan and Hardeep Singh. The challenges in defining and measuring diagnostic error. *Diagnosis*, 2(2):97–103, 6 2015.

[41] Harvey V. Fineberg. *Retooling for an Aging America*. National Academies Press, 8 2008.

[42] Surveillance for sensory impairment, activity limitation, and health-related quality of life among older adults–United States, 1993-1997. Accessed May 30, 2023 https://pubmed.ncbi.nlm.nih.gov/10634273/, 12 1999.

[43] GE HealthCare. Patient monitoring. https://www.gehealthcare.com/products/patient-monitoring. Accessed on May 30, 2023.

[44] REED M. GARDNER and M. MICHAEL SHABOT. Patient-monitoring systems.

[45] JR Wilkinson, S Walrond, LJ Ells, and CD Summerbell. Surveillance and monitoring. *Obesity reviews*, 8:23–29, 2007.

[46] National Academies Press (US). Surveillance and Monitoring. Accessed May 30, 2023 https://www.ncbi.nlm.nih.gov/books/NBK100180/, 2000.

[47] Ajitesh Kumar. Digital Healthcare Technology Innovations: Examples - Data Analytics. Accessed May 30, 2023 https://vitalflux.com/digital-healthcare-technology-innovations-examples/#Telemedicine_Telehealth, 6 2022.

[48] Shardul Bhatt. Top 12 Most Popular Healthcare Applications Examples. Accessed May 30, 2023 https://www.botreetechnologies.com/blog/most-popular-healthcare-applications-examples/, 3 2023.

[49] A Luthra. *The ECG Made Easy*. Jaypee Brothers, Medical Publisher, Delhi, India, 2012.

[50] Foteini Agrafioti, Dimitrios Hatzinakos, and Adam K. Anderson. ECG Pattern Analysis for Emotion Detection. *IEEE Transactions on Affective Computing*, 3(1):102–115, 1 2012.

[51] Ary Louis Goldberger. *Clinical Electrocardiography*. Elsevier Gezondheidszorg, Maarssen, Pays-Bas, 2006.

[52] Jwan Najeeb Saeed and Siddeeq Y. Ameen. Smart Healthcare for ECG Telemonitoring System. *JOURNAL OF SOFT COMPUTING AND DATA MINING*, 2(2), 10 2021.

[53] Aaron Frederick Bulagang, Ng Seik Weng, James Mountstephens, and Jason Teo. A review of recent approaches for emotion classification using electrocardiography and electrodermography signals. *Informatics in Medicine Unlocked*, 20:100363, 1 2020.

[54] Divya Sri Rajeswari, Hitha Shree J, Ananya L.N. Simha, Maria Nuzhath Subhani, Shivaleelavathi B G, and Veeramma Yatnalli. A review on remote health monitoring sensors and their filtering techniques. *Global transitions proceedings*, 2(2):392–401, 8 2021.

[55] Fizar Ahmed. An internet of things (iot) application for predicting the quantity of future heart attack patients. *International Journal of Computer Applications*, 164(6):36–40, 2017.

[56] Normalization.

[57] British Heart Foundation. Pulse rate. Accessed May 30, 2023 https://www.bhf.org.uk/informationsupport/heart-matters-magazine/medical/ask-the-experts/pulse-rate.

[58] James G. Fujimoto, Costas Pitris, Stephen A. Boppart, and Mark E. Brezinski. Optical Coherence Tomography: An Emerging Technology for Biomedical Imaging and Optical Biopsy. *Neoplasia*, 2(1-2):9–25, 1 2000.

[59] A Gh Podoleanu. Optical coherence tomography. *Journal of microscopy*, 247(3):209–219, 2012.

[60] Samuel BSc Mckenzie. What is Optical Coherence Tomography? Accessed May 30, 2023 https://www.news-medical.net/health/What-is-Optical-Coherence-Tomography.aspx, 10 2018.

[61] Alex DelVecchio. IoMT (Internet of Medical Things) or healthcare IoT, 6 2019.

[62] Mohamed Akram Khelili, Sihem Slatnia, Okba Kazar, Abdelhak Merizig, and Seyedali Mirjalili. Deep learning and metaheuristics application in internet of things:: A literature review. 2023.

[63] J. Smith and K. Jones. Designing an internet of medical things (iomt) architecture for healthcare applications. *IEEE Access*, 8:226577–226590, 2020.

[64] C. Jones and S. Lee. Network connectivity in the internet of medical things (iomt). *Journal of Medical Devices*, 7(3):1–6, 2019.

[65] M. Johnson, A. Smith, and K. Lee. The internet of medical things (iomt): Applications, benefits, and challenges. *Journal of Healthcare Information Management*, 32(4):23–30, 2018.

[66] Kfalcone. Chapter 12: Cardiovascular Disease. Accessed May 30, 2023 https://pressbooks.pub/introtohealth/chapter/chapter-12-cardiovascular-disease/.

[67] Paul A. Iaizzo. *Handbook of Cardiac Anatomy, Physiology, and Devices.* 12 2005.

[68] Home Health UK. Heart diagram. Accessed May 30, 2023 https://homehealth-uk.com/heart-diagram/, 2022. January 24.

[69] Cleveland Clinic Medical Professional. Blood Vessels. Accessed May 30, 2023 https://my.clevelandclinic.org/health/body/21640-blood-vessels.

[70] Adam Felman. Everything you need to know about heart disease. Accessed May 30, 2023 https://www.medicalnewstoday.com/articles/237191#causes-and-risk-factors, 7 2021.

[71] K. V. Divya, Akash Sirohi, Sagar Dhanraj Pande, and Rahul Malik. An iomt assisted heart disease diagnostic system using machine learning techniques. 2020.

[72] David H Bennett. *Cardiac Arrhythmias 7th Edition: Practical Notes on Interpretation and Treatment.* CRC Press, 9 2006.

[73] Mohammad Monirujjaman Khan and Fahad Algarni. A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS. *IEEE Access*, 8:122259–122269, 7 2020.

[74] NikhleshK Singh and Geetika Kaur. Inflammation and retinal degenerative diseases. *Neural Regeneration Research*, 18(3):513, 1 2023.

[75] Retina: Anatomy, Function Common Conditions. Accessed May 30, 2023 https://my.clevelandclinic.org/health/body/22694-retina-eye.

[76] FELLAH KAOUTHAR MANAR. Deep learning approach for early diabetic retinopathy diagnosis. 2022.

[77] Retinal diseases - Symptoms and causes. Accessed May 30, 2023 https://www.mayoclinic.org/diseases-conditions/retinal-diseases/symptoms-causes/syc-20355825, 1 2022.

[78] Jerzy Z Nowak. Age-related macular degeneration (amd): pathogenesis and therapy. *Pharmacological reports*, 58(3):353, 2006.

[79] Jayakrishna Ambati and Benjamin J. Fowler. Mechanisms of age-related macular degeneration. *Neuron*, 75(1):26–39, 2012.

[80] Diabetic macular edema. Accessed May 30, 2023 https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/multimedia/diabetic-macular-edema/img-20124558, 2 2023.

[81] What Are Drusen? Accessed May 30, 2023 https://www.aao.org/eye-health/diseases/what-are-drusen, 3 2022.

[82] Macular Degeneration. Accessed May 30, 2023 https://eyesight.org/macular-degeneration/, 1 2022.

[83] Agnes Moors. Theories of emotion causation. *A review; In Cognition and Emotion, 1st ed.*, Volume 1(pp. 1-37.), 1 2010.

[84] Iris B. Mauss and Michael E. Robinson. Measures of emotion: A review. *Cognition Emotion*, 23(2):209–237, 2 2009.

[85] Somchanok Tivatansakul and Michiko Ohkura. *International Journal of Affective Engineering*, (2):51–61, 1.

[86] Marwan Dhuheir, Abdullatif Albaseer, Emna Baccour, Aiman Erbad, Mohamed Abdallah, and Mounir Hamdi. Emotion Recognition for Healthcare Surveillance Systems Using Neural Networks: A Survey. *Communications and Mobile Computing*, 6 2021.

[87] Dimitrios Ververidis and Constantine Kotropoulos. Emotional speech recognition: Resources, features, and methods. *Speech Communication*, 48(9):1162–1181, 9 2006.

[88] Haoqi Li, Brian R. Baucom, and Panayiotis G. Georgiou. Linking emotions to behaviors through deep transfer learning. *PeerJ*, 6:e246, 1 2020.

[89] M. J. Hasnul, Nor Azlina Ab Aziz, Salem Alelyani, Mohamed Mohana, and Azlan Abdul Aziz. Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare—A Review. *Sensors*, 21(15):5015, 7 2021.

[90] Murugappan Murugappan, Nagarajan Ramachandran, and Yaacob Sazali. Classification of human emotion from EEG using discrete wavelet transform. *Journal of Biomedical Science and Engineering*, 03(04):390–396, 4 2010.

[91] Walter B. Cannon. The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *American Journal of Psychology*, 39(1/4):106, 12 1927.

[92] Yucel Cimtay, Erhan Ekmekcioglu, and Seyma Caglar-Ozhan. Cross-Subject Multimodal Emotion Recognition Based on Hybrid Fusion. *IEEE Access*, 8:168865–168878, 9 2020.

[93] D. Nikolova, P. Georgieva, P. Petkova, and A. Manolova. ECG-based emotion recognition: Overview of methods and applications. . *In Proceedings of the ANNA 2018—Advances in Neural Networks and Applications, St,* (pp. 118–122.), 9 2018.

[94] Christos D. Katsis, Nikolaos S. Katertsidis, G. Ganiatsas, and Dimitrios I. Fotiadis. Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach. *IEEE transactions on systems, man, and cybernetics*, 38(3):502–512, 5 2008.

[95] Guillaume Chanel, Sunny Avry, Gaëlle Molinari, Mireille Bétrancourt, and Thierry Pun. Multiple users' emotion recognition: Improving performance by joint modeling of affective reactions. *Affective Computing and Intelligent Interaction*, 10 2017.

[96] Hatice Gunes and Maja Pantic. Automatic, Dimensional and Continuous Emotion Recognition. *International Journal of Synthetic Emotions*, 1(1):68–99, 1 2010.

[97] Anna Kolakowska, Agnieszka Landowska, Mariusz Szwoch, Wioleta Szwoch, and Michał Wróbel. Emotion Recognition and Its Applications. *Advances in intelligent systems and computing*, pages 51–62, 1 2014.

[98] Tengfei Song, Wenming Zheng, Cheng Lu, Yuan Zong, Xilei Zhang, and Zhen Cui. MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition. *IEEE Access*, 7:12177–12191, 1 2019.

[99] Robert Plutchik. A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION. *Elsevier eBooks*, pages 3–33, 1 1980.

[100] Margaret M. Bradley and Peter Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 3 1994.

[101] Albert Mehrabian. Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression. *Journal of Psychopathology and Behavioral Assessment*, 19(4):331–357, 12 1997.

[102] Siao Zheng Bong, M Murugappan, and Sazali Yaacob. Analysis of electrocardiogram (ecg) signals for human emotional stress classification. In *Trends in Intelligent Robotics, Automation, and Manufacturing: First International Conference, IRAM 2012, Kuala Lumpur, Malaysia, November 28-30, 2012. Proceedings*, pages 198–205. Springer, 2012.

[103] Saskia Koldijk, Maya Sappelli, Suzan Verberne, Mark A. Neerincx, and Wessel Kraaij. The SWELL Knowledge Work Dataset for Stress and User Modeling Research. 11 2014.

[104] Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven. Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. 10 2018.

[105] View of Identifying Retinal Diseases on OCT Image Based on Deep Learning. Accessed May 30, 2023 https://online-journals.org/index.php/i-joe/article/view/33639/12359.

[106] Jen Hong Tan, Sulatha V Bhandary, Sobha Sivaprasad, Yuki Hagiwara, Akanksha Bagchi, U Raghavendra, A Krishna Rao, Biju Raju, Nitin Shridhara Shetty, Arkadiusz Gertych, et al. Age-related macular degeneration detection using deep convolutional neural network. *Future Generation Computer Systems*, 87:127–135, 2018.

[107] Philippe M Burlina, Neil Joshi, Michael Pekala, Katia D Pacheco, David E Freund, and Neil M Bressler. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA ophthalmology*, 135(11):1170–1176, 2017.

[108] T Vivekanandan and N Ch Sriman Narayana Iyengar. Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Computers in biology and medicine*, 90:125–136, 2017.

[109] Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava. Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, 7:81542–81554, 2019.

[110] Mohammad Kachuee, Shayan Fazeli, and Majid Sarrafzadeh. Ecg heartbeat classification: A deep transferable representation. In *2018 IEEE international conference on healthcare informatics (ICHI)*, pages 443–444. IEEE, 2018.

[111] Deger Ayata, Yusuf Yaslan, and Mustafa E. Kamasak. Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems. *Journal of Medical and Biological Engineering*, 40(2):149–157, 4 2020.

[112] Byung-Gee Kim and Sungho Jo. Deep Physiological Affect Network for the Recognition of Human Emotions. *IEEE Transactions on Affective Computing*, page 1, 4 2020.

[113] Time Series Classification Website. Accessed May 30, 2023 http://timeseriesclassification.com/description.php?Dataset=ECG5000.

[114] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[115] Klaus Nordhausen. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition by Trevor Hastie, Robert Tibshirani, Jerome Friedman. *International Statistical Review*, 77(3):482, 12 2009.

[116] A Review of Machine Learning and Deep Learning Applications. https://ieeexplore.ieee.org/abstract/document/8697857, August 2018.

[117] What is Deep Learning? | IBM. Accessed May 30, 2023 https://www.ibm.com/topics/deep-learning.

[118] Zerari Achouak. Pediatric Bone Age Assessment from Hand X-ray using Deep Learning Approach, 6 2022.

[119] Xiangbin Liu, Liping Song, Shuai Liu, and Yudong Zhang. A Review of Deep-Learning-Based Medical Image Segmentation Methods. *Sustainability*, 13(3):1224, 1 2021.

[120] Mohammed Imran and Sarah A. Alsuhaibani. *A Neuro-Fuzzy Inference Model for Diabetic Retinopathy Classification*. 1 2019.

[121] Djihane Houfani, Sihem Slatnia, Okba Kazar, Ikram Remadna, Hamza Saouli, Guadalupe Ortiz, and Abdelhak Merizig. An improved model for breast cancer diagnosis by combining pca and logistic regression techniques. *International Journal of Computing and Digital Systems*, 13(1):701–716, 2023.

[122] Weihong Xu, Zhongfeng Wang, Xiaohu You, and Chuan Zhang. Efficient fast convolution architectures for convolutional neural network. *2017 IEEE 12th International Conference on ASIC (ASICON)*, pages 904–907, 2017.

[123] Naveen. What is Convolutional Neural Network? *Intellipaat Blog*, 3 2023.

[124] Saeed Iqbal, Adnan N. Qureshi, Jianqiang Li, and Tariq Mahmood. On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks. *Archives of Computational Methods in Engineering*, 4 2023.

[125] Manish Chablani. DenseNet - Towards Data Science. *Towards Data Science*, July 2019.

[126] Karen Simonyan. Very Deep Convolutional Networks for Large-Scale Image Recognition, 9 2014.

[127] Gaudenz Boesch. VGG Very Deep Convolutional Networks (VGGNet) 8211; What you need to know. *viso.ai*, 3 2023.

[128] JalFaizy Shaikh. Deep Learning in the Trenches: Understanding Inception Network from Scratch. *Analytics Vidhya*, 5 2020.

[129] Richard Gall and Richard Gall. What is LSTM? *Packt Hub*, 9 2022.

[130] Enes Zvornicanin and Enes Zvornicanin. Differences Between Bidirectional and Unidirectional LSTM | Baeldung on Computer Science. Accessed June 10, 2023 https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm, 6 2023.

[131] Welcome to Python.org. Accessed May 30, 2023 https://www.python.org/about/, 5 2023.

[132] Google Cloud. Encoder-decoder architecture. Accessed May 30, 2023 https://www.coursera.org/learn/encoder-decoder-architecture, n.d. Online course.

[133] Lisa Torrey and Jude Shavlik. Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pages 242–264. IGI global, 2010.

[134] Patrick Ferdinand Christ. Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks, 2 2017.

[135] Apprenez à construire un CNN et gagnez du temps avec le Transfer Learning. Accessed May 30, 2023 https://openclassrooms.com/en/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5088816-apprenez-a-construire-un-cnn-et-gagnez-du-temps-avec-le-transfer-learning,.

[136] Marco Colizzi, Antonio Lasalvia, and Mirella Ruggeri. Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care? *International Journal of Mental Health Systems*, 14(1), 3 2020.

[137] World Health Organization: WHO. WHO reveals leading causes of death and disability worldwide: 2000-2019. 12 2020.

[138] Eric A. Swanson and James G. Fujimoto. The ecosystem that powered the translation of oct from fundamental research to clinical and commercial impact. *Biomed. Opt. Express*, 8(3):1638–1664, Mar 2017.

[139] Bo Sun and Zihuai Lin. Emotion recognition using machine learning and ecg signals. *arXiv preprint arXiv:2203.08477*, 2022.

[140] M. J. Hasnul, Nor Azlina Ab Aziz, Salem Alelyani, Mohamed Mohana, and Azlan Abdul Aziz. Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare—A Review. *Sensors*, 21(15):5015, 7 2021.

[141] Kaggle. Heart failure prediction dataset. https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction, September 2021. Accessed on January 1, 2023.

[142] ECG Heartbeat Categorization Dataset. Accessed May 30, 2023 https://www.kaggle.com/datasets/shayanfazeli/heartbeat, 5 2018.

[143] Javed Rashid, Maryam Ishfaq, Ghulam Ali, Muhammad Saeed, Mubasher Hussain, Tamim Alkhalifah, Fahad Alturise, and Noor Samand. Skin Cancer Disease Detection Using Transfer Learning Technique. *Applied sciences*, 12(11):5714, 6 2022.

[144] Automatic Cough Detection in Acoustic Signal using Spectral Features. *IEEE Conference Publication | IEEE Xplore*, 7 2019.

[145] Raghavendra Kumar, Pardeep Kumar, and Yugal Kumar. Time Series Data Prediction using IoT and Machine Learning Technique. *Procedia Computer Science*, 167:373–381, 1 2020.

[146] Retinal OCT Images (optical coherence tomography). Accessed May 30, 2023 https://www.kaggle.com/datasets/paultimothymooney/kermany2018, 3 2018.

[147] Kizito Nkurikiyeyezu, Kana Shoji, Anna Yokokubo, and Guillaume Lopez. Thermal Comfort and Stress Recognition in Office Environment. 2 2019.

[148] Marek Malik, J. Thomas Bigger, A. John Camm, Robert E. Kleiger, Alberto Malliani, Arthur J. Moss, and Peter J. Schwartz. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17(3):354–381, 3 1996.

[149] Soeren Kroell. Main HRV parameters - Part 1: Time-related parameters. *Main HRV parameters - Part 1: Time-related parameters*, 7 2020.

[150] Welcome to Python.org. Accessed May 30, 2023 https://www.python.org/about/, 5 2023.

[151] TensorFlow. Accessed May 30, 2023 https://www.tensorflow.org/?hl=fr.

[152] Keras Team. Keras documentation: Code examples. Accessed May 30, 2023https://keras.io/examples/.

[153] scikit-learn: machine learning in Python — scikit-learn 0.16.1 documentation. Accessed May 30, 2023 https://scikit-learn.org/.

[154] NumPy. Accessed May 30, 2023 https://numpy.org/.

[155] pandas - Python Data Analysis Library. Accessed May 30, 2023 https://pandas.pydata.org.

[156] Matplotlib —Visualization with Python. Accessed May 30, 2023 https://matplotlib.org.

[157] Documentation — NeuroKit 0.2.1 documentation. Accessed May 30, 2023 https://neurokit2.readthedocs.io/en/legacy_docs/.

[158] Kaggle: Your Machine Learning and Data Science Community. Accessed May 30, 2023 https://www.kaggle.com,.

[159] Download Android Studio App Tools - Android Developers. Accessed May 30, 2023 https://developer.android.com/studio.

[160] Software. Accessed May 30, 2023 https://wiki-content.arduino.cc/en/software.

[161] Raspberry Pi Ltd. Raspberry Pi OS – Raspberry Pi. Accessed May 15, 2023 https://www.raspberrypi.com/software/.

[162] Firebase. Accessed May 30, 2023 https://firebase.google.com.

[163] Flutter. Accessed May 30, 2023 https:https://flutter.dev.

[164] Kotlin. Accessed May 30, 2023 https://developer.android.com/kotlin?hl=fr.

[165] ONNX. Accessed May 30, 2023 https://onnx.ai/.

[166] NodeMcu – An open-source firmware based on ESP8266 wifi-soc. Accessed May 30, 2023 https://www.nodemcu.com/index_en.html.

[167] Mixed-signal and digital signal processing ICs | Analog Devices. Accessed May 30, 2023 https://www.analog.com/en/index.html.

[168] WatElectronics. Pulse Sensor : Types, Pin Diagram, Circuit Its Applications. Accessed May 15, 2023 https://www.watelectronics.com/pulse-sensor/, 6 2022.

[169] Kuntoro Adi Nugroho. A Comparison of Handcrafted and Deep Neural Network Feature Extraction for Classifying Optical Coherence... *ResearchGate*, 9 2018.

[170] Terri Hoare. Classification of retinal pathology from OCT images using a parametric tuned CNN, 2020.

[171] Siao Zheng Bong, Murugappan Murugappan, and Sazali Yaacob. *Analysis of Electrocardiogram (ECG) Signals for Human Emotional Stress Classification.* Springer Science+Business Media, 11 2012.

[172] A S Anusha, Joy Idiculla Jose, S P Preejith, Joseph Jayaraj, and Sivaprakasam Mohanasankar. Physiological signal based work stress detection using unobtrusive sensors. *Biomedical Physics Engineering Express*, 4(6):065001, 9 2018.

[173] Patrícia J. Bota, Cheng Wang, Ana Fred, and Hugo Manuel Ribeiro Dias Da Silva. A Review, Current Challenges, and Future Possibilities on Emotion Recognition Using Machine Learning and Physiological Signals. *IEEE Access*, 7:140990–141020, 9 2019.

[174] Kizito Nkurikiyeyezu, Kana Shoji, Anna Yokokubo, and Guillaume Lopez. Thermal Comfort and Stress Recognition in Office Environment. 2 2019.

[175] Pritam Sarkar and Ali Etemad. Self-Supervised Learning for ECG-Based Emotion Recognition. 5 2020.

[176] Sihem Nita, Salim Bitam, Matthieu Heidet, and Abdelhamid Mellouk. A new data augmentation convolutional neural network for human emotion recognition based on ECG signals. *Biomedical Signal Processing and Control*, 75:103580, 5 2022.