



# Combination of different physiological signals for the diagnosis of heart diseases

Egyptian E-Learning University Graduation Project 2023-2024

By:

Abdallah Magdy Fouad Habib

Osama Ahmed Mosleh

Hossam Hassan Mohamed Mohamed

AbdulRahman Hassan Abdulaal Rashwan

Fares Reda Ebrahim

Ahmed Atef Ahmed

Alaa Atia El-sayed Atiaa

AbdElRahman Hisham Mohamed

Supervised by:

Dr. Mahmoud Bassiouni

T.A. Amany

Semester 2023 / 2024

## Contents

<b>1. Introduction.....</b>	<b>1</b>
1.1. Heart.....	1
1.2. Type of Signals .....	9
1.3. Electrocardiogram (ECG).....	11
1.4. Phonocardiogram (PCG).....	13
1.5. Photoplethysmogram (PPG) .....	14
1.6. Arterial Blood Pressure (ABP).....	16
<b>2. Literature Review .....</b>	<b>Error! Bookmark not defined.</b>
2.1. Related Work based on ECG for Heart Diseases Diagnosis .....	<b>Error! Bookmark not defined.</b>
2.2. Related Work based on PCG for Heart Diseases Diagnosis .....	<b>Error! Bookmark not defined.</b>
2.3. Related Work based on PPG for Heart Diseases Diagnosis .....	<b>Error! Bookmark not defined.</b>
2.4. Combination of Heart Signals for the diagnosis of Diseases .....	<b>Error! Bookmark not defined.</b>
<b>3. Deep learning .....</b>	<b>34</b>
3.1. Why deep learning? .....	35
3.2. DL Properties and Dependencies .....	36
3.3. Types of Deep Learning networks.....	38
3.3.1. Recursive neural networks.....	38
3.3.2. Recurrent neural networks .....	38
3.3.3. Convolutional neural networks .....	39
3.3.4. Benefits of employing CNNs .....	40
3.3.5. CNN layers .....	41
3.3.6. Regularization to CNN .....	44
3.4. Optimizer selection .....	45
3.4.1. gradient-based learning optimization algorithm .....	46
3.5. CNN architectures.....	47
3.5.1. ResNet.....	47
3.5.2. Inception: ResNet and Inception-V3/4.....	50
3.5.3. Xception Net .....	51
<b>4. Methodology .....</b>	<b>52</b>
4.1. Dual Approach Planning .....	<b>Error! Bookmark not defined.</b>
4.1.1. ECG Image Processing .....	<b>Error! Bookmark not defined.</b>

4.1.2.	Digitization of ECG Lead.....	Error! Bookmark not defined.
4.2.	Data Gathering .....	Error! Bookmark not defined.
4.3.	Preprocessing.....	Error! Bookmark not defined.
4.3.1.	Filtering.....	Error! Bookmark not defined.
4.3.2.	Cropping & Resizing .....	Error! Bookmark not defined.
4.3.3.	Masking .....	Error! Bookmark not defined.
4.4.	Digitization .....	Error! Bookmark not defined.
4.4.1.	PaperECG .....	Error! Bookmark not defined.
4.4.2.	Automating Script .....	Error! Bookmark not defined.
4.5.	Normalization.....	Error! Bookmark not defined.
4.6.	Digitization Plotting .....	Error! Bookmark not defined.
4.6.1.	CWT Transformation.....	Error! Bookmark not defined.
4.6.2.	Scalogram .....	Error! Bookmark not defined.
4.7.	Training.....	Error! Bookmark not defined.
4.7.1.	Model Selection .....	Error! Bookmark not defined.
4.7.2.	Libraries Needed .....	Error! Bookmark not defined.
4.7.3.	Data Organization.....	Error! Bookmark not defined.
4.7.4.	Preprocessing.....	Error! Bookmark not defined.
4.7.5.	Model Training .....	Error! Bookmark not defined.
4.7.6.	Confusion Matrix Generation .....	Error! Bookmark not defined.
4.7.7.	5-Fold Cross Validation.....	Error! Bookmark not defined.
<b>5.</b>	<b>Experiment Results .....</b>	Error! Bookmark not defined.
5.1.	Image Processing Approach Training Results.....	Error! Bookmark not defined.
		Error! Bookmark not defined.
5.2.	Signal Processing Approach Training Results:.....	Error! Bookmark not defined.
<b>6.</b>	<b>Mobile Application .....</b>	Error! Bookmark not defined.
6.1.	Technologies Used .....	Error! Bookmark not defined.
6.2.	Included Libraries .....	Error! Bookmark not defined.
6.2.1.	Picasso .....	Error! Bookmark not defined.
6.2.2.	Zoomage .....	Error! Bookmark not defined.
6.2.3.	Image Cropper.....	Error! Bookmark not defined.
6.2.4.	Volley – File Upload Request.....	Error! Bookmark not defined.
6.3.	Basic Application Structure and Flow .....	Error! Bookmark not defined.
6.4.	Server-Side Application .....	Error! Bookmark not defined.
6.5.	Mobile Implementation.....	Error! Bookmark not defined.

6.5.1.	Main Activity .....	Error! Bookmark not defined.
6.5.2.	Align Activity .....	Error! Bookmark not defined.
6.5.3.	Result Activity .....	Error! Bookmark not defined.
<b>7.</b>	<b>References .....</b>	Error! Bookmark not defined.

# **Introduction**

# **1. Introduction**

Heart diseases are considered one of the most of issues that can cause critical circumstances and in most of the cases it can lead to the death of several people. Dealing with these diseases in terms of accurate diagnosis and detection is one of the major tasks. Therefore, there exist several signals that can be used to diagnose complex heart diseases such as Electrocardiogram (ECG), Phonocardiogram (PCG), and Photoplethysmogram (PPG).

There exist various systems that depend on a single signal for the diagnosis of heart diseases. Some of these systems can rely on machine learning approaches (ML) and other systems depend on deep learning approaches (DL) in the process of feature extraction and diagnosis. It can be seen that various heart signals can contribute in increasing the performance of diagnosing heart diseases. Also, few numbers of systems exist that can combine more than one signal for the diagnosis of heart diseases. Therefore, the main contribution of the project is to use or apply more than one signal combined together to diagnose heart diseases.

## **1.1. Heart**

Your muscular heart, the main organ in your cardiovascular system, is vital for life. Its parts work together to move blood through your body in a coordinated way. It constantly sends oxygen to your cells and takes away waste. Many conditions can affect this organ and keep it from working well.

The heart is a fist-sized organ that pumps blood throughout your body. It's your circulatory system's main organ. Muscle and tissue make up this powerhouse organ. Your heart contains four muscular sections (chambers) that briefly hold blood before moving it. Electrical impulses make your heart beat, moving blood through these chambers. Your brain and nervous system direct your heart's function. Myocardial injury (myocarditis, cytokine storm), acute coronary syndrome, and arrhythmia.

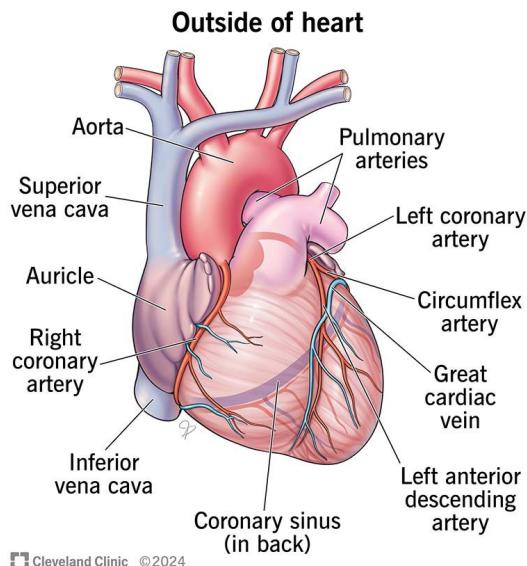


Figure. 1.1 Heart muscular organ that pumps blood to your body.

Your heart's main function is to move blood throughout your body. Blood brings oxygen and nutrients to your cells. It also takes away carbon dioxide and other waste so other organs can dispose of them.

Your heart also:

- Controls the rhythm and speed of your heart rate.
- Maintains your blood pressure.

Your heart works with these body systems to control your heart rate and other body functions:

- **Nervous system:** Your nervous system helps control your heart rate. It sends signals that tell your heart to beat slower during rest and faster during stress.
- **Endocrine system:** Your endocrine system sends out hormones. These hormones tell your blood vessels to constrict or relax, which affects your blood pressure. Hormones from your thyroid gland can also tell your heart to beat faster or slower.

The parts of your heart are like the parts of a building. Your heart anatomy includes:

Walls.

Chambers that are like rooms.

Valves that open and close like doors to the rooms.

Blood vessels like plumbing pipes that run through a building.

An electrical conduction system like electrical power that runs through a building

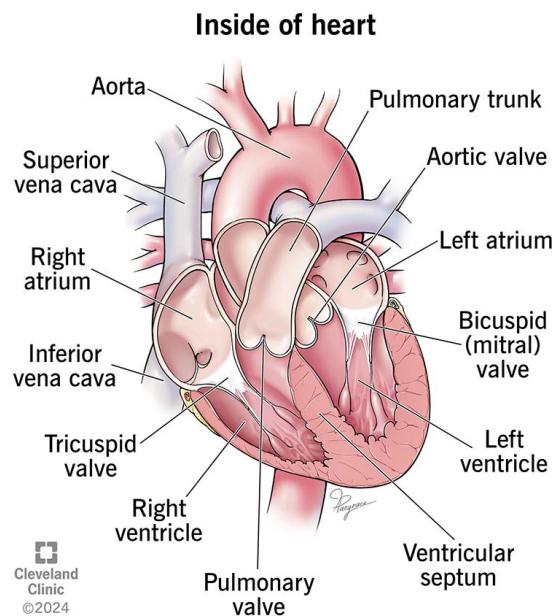


Figure. 1.2 Blood moves through chambers inside your heart.

Your heart walls are the muscles that contract (squeeze) and relax to send blood throughout your body. A layer of muscular tissue called the septum divides your heart walls into the left and right sides.

Your heart walls have three layers:

Endocardium: Inner layer.

Myocardium: Muscular middle layer.

Epicardium: Protective outer layer.

The epicardium is one layer of your pericardium. The pericardium is a protective sac that covers your entire heart. It produces fluid to lubricate your heart and keep it from rubbing against other organs

Your heart has four separate chambers. You have two chambers on the top (atrium, plural atria) and two on the bottom (ventricles), one on each side of your heart.

**Right atrium:** Two large veins deliver oxygen-poor blood to your right atrium. The superior vena cava carries blood from your upper body. The inferior vena cava brings blood from your lower body. Then the right atrium pumps the blood to your right ventricle.

**Right ventricle:** The lower right chamber pumps the oxygen-poor blood to your lungs through the pulmonary artery. The lungs reload the blood with oxygen.

**Left atrium:** After the lungs fill your blood with oxygen, the pulmonary veins carry the blood to the left atrium. This upper chamber pumps the blood to your left ventricle.

**Left ventricle:** The left ventricle is slightly larger than the right. It pumps oxygen-rich blood to the rest of your body.

Your heart valves are like doors between your heart chambers. They open and close to allow blood to flow through. They also keep your blood from moving in the wrong direction.

### Atrioventricular valves

The atrioventricular (AV) valves open between your upper and lower heart chambers. They include:

- **Tricuspid valve:** Door between your right atrium and right ventricle.
- **Mitral valve:** Door between your left atrium and left ventricle.

### Semilunar valves

Semilunar (SL) valves open when blood flows out of your ventricles. They include:

- **Aortic valve:** Opens when blood flows out of your left ventricle to your aorta (artery that carries oxygen-rich blood to your body).

- **Pulmonary valve:** Opens when blood flows from your right ventricle to your pulmonary arteries (the only arteries that carry oxygen-poor blood to your lungs).

## Blood vessels

Your heart pumps blood through three types of blood vessels:

- Arteries carry oxygen-rich blood from your heart to your body's tissues. The exception is your pulmonary arteries, which go to your lungs.
- Veins carry oxygen-poor blood back to your heart.
- Capillaries are small blood vessels where your body exchanges oxygen-rich and oxygen-poor blood.

## Coronary arteries

Your heart receives nutrients through a network of coronary arteries. These arteries run along your heart's surface. They serve the heart itself and include the:

- **Left coronary artery:** Divides into two branches (the circumflex artery and the left anterior descending artery).
- **Circumflex artery:** Supplies blood to the left atrium and the side and back of the left ventricle.
- **Left anterior descending artery (LAD):** Supplies blood to the front and bottom of the left ventricle and the front of the septum.
- **Right coronary artery (RCA):** Supplies blood to the right atrium, right ventricle, bottom portion of the left ventricle and back of the septum.

Your heart's conduction system is like the electrical wiring of a building. It controls the rhythm and pace of your heartbeat. Signals start at the top of your heart and move down to the bottom. Your conduction system includes:

**Sinoatrial (SA) node:** Sends the signals that make your heart beat.

**Atrioventricular (AV) node:** Carries electrical signals from your heart's upper chambers to its lower ones.

**Left bundle branch:** Sends electric impulses to your left ventricle.

**Right bundle branch:** Sends electric impulses to your right ventricle.

**Bundle of His:** Sends impulses from your AV node to the Purkinje fibers.

**Purkinje fibers:** Make your heart ventricles contract and pump out blood.

Your heart is in the front of your chest. It sits slightly behind and to the left of your sternum (breastbone), which is in the middle of your chest.

Your heart is slightly on the left side of your body. It sits between your right and left lungs. The left lung is slightly smaller to make room for the heart in your left chest. Your rib cage protects your heart.

Your heart looks a little bit like an upside-down pyramid with rounded edges. Large blood vessels go into and out of your heart to bring blood into and away from your heart. They connect your heart to the rest of your body, which it supplies with blood and oxygen

Everyone's heart is a slightly different size. Generally, your heart is about the same size as your fist. On average, an adult's heart weighs about 10 ounces. Your heart may weigh a little more or a little less, depending on your body size and sex.

Heart conditions are among the most common types of disorders. In the United States, heart disease is the leading cause of death.

### **Common conditions that affect your heart include:**

- **Arrhythmia:** A heartbeat that's too fast, too slow or beats with an irregular rhythm.
- **Cardiomyopathy:** Unusual thickening, enlargement or stiffening of your heart muscle.
- **Congestive heart failure:** Your heart is too stiff or too weak to properly pump blood throughout your body.
- **Coronary artery disease:** Plaque buildup that leads to narrow coronary arteries.

- **Diabetes:** Your blood sugar is higher than it should be.
- **Heart attack (myocardial infarction):** A sudden coronary artery blockage that cuts off oxygen to part of your heart muscle.
- **Heart valve disease:** A valve in your heart isn't working right.
- **High blood pressure:** Your blood is pushing too hard against your artery walls.
- **High cholesterol:** Your blood has too many fats in it.
- **Pericarditis:** Inflammation in your heart's lining (pericardium).

### **Common signs or symptoms of heart conditions**

Symptoms of heart conditions include:

- Chest pain.
- Heart palpitations.
- Dizziness.
- Shortness of breath.
- Fatigue.
- Swelling in your lower body.

### **Common tests to check the health of your heart**

Tests to check your heart health include:

- Blood pressure measurement.
- Electrocardiogram (EKG).
- Echocardiogram.
- Chest X-ray.
- Blood tests.
- Cardiac catheterization.

- Computed tomography (CT).
- Heart MRI (magnetic resonance imaging).
- Stress test.

## 1.2. Type of Signals

### ➤ Electrocardiogram (ECG or EKG)

Description: An ECG measures the electrical activity of the heart over a period of time. Electrodes are placed on the skin to detect these electrical signals, which are then recorded and displayed as a graph.

Components:

- P wave: Atrial depolarization.
- QRS complex: Ventricular depolarization.
- T wave: Ventricular repolarization.
- PR interval: Time between the onset of atrial depolarization and the onset of ventricular depolarization.

Uses:

- Diagnosing arrhythmias, heart attacks, and other heart conditions.
- Monitoring heart health during physical exams or in critical care.

### ➤ Phonocardiogram (PCG)

Description: A PCG records the sounds produced by the heart, specifically the sounds of the heart valves opening and closing. It is often used in conjunction with a stethoscope but can be recorded for detailed analysis.

Components:

- S1 and S2 sounds: Represent the "lub-dub" sounds of the heart, corresponding to the closure of the atrioventricular and semilunar valves, respectively.
- S3 and S4 sounds: Additional heart sounds that may indicate heart disease.

Uses:

- Diagnosing valve disorders, such as stenosis or regurgitation.
- Detecting abnormalities in heart sounds that may indicate various cardiac conditions.

➤ **Photoplethysmogram (PPG)**

Description: A PPG measures blood volume changes in the microvascular bed of tissue using a light source and a photodetector. It is often used in devices like pulse oximeters and fitness trackers.

Components:

- AC component: Pulsatile component related to the heartbeat.
- DC component: Non-pulsatile component related to the baseline level of blood flow.

Uses:

- Monitoring heart rate and blood oxygen levels.
- Assessing vascular health and blood flow.
- Used in wearable health devices for continuous monitoring.

➤ **Atrial Blood Pressure (ABP)**

Description: ABP measures the pressure within the arteries. It provides a continuous and real-time indication of cardiovascular function, often measured invasively using a catheter inserted into an artery.

Components:

- Systolic pressure: The peak pressure in the arteries during the contraction of the heart muscles.
- Diastolic pressure: The lowest pressure in the arteries when the heart muscles are relaxed.
- Mean arterial pressure (MAP): The average pressure in a patient's arteries during one cardiac cycle.

Uses:

- Monitoring critically ill patients in intensive care units.
- Assessing the effectiveness of treatments for hypertension and other cardiovascular conditions.

- Providing real-time data during surgeries and other medical procedures.

These heart signals collectively offer a comprehensive view of heart function and cardiovascular health, each providing unique and valuable insights into different aspects of the cardiovascular system.

### **1.3. Electrocardiogram (ECG)**

An electrocardiogram (ECG) is simply a recording of the electrical activity generated by the heart [1]. Sample ECG signals associated with a common cardiac cycle are illustrated in Fig. 3 [2,3]. The ECG is an effective non-invasive tool for various biomedical applications such as measuring the heart rate, examining the rhythm of heartbeats, diagnosing heart abnormalities, emotion recognition and biometric identification. One of the major fields in which ECG analysis is required is the diagnosis of cardiovascular diseases. As reported by the World Health Organization, cardiovascular diseases are the main reason for deaths worldwide. Among the cardiovascular diseases, cardiac arrhythmias are the most common, and as a result, their precise classification has been of great interest in biomedical studies [4].

One of the most effective tools for identifying arrhythmias is ECG signal exploration [5]. The investigation of individual ECG beat characteristic shapes, morphological features, and spectral possessions can provide meaningfully correlated clinical information for the automatic recognition of an ECG pattern. However, automated classification of ECG beats is a difficult problem because the morphological and temporal features of the ECG signals include noteworthy dissimilarities for different patients under different physical circumstances [6]. The main problem for diagnosing heart diseases with ECGs is that an ECG signal can vary for each person, and sometimes different patients have separate ECG morphologies for the same disease.

Moreover, two different diseases could have approximately the same properties on an ECG signal. These problems cause some difficulties for the problem of heart disease diagnosis [5,7,8]. To detect abnormalities of the heartbeat, the electrical signal of each heartbeat must be analyzed. Therefore, the process of analyzing long-term ECG records, especially for bedside monitoring or wearable online health care monitoring, can be very troublesome for

a person, and it is very time-consuming. Furthermore, some personal errors can occur throughout an ECG analysis due to fatigue, [9] and the interpretation of the signal requires deep knowledge [10].

Therefore, computer-assisted methods that provide automatic ECG analysis are utilized. The use of ECG analysis in fields other than the diagnosis of cardiovascular diseases has also increased substantially. Many researchers have used ECG signals for emotion recognition, especially for stress level detection in addition to many other signals such as the electroencephalogram, skin temperature, blood pressure, electromyogram, heart rate variability, cortisol levels, and thermal imaging features. Researchers measure ECG signals at different critical moments (stress situations), such as during an oral exam, after a holiday for students, in office environments for office workers, and during a driving task for drivers. The results of these studies reveal that ECG features are useful at distinguishing the characteristics between different mental workloads and stress levels as well [1].

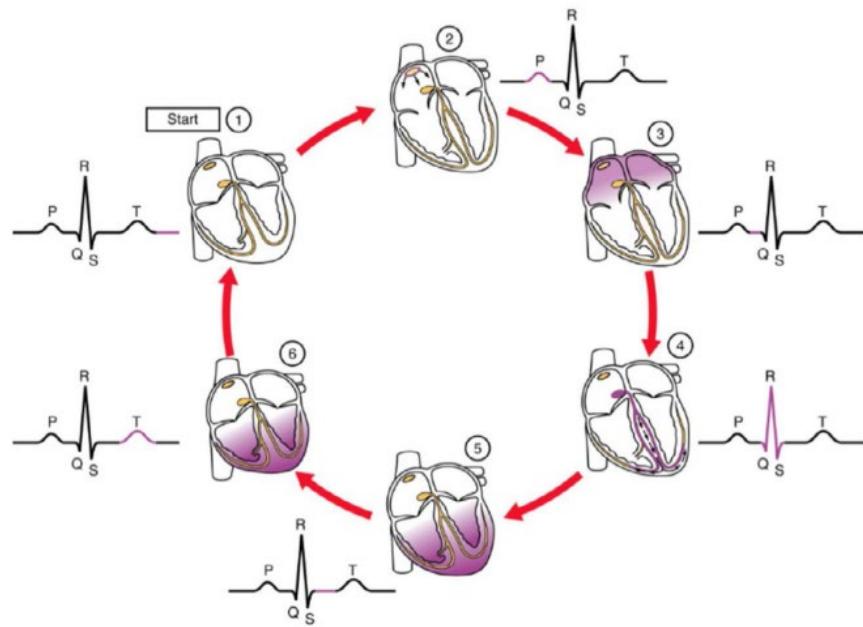


Figure. 1.3 A sketch of a common cardiac cycle with the associated waves of an ECG signal (one-lead)

#### 1.4. Phonocardiogram (PCG)

Four main chambers compose the human heart. These are the right and left atrium and right and left ventricle as shown in Fig. 4. The heart's right part is much smaller with less myocardium in their heart wall. Based on the two circulatory loops' size, there is a variation and difference in the main function of the left and right sides. Blood is pumped to the extremities of the human body by the heart's left side, while the human heart's right side operates as a pulmonary circulation to the lungs [12].

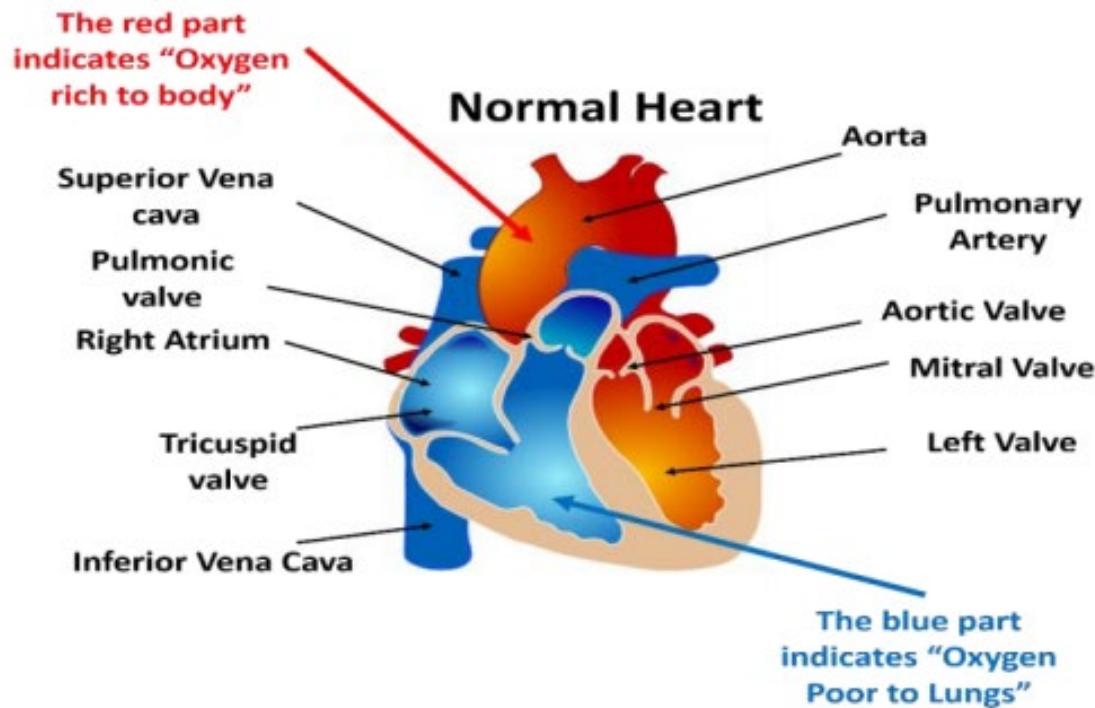


Figure. 1.4 Normal Heart Structure

The functionality mechanism generates acoustic signals and vibrations that can be gained and obtained over the wall chest. The normal heart sounds are defined as "lubb" and "dupp", and they are caused by pushing blood on the valves of the human heart as sh The "lubb" sound is known by S1 and it comes first in the human heartbeat and it is longer than that of the two heart sounds. Lubb is generated by closing the AV valves that are located at the beginning of the ventricular systole. The "dupp" sound is also known as S2 and it comes next in the human heartbeat, and it is a shorter and sharper sound that is resulted from the closing of the semilunar valves at the end of the ventricular systole. The pattern of S1-S2 or lubb-dupp is repeated in the heart. Some problems or different sounds such as

gurgling or liquid rushing in the heart may indicate problems in the heart that cause defects in the ventricular or atrial or leakage in the valves. Heart sounds can be used in a lot of applications for diagnosis.

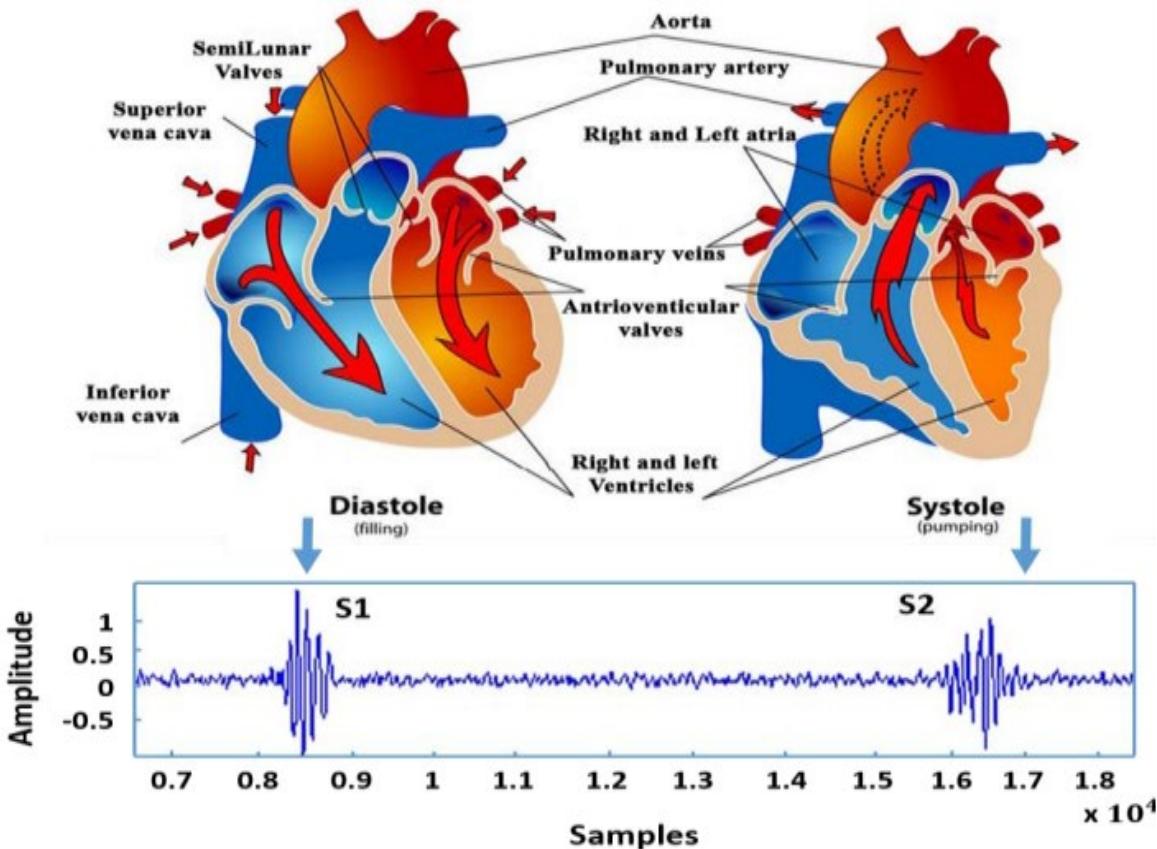


Figure. 1.5 (a) The diastole (S1) and systole (S2) cycle in the heart [14]

## 1.5. Photoplethysmography (PPG)

PPG technology measures the change in blood volume in the tissue during a heart cycle using a light source. This volumetric measurement provides important information regarding the cardiovascular system. A PPG sensor mainly consists of two electronic components, a light emitter and a light intensity sensing component. Typically, LED is used as a light emitter and a photodetector to detect (sense) the change in light intensity [13]. A PPG pulse corresponding to one heartbeat includes the systolic and diastolic phases. During the systolic phase, the volume of blood in arteries is more; this is because during this phase heart contracts and pushes oxygen-rich blood to all the tissues and organs. The

systolic phase causes more light is absorbed by the blood cells. Therefore, the amount of light detected by the photodetector during the systolic phase is low. During the diastolic phase, the blood has flown back into the heart. Therefore, during the diastole phase, the light detected by the photodetector increases due to a decrease in the blood volume. Depending upon application and sensor placement, PPG can be used either in transmissive mode or in refection mode, as shown in Fig. 6 [14].

When a photodetector and LED are placed on parallel sides of a finger to detect the transmitted light, this mode is known as a transmissive mode. In transmissive mode, the probe is in a projection that the photodetector and LED face each other with a layer of tissues between them [15]. Detection in transmissive mode depends upon transmission of light from body parts, so thin structures like the earlobe and finger are preferred in this mode. When both photodetector and LED are placed on the same side of a finger to detect the reflected light, it is a reflective mode. In refection mode, both the sensors are placed next to each other with an approximate spacing of 3 cm.

Therefore, refection mode can use anybody site like the forehead and wrist. Choice of the site to place PPG sensors depends on the patient's blood perfusion, comfortability of the subject, and application [16]. The role of the photodetector is to detect and quantify the light absorbed during pulsatile and non-pulsatile few [17]. During pulsatile few, light is absorbed by the change in blood few inside the arteries, which is synchronous with a heartbeat. During the non-pulsatile few, light is absorbed by background tissues. Therefore, a photodetector detects the volumetric change in blood few in arteries by detecting the light intensity difference [18].

Measurement of this change in light intensity thus helps to analyze the functionality of the heart. A PPG signal mainly consists of AC and DC components. AC component in the PPG output waveform indicates the change in light intensity during the systolic and diastolic phase due to the blood in arteries [19]. The steady DC part of the PPG waveform indicates the light absorbed by tissues, skin, and bone, as shown in Fig. 7. Analysis of the DC component provides valuable information regarding venous blood few, respiration, and

thermoregulation. Variation in light intensity detected due to arterial blood flow is around 1% only, which provides information on the heart's functionality [20].

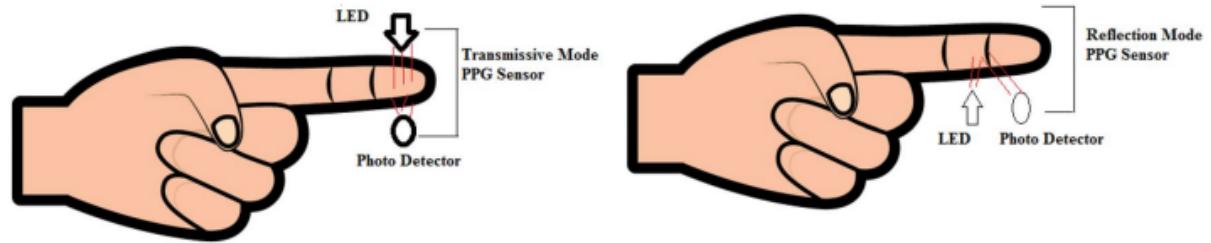


Figure 1.6 Placement of sensor in transmissive mode PPG (left) and reflection mode PPG (right)

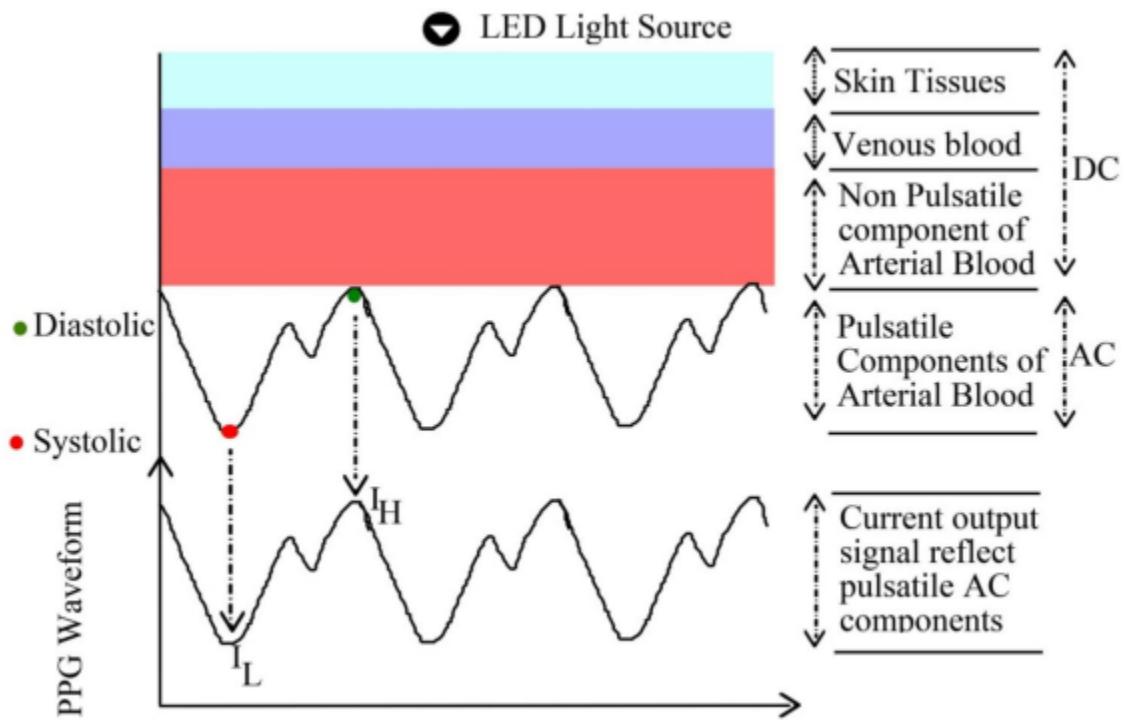


Figure 1.7 Variation in light intensity during pulsatile and non-pulsatile flow

## 1.6. Atrial Blood Pressure (ABP)

Arterial Blood Pressure (ABP) is an invasive continuous BP measurement method that has been widely accepted as the golden standard [21]. However, since the mechanism of ABP is to insert a catheter into an artery to conduct real-time BP monitoring, it is highly sensitive

to body movement, such as position changes, and it is relative to the accessed artery, which may bring risks of complications to the patient like infection [22].

Techniques such as cuff-based ABP measurement devices have been widely used for monitoring ABP. However, such measurements might cause discomfort to patients during the inflation and deflation of the cuff. This can affect the accuracy and introduce higher levels of uncertainty [23-25]. Thus, researchers and clinicians are interested in non-invasive approaches [26, 27] to measure ABP.

# **Literature Review**

## **2. Literature Review**

Several Studies have presented different methodologies for the diagnosis of heart diseases. Each study used a specific heart signal for the diagnosis and for the detection of various heart diseases. In this chapter, the related work considering the usage of ECG, PCG, PPG, and the combination between them is presented to show the performance in the detection of heart diseases.

### **2.1. Related work on ECG for heart diseases diagnosis**

Several studies have been proposed for the diagnosis of heart diseases using ECG signals. Some studies used 1-D ECG signals to detect abnormal cases, while other studies converted the 1-D ECG signal to 2D form using a method based on frequency and wavelet transforms. All the studies presented for the diagnosis of heart diseases using ECG in the last three years are based on deep learning as it showed higher performance in terms of diagnosis accuracy over other approaches.

In 2020, a methodology was proposed by Mazaheri and Khodadadi [31] for the diagnosis of 7 types of arrhythmias using machine learning approaches. The data are collected from the MIT-BIH arrhythmia dataset and classes used for diagnosis is normal sinus rhythm (NSR), atrial premature beat (APB), premature ventricular contraction (PVC), left bundle branch block (LBBB), right bundle branch block (RBBB), AF, and second degree of heart block (SDHB). The number of fragments collected from NSR, PVC, and LBBB is 31, while the APB, RBBB, and AFIB categories are 30, whereas 10 fragments are collected from the SDHB category. The pre-processing of the ECG fragments is based on gain and fixed element reduction, rescaling, and standardization. The feature extraction is based on three different sets relying on morphological features, frequency domain features, and nonlinear features. Several optimization techniques are applied for the selection of the most important features. The classification is performed using 5 classifiers which are k-nearest neighbor (KNN), feed-forward neural network (FF net), fitting neural network (Fit net), radial basis function neural network (RBFNN), and pattern recognition network (Pat net). The highest accuracy was achieved using the FF net classifier and the combination of the time, frequency, nonlinear, and morphological extracted features.

A study was also presented by Khalil and Adib [32] for the diagnosis of several arrhythmias using ECG signals. The data are gathered from the MIT-BIH arrhythmia, and the arrhythmias diagnosed in this study are 13200 (NSR), 7100 (PVC), 7200 (LBBB), 8000 (RBBB), 2500 (APC), and 7000 (PAC) fragments. The ECG signals are normalized as a pre-processing step. Then, the features are extracted using a developed model defined by ML-WCNN. The model takes the ECG fragment as an input and then the fragment is input to stationary wavelet transform (SWT) for decomposition. Then, the first 5 details and the fifth approximation are extracted and input to 1D convolutional layers, and also the ECG fragment is input in parallel to the 1D convolutional layer. Then, the output of the 1 D convolutional layers is input to batch normalization layers. The output of the batch normalization layer is input to 1D max pooling layers and the output of the max pooling layers is input to 1D global average pooling layers. The results of the global average pooling layers are concatenated or maximized and the result is input to 2 dense layers, 1 dropout, and ending with a final dense layer. The performance achieved higher results in accuracy using WCNN based on the maximum layer.

Moreover, a study was provided by Huang et al. [33] for the diagnosis of several heart diseases. The dataset applied in this study is the MIT-BIH arrhythmia, and 5 types of ECG arrhythmia are used for diagnosis. The types are NSR, LBBB, RBBB, PVC, and APC. About 540 samples are extracted from NSR, LBBB, RBBB, and APC records, while a number of 360 samples are extracted from PVC records. For pre-processing, the ECG samples are passed as an input to wavelet packet transform (WPT) and then the ECG signals are reconstructed in a range from 11.25- 22.5 Hz). Then, the reconstructed signal is passed to the FCResnet model. This model consists of three submodules which are the fast down-sampling convolutional module, the residual convolutional module, and the classification module. The first module consists of 2 (1D) convolutional layers, 2 dropout layers, and 2 batch normalization layers. The second module consists of 3 (1D) convolutional layers, 3 batch normalization layers, 1 dropout layer, and 1 max pooling layer. Finally, the classification module consists of 1 convolutional layer, 1 flattened layer, 1 dropout layer, and 3 dense layers. Two dense layers are combined with ReLU activation layers, and the last layer is combined with the Softmax classifier. Another study was presented by Xu et al. [34] for the diagnosis of 5 ECG arrhythmia which are non-ectopic beats, supraventricular ectopic

beats, ventricular ectopic beats, fusion beats, and unknown beats. The ECG records are collected from the MIT-BIH arrhythmia dataset. The ECG records are de-noised using downsampling from 300 Hz and 360 Hz to 125 Hz, followed by a normalization of the records from 0 to 1. Then, the R and T peaks are detected and the beats are extracted ending with zero padding. Then, the features are extracted using a proposed model consisting of 2 convolutional layers, 2 bi-directional long short-term memory (Bi-LSTM), 4 main residual blocks, and 2 fully connected layers. Each residual module operates in the form of a squeeze-and-excitation network (SENet). Finally, the beats are classified using the Softmax classifier. It is important to mention that the model is trained on a 2017 CinC PhysioNet dataset and the learning is transferred to the MIT-BIH arrhythmia dataset.

A study was presented by Shaker et al. [35] for the diagnosis of several ECG heartbeats. The data was collected from the MIT-BIH arrhythmia dataset, and 15 types of ECG classes are obtained from 5 main categories that are classified in this study. The categories are N, S, V, F, and Q, while the classes are NSR, LBBB, RBBB, APC, AP, NE, BAP, NP, PVC, VE, VF, VFN, FPN, and UN. The total number of heartbeats is 102098 obtained from the former classes. To balance the former classes the generative adversarial network (GAN) is applied. A deep learning approach is applied to extract features. The model consists of 3 inception models, followed by 3 fully connected layers. The inception model consists of 4 (1D) convolutional layers ending with a filter concatenation and max pooling layer. Finally, the diagnosis result is based on the Softmax classifier. In 2021, a methodology was presented by Eltrass et al. [36] for the diagnosis of 3 ECG types which are NSR, AF, and CHF. The ECG signals are collected from three datasets which are the MIT-BIH ARR database, MIT-BIH normal database, and the BIDMC CHF. The number of records collected is 162 and each record is divided into 6 fragments. Each fragment is 10,000 samples forming a total of 972 fragments. The ECG signals are filtered using normalization and a multi-stage kernel adaptive filter to remove the baseline wander and power line interference. Then, the 1D ECG signals are divided into fragments, and each fragment converts to a 2D form using CQ-NSGT. Then, the 2D images are input to Alexnet for feature extraction and then the classification is performed using the Softmax classifier. A comparison is performed between the performance of the CQ-NSGT with Alexnet and CWT with Alexnet. The results showed the performance of the CQ-NSGT with Alexnet with higher accuracy of 5.93%.

Another study was presented by Rath et al. [37] for the diagnosis of normal and abnormal heart diseases. The subjects are collected from two databases which are MIT-BIH arrhythmia and PTB. The number of ECG records is unbalanced in each dataset and to balance them generative adversarial network (GAN) is used to generate other records. The GAN forms 428 and 72 records for training and testing concerning the MIT-BIH arrhythmia, while the PTB database after applying GAN contains 551 for training and 149 for the test. A deep learning model based on (GAN-LSTM) was used for feature extraction. In other words, the original data and the produced GAN ECG data are input into 3 LSTM layers, 2 dropout layers, and 1 Softmax activation function. The performance of the GAN-LSTM achieved the highest accuracy on the two datasets. A comparison is applied between the GAN-LSTM model with SVM, LSTM, NB, MLP, and GAN discriminator, and the GAN-LSTM reached the highest accuracy in comparison with other methods.

Another study was provided by Panganiban et al. [38] for the diagnosis of several abnormal heart diseases relying on ECG signals. The data were obtained from four main datasets which are MIT-BIH AF, PAF prediction challenge, Challenge 2015 Training set, and Fantasia databases. The records collected from the former database include AF, normal, bradycardia, BBB, and tachycardia. The records are divided into fragments and each fragment is passed to CWT to generate the spectrogram of the fragments. Then, the spectrogram images are passed to the Inception V3 model for feature extraction and softmax for classification. Two main experiments (Binary and multi-classes) are performed on the records obtained from the datasets, and the accuracy of the model on the multi-class is higher than that of the binary class experiment.

Another study was presented by Qin et al. [39] for the diagnosis of various ECG diseases. The ECG records are collected from two main datasets. The first one is the 2018 China Physiological Signal Challenge's training set and it consists of 9 categories forming 5850 records. The nine categories are normal, AF, I-AVB, LBBB, RBBB, PAC, PVC, ST-segment depression, and ST-segment elevated (STE). The second dataset is the 2018 China Physiological Signal Challenge and it consists of 73 categories forming 291 recordings. The recordings are filtered using Butterworth low pass filter, and the features are extracted using a deform-CNN model. The deform model starts with 2 convolutional layers, and 1

deformable convolution, followed by several sublayers. Each sublayer consists of 2 convolutional layers, 1 max-pooling layer, then convolutional in the x and y direction ending with bilinear interpolation. Finally, the output of the last sublayer is input to 2 dense layers and the classification is using the softmax classifier. The data are divided into 60% training, 20% validation, and 20% testing. The performance of the model achieved a higher accuracy on the second dataset.

Tyagi and Mehra [40] applied a methodology for the diagnosis of 16 ECG classes obtained from the MIT-BIH arrhythmia dataset. The classes are NSR, LBBB, RBBB, APC, PVC, PACE, atrial premature heal (AP), ventricular flutter (VFR), a fusion of ventricular and normal beats, blocked atrial premature beat (BAP), nodal, a fusion of paced and normal beta (FPN), ventricular escape beat (VE), nodal premature beat (NP\_J), atrial escape beat (AE-e), and Linclassifkable beat (LK). The number of heartbeats collected from the 16 classes is 110100. They are filtered using DWT based on a haar wavelet and a smoothing algorithm is applied to remove the remaining noises. The R peaks and the QRS fragments are segmented and the RR interval is captured from the records. The features are extracted using a CNN model relying on 2 convolutional layers, 2 pooling layers, and 2 fully connected layers. The parameter of the CNN model is optimized using the Grasshopper optimization algorithm (GOA). Finally, the classification is performed using a softmax classifier.

In 2022, various studies based on deep learning are provided for the diagnosis of several heart diseases. Dixit and Kala [41] presented a study for the diagnosis of normal and congestive heart failure (CHF) ECG classes. The data were collected from five datasets which are congestive heart failure RR interval (CHF-RR), BIDMC congestive heart failure (BIDMC-CHF), normal sinus rhythm RR interval (NSR-RR), Fantasia, and MIT-BIH (NSR). The ECG records are divided into 3 types of fixed time length fragments. The 3 fragment types are 500, 1000, and 2000 time durations in seconds. The first type has a total of 25498 fragments, while the second type has a total of 12698 fragments, and the third type has a total of 6347 fragments. The fragments are de-noised using a normalization process. Then, a proposed DL model named CNN-LSTM-AM is used for feature extraction and classification. The CNN part involves a convolution layer, and pooling layer, while the LSTM part includes two Bi-LSTM layers ending with an attention layer. The performance using

fragments relying on 2000 time durations achieved the highest accuracy. A study was presented by Monedro et al. [42] for the diagnosis of 13 types of ECG classes. The class types are complete arrhythmia due to AF, 1dAVb, wolf Parkinson, long QT, short QT, Nodal, Incomplete RBB, Complete RBB, incomplete RBB with narrow QRS, sinus tachycardia, bradycardia, arrhythmia, and other cardiac arrhythmias. The dataset was obtained from a pricing investment company. The number of records collected is 284,000 and they are recorded in an SCP format. The records are filtered using the 2nd order butter worth filter to remove the common noises below 0.5 Hz, and then the records are passed to DWT to soften the signal and eliminate the invalid frequencies. A set of rules are added to identify the P, Q, R, S, and T waves. Then, a rule-based decision system is provided to determine the diagnosis type. Finally, the decision tree classifier is used to supply a diagnosis accuracy.

Anand et al., 2022 [43] diagnosed several heartbeats based on ECG signals. The data were collected from two datasets which are PTB-XL and Arrhythmia. The PTB-XL dataset consists of 5 classes which are conduction disturbance (CD), hypertrophy (HYP), MI, normal ECG (norm), and ST/T change (STTC). The arrhythmia dataset consists of 4 classes which are AF, grouped supraventricular tachycardia (GSVT), sinus bradycardia (SB), and sinus rhythm (SR). The records are filtered using the normalization process. Then, a model was proposed based on ST-CNN-GAP-5. This model consists of 6 (2D) convolutional layers, 7 batch normalization, 7 ReLU activation layers, 5 (2D) max pooling, 1 (2D) global average pooling, 1 dense layer, 1 flatten layer, and 1 dropout layer. Finally, the classification process is based on a sigmoid classifier. The performance of the model on the arrhythmia dataset is higher in accuracy than PTB-XL.

Rath et al. [44] presented a study for the diagnosis of two main ECG classes which are normal and abnormal. The authors collected the data from two main datasets which are PTB-ECG and MIT-BIH. The number of collected records from the first dataset is 200 divided into 140 training and 60 validation, while the records of the second dataset are 268 divided into 188 training and 80 validation. The records are normalized before any further processing. Then, 4 main deep learning models are applied for feature extraction. The first is based on an auto encoder (AE), radial bias network (RFBN), self-organizing map (SOM), and restricted Boltzmann machine (RBM). The highest performance was achieved using the

ensemble of the AE and SOM. Finally, the accuracy resulting from the first dataset was higher than that of the second dataset.

A study was also presented by Shin et al. [45] for the diagnosis of different ECG heartbeats. The authors obtained the data from the MIT-BIH database. It holds about 17 classes or ECG categories with a 10-sec duration obtained from 45 participants. The number of classes used for diagnosis in this study is 4 and they are 283 NSR, 135 AF, 133 PVC, and 103 LBBB records. Each record is passed to a Butterworth notch and moving-average filters for denoising and then the records are passed to a matching pursuit algorithm to balance the data. Then, the wavelet transform is used on the ECG records to obtain the scalogram. The scalogram images are input to the MobileNetV2-BiLSTM structure for feature extraction. Finally, the features are input to Softmax for classification. Table 3.1 presents the most recent studies that are published in the last 3 years for the diagnosis of heart diseases. The table shows the whole methodology presented in each study in terms of data acquisition, filtration of the ECG signals, feature extraction, classification, and final diagnosis performance.

**Table 2.1** Literature review of the studies designed for the diagnosis of heart diseases based on ECG signals

<b>Authors &amp; Year</b>	<b>Dataset &amp; classes &amp; Pre-processing &amp; Feature Extraction</b>	<b>Classification</b>	<b>Results</b>
Mazaher and khodadadi [31] 2020	MIT-BIH Arrhythmia 7 classes 71 F Gain and fixed element reduction + rescaling and standardization + Morphological + Frequency + Non-linear features	KNN, FF net, Fit net, RBFNN, Pat net	10 Fold cross-validation A = 98.75% using Pat net
Khalil and Adib [32] 2020	MIT-BIH Arrhythmia 6 classes 45000 F Normalization ML-WCNN (1D-CNN + SWT)	Softmax	10 Fold cross validation A = 99.75%
Huang et al. [33] 2020	MIT-BIH Arrhythmia 5 classes 2520 Sam Wavelet packet transform FCResnet (Fast down-sampling convolution + residual module * 3 + convolution + classification module )	Softmax	Train: 2100 Sam Test: 420 Sam A = 98.79%

Xu et al. [34] 2020	The training set of the 2017 PhysioNet /CinC Challenge + MIT-BIH Arrhythmia 5 classes Down sampling + Normalization + T episode and R-peaks detection + Beat extraction + Zero padding + Combination of CNN and RNN architectures + SENet	Softmax	Train: 800 R Remaining records for the test A = 95.90% SEN = 95.90% SPEC = 96.34%
Shaker et al. [35] 2020	MIT-BIH Arrhythmia 15 classes 102098 HB Butterworth band pass filter + GAN Deep learning approach based on the Inception Module	Softmax	Train: 21077 HB Test: 81021 HB P = 90. 0% A = 98.0% SPEC = 97.4% SEN = 97.7%
Eltass et al. [36] 2021	MIH-BIH ARR dataset + MIT-BIH NSR dataset + BIDMC CHF dataset 3 classes 972 F Normalization + multi-stage kernel adaptive filter design CNN (Alexnet) + CQ-NSGT + MLP	Softmax	k- fold cross validation A = 98.82% SPEC = 99.21% SEN = 98.87% P = 99.20%
Rath et al. [37] 2021	MIT-BIH arrhythmia 2 classes 500 S PTB-ECG dataset 2 classes 700 S Normalization + SVM + NB + MLP + GAN + GAN- LSTM Model	LSTM	GAN-LSTM Model For MIT-BIH: A = 99.2 % F1-Score = 98.7% AUC = 0.984 For PTB: A = 99.4 % F1-Score = 99.3% AUC = 0.995
Panganiban, et al. [38] 2021	MIT-BIH AF, PAF Prediction Challenge, PTB, Challenge 2015Training set, and Fantasia databases 5 classes 323 Records Records are cut and converted to text files in the form of samples in CSV files CWT + InceptionV3	Softmax	Binary class: Total : 4248 Spectrogram Images A = 98.73% PPV: 96.83% Multi-class: Total: 4248 Spectrogram Images A= 97.33% F1-Score = 99.21%
Qin et al. [39] 2021	CPSC-2018 9 classes 5850 R CPSC-Extra 73 classes 291 R Butterworth low pass filter + Deform-CNN	Softmax	Training = 60% Validation = 20% Testing = 20% For CPSC-2018: A = 86.3% For CPSC-Extra A = 89.7%
Tyagi and Mehra [40] 2021	MIT-BIH Arrhythmia Database 16 classes 110100 HB DWT CNN-GOA	Softmax	Training = 73006 HB Test = 37094 HB A = 99.58% Error Rate = 0.42%

Dixit and Kala [41] 2022	CHF-RR, BIDMC-CHF, Fantasia, MIT-BIH-NSR datasets 2 class 25498 F Normalization + DCNN-LSTM- AM	LSTM	A = 99.52% SEN = 99.31% SPEC = 99.28% F-Score = 98.94% AUC = 99.9%
Monerdero [42] 2022	ECG database provided by Preving Investment company 13 classes 284000 R +2 <sup>nd</sup> order Butterworth bandpass filter + DWT relying on Mallat's pyramid algorithm Identification of R, P, Q, S, T waves + extracting mean, median, min, and max from the waves + Rule-based system for diagnosis	Decision Tree	A = 80 %
Anand et al. [43] 2022	PTB-XL dataset 5 classes 27826 R + Arrhythmia dataset 4 classes 10588 R + Normalization + ST-CNN-GAP-5	Sigmoid	PTB-XL: A = 89.73% AUC = 93.41% Arrhythmia: A = 95.84% AUC = 99.46%
Rath, A. et al. [44] 2022	PTB-ECG 2 classes 200 R MIT-BIH 2 classes 268 R Normalization + AE, RBFN, SOM, RBM + Ensemble of AE and SOM	SOM	PTB-ECG: Train = 140 Validation= 60 SOM-AE A = 0.992% F1-Score = 98.6% AUC = 99.5% MIT-BIH: Train = 188 Validation = 80 SOM-AE A = 0.984% F1-Score = 97.1% AUC = 99.7%
Shin et al. [45] 2022	MIT-BIH database 4 classes 654 R Resizing + digital filter + Data Augmentation + Scalogram MobileNetV2 + BiLSTM	Softmax	A = 91.7% SEN = 92.0% SPEC = 91.0% Precision = 92.0% F1-Score = 92.0%

S: Subjects Sam: Samples F: Fragments R: Records HB: Heartbeats

## 2.2. Related work on PCG for heart diseases diagnosis

Current research in this field is focused on the development of suitable algorithms, which in the future may lead to development of an intelligent stethoscope. Because of the nature

of PCG signals and undesired noise during examination, it is important to divide the diagnosis process into two steps. The first being the processing of original signals aimed to extract features, which would help to distinguish all of the types of signals, and the second associated with the process of signal classification. There were several successful attempts to develop such systems [46-57], where the majority of works were focused on the application of techniques based on Artificial Neural Networks (ANN) and Support Vector Machines (SVM).

One of the earliest heart valve disease detection systems based on ANN was developed by Turkoglu, Arslan and Ilkay [46], who used wavelet entropy and short-time Fourier transform to determine specific features of heart signals, consequently obtaining classification accuracy of 94% for normal heart sounds, and 95.9% for pathological ones. Wavelet analysis of the PCG signal in combination with homomorphic filtering and K-means clustering method was presented by Gupta et al. (leading to 97% accuracy in distinguishing two abnormal and one normal heart states) [47]. A multilevel wavelet decomposition with a multilayer perceptron trained by a back-propagation algorithm achieved 94.42% of accuracy in identifying four heart states [48]. Other works include the use of multivariate matching pursuit to model murmurs and classifying them with a three-layer feed-forward perceptron network with 92.5% of accuracy (distinguishing normal from abnormal heart states) [49] or a combination of detection of characteristic heart features (activity, complexity, mobility and spectral peaks) with ANN, providing a rate of 98% in identification, however able to distinguish only three of them [50].

Another group of methods employ the SVM as the main classifier. An approach for heart sounds identification presented by Wu et al. ensured 95% of accuracy using wavelet transform to extract the envelope of PCG signals [51]. However, the authors were able to distinguish only normal from abnormal heart states. The same results were achieved by Jiang and Choi [52] who developed a system for in-home use, however, this system was proven only by a case study. A diagnosis system based on principle component analysis connected with an adaptive network was developed by Avci and Turkoglu [53]. In this case the system ensured 96% accuracy in classification of normal and 93.1% of two abnormal heart states. Later, Avci improved this system and developed genetic Support Vector

Machines, which gave 95% of accuracy [54]. There are also examples of using wavelet transforms and short time Fourier transform methods for feature extraction [55] or using wavelet decomposition to distinguish two out of five types of heart states with up to 93.42% accuracy [56].

As per the literature study, many workers are researching in this field. Many researchers have described different algorithms and categorization methods for cardiac sounds. Randhawa and Singh [57] surveyed methods for automatically analyzing and classifying cardiac sounds. Gupta et al. [58] worked on identifying S1 & S2 Heart Sounds using a Stacked Auto Encoder with the combinatory feature for categorizing fundamental Heart Sounds. Then, the CNN approach is also used. This method has restrictions on analyzing and detecting S3 Heart Sounds. Lecun et al. [59] proposed an innovative algorithm to detect heart and lung sounds (HLS) based on the variational mode decomposition (VMD) of PCG signal. It is helpful to detect the third Heart Sound. However, this algorithm has limitations in real-time environment applications. Roy and Roy [60] proposed a new method of extracting PCG signals from noise through sparse recovery analysis. Barma et al. [61] did focus on the detection of second Heart Sound, which engages with computing the time duration and the energy of normalized instantaneous frequencies of A2s (Aortic valve) and P2s (Pulmonary valve), but it could not categorize cardiac sounds.

Tavel [62] and Muduli et al. [63] did work on Heart Sound analysis using the discrete wavelet transform (DWT) algorithm. It has specific restrictions in real-time analysis. Roy et al. [64] made PCG signal analysis using Shannon Energy Envelop and DWT method, but it cannot classify the Heart Sounds efficiently. Mishra et al. [65] made heart defect analysis and classification using Pattern Recognition Technique, but the method is based on echo imaging. Mishra et al. [66] made PCG signal analysis using classification by feature extraction but not in the real-time analytical process. El-Segaier et al. [67] made Heart Sound analysis and its categorization using a fuzzy classifier-based soft computing method. The laboratory study was done on a readily available online Heart Sound repository. It was an offline study and not validated with human volunteers. Nygaard et al. [68] studied the classification of PCG signals. Anju and Sanjay Kumar [69] did a segmentation and categorization of PCG signals for Heart Sound analysis. There is no

significant work on developing real-time screening of valvular diseases. Buchanna et al. [70] studied the classification of EEG signals using fractal analysis and support vector regression. Ye et al. [71] analyzed multimodal medical images using semantic segmentation maps generated through deep learning.

### **2.3. Related work on PPG for heart diseases diagnosis**

Sukor et al. [72] discussed the presence and removal of signal artifacts in PPG signals using the waveform morphology method and achieved an accuracy of  $83 \pm 11\%$  in the detection of CVD. CVDs are becoming one of the top causes of death worldwide. Approximately 16.7 million deaths worldwide, or 29.2% of all deaths, have occurred as a result of the different types of CVDs over the past ten years. All things considered, doctors and wellbeing experts advocate for the early discovery and treatment of potential side effects to guarantee the shirking of an out-and-out heart-related disease [73]. The main disadvantage of PPG signals is that they are highly corrupted by noise components, such as skin artifacts and motion artifacts. Hence, the preprocessing of PPG signals to attain cleaning signals is, itself, a broad area of research. Tun [74] designed an FIR filter for a heart rate detection system, which helps to minimize motion artifacts in a light-based measuring system. The reduction in artifacts in PPG signals using an AS-LMS adaptive filter was discussed by Ram et al. [75] and Luke et al. [76] presented a methodology using efficient signal processing algorithms for artifact removal in PPG signals and attained an SNR value of 41.52 db. From the analysis of PPG signals, the common parameters extracted are the pulse transit time, heart rate, stiffness index, and pulse wave velocity. Large artery stiffness and pulse width, pulse interval, systolic amplitude, augmentation index, and peak-to-peak interval are certain characteristics extracted from PPG signals [77]. Shintomi et al. [78] used mobile and wearable sensor equipment to measure the effectiveness of the heartbeat error and compensation strategies on heart rate variability (HRV). PPG is frequently used instead of an electrocardiogram (ECG) to assess heart rate in wearable devices. However, there are inherent differences between PPG and ECG due to the fact that PPG is affected by body motions, vascular stiffness, and blood pressure variations. Using ECG and PPG readings obtained from 28 people, the methods described in [79] determined how these errors affected the analysis of HRV. The assessment's

findings demonstrate that the error compensation method enhances the precision of HRV analysis in both the time and frequency domains as well as in nonlinear analysis. When compared to recent ECG observing systems, PPG signal measurements are more accessible and require less hardware for signal acquisition [80]. PPG does not require a reference signal, making it possible to integrate PPG sensors with wristbands. As a result, it can be employed in a variety of studies for the investigation and diagnosis of CVDs [80]. Moshawrab et al. [81] discussed how CVDs can be detected and predicted using smart wearable devices. In addition, the authors indicated that a review of the evaluation of the development and usage of smart wearable devices for the management of CVD demonstrates the high efficacy of such smart wearable devices.

Certain works associated with PPG signal exploration and their use in various biomedical areas are provided below. For approximately three decades, research on PPG signals with respect to CVDs has been very interesting, and many research outcomes have been reported. Allen [82] examined PPG and how it could be used in clinical physiological measurements through the clinical monitoring of the human body, autonomic function, and vascular assessment. Sunil Kumar and Harikumar [83] used parameters, such as independent component analysis (ICA), principal component analysis (PCA), entropy, and mutual information (MI) to analyze the PPG signals. PPG has proven to be the most capable method for the early screening of heart-related diseases. Almarshad et al. [84] disclosed the diagnostic properties of PPG signals in addition to their prospective clinical applications in healthcare and assessed the possible effects of PPG signals on the screening, monitoring, diagnosis, and fitness of inpatients and outpatients. Yousefi et al. [85] developed a technique to automatically detect premature ventricular contraction (PVC) from the extracted features of PPG signals using higher-order statistics (HOS) and the chaotic method for the KNN classifier, with a classification accuracy of 95.5%. Cardiac health supervision based on PPG through smart phones or wearable sensors has the potential to attain high accuracy with fewer false alarms, as mentioned by Ukil et al. [86]. With the help of the ensemble method, Almanifi et al. [87] used a computer-vision approach to find human activities using PPG signals. The accuracy of the PPG was 88.91%, which shows that wrist PPG data could be used with the ensemble method in human activity

recognition (HAR) to make accurate detections. The time domain analysis of a PPG signal and its second derivative was performed by Paradkar et al. [88] to extract the features from the PPG signal, and these extracted features were classified by the SVM classifier to detect coronary artery disease (CAD) with a sensitivity of 85% and a specificity of 78%. Neha et al. [89] investigated the detection of arrhythmias using PPG signals. In this analysis, a low-pass Butterworth filtering method was used to remove artifacts, and the extracted features were applied to various machine learning algorithms to classify normal and abnormal pulses. The results show that the SVM classifier had a high accuracy of 97.674% in identifying arrhythmia pulses. Prabhakar et al. [90] analyzed metaheuristic-based optimization algorithms as dimensionality reduction techniques and then applied various classifiers to these dimensionally reduced values for the classification of CVD. The results in [90] show that, for chi-squared probability density function (PDF) optimized values, the artificial neural network (ANN) classifier attained a maximum classification accuracy of 99.48% for normal subjects, and the logistic regression classifier produced a maximum classification accuracy of 99.48% for CVD cases. In order to analyze CVD patients utilizing PPG signals, Sadad et al. [91] applied various machine learning techniques and a deep learning model to create a system that helps doctors with continuous monitoring, and they obtained an accuracy of 99.5%.

## 2.4. Combination of Heart Signals for the diagnosis of Diseases

A study presented by Fatima Chakir et al. [92] on recognizing cardiac abnormalities using synchronized ECG and PCG signals aimed to enhance the diagnosis of heart conditions by leveraging the combined data from these two types of signals. The methodology involved several critical steps: data collection from the PhysioNet/Computing in Cardiology Challenge 2016, processing and segmentation of PCG recordings to identify S1 and S2 heart sounds, and detecting the QRS complex in ECG signals using the Pan-Tompkins algorithm. The researchers implemented custom instructions to refine QRS detection and enhance the identification of heart sound components. Ten relevant biomarkers were extracted from the synchronized signals, including systole and diastole lengths, R wave amplitudes, and the zero-crossing rate of heart sound signals. These features were then used in various classification models, including Naive Bayes, K-Nearest Neighbors,

Support Vector Machine (SVM), Random Forest, and Linear Discriminant Analysis, to diagnose heart abnormalities. The study found that combining ECG and PCG signals significantly improved diagnostic performance compared to using PCG alone. Specifically, the SVM classifier achieved the highest accuracy (92.5%), AUC (0.95), and specificity (92.86%) when using the combined features, demonstrating the superiority of this integrated approach for heart condition assessment.

Also another study presented by Roghayyeh Arvanaghi et al. [93] applied a fusion of ECG and Atrial Blood Pressure (ABP) signals introduced a novel wavelet-based method to enhance the classification of cardiac arrhythmias. This study utilized data from the MIMIC PhysioNet database, focusing on recordings from critical care units. The researchers applied Discrete Wavelet Transform (DWT) using the Daubechies mother wavelet to decompose and denoise ECG and ABP signals. The fusion process involved combining the approximation coefficients of the ECG signal with the detail coefficients of the ABP signal, followed by extracting frequency domain features such as maximum frequency, average periodogram, mean frequency, and median frequency from the fused signal. These features were statistically evaluated using the Wilcoxon rank sum test to assess their significance in distinguishing between normal and abnormal cardiac conditions. The classification was performed using a Multilayer Perceptron Neural Network (MLPNN) with an optimized structure of three hidden layers. The results demonstrated that the fusion technique markedly improved classification accuracy across all tested categories of cardiac arrhythmias. For instance, the accuracy for classifying two classes (healthy and unhealthy) reached 96.6% with the fused signal, compared to 88.7% using only the ECG signal. The study concluded that integrating ECG and ABP signals through the proposed fusion method significantly enhances the precision of cardiac arrhythmia classification, highlighting the potential of multimodal signal analysis in clinical diagnostics.



# Deep Learning

## 3. Deep learning

Deep learning (DL) is a branch of machine learning (ML) and artificial intelligence (AI) is nowadays considered as a core technology of today's Fourth Industrial Revolution. Due to its learning capabilities from data, DL technology originated from artificial neural network (ANN), has become a hot topic in the context of computing, and is widely applied in various application areas like healthcare, visual recognition, text analytics, cybersecurity, and many more. However, building an appropriate DL model is a challenging task, due to the dynamic nature and variations in real-world problems and data. Moreover, the lack of core

understanding turns DL methods into black-box machines that hamper development at the standard level. This article presents a structured and comprehensive view on DL techniques including a taxonomy considering various types of real-world tasks like supervised or unsupervised. In our taxonomy, we consider deep networks for supervised or discriminative learning, unsupervised or generative learning as well as hybrid learning and relevant others. We also summarize real-world application areas where deep learning techniques can be used. Finally, we point out ten potential aspects for future generation DL modeling with research directions. Overall, this article aims to draw a big picture on DL modeling that can be used as a reference guide for both academia and industry professionals.

Nowadays, Deep Learning technology is considered as one of the hot topics within the area of machine learning, artificial intelligence as well as data science and analytics, due to its learning capabilities from the given data. Many corporations including Google, Microsoft, IBM, etc., study it actively as it can provide significant results in different classification and regression problems and datasets. In terms of working domain, DL is considered as a subset of ML and AI, and thus DL can be seen as an AI function that mimics the human brain's processing of data. The worldwide popularity of "Deep learning" is increasing day by day. Deep learning differs from standard machine learning in terms of efficiency as the volume of data increases

DL is derived from the conventional neural network but considerably outperforms its predecessors. Moreover, DL employs transformations and graph technologies simultaneously to build up multi-layer learning models. The most recently developed DL techniques have obtained good outstanding performance across a variety of applications, including audio and speech processing, visual data processing, natural language processing (NLP).

### **3.1. Why deep learning?**

- Universal Learning Approach: Because DL can perform in approximately all application domains, it is sometimes referred to as universal learning.
- Robustness: In general, precisely designed features are not required in DL techniques. Instead, the optimized features are learned in an automated fashion

related to the task under consideration. Thus, robustness to the usual changes of the input data is attained.

- Generalization: Different data types or different applications can use the same DL technique, an approach frequently referred to as transfer learning (TL) which explained in the latter section. Furthermore, it is a useful approach in problems where data is insufficient.
- Scalability: DL is highly scalable. ResNet, which was invented by Microsoft, comprises 1202 layers and is frequently applied at a supercomputing scale. Lawrence Livermore National Laboratory (LLNL), a large enterprise working on evolving frameworks for networks, adopted a similar approach, where thousands of nodes can be implemented.

#### Understanding Various Forms of Data for Deep learning models

As DL models learn from data, an in-depth understanding and representation of data are important to build a data-driven intelligent system in a particular application area. In the real world, data can be in various forms, which typically can be represented as below for deep learning modeling:

- Sequential Data Sequential data: is any kind of data where the order matters, i.e., a set of sequences. It needs to explicitly account for the sequential nature of input data while building the model. Text streams, audio fragments, video clips, time-series data, are some examples of sequential data.
- Image or 2D Data: A digital image is made up of a matrix, which is a rectangular array of numbers, symbols, or expressions arranged in rows and columns in a 2D array of numbers. Matrix, pixels, voxels, and bit depth are the four essential characteristics or fundamental parameters of a digital image.
- Tabular Data: A tabular dataset consists primarily of rows and columns. Thus, tabular datasets contain data in a columnar format as in a database table. Each column (field) must have a name and each column may only contain data of the defined type. Overall, it is a logical and systematic arrangement of data in the form of rows and columns that are based on data properties or features. Deep learning models can learn efficiently on tabular data and allow us to build data-driven intelligent systems.

### 3.2. DL Properties and Dependencies

A DL model typically follows the same processing stages as machine learning modeling. we have shown a deep learning workflow to solve real-world problems, which consists of three processing steps, such as data understanding and preprocessing, DL model building, and training, and validation and interpretation. However, unlike the ML modeling, feature extraction in the DL model is automated rather than manual. K-nearest neighbor, support vector machines, decision tree, random forest, naive Bayes, linear regression, association rules, k-means clustering, are some examples of machine learning techniques that are commonly used in various application areas. On the other hand, the DL model includes convolution neural network, recurrent neural network, autoencoder, deep belief network, and many more, discussed briefly with their potential application areas in Section 3. In the following, we discuss the key properties and dependencies of DL techniques, that are needed to consider before started working on DL modeling for real-world applications.

- Data Dependencies Deep learning is typically dependent on a large amount of data to build a data-driven model for a particular problem domain. The reason is that when the data volume is small, deep learning algorithms often perform poorly. In such circumstances, however, the performance of the standard machine-learning algorithms will be improved if the specified rules are used.
- Hardware Dependencies the DL algorithms require large computational operations while training a model with large datasets. As the larger the computations, the more the advantage of a GPU over a CPU, the GPU is mostly used to optimize the operations efficiently. Thus, to work properly with the deep learning training, GPU hardware is necessary. Therefore, DL relies more on high-performance machines with GPUs than standard machine learning methods.
- Feature Engineering Process Feature engineering is the process of extracting features (characteristics, properties, and attributes) from raw data using domain knowledge. A fundamental distinction between DL and other machine-learning techniques is the attempt to extract high-level characteristics directly from data. Thus, DL decreases the time and effort required to construct a feature extractor for each problem.
- Model Training and Execution time in general, training a deep learning algorithm takes a long time due to a large number of parameters in the DL algorithm; thus, the model training process takes longer. For instance, the DL models can take more than one week to complete a training session, whereas training with ML algorithms takes relatively little time, only seconds to hours. During testing, deep learning algorithms take extremely little time to run, when compared to certain machine learning methods.

- Black-box Perception and Interpretability Interpretability is an important factor when comparing DL with ML. It is difficult to explain how a deep learning result was obtained, On the other hand, the machine-learning algorithms, particularly, rule-based machine learning techniques provide explicit logic rules (IF-THEN) for making decisions that are easily interpretable for humans. For instance, in our earlier works, we have presented several machines learning rule-based techniques, where the extracted rules are human-understandable and easier to interpret, update or delete according to the target applications.

### **3.3. Types of Deep Learning networks**

#### **3.3.1. Recursive neural networks**

RvNN can achieve predictions in a hierarchical structure also classify the outputs utilizing compositional vectors. Recursive auto-associative memory (RAAM) is the primary inspiration for the RvNN development. The RvNN architecture is generated for processing objects, which have randomly shaped structures like graphs or trees. This approach generates a fixed-width distributed representation from a variable-size recursive-data structure. The network is trained using an introduced backpropagation through structure (BTS) learning system. The BTS system tracks the same technique as the general-back propagation algorithm and can support a treelike structure. Auto-association trains the network to regenerate the input-layer pattern at the output layer. RvNN is highly effective in the NLP context introduced RvNN architecture designed to process inputs from a variety of modalities. These authors demonstrate two applications for classifying natural language sentences: cases where each sentence is split into words and nature images, and cases where each image is separated into various segments of interest. RvNN computes a likely pair of scores for merging and constructs a syntactic tree. Furthermore, RvNN calculates a score related to the merge plausibility for every pair of units. Next, the pair with the largest score is merged within a composition vector. Following every merge, RvNN generates (a) a larger area of numerous units, (b) a compositional vector of the area, and (c) a label for the class (for instance, a noun phrase will become the class label for the new area if two units are noun words). The compositional vector for the entire area is the root of the RvNN tree structure. RvNN has been employed in several applications.

#### **3.3.2. Recurrent neural networks**

RNNs are a commonly employed and familiar algorithm in the discipline of DL. RNN is mainly applied in the area of speech processing and NLP contexts. Unlike conventional networks, RNN uses sequential data in the network. Since the embedded structure in the sequence of the data delivers valuable information, this feature is fundamental to a range of different applications. For instance, it is important to understand the context of the sentence in order to determine the meaning of a specific word in it. Thus, it is possible to consider the RNN as a unit of short-term memory, where  $x$  represents the input layer,  $y$  is the output layer, and  $s$  represents the state (hidden) layer. For a given input sequence, a typical unfolded RNN introduced three different types of deep RNN techniques, namely “Hidden-to-Hidden”, “Hidden-to-Output”, and “Input-to-Hidden”. A deep RNN is introduced that lessens the learning difficulty in the deep network and brings the benefits of a deeper RNN based on these three techniques.

However, RNN’s sensitivity to the exploding gradient and vanishing problems represent one of the main issues with this approach. More specifically, during the training process, the reduplications of several large or small derivatives may cause the gradients to exponentially explode or decay. With the entrance of new inputs, the network stops thinking about the initial ones; therefore, this sensitivity decays over time. Furthermore, this issue can be handled using LSTM. This approach offers recurrent connections to memory blocks in the network. Every memory block contains a number of memory cells, which have the ability to store the temporal states of the network. In addition, it contains gated units for controlling the flow of information. In very deep networks, residual connections also have the ability to considerably reduce the impact of the vanishing gradient issue which explained in later sections. CNN is considered to be more powerful than RNN. RNN includes less feature compatibility when compared to CNN.

### **3.3.3. Convolutional neural networks**

In the field of DL, the CNN is the most famous and commonly employed algorithm. The main benefit of CNN compared to its predecessors is that it automatically identifies the relevant features without any human supervision. CNNs have been extensively applied in a range of different fields, including computer vision, speech processing, Face Recognition, etc. The

structure of CNNs was inspired by neurons in human and animal brains, similar to a conventional neural network. More specifically, in a cat's brain, a complex sequence of cells forms the visual cortex; the CNN simulates this sequence. Goodfellow et al. identified three key benefits of the CNN: equivalent representations, sparse interactions, and parameter sharing. Unlike conventional fully connected (FC) networks, shared weights and local connections in the CNN are employed to make full use of 2D input-data structures like image signals. This operation utilizes an extremely small number of parameters, which both simplifies the training process and speeds up the network. This is the same as in the visual cortex cells. Notably, only small regions of a scene are sensed by these cells rather than the whole scene (i.e., these cells spatially extract the local correlation available in the input, like local filters over the input).

A commonly used type of CNN, which is similar to the multi-layer perceptron (MLP), consists of numerous convolution layers preceding sub-sampling (pooling) layers, while the ending layers are FC layers. An example of CNN architecture for image classification is illustrated in the Figure below.

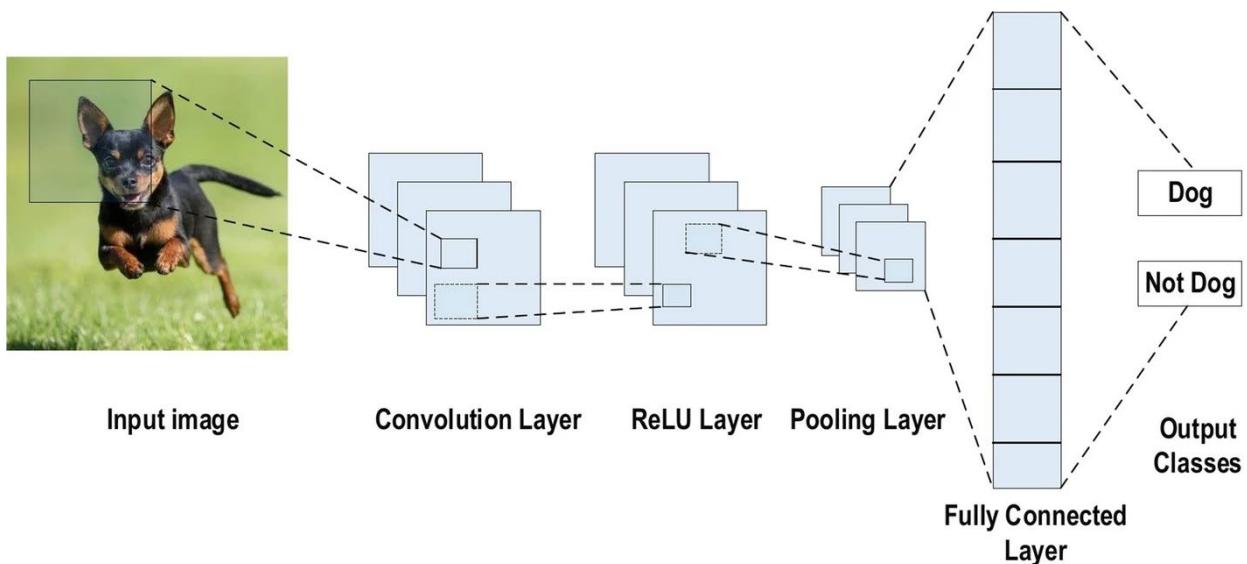


Figure. 3.1 Main architecture of the CNN

### 3.3.4. Benefits of employing CNNs

The benefits of using CNNs over other traditional neural networks in the computer vision environment are listed as follows:

- The main reason to consider CNN is the weight sharing feature, which reduces the number of trainable network parameters and in turn helps the network to enhance generalization and to avoid overfitting.
- Concurrently learning the feature extraction layers and the classification layer causes the model output to be both highly organized and highly reliant on the extracted features.
- Large-scale network implementation is much easier with CNN than with other neural networks.

### 3.3.5. CNN layers

The CNN architecture consists of a number of layers (or so-called multi-building blocks).

Each layer in the CNN architecture, including its function, is described in detail below.

1. Convolutional Layer: In CNN architecture, the most significant component is the convolutional layer. It consists of a collection of convolutional filters (so-called kernels). The input image, expressed as N-dimensional metrics, is convolved with these filters to generate the output feature map.
  - Kernel definition: A grid of discrete numbers or values describes the kernel. Each value is called the kernel weight. Random numbers are assigned to act as the weights of the kernel at the beginning of the CNN training process. In addition, there are several different methods used to initialize the weights. Next, these weights are adjusted at each training era; thus, the kernel learns to extract significant features.
  - Convolutional Operation: Initially, the CNN input format is described. The vector format is the input of the traditional neural network, while the multi-channelled image is the input of the CNN. For instance, single-channel is the format of the gray-scale image, while the RGB image format is three-channelled. To understand the convolutional operation, let us take an example of a  $4 \times 4$  gray-scale image with a  $2 \times 2$  random weight-initialized kernel. First, the kernel slides over the whole image horizontally and vertically. In addition, the dot product between the input image and the kernel is determined, where their corresponding values are multiplied and then summed up to create a single scalar value, calculated concurrently. The whole process is then

repeated until no further sliding is possible. Note that the calculated dot product values represent the feature map of the output

2. Pooling Layer: The main task of the pooling layer is the sub-sampling of the feature maps. These maps are generated by following the convolutional operations. In other words, this approach shrinks large-size feature maps to create smaller feature maps. Concurrently, it maintains the majority of the dominant information (or features) in every step of the pooling stage. In a similar manner to the convolutional operation, both the stride and the kernel are initially size-assigned before the pooling operation is executed. Several types of pooling methods are available for utilization in various pooling layers. These methods include tree pooling, gated pooling, average pooling, min pooling, max pooling, global average pooling (GAP), and global max pooling. The most familiar and frequently utilized pooling methods are the max, min, and GAP pooling.
3. Activation Function (non-linearity) Mapping the input to the output is the core function of all types of activation function in all types of neural network. The input value is determined by computing the weighted summation of the neuron input along with its bias (if present). This means that the activation function makes the decision as to whether or not to fire a neuron with reference to a particular input by creating the corresponding output.

Non-linear activation layers are employed after all layers with weights (so-called learnable layers, such as FC layers and convolutional layers) in CNN architecture. This non-linear performance of the activation layers means that the mapping of input to output will be non-linear; moreover, these layers give the CNN the ability to learn extra-complicated things. The activation function must also have the ability to differentiate, which is an extremely significant feature, as it allows error back-propagation to be used to train the network. The following types of activation functions are most commonly used in CNN and other deep neural networks.

- Sigmoid: The input of this activation function is real numbers, while the output is restricted to between zero and one. The sigmoid function curve is S-shaped
- Tanh: It is similar to the sigmoid function, as its input is real numbers, but the output is restricted to between – 1 and 1

- ReLU: The mostly commonly used function in the CNN context. It converts the whole values of the input to positive numbers. Lower computational load is the main benefit of ReLU over the others
4. Fully Connected Layer: Commonly, this layer is located at the end of each CNN architecture. Inside this layer, each neuron is connected to all neurons of the previous layer, the so-called Fully Connected (FC) approach. It is utilized as the CNN classifier. It follows the basic method of the conventional multiple-layer perceptron neural network, as it is a type of feed-forward ANN. The input of the FC layer comes from the last pooling or convolutional layer. This input is in the form of a vector, which is created from the feature maps after flattening.
5. Loss Functions: The previous section has presented various layer-types of CNN architecture. In addition, the final classification is achieved from the output layer, which represents the last layer of the CNN architecture. Some loss functions are utilized in the output layer to calculate the predicted error created across the training samples in the CNN model. This error reveals the difference between the actual output and the predicted one. Next, it will be optimized through the CNN learning process.

However, two parameters are used by the loss function to calculate the error. The CNN estimated output (referred to as the prediction) is the first parameter. The actual output (referred to as the label) is the second parameter. Several types of loss function are employed in various problem types. The following concisely explains some of the loss function types.

- **Cross-Entropy or SoftMax Loss Function:** This function is commonly employed for measuring the CNN model performance. It is also referred to as the log loss function. Its output is the probability  $p \in \{0,1\}$ . In addition, it is usually employed as a substitution of the square error loss function in multi-class classification problems. In the output layer, it employs the SoftMax activations to generate the output within a probability distribution.
- **Euclidean Loss Function:** This function is widely used in regression problems. In addition, it is also the so-called mean square error

- **Hinge Loss Function:** This function is commonly employed in problems related to binary classification. This problem relates to maximum-margin-based classification; this is mostly important for SVMs, which use the hinge loss function, wherein the optimizer attempts to maximize the margin around dual objective classes

### 3.3.6. Regularization to CNN

For CNN models, over-fitting represents the central issue associated with obtaining well-behaved generalization. The model is entitled over-fitted in cases where the model executes especially well on training data and does not succeed on test data (unseen data) which is more explained in the latter section. An under-fitted model is the opposite; this case occurs when the model does not learn a sufficient amount from the training data. The model is referred to as “just-fitted” if it executes well on both training and testing data. Various intuitive concepts are used to help the regularization to avoid over-fitting.

- Dropout: This is a widely utilized technique for generalization. During each training epoch, neurons are randomly dropped. In doing this, the feature selection power is distributed equally across the whole group of neurons, as well as forcing the model to learn different independent features. During the training process, the dropped neuron will not be a part of back-propagation or forward-propagation. By contrast, the full-scale network is utilized to perform prediction during the testing process.
- Drop-Weights: This method is highly similar to dropout. In each training epoch, the connections between neurons (weights) are dropped rather than dropping the neurons; this represents the only difference between drop-weights and dropout.
- Data Augmentation: Training the model on a sizeable amount of data is the easiest way to avoid over-fitting. To achieve this, data augmentation is used. Several techniques are utilized to artificially expand the size of the training dataset. More details can be found in the latter section, which describes the data augmentation techniques.
- Batch Normalization: This method ensures the performance of the output activations. This performance follows a unit Gaussian distribution. Subtracting the mean and

dividing by the standard deviation will normalize the output at each layer. While it is possible to consider this as a pre-processing task at each layer in the network, it is also possible to differentiate and to integrate it with other networks. In addition, it is employed to reduce the “internal covariance shift” of the activation layers. In each layer, the variation in the activation distribution defines the internal covariance shift. This shift becomes very high due to the continuous weight updating through training, which may occur if the samples of the training data are gathered from numerous dissimilar sources (for example, day and night images). Thus, the model will consume extra time for convergence, and in turn, the time required for training will also increase. To resolve this issue, a layer representing the operation of batch normalization is applied in the CNN architecture.

### **3.4. Optimizer selection**

This section discusses the CNN learning process. Two major issues are included in the learning process: the first issue is the learning algorithm selection (optimizer), while the second issue is the use of many enhancements (such as Adam, Adagrad, and momentum) along with the learning algorithm to enhance the output.

Loss functions, which are founded on numerous learnable parameters (e.g., biases, weights, etc.) or minimizing the error (variation between actual and predicted output), are the core purpose of all supervised learning algorithms. The techniques of gradient-based learning for a CNN network appear as the usual selection. The network parameters should always update through all training epochs, while the network should also look for the locally optimized answer in all training epochs in order to minimize the error.

The learning rate is defined as the step size of the parameter updating. The training epoch represents a complete repetition of the parameter update that involves the complete training dataset at one time. Note that it needs to select the learning rate wisely so that it does not influence the learning process imperfectly, although it is a hyper-parameter.

Gradient Descent or Gradient-based learning algorithm: To minimize the training error, this algorithm repetitively updates the network parameters through every training epoch. More

specifically, to update the parameters correctly, it needs to compute the objective function gradient (slope) by applying a first-order derivative with respect to the network parameters. Next, the parameter is updated in the reverse direction of the gradient to reduce the error. The parameter updating process is performed through network back-propagation, in which the gradient at every neuron is back-propagated to all neurons in the preceding layer.

### **3.4.1. gradient-based learning optimization algorithm**

- Batch Gradient Descent: During the execution of this technique, the network parameters are updated merely one time behind considering all training datasets via the network. In more depth, it calculates the gradient of the whole training set and subsequently uses this gradient to update the parameters. For a small-sized dataset, the CNN model converges faster and creates an extra-stable gradient using BGD. Since the parameters are changed only once for every training epoch, it requires a substantial number of resources. By contrast, for a large training dataset, additional time is required for converging, and it could converge to a local optimum (for non-convex instances).
- Stochastic Gradient Descent: The parameters are updated at each training sample in this technique. It is preferred to arbitrarily sample the training samples in every epoch in advance of training. For a large-sized training dataset, this technique is both more memory-effective and much faster than BGD. However, because it is frequently updated, it takes extremely noisy steps in the direction of the answer, which in turn causes the convergence behavior to become highly unstable.
- Mini-batch Gradient Descent: In this approach, the training samples are partitioned into several mini-batches, in which every mini-batch can be considered an under-sized collection of samples with no overlap between them. Next, parameter updating is performed following gradient computation on every mini-batch. The advantage of this method comes from combining the advantages of both BGD and SGD techniques. Thus, it has a steady convergence, more computational efficiency and extra memory effectiveness. The following describes several enhancement techniques in gradient-based learning algorithms (usually in SGD), which further powerfully enhance the CNN training process.

- Momentum: For neural networks, this technique is employed in the objective function. It enhances both the accuracy and the training speed by summing the computed gradient at the preceding training step, which is weighted via a factor  $\lambda$  (known as the momentum factor). However, it therefore simply becomes stuck in a local minimum rather than a global minimum. This represents the main disadvantage of gradient-based learning algorithms. Issues of this kind frequently occur if the issue has no convex surface (or solution space).
- Adaptive Moment Estimation (Adam): It is another optimization technique or learning algorithm that is widely used. Adam represents the latest trends in deep learning optimization. This is represented by the Hessian matrix, which employs a second-order derivative. Adam is a learning strategy that has been designed specifically for training deep neural networks. More memory efficient and less computational power are two advantages of Adam. The mechanism of Adam is to calculate adaptive LR for each parameter in the model. It integrates the pros of both Momentum and RMSprop. It utilizes the squared gradients to scale the learning rate as RMSprop and it is similar to the momentum by using the moving average of the gradient.

### **3.5. CNN architectures**

Over the last 10 years, several CNN architectures have been presented. Model architecture is a critical factor in improving the performance of different applications. Various modifications have been achieved in CNN architecture from 1989 until today. Such modifications include structural reformulation, regularization, parameter optimizations, etc. Conversely, it should be noted that the key upgrade in CNN performance occurred largely due to the processing-unit reorganization, as well as the development of novel blocks. In particular, the most novel developments in CNN architectures were performed on the use of network depth. In this section, we review the most popular CNN architectures, beginning from the AlexNet model in 2012 and ending at the High-Resolution (HR) model in 2020. Studying these architectures' features (such as input size, depth, and robustness) is the key to help researchers to choose the suitable architecture for their target task.

#### **3.5.1. ResNet**

He et al. developed ResNet (Residual Network), which was the winner of ILSVRC 2015. Their objective was to design an ultra-deep network free of the vanishing gradient issue, as compared to the previous networks. Several types of ResNet were developed based on the number of layers (starting with 34 layers and going up to 1202 layers). The most common type was ResNet50, which comprised 49 convolutional layers plus a single FC layer. The overall number of network weights was 25.5 M, while the overall number of MACs was 3.9 M. The novel idea of ResNet is its use of the bypass pathway concept, as shown in Figure below, which was employed in Highway Nets to address the problem of training a deeper network in 2015. This is illustrated in Figure below, which contains the fundamental ResNet block diagram. This is a conventional feedforward network plus a residual connection.

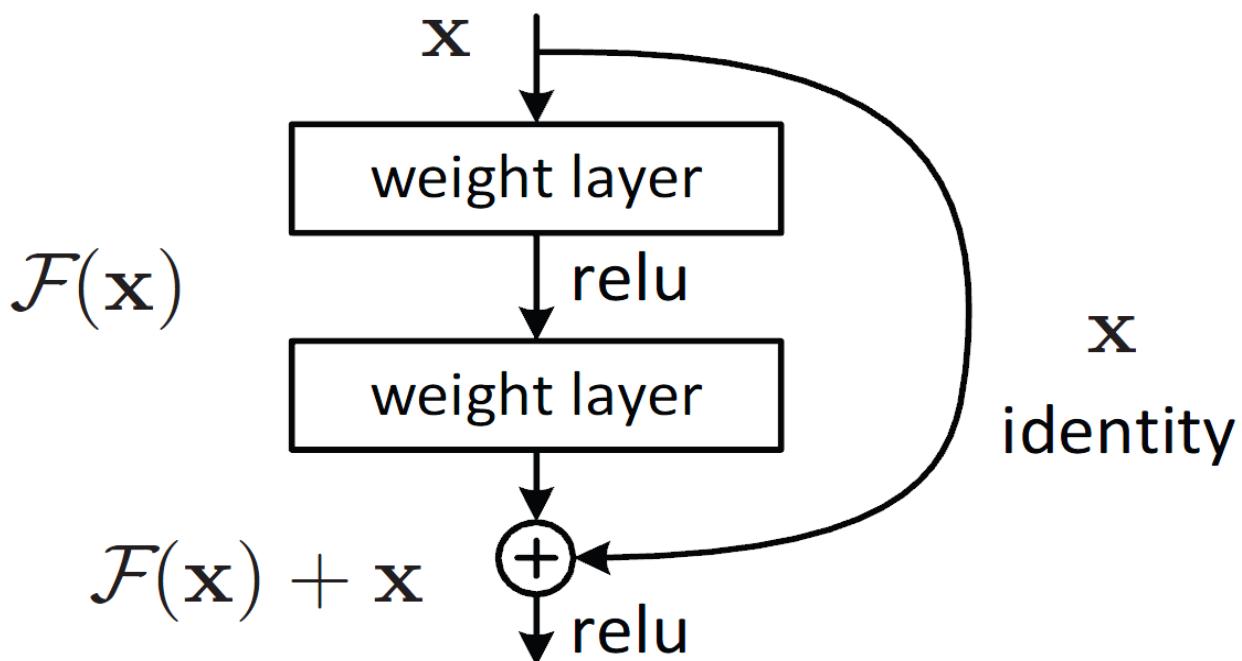


Figure 3.2 ResNet basic Block

In comparison to the highway network, ResNet presented shortcut connections inside layers to enable cross-layer connectivity, which are parameter-free and data-independent. Note that the layers characterize non-residual functions when a gated shortcut is closed in the highway network. By contrast, the individuality shortcuts are never closed, while the residual information is permanently passed in ResNet. Furthermore, ResNet has the potential to

prevent the problems of gradient diminishing, as the shortcut connections (residual links) accelerate the deep network convergence. ResNet was the winner of the 2015-ILSVRC championship with 152 layers of depth; this represents 8 times the depth of VGG and 20 times the depth of AlexNet. In comparison with VGG, it has lower computational complexity, even with enlarged depth.

### 3.5.2. Inception: ResNet and Inception-V3/4

Szegedy et al. proposed Inception-ResNet and Inception-V3/4 as upgraded types of Inception-V1/2. The concept behind Inception-V3 was to minimize the computational cost with no effect on the deeper network generalization. Thus, Szegedy et al. used asymmetric small-size filters ( $1 \times 5$  and  $1 \times 7$ ) rather than large-size filters ( $7 \times 7$  and  $5 \times 5$ ); moreover, they utilized a bottleneck of  $1 \times 1$  convolution prior to the large-size filters. These changes make the operation of the traditional convolution very similar to cross-channel correlation. Previously, Lin et al. utilized the  $1 \times 1$  filter potential in NIN architecture. Subsequently, utilized the same idea in an intelligent manner. By using  $1 \times 1$  convolutional operation in Inception-V3, the input data are mapped into three or four isolated spaces, which are smaller than the initial input spaces. Next, all of these correlations are mapped in these smaller spaces through common  $5 \times 5$  or  $3 \times 3$  convolutions. By contrast, in Inception-ResNet, Szegedy et al. bring together the inception block and the residual learning power by replacing the filter concatenation with the residual connection. Szegedy et al. empirically demonstrated that Inception-ResNet (Inception-4 with residual connections) can achieve a similar generalization power to Inception-V4 with enlarged width and depth and without residual connections. Thus, it is clearly illustrated that using residual connections in training will significantly accelerate the Inception network training. Figure below shows The basic block diagram for Inception Residual unit.

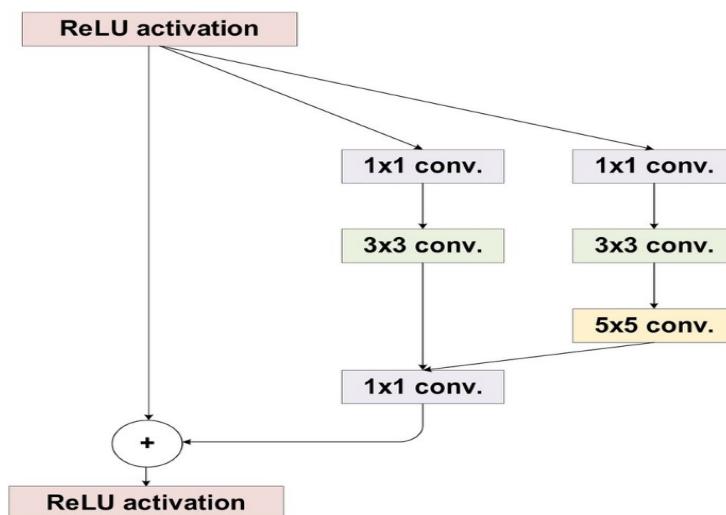
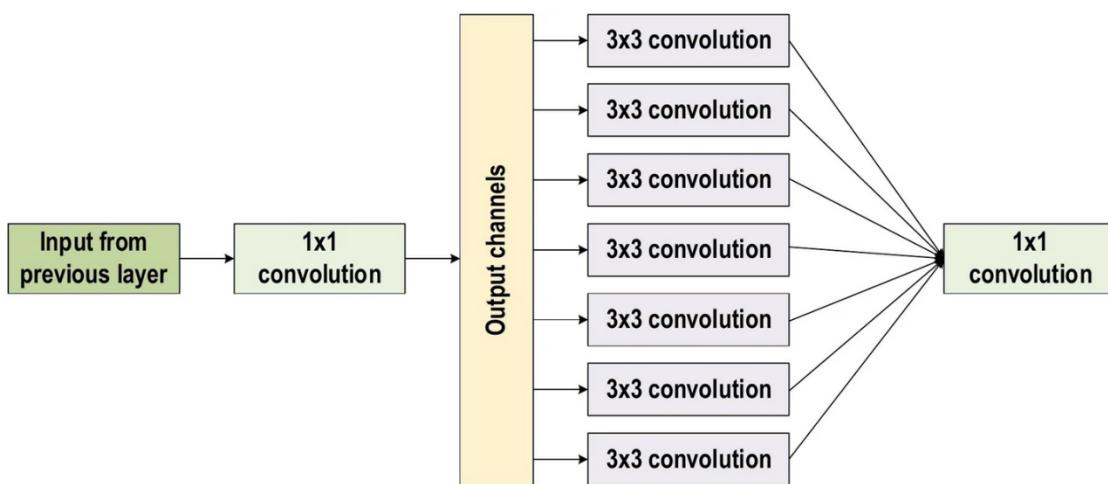


Figure 3.3 Basic block diagram for Inception Residual unit

### 3.5.3. Xception Net

Extreme inception architecture is the main characteristic of Xception. The main idea behind Xception is its depthwise separable convolution. The Xception model adjusted the original inception block by making it wider and exchanging a single dimension ( $3 \times 3$ ) followed by a  $1 \times 1$  convolution to reduce computational complexity. Figure 24 shows the Xception block architecture. The Xception network becomes extra computationally effective through the use of the decoupling channel and spatial correspondence. Moreover, it first performs mapping of the convolved output to the embedding short dimension by applying  $1 \times 1$  convolutions. It then performs  $k$  spatial transformations. Note that  $k$  here represents the width-defining cardinality, which is obtained via the transformations number in Xception. However, the computations were made simpler in Xception by distinctly convolving each channel around the spatial axes. These axes are subsequently used as the  $1 \times 1$  convolutions (pointwise convolution) for performing cross-channel correspondence. The  $1 \times 1$  convolution is utilized in Xception to regularize the depth of the channel. The traditional convolutional operation in Xception utilizes a number of transformation segments equivalent to the number of channels; Inception, moreover, utilizes three transformation segments, while traditional CNN architecture utilizes only a single transformation segment. Conversely, the suggested Xception transformation approach achieves extra learning efficiency and better performance but does not minimize the number of parameters.

Figure 3.4 Xception basic Block



# **Methodology**

## **4. Introduction**

The project presents two main methodologies for the diagnosis of heart diseases based on the combination of two heart signals. The first methodology is based on detecting heart diseases using the combination of ECG and ABP signals, and the second methodology is based defining normal and abnormal heart diseases based on ECG and PCG signals. The two methodologies are based on a set of phases. These steps are based on data acquisition, pre-processing, segmentation, feature extraction, and classification.

### **4.1. ECG-ABP Diagnosis and ECG-PCG Diagnosis Methodologies**

As mentioned before, the proposed methodology consists of five main phases. In the data acquisition phase, the data are collected from an online available dataset. This dataset holding ECG and ABP records for different subjects. In the pre-processing phase, the signals suffer from three common noises and they are filtered using a pre-processing chain. The segmentation phase depends on segmenting the ECG heartbeats and select its corresponding ABP hearbeats. Then, the feature extraction is based on deep learning approach and the classification is based on three main classifiers naïve bayes, k-nearest neighbor (KNN), and support vector machine (SVM).

#### **4.3.1 Data acquisition**

All of the data employed in this study were taken from the MIMIC of the physioNet database. MIMIC is the Multi Parameter database in physioNet. MGH waveform database is one of the sets which categorized in the Multi Parameter databases and ECG databases. These data include the collection of 250 recordings and each contains 3 lead ECGs, ABP, Pulmonary Arterial Pressure (PAP), Central Venous Pressure (CVP), respiration and airway CO<sub>2</sub> signals from patients in critical care units. Some recordings include intracranial, left atrial, ventricular and intra aortic pressure waveforms. The database comprises a broad spectrum of physiologic and pathophysiologic states.

Individual recordings vary in length from 12 to 86 minutes, and in the most cases are about an hour long. In this paper, we used ECG lead I and blood pressure records. In addition, the normalized data are available in the dataset. Therefore, the feature extraction procedure was employed on the normalized data. In the physioNet, the signal of patients, regardless of the type of disorders, numbered from 1 to 250. In this project, five types of cardiac arrhythmia have been used from this database. These arrhythmias consist of normal sinus rhythm, Sinus tachycardia, atrial fibrillation, normal sinus rhythm with left bundle branch block, and normal sinus rhythm with right bundle branch block.

In terms of the ECG and PCG methodology the PCG and ECG records exploited a sample of 100 recordings of PCG and ECG signals from dataset A of the 2016 PhysioNet/CinC Challenge data base. We have split this sample into two: a subset of 60% constitutes the training set, and the remaining 40% forms the test set. Note that the two cardiac signals ECG and PCG are by default synchronized, so the simultaneous acquisition of the ECG and PCG is the way to maintain this synchronization

### 4.3.2 Pre-processing

Preprocessing is a process of de-noising, filtering the signal by removing the most common noises that can appear and delivering it in a pure shape to extract features from it. There are three types of noises in the ECG signal: power-line noise, high-frequency noise and baseline drift. As a result of a series of experiments, the following combination of methods has been selected for the preprocessing phase. Firstly, baseline drift correction is done using wavelet decomposition. Donoho and Johnstone proposed the universal ‘VisuShrink’ threshold given by

$$Thr = \sigma \sqrt{2 \log(N)} \quad (4.1)$$

where N is the number of data points and  $\sigma$  is an estimate of the noise level. The wavelet-based de-noising process is summarized as follows: the resulting discrete wavelet transform (DWT) detail coefficients are thresholded by shrinkage (soft) strategy. Reconstructing the original sequence from the thresholded wavelet detail coefficients leads to removal of

baseline drift. Baseline drift correction is done by using wavelet decomposition with wavelet name db8 with  $N = 9$ .  $N$  is the level number, and a soft threshold=4.29 is used. Secondly, an adaptive band stop filter fairly well suppresses power-line noise with  $W_s = 50$  Hz, where the  $W_s$  is stop band corner frequency.

Then, a low-pass Butterworth filter is used with  $W_p = 40$  Hz, where  $W_p$  is the pass band corner frequency,  $W_s = 60$  Hz,  $R_p = 0.1$  dB.  $R_p$  is the pass band ripple and  $R_s = 30$  dB is the attenuation in the stop band to remove the remaining noise components, caused by possible highfrequency distortions. The last step is smoothing the signal with  $N = 5$ , where  $N$  is the smoothing value to produce the preprocessed signal.

#### **4.3.3 Segmentation**

Our segmentation starts with detecting QRS by using some filters, starting with cancelation of DC shift and normalization, low-pass filtering, high-pass filtering, derivative filter, squaring and moving window integration to produce a vector of zeros and ones with the same size as the preprocessed signal. The ones in this vector determine the QRS interval in the whole signal.

This vector is divided into two vectors: left and right, determining the start of the Q peaks and S peaks, respectively, in the whole signal. Then, a window of 60 samples is created before the left (Q peak) and the max sample in it are gotten to obtain P peak. Another window of 125 samples is created after the right (S peak) and the max values in the samples are gotten to obtain T peak. R peak is detected by finding the max value between the left and the right. Q is detected by selecting the min values between the left vector, starting of Q peaks and R locations. S is detected by selecting the min values between the R locations and right vector, starting of S peaks. Therefore, we have detected the P, Q, R, S ,T locations and peaks, as well as the R interval (samples from Q to S).

After the peaks and the locations have been detected, we need to determine the most informative fragment in the signal that leads to a significant effect in the last classification outcome. The most important and effective component is QRS complex, while P and T are considered the most uncertain and suspicious components. P wave has low amplitude and

can be affected by noise, while T wave has a dynamic location and its position always depends on the heart rate. Although T and P can enhance and provide some useful information to be used for improving the system accuracy, they will also enhance some questioning to the extraction, or the processing techniques. In order to determine the power of T and P waves, it is decided to select four informative fragments to be used: QRS, P-QRS, QRS-T and P-QRS-T, as shown in Fig. 4. Those four fragments are considered the most informative ones. In order to choose which fragment is considered to be the most informative, an experiment is done to find this. Finally, from this experiment, P-QRS-T fragment is selected as the most informative fragment, producing a better performance than the other fragments.

We want to select the most discriminant fragments that most describe the signal and that can represent the signal in the identification process. Subsequently, a lot of checks have been made on each P-QRS-T fragment in the signal to determine whether to take this fragment, or leave it. Firstly, for each P-QRS-T fragment the R interval in it is checked. In other words, the number of samples of the R interval in P-QRS-T fragment must be greater than 30 samples and less than 70 samples. Otherwise, the P-QRS-T fragment is rejected. Next, for each P-QRS-T fragment, we calculate some amplitudes, distances and means, such as RQ, RS for amplitudes, PR, RT, QS, RQ, RS for distances RQ, RS, PR, RT, QS, RQ, RS for means and the mean of R interval. For each two successive P-QRS-T fragments RR distances, mean and the median of RR distances are calculated. For the mean and the median of RR distance, we choose the minimum value between them and it will be the RR thresholding, as it shows better performance experimentally. The last step is selecting the most similar P-QRS-T fragments.

This is done by putting some restrictions, conditions and weight sum for the ECG fragments. For each two successive P-QRS-T fragments if RR location is less than  $0.9 * RR$  thresholding, then the first P-QRS-T fragment is chosen. Otherwise, the first P-QRS-T fragment is rejected. While the condition of thresholding is satisfied for each two successive P-QRS-T fragments, starting from the ECG signal, we create successive conditions starting by: "If the R interval of the first P-QRS-T fragment, subtracted from the mean of the R interval, is less than the R interval of the second P-QRS-T fragment, subtracted from the

mean of the R interval, then increase the weight sum of the first P-QRS-T fragment by 0.3. Otherwise, increase the second P-QRS-T fragment.”

The same is made for RT distance in each two successive fragments by weight 0.3, PR distance, RQ amplitude, RS amplitude, QS distance by higher weight, equal to 0.75, as these amplitudes and distances are more effective experimentally. At the end, the fragment that gives a higher weight is selected from the ECG signal. We pick the highest weight P-QRS-T fragments representing the signal. The P-QRS-T fragment length is fixed at 281 samples for each cardiac cycle, regardless of the actual lengths of PR, QRS, and QT intervals. 281 samples (110 samples to the left of R peak and 140 samples to the right) are extracted and analyzed. The ABP heartbeat is selected starting from the sample corresponding to the R peak in the ECG to the 250 sample from the ABP heartbeat. In terms of the ECG and PCG fragments, the 390 samples (250 samples to the left of the R peaks and 140 samples to the right) are extracted from the ECG signals. The PCG signals are extracted based on 200 samples before the R peak and 1000 samples after the R peak from the PCG signal to capture S1 and S2.

#### 4.3.4 Feature Extraction

Recently, TCN (Temporal Convolutional Network) has found application in various domains, including probabilistic prediction, traffic forecasting, sound event detection, and more. Initially introduced by Le et al. [95, 96], TCN demonstrated exceptional performance in tasks such as video action segmentation, prediction, and classification. TCN operates through two primary steps. First, it computes low-level features using CNN (Convolutional Neural Network) models, which encode spatial-temporal information. Second, these low-level features are passed into a model, which can be either a CNN or an RNN (Recurrent Neural Network), to capture high-level temporal patterns.

The proposed TCN in this context takes input data from the pre-processed pixels in the form of sequential features. These inputs are then transformed into a probability distribution. Subsequently, they are processed through four stacks of residual blocks. Each residual block consists of two key components: two dilated causal convolution layers, two weight normalization layers, two dropout layers, one ReLU activation layer, and an optional

convolutional layer. Notably, the first residual block differs by containing three dilated causal convolution layers [18].

**Dilated Causal Convolutional Layer:** In the TCN model, it can take an input sequence of a specific length and generate an output of the same length. It's termed "causal" because activations produced at a particular time step cannot depend on activations from future time steps. The input to the TCN defined by  $Y = [y_1, y_2, \dots, \dots, y_i]$  and a filter  $f: \{0, \dots, k - 1\}$ . The dilated causal convolutional operation on the  $i$  the point of  $Y$  is defined using the following equation:

$$C(y_i) = \sum_{a=0}^{k-1} f(a) - y_{i-a.d} \quad (4.2)$$

Where  $d$  is the dilation factor,  $k$  is the filter size,  $i - a.d$  represents the direction of the past. In other words, the first layer maintains  $Y$  as the input sequence, whereas in higher layers  $Y$  represents the output of the former layer. For each dilation convolutional layer there exist a dilation factor and this factor increases exponentially by 2.

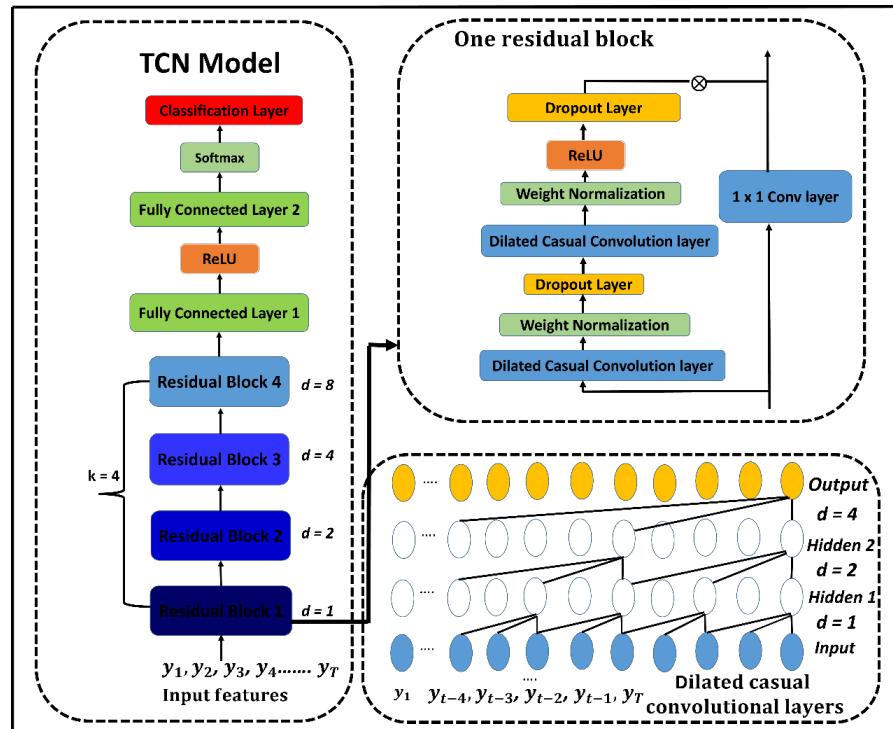


Figure 4.1 Entire TCN Architecture

Weight Normalization Layer (WN): This layer is applied to each dilated convolutional layer. Its primary purpose is to separate the weight direction from its magnitude. To achieve this, weights are normalized using a specific learning scaling parameter. The equation of the weight normalization operation is defined using the following equation:

$$o_j = s_j \frac{W_j * x}{\|W_j\|_F + \epsilon} + j \quad (4.3)$$

$x$  is the input of the WN layer,  $o$  is the output of the WN layer,  $s_j$  is defined as the scale,  $j$  determines the bias,  $\epsilon$  is a constant that is used for numerical stability,  $W_j$  and  $\|W_j\|_F$  the layer's weight and the Frobenius norm of the weights for the output channel  $j$  respectively, and  $*$  is the convolution operator.

The input to the residual block undergoes a  $1 \times 1$  optional convolutional layer and is then combined with the output of the residual block. This layer's purpose is to be applied when the number of channels in the output and input don't match. The same procedure is repeated for the remaining residual blocks. Finally, after executing all four blocks, the output of the fourth block is passed through two fully connected layers, one ReLU activation layer, one Softmax layer, and a classification layer. Figure 4.1 provides an overview of the key parameters of the TCN, including the number of blocks, dilation factors, and input channel numbers. It also details the parameters of each dilated causal convolutional layer, including weight, bias, stride, and padding. Moreover, the figure outlines the parameters of the last fully connected layer and the  $1 \times 1$  convolutional layer. Importantly, the length of our output matches the length of the input.

### 4.3.5 Classification

#### 4.3.5.1 K-Nearest Neighbor (KNN)

The K-Nearest Neighbors (KNN) algorithm is a non-parametric, instance-based learning method used for classification and regression tasks. It works based on the assumption that similar data points tend to belong to the same class or have similar output values. In the context of classification, KNN assigns a class label to a new data point based on the majority class among its  $k$  nearest neighbors. One of the key components of the KNN algorithm is

the distance metric used to measure the similarity between data points. The most commonly used distance metric in KNN is the Euclidean distance, although other distance metrics such as Manhattan, Minkowski, or Cosine similarity [30].

In the KNN algorithm with Euclidean distance, when a new data point is to be classified, the distances between the new data point and all other data points in the training set are computed. The  $k$  nearest neighbors of the new data point are then identified based on the smallest Euclidean distances. Finally, the majority class among the  $k$  nearest neighbors is assigned as the predicted class label for the new data point.

KNN with Euclidean distance is intuitive and easy to implement, making it a popular choice for classification tasks, especially in cases where the decision boundary is irregular or nonlinear. However, it has some limitations, such as computational inefficiency with large datasets and sensitivity to the choice of the number of neighbors ( $k$ ) and the distance metric. Additionally, it assumes that all features contribute equally to the distance calculation, which may not always be appropriate for all datasets. Despite these limitations, KNN with Euclidean distance remains a valuable tool in the machine learning toolkit, particularly for its simplicity and effectiveness in certain scenarios.

#### 4.3.5.2 Multi-layer Perceptron (MLP) Classifier

MLP is a feedforward ANN that maps the inputs onto a set of output categories. ANNs consist of neurons that operate in parallel, i.e., parallel processing, which is an important advantage. The classification function of the neural network is dependent on the product of the weights between and the biases within the neurons. The training is based on backpropagation. The network weights are adapted and updated until the operation produces the desired output [31]. Mathematically, the interval activities of the neurons in the network are given by:

$$Y_k = f\left(\sum_{j=1}^{n_k} w_{kj} \left(f \sum_{i=1}^d w_{ji} x_i + w_{j0}\right) + w_{k0}\right) \quad (4.4)$$

where  $x_i$  represent features fed to the input layer;  $Y_k$ , the final result of the neuron; and  $j$ , the index of the neuron. The weight between the input ( $i$ ) and the hidden layer neuron ( $j$ ) is  $w_{ji}$ ,

while  $W_{kj}$  is the weight between the  $j$  and the output ( $k$ ).  $w_{j0}$ ,  $w_{k0}$ , and  $n_k$  are the biases of the hidden, output, and the last neurons, respectively.  $f(.)$  is the activation function for the hidden and output layers, which are log-sigmoid for the hidden and linear activation layers of the output layer. The full network has 28 layers consisting of 25 CNN layers followed by three layers of the ANN, which compose the MLP. The CNN performs feature extraction with momentum backpropagation. The features are input to the MLP for classification. There is an input layer, where the number of neurons corresponds to the number of features output by the CNN, i.e., 1,000. The hidden layer contains 2,000 neurons, which is the product of the number of classes (4) and the number of features (1,000) divided by 2. The output layer consists of four neurons, which yield the final classification results for the four classes. There are logsig activation functions in the hidden units, while the output unit contains a linear function. Cross-entropy loss function was used to measure the error between the actual and the desired output.

#### 4.3.5.3 Support-Vector Machine (SVM) Classifier

In this hybrid model, CNN is combined with the efficient shallow SVM classifier, i.e., SVM replaces the classification layer of the CNN. The SVM classifier was developed for binary classification [32] and aims to find the optimal hyperplane  $f(w, x) = w \cdot x + b$  that separates two classes in a given dataset. SVM learns the parameters  $w$  and  $b$  by solving an optimization problem, as given by:

$$\begin{aligned} & \text{Minimize: } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^p E_i \\ & \text{s.t. } \begin{cases} y_i(w \cdot x + b) \geq 1 - E_i \\ E_i \geq 0, i = 1, \dots, p \text{ (No. of Samples)} \end{cases} \end{aligned} \quad (4.5)$$

where  $\|w\|^2$  represent the Euclidean norm;  $C$ , the penalty parameter that governs maximization of margin and minimization of classifications error;  $E$ , the cost function; and  $y$ , the actual label. The best hyperplane is determined by minimizing the cost function while taking into account the two conditions of maximizing the margin and minimizing the error. This optimization problem can be expressed as a quadratic programming problem subject to constraints, and the solution for the discriminant function can be expressed as:

$$f(x) = \sum_i \alpha_i y_i k(x_i \cdot x) + b \quad (4.6)$$

where  $k(x_i \cdot x)$  represents the kernel function that maps the input data to a higher dimensional space for maximization of the margin; and  $\alpha_i$  the Lagrange Multipliers. In our case,  $k(x_i \cdot x)$  is the polynomial of degree  $d$ , i.e.,  $k(x_i \cdot x) = (1 + x_i \cdot x)^d$ . To build a multiclass classifier based on SVM, error-correcting output code (ECOC) was applied, which reduced the multiclass classification problem to a set of binary problems using the one-against-one (OAO) or one-against-all (OAA) strategies. In our case, we used OAO, which would involve four classes with SVM based on six datasets. Finally, the four classifiers were combined using the ECOC pattern, the strategy of the combination being based on a one-versus-one coding design.

# **Mobile Application**

## **5. Introduction**

Smart phones today have become increasingly popular with the general public for its diverse abilities like navigation, social networking, and multimedia facilities to name a few. These phones are equipped with high-end processors, high-resolution cameras, built-in sensors like accelerometer, orientation-sensor, light-sensor, and much more. According to comScore survey, 77.3% of US adults use smart phones in their daily lives.

Motivated by the capability of smart phones and their extensive usage, I focused on utilizing them for biomedical applications. In this thesis, I present a new application for a smart phone to quantify the vital signs such as heart rate, respiratory rate and blood pressure with the help of its built-in sensors. Using the camera and a microphone, I have shown how the blood pressure and heart rate can be determined for a subject.

People sometimes encounter minor situations like fainting or fatal accidents like car crash at unexpected times and places. It would be useful to have a device which can measure all vital signs in such an event.

The second part of this thesis demonstrates a new mode of communication for next generation 9-1-1 calls. In this new architecture, the call-taker will be able to control the multimedia elements in the phone from a remote location.

This would help the call-taker or first responder to have a better control over the situation. Transmission of the vital signs measured using the smart phone can be a life saver in critical situations.

In today's voice oriented 9-1-1 calls, the dispatcher first collects critical information (e.g., location, call-back number) from caller, and assesses the situation.

Meanwhile, the dispatchers constantly face a "60-second dilemma"; i.e., within 60 seconds, they need to make a complicated but important decision, whether to dispatch and, if so, what to dispatch. The dispatchers often feel that they lack sufficient information to make a confident dispatch decision. This remote-media-control described in this system will be able to facilitate information acquisition and decision-making in emergency situations within the 60-second response window in 9-1-1 calls using new multimedia technologies.

Is an android Application that can estimate Heart rate, Blood pressure, Respiration rate and Oxygen rate from only the camera of the mobile without using any extra sensors.

## **5.1 Heart Rate (HR):**

A pulse is a sudden burst of blood to the circulatory system when the walls of the heart contract. Heart rate or pulse rate is defined as the number of heart beats or pulses in a minute. The human heart comprises the atrium and the ventricles, which coordinate to form a complete pumping action. Approximately 2000 gallons of blood is pumped by the heart every day.

A Heart beat cycle consists of two components, namely systole and diastole. Systole occurs when there is an electrical impulse generated by the Sinoatrial (SA) Node, triggering the heart to contract. Diastole occurs when the heart is relaxed. Systole and diastole alternate each other to produce a heart beat. The heart rate is not just about how fast the heart is beating; it is a regulatory mechanism for delivering oxygen to the muscles to keep up the demand.

Acoustically, the heart rate is measured by listening to the heart beats, which are amplified through the use of a stethoscope. Usually the number of beats for a small interval of time, say 10 seconds, is observed and obtained for a minute by multiplying with 6.

In the same way, the pulse felt at the wrist and neck can be measured and directly related to the heart rate. Precise method of determining pulse rate involves the use of an electrocardiography (ECG or EKG), pulse oximetry, etc. Shelley [4] discussed about the effectiveness of pulse oximetry in the detection of pulse even under noisy conditions where the use of stethoscope is hopeless.

There are many commercial heart rate monitors available in the market which use two tiny electrode strips to find the heart rate, the same way an ECG works. These electrodes are generally attached to some fitness gear or costume, displaying the measurements on a screen.

ECG uses the electrical activity of the heart over a period of time, measured through the electrodes connected to the skin. These electrodes induce a tiny current of a few  $\mu\text{A}$  into the body and detect electrical changes caused by the heart during each heartbeat. These changes are captured, amplified and delivered as an output. The model works on the principle that, every heart beat pertains to a rush of blood in the blood vessels, even in the capillaries at the finger-tips. Whenever the capillaries are rich in blood during a systolic pulse, more light is getting absorbed by the blood, leading to low.

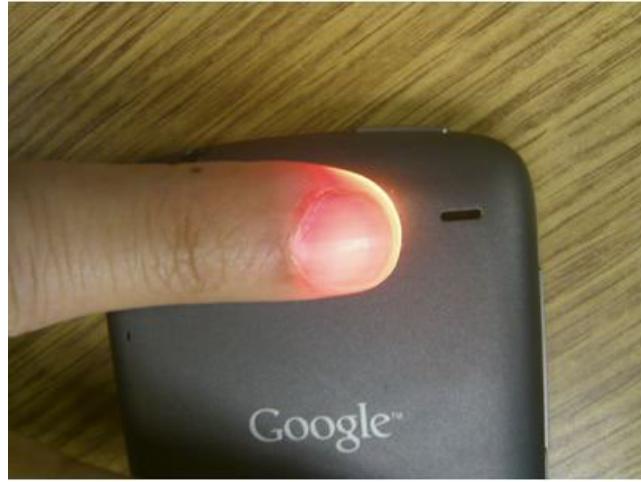


Figure 5.1 Method of placing the finger over the camera for heart rate measurement

Reactive index and darker frame intensities. Likewise, during a diastolic pulse, most of the light gets reflected leading to bright frames. This change in intensity of light which can pass through the finger creates an alternative pattern of waves similar to a pulse. These changes in intensity with time gives the heart rate of a person.

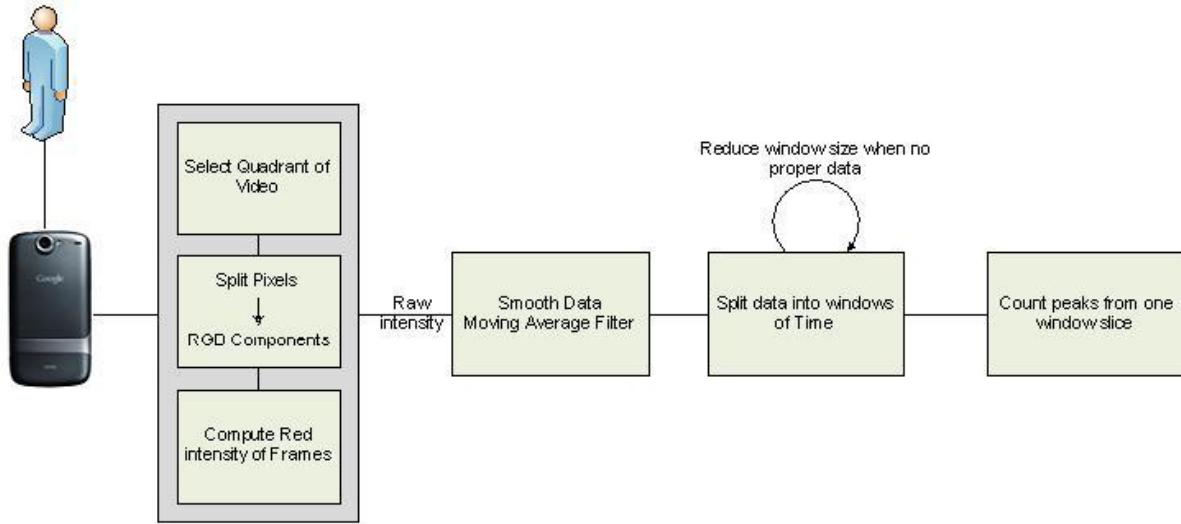


Figure 5.2 Architecture of heart rate system

In the proposed method, I record a video of short duration, with the finger placed over the lens of the mobile camera. The ash is turned ON, so that adequate amount of light can reach the finger for proper measurement. For this experiment, I developed an application for Nexus One [5] to keep the LED ash consistently ON while recording video from camera.

I discarded the first three seconds of data from the camera, since the CMOS sensor of the camera tries to focus when turned ON. Also, the camera doesn't need to be focused, as the results rely only on the amount of light entering the video feed. It was generally hard to Detect the fluctuations in the frames unless the pulses are distinct. A similar methodology was used by Banitsas [6] with a slightly different approach in the analysis of video frames.



Figure 5.3 Area under analysis

To measure the heart rate:

- Using a wake lock mechanism to Make sure that device doesn't go to sleep mode until finish HR processes" Wakelocks are power-managing software mechanisms, which make sure that your Android device doesn't go into deep sleep (which is the state that you should strive for), because a given app needs to use your system resources".
- Open device camera and set orientation to 90 degree.
- Turn on Flash Light of Camera and start record frames for 30 Sec.
- Store system time as a start time for the analyzing process.
- Getting frames data "PPG signal" of Green and Red Colors from the camera by applying image processing method on each frame to get the RGB intensities and start the heartbeat process.
- Count number of frames in 30 seconds.
- Check if we got a good red intensity to process.
- Calculating the sampling frequency using FFT Algorithm.
- Make sure that if the heart rate from red and green intensities are reasonable.
- Take the average between them, otherwise take the green or red if one of them is good.
- Get the HR value, Then Storing it in the user's History log.

Working of the system comprises of six functional modules. Initially the video frames were split into four quadrants and only the first quadrant Figure 7.3 was considered for analysis, since I observed most of the changes and actuations are predominant in that region. Every Pixel information on each frame was split into individual Red(R), Blue (B) and Green (G) components. In most samples I observed, the prominent color applied only to R with the others tending to zero in every frame, hence difference in the red channel (Rc) intensity to that of all the channels of a frame was negligible. For accuracy of plots, I have considered only the Rc in video frames. The average intensity of pixels for every frame was calculated as its frame intensity. The raw intensity values were filtered with a moving average filter to

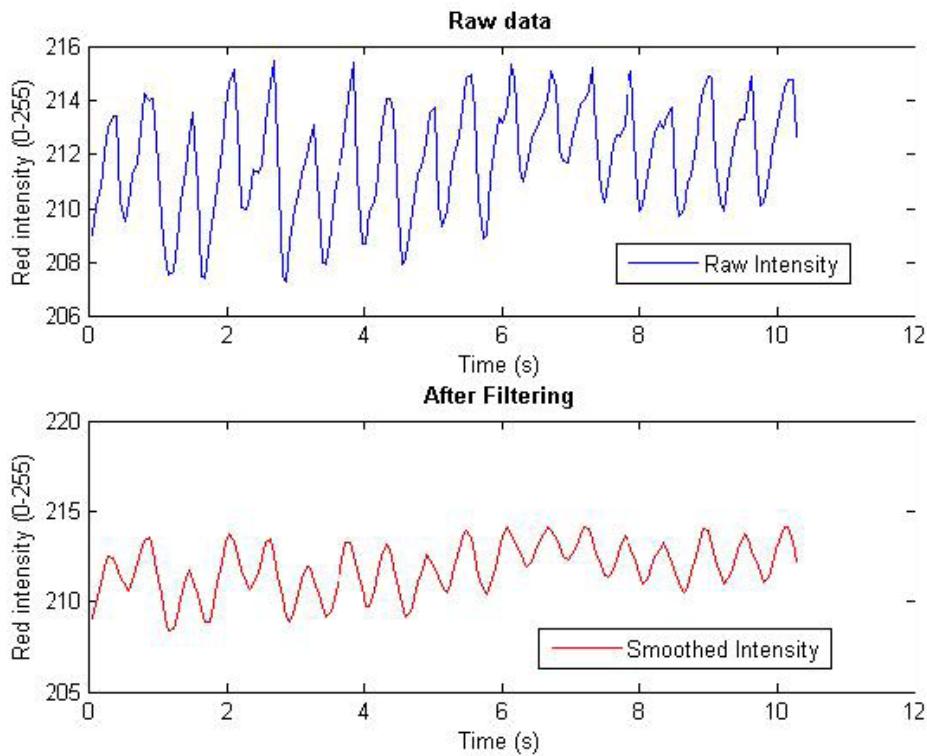


Figure 7.4 Filtered Data for analysis

Remove rough peaks from the graph for easier identification of peaks. Figure 5.4 shows the filtered results from the raw data obtained from finger pulse. The entire frame was split into Windows of fixed length ( $W_t$ ), for determination of peaks occurring at equal intervals of time  $I_s$  is calculated by determining the number of peaks ( $n$ ) in the window and multiplying the peak count with the window length as given by the equation (7.1).

$$HR = n * 60 = W_t \quad (5.1)$$

**Heart Rate Accuracy with Proposed Method.** This method showed encouraging results with high percentage of accuracy. The collected data was validated with a commercial heart rate monitor available at a Fitness center. To prove the effectiveness of the proposed system, I induced higher heart rate to the subject with excessive physical activity. From Table 5.1 it can be seen that the method gives a high.

Table 7.1 Accuracy of results at varying heart rate for a single subject

Actual HR	5 sec Window		10 sec Window	
	Value	Accuracy	Value	Accuracy
102	108	94.11%	102	100.0%

108	96	88.89%	102	94.44%
114	108	94.74%	114	100.0%
132	132	100.0%	132	100.0%
154	144	93.51%	150	97.40%

Percentage of accuracy from the obtained data in finding the HR of the person. It can also be seen that, with the increase in window size for analysis, the error propagation decreases to a very minimum. For better accuracy of the result, the user should hold the finger over the camera lens for a longer time. Table 5.2 shows the expected error in the system when varying the window size. For less error in the results for considered samples, the window size should be kept large. Even more, the accuracy of the heart rate for a full minute of data is precise with the actual heart rate measured manually, with a 100 % accuracy all the time. Based on the need, it is possible to measure every single heart beat with precision.

Table 5.2 Expected error from the system based on window size

Window Size (Sec)	Expected Error	Error %
5	$\pm 12$	0-11.1
10	$\pm 6$	0-5.6
15	$\pm 4$	0-3.7
20	$\pm 3$	0-2.8
30	$\pm 2$	0-1.8

## 5.2 Respiratory Rate (RR):

Respiration is a physical process which involves exchange of oxygen and carbon dioxide. The main function of the respiratory system is to provide ample oxygen to meet the energy production requirement of the body. A normal ventilation is an automatic, seemingly involuntary action comprising the expansion of the chest cage during inhalation and contraction during exhalation.

A person's respiratory rate is the number of breaths you take per minute. The normal respiration rate for an adult at rest is 12 to 20 breaths per minute. A respiration rate under 12 or over 25 breaths per minute while resting is considered abnormal.

The usual practice for measuring respiratory rate, involves a watch with a second hand or a stop watch. The number of times, the chest moves upward is counted when the person is in a seated position or lying at for a full 60 seconds or count for 30 seconds and multiplied by two. It is considered optimal to place the hand on the upper chest to feel it rise and fall.

Each rise and fall counts as one cycle of respiration. Also, while calculating the respiratory rate, a note has to be made about the effort in breathing by the person and the type of breathing.

The specific objective of this work was to determine the feasibility of extracting HR and RR from a cell phone camera. Figure 6.5 explains the entire process. The method started with a subject placing his finger over the cell phone camera without pressing down any additional force and the smartphone camera's flash turned on. The subject was instructed to breath at a controlled breathing rate of 6, 10, 12 breathes per minutes (bpm), which corresponded to 0.1, 0.16 and 0.2 Hz. by providing a breathing tone, the subject was guided to inhale and exhale at appropriate times to maintain the desired breathing frequency. Also, three sets of uncontrolled data were also recorded from the same subject, as follows: breathing under regular pace, sit-ups and walking. For the first phase of this experiment, a smartphone was used to record a video of the light absorbed by the index tissue. The length of the videos was four minutes for each data setting.

The videos were recording at a sampling rate of 29.97 fps with an image pixel density of 1280 x 720. After the videos were recorded, MATLAB was used to analyze the data gathered. First, the VideoReading tool offered in MATLAB was used to extract frames from the recoded video. Next, the RGB components were extracted from every frame. However, for this study the only the green values were used to compute the average of light absorbed by the finger tissue in every frame. Thus, the PPG signal is acquired and used to extract HR and RR.

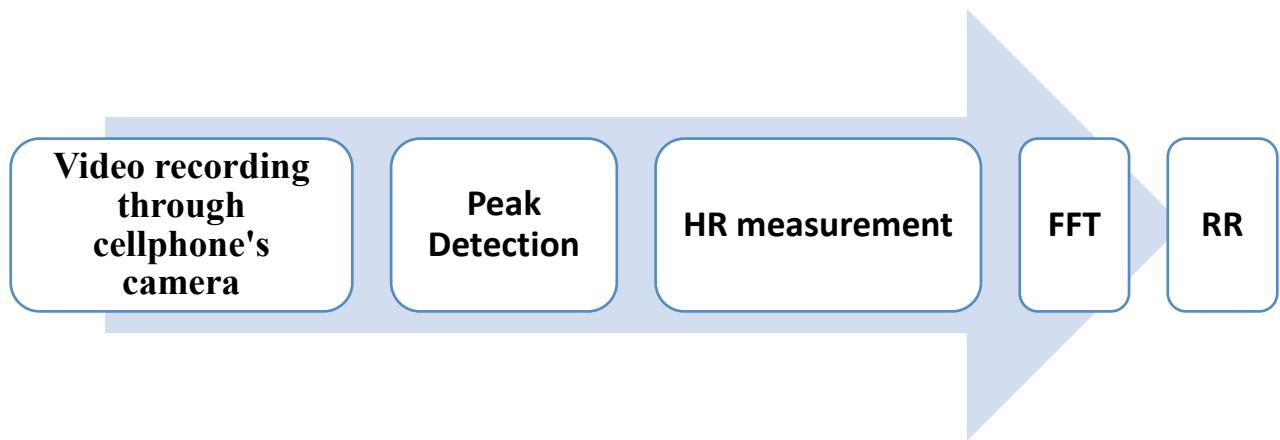


Figure 7.5 Overview of the method.

To measure the Respiratory Rate:

- Using a wake lock mechanism to Make sure that device doesn't go to sleep mode until finish RR processes" Wakelocks are power-managing software mechanisms, which make sure that your Android device doesn't go into deep sleep (which is the state that you should strive for), because a given app needs to use your system resources".
- Open device camera and set orientation to 90 degree.
- Turn on Flash Light of Camera and start record frames for 30 Sec.
- Store system time as a start time for the analyzing process.
- Getting frames data "PPG signal" of Green and Red Colors from the camera by applying image processing method on each frame to get the RGB intensities and start the heartbeat process.
- Count number of frames in 30 seconds.
- Check if we got a good red intensity to process.
- Calculating the sampling frequency using FFT Algorithm.
- A band pass filter must be applied after (FFT) from 0.1Hz to 0.4Hz with 0.2Hz center frequency to get the RR. (Fft2).
- Make sure that if the respiration rate from red and green intensities are reasonable.
- Take the average between them, otherwise take the green or red if one of them is good.

- Get the RR value, Then Storing it in the user's History log.

After acquiring HR from the PPG signal, the next step was extracting BR from the HR in the spectrum domain. This is possible because respiration rate modulates amplitude and frequency of a signal. Before spectral analysis, the HR signal was interpolated in order to address the issue of irregular sampling from the cellphone and because R-wave are not equidistantly timed events. After this, the fast Fourier transform (FFT) of the HR was computed. We observed that the FFT plots showed a clear harmonic peak at the frequencies, which corresponded to the respective breathing rate. Figure 7.6 shows the FFT of the HR signal.

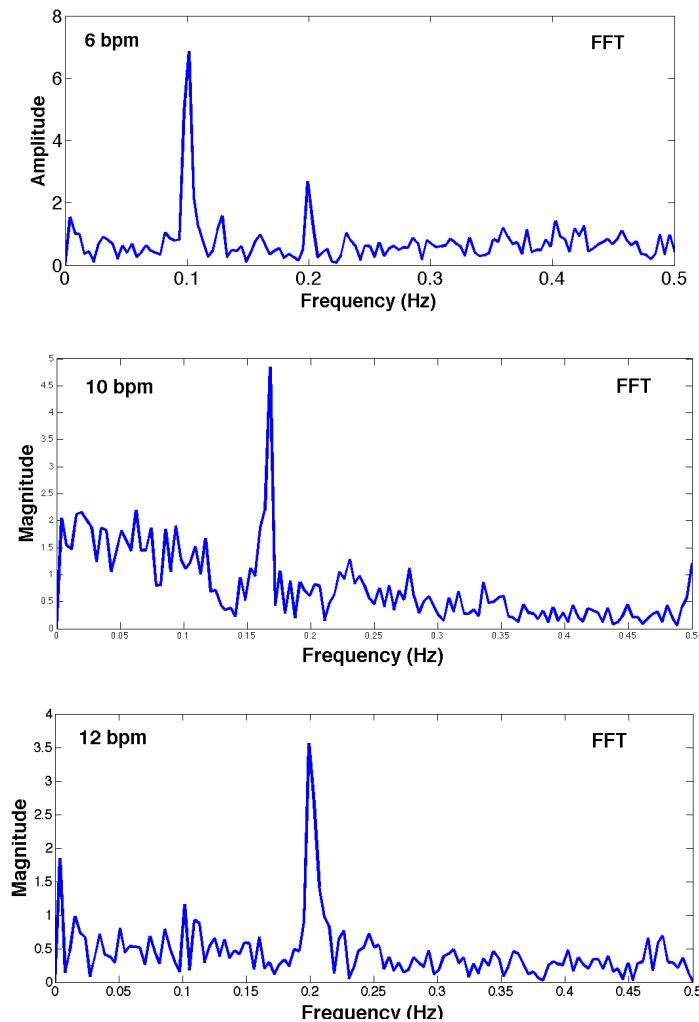


Figure 5.6 Represents the fast Fourier transform plots. Each harmonic peak corresponded to the respective breathing frequency, in which the samples were collected. (a) Breathing frequency of 0.1 Hz corresponds to 6bpm, (b) 0.16 corresponds to 10bpm and (c) 0.2Hz corresponds to 12bpm.

### **5.3 Blood Pressure (BP):**

Measurement of vital signs or parameters in a human is an arduous task when sudden dizziness or fainting could occur during unexpected situations. The only information that we are aware of is that these are the most common symptoms of low blood pressure. However, to check a person's blood pressure in such situations, we need for a portable, convenient device or apparatus. Despite the availability of digital wrist and arm blood-pressure meters, most people do not carry these devices during their daily travel to work place, gymnasiums, recreating facilities, etc.

Blood pressure, the amount of force applied on the walls of the arteries when the blood is forced throughout the body, depends on factors such as the amount of blood in the body, the pumping rate of the heart, the flexibility of the arterial walls, and resistance to blood flow due to the size of the arteries. A continuous variation of blood pressure occurs in a human due to factors such as physical activity, medication, anxiety, and emotions.

The body has unique mechanisms to regulate a person's blood flow; whenever a person's blood pressure drops, the heart rate increases to pump more blood and the arterial walls contract to increase the blood pressure. The blood pressure is given by two numbers measured in millimeters of mercury (mmHg).

- SV - Stroke Volume: The Volume of blood pumped from the heart in one beat.
- Pm - Mean Arterial pressure: Medical term to describe the average blood pressure of an individual.
- Ps - Systolic Pressure: The blood pressure when the heart is contracting.
- Pd - Diastolic Pressure: The blood pressure when the heart is relaxing.
- Pp - Pulse Pressure: The difference between the systolic and diastolic pressure.
- HR - Heart Rate/Pulse rate: The number of heart beats per minute.
- Z - Impedance to blood flow: The total opposition of blood in the vascular system for blood flow.
- R - Resistance to blood flow: The measure of blood opposition in the blood vessel for steady flow.
- Na - Difference in video intensity within a single systole and diastole.

- Sa - Peak amplitude in video intensity when the cuff pressure equals systolic pressure.
- Da - Peak amplitude in video intensity when the cuff pressure equals diastolic pressure.
- ET - Ejection Time: Time taken for the opening and closure of the aortic valve.
- BSA - Body Surface Area: Physiological term to measure or calculate the surface area of the human body.
- Q - Cardiac Output: The volume of blood being pumped by the heart in a time interval of one minute.
- V TT - Vascular Transit Time: The transmission delay of the first heart sound to be felt as a pulse in finger.

This estimation method follows a pattern similar to the auscultation techniques to listen for the internal sounds of the body, where a camera sensor is used to view the blood flow in the arteries.

The pulse pressure is computed by correlating the time difference between the occurrence of systolic and diastolic pulse in the graph. I formulated a regression equation based on the pressure readings on the meter to the time it takes to reach the value. Figure 7.7 (a) shows the curve fit to the data for a decrease in cuff pressure on the automatic cuff system. By applying the regression equation,

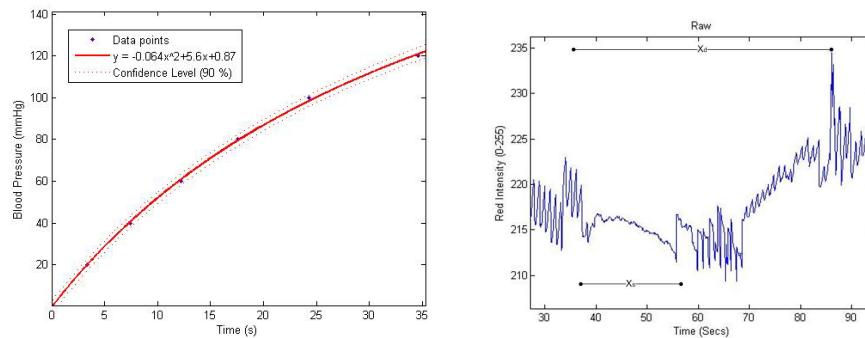


Figure 6.7 (a)Regression relation for the pressure value on the meter, corresponding to the time it takes to reach that value from full inflation of cuff (b) Determination of  $X_s$ ,  $X_d$  which are the time taken to reach systolic and diastolic pressure on the cuff respectively.  $y_s$  and  $y_d$  were determined from  $X_s$ ,  $X_d$  in 3(b)

based on the regression equation derived in (7.2)(a)

I can determine the blood pressure at a certain instance of time.

$$y = -0.064x^2 + 5.6x + 0.87 \quad (7.2)$$

The proposed approach utilizes only a mobile phone and an automatic cuffing system; hence, the pulse pressure has to be determined using a naive approximation method. The results obtained through the phone along with the cuff mechanism clearly show the distinctive regions of the systolic and diastolic phase consistently; therefore, this information from the graph can be used to predict the pulse pressure. The pulse pressure is obtained by taking the change in pressure between systolic and diastolic time. Therefore, the pulse pressure is calculated by

$$P_p = y_d - y_s \quad (7.3)$$

Where  $y_d$  and  $y_s$  are calculated based on the regression equation using  $x_d$  and  $x_s$  from the Figure 5.7 From the systolic and diastolic peaks observed, the pulse pressure values were calculated using equation (5.3). The regression equation changes for differing cuffs and will be linear if the cu\_ pressure reduced steadily and gradually.

Even though the measured values of pulse pressure are approximate, the estimated values reveal closeness to actual values for some datasets. The person's heart rate was determined by taking the video sample of around 15 - 30 seconds before inflating the cuff. Now having obtained the pulse pressure, I can calculate the stroke volume from equation 8 by applying the person's corresponding age and weight. When stroke volume is considered as a constant over a period of time, the pulse pressure must stay constant since stroke volume and pulse pressure are directly related.

Table 6.3 shows the accuracy of finding the pulse pressure with the proposed method, conducted on five subjects at different times. I was able to observe that, the calculated accuracy of pulse pressure varied between 50 - 100 % over the datasets.

Having the pulse pressure, I can calculate the mean arterial pressure for the subject using the equation

$$P_m = HR \cdot SV \cdot Z \quad (5.4)$$

Table 6.3 Estimated Accuracy of Pulse pressure for five different subjects

Subject	Measured P <sub>p</sub>	Actual P <sub>p</sub>	Accuracy %
Sample 1	30	20	50
Sample 2	22	21	95.2
Sample 3	30	24	75
Sample 4	36	36	100
Sample 5	44	48	91.67

Impedance [17] is a major factor in the case of impedance cardiography technique for determining a person's blood pressure.

$$Z = \frac{R}{1+iwRc} \quad (7.5)$$

Due to unavailability of impedance graphs, I created a crude relation for describing the resistance of blood flow with the amplitude of the waves. Now, the impedance factor Z has to be determined to calculate the mean arterial pressure. Because Z is calculated only for impedance cardiograph, an alternative method was required to determine the impedance factor for a dataset. In my proposed system, data is normalized and a crude value of Z is calculated using

$$Z = \frac{Na}{Sa - Da} \quad (5.6)$$

where Sa, Da denote peak amplitude in the video frame at the systolic and diastolic time instants respectively, and Na represents the difference in video intensity during one systolic and diastolic cycle. The values of Na, Sa, and Da are obtained from the Figure 5.8. Calculation of Z is based on an assumption that, the ratio between a normal blood flow without external resistance to the difference between the peak amplitudes with the maximum.

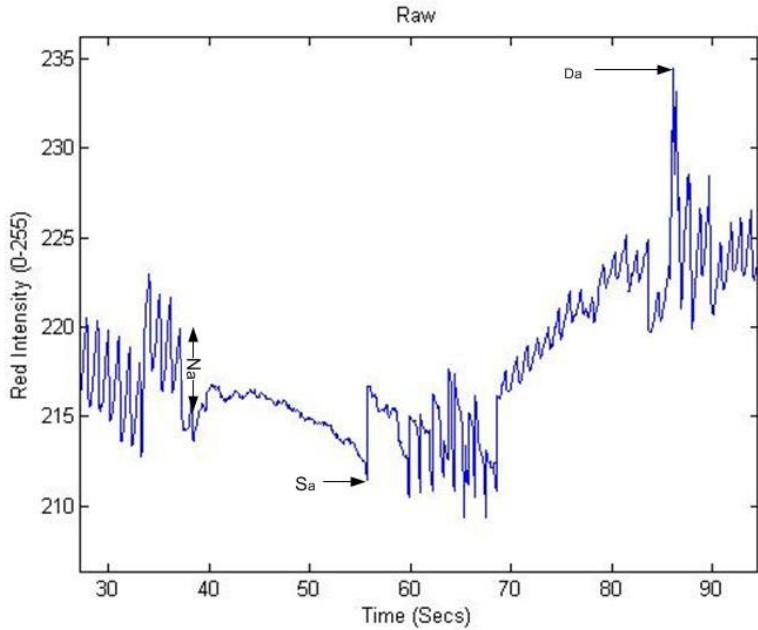


Figure 7.8 Estimation of Z. Values of Na, Sa, Da are obtained as shown

substituted in equation (7.6) external resistance gives the impedance factor for the flow. However, currently, no test equipment exist to support my assumption of the impedance factor theory.

A person's stroke volume (SV) is determined based on the approximate amount of blood in the person and their pulse pressure. Equation (6.7) [18] gives the value of stroke volume (the blood pumped by heart in each beat).

$$SV \text{ (mL)} = (0.013 * \text{Bodyweight(kgs)} - 0.007 * \text{Age(years)} - 0.004 * \text{HR} + 1.307) * Pp \quad (5.7)$$

With the help of the mean arterial pressure and pulse pressure, the systolic and diastolic pressure can be determined. In medical terms, pulse pressure is determined from the impedance cardiogram during a continuous pressure measurement.

$$Ps = Pm + 2/3 Pp \quad (5.8)$$

$$Pd = Pm + Pp / 3 \quad (5.9)$$

Where  $P_m$  and  $P_p$  are mean arterial and pulse pressure, respectively. Having values for the necessary variables, I can estimate the systolic and diastolic pressure by substituting these values in the equation (5.8) and (5.9). Based on my assumptions and analysis, I determined the blood pressure values (tabulated in Table 4.2).

Table 5.4 Accuracy of  $P_s$  and  $P_d$  for five different individuals

Subject	Measured		Actual		Mean Accuracy %
	$P_s$ (mmHg)	$P_d$ (mmHg)	$P_s$ (mmHg)	$P_d$ (mmHg)	
1	90	60	85	65	91
2	76	54	87	66	84.5
3	74	44	89	65	75.5
4	106	70	102	66	96.1
5	130	86	123	75	89.7

From Table 5.4 it can be seen that the pressure values are imprecise, but a person's bp level (high, medium, low) can still be determined. However, the magnitude of error increases rapidly, when the determined values of the variables are erroneous. I validated my measurements with the help of a commercial blood pressure meter, but this cannot guarantee a high level of accuracy. All the tests were done multiple times to determine the consistency of the plots.

- Using a wake lock mechanism to Make sure that device doesn't go to sleep mode until finish BP processes "Wakelocks are power-managing software mechanisms, which make sure that your Android device doesn't go into deep sleep (which is the state that you should strive for), because a given app needs to use your system resources".
- Open device camera and set orientation to 90 degree.
- Turn on Flash Light of Camera and start record frames for 30 Sec.
- Store system time as a start time for the analyzing process.
- Getting frames data "PPG signal" of Green and Red Colors from the camera by applying image processing method on each frame to get the RGB intensities and start the heartbeat process.

- Count number of frames in 30 seconds.
- Check if we got a good red intensity to process.
- Calculating the sampling frequency using FFT Algorithm.
- Make sure that if the heart rate from red and green intensities are reasonable.
- Take the average between them, otherwise take the green or red if one of them is good.
- Perform some medical equations to calculate the value of BP Using some values such as heart rate and some user data like age, weight, height and gender.
- Storing the BP value in the user's History log.

## **5.4 Oxygen Rate:**

Human life relies on the oxygen level in the blood. Normally, a small fraction of molecular oxygen transported by hemoglobin is dissolved in healthy people's blood. Hence, assessing oxygen saturation ( $\text{SpO}_2$ ) – the fraction of oxygen-saturated hemoglobin relative to total hemoglobin – is critical to indicate the health status of human brain, heart, and respiratory system. Referred to as peripheral oxygen saturation ( $\text{SpO}_2$ ), pulse oximetry is a common physiological measurement of  $\text{SpO}_2$  in both in-hospital environment and at in-home healthcare. Such the  $\text{SpO}_2$  monitor is done using a noninvasive pulse oximeter by emitting light at specific wavelengths into an area of the body (e.g. finger, toe, earlobe, etc.). While FD Approved pulse oximeters provide a reliable  $\text{SpO}_2$  level, they have a number of short comings including (1) high price, (2) large probe, (3) ill-fitting finger, and (4) external device carry requirement.

Recently, mobile applications for  $\text{SpO}_2$  estimation have been widely deployed and address those problem in different ways. Existing work commonly made use of build-in flashlight and camera on smart phones to predict the oxygen level. Particularly, users touch their finger on the camera surface to capture the reflected lights coming from a particular light source. On the other hand, users can simply carry their smart phone installed with the Health Watcher app while walking. By analyzing their walking gait, this app can passively predict their  $\text{SpO}_2$  level. However, all of them provide a low accurate  $\text{SpO}_2$  level that is absolutely not intended for use with medical quality.

This work proposes a novel practical oxygen saturation sensing system, which has the potential to accurately provide the SpO<sub>2</sub> level. In details, the system includes an add-on containing multiple filters and clapped to the smart phone as simply as using a phone case. By leveraging the advancement of 3D printing technology, the add-on is really low-cost and lightweight. Due to limitations of camera hardware, the reflected light captured by the device needs to be further processed to obtain usable PPG signals. Thus, our system takes the recorded frames and processes to acquire a clear PPG signal. The high-quality signal is finally input into a non-linear calibrated model to obtain the SpO<sub>2</sub> level.

## 5.5 Challenges

The system design addresses the following challenges.

- In existing apps, color channels have been used to substitute for the use of LED Red and Infrared lights in commercial pulse oximeters. Even though such applications bring a simple design to the hardware, their accuracy is not high due to a short distance among visible wavelengths. This phenomenon can be reasonably explained using a relational graph between the SpO<sub>2</sub> and the choice of light sources. Accordingly, our main goal aims at keeping one in the range of Red and pushing the other closely to Near-Infrared. Thought that the problem can be solved by adding multiple LEDs to the front or back of the smartphone, it, however, seems to be displeasing where it will cost an amount spent for extra components.
- The second problem arises when people sometimes place their finger outside the camera region. As a result, the average intensity includes non-pulsatile pixels that possibly increases the prediction error. In addition, mobile devices nowadays come with various designs for the camera and technical implementation for its lens integrated in the image sensor and the flashlight.
- Recall that the absorption ratio is calculated following the change of oxygen hemoglobin in the circulation corresponding to the minimum and maximum extrema in the PPG signal. Diacrotic notch, on the other hand, is a certain stage in the circulation system that can cause a sudden closure of aortic valve to produce an almost flat region in the middle of cycle. Despite of the short period of this phase, it still accumulates to the prediction error. Therefore, it should be profoundly removed also with other possible noise and distortion.

Our system design, which aims to extract the Red and IR from a white light using a 3D add-on, is efficient in term of time alignment and wavelength specification.

**Spatial modulation:** Light rays from flash-light bounce off our finger and penetrate through the filters set in front of camera lenses. The goal of our mounted camera add-on is to (1) assist the reflected lights toward the camera lenses (2) extract explicitly the IR and Red at the same time stamps. The add-on component not only limits the range of wavelengths but also assists the lights not to be off the camera lenses region, which is the main cause in the reflectance mode.

**Raw PPG acquisition based adaptive ROIs:** Under our specific screen division, trivial adaptive regions selection such as intensity based or taking the image centers are inadequate. Therefore, we are motivated to use a spatial sliding window which satisfies (1) spatial stability and (2) strong temporal variation. While spatial stability represented by the variance of intensities, the second condition estimates the variance of average intensities in one specific amount of time. Shifting a window by one to another pixel and estimate these statistical parameters would not be computationally efficient due to the overlapping patches. Therefore, to avoid this kind of recalculation issue, we approach to use the Integral Image with a suitable modification of the mean and variance calculation. The output PPG undergoes a linear phase filter that helps to remove unwanted frequencies except heart rates. A window size is chosen as 6 seconds as sufficient to gain enough number of peaks and troughs for estimating values R. In each window, a dicrotic removal refines the location of local extrema by cleaning all the parts of dicrotic. Those problems coming from light scattering, image intensity versus real-intensity can be handle through a calibration at training.

**Non-linear calibration:** In our opinion, the SpO<sub>2</sub> should follow the Beer-lambert equation. The values of absorption parameters are predicted through a non-linear regression model. In general, our system consists of 4 main stages that processes the raw signal, camera frame, to predict oxygen level.

(1) Extracted PPG signals from Red and IR filter are denoted as s<sub>r</sub> and s<sub>ir</sub>. Each point t in PPG signal is calculated by the mean value at frame t of specific channel.

$$s_r(t) = \mu(\text{InpV id}\{U_r^t, t\}) \quad (5.10)$$

$$s_{ir}(t) = \mu(\text{InpV id}\{U_{ir}^t, t\})$$

(5.11)

with  $t = 1, T$  and  $U$  denotes for the region of interests in specific time.

(2) A bandpass filter is utilized to remove components that are not related to the pulsatile signals. We define the signal frequencies are from 40 bpm to 230 bpm following the range of human heart beat.

(3) Dicrotic notch is the natural factor that appears in the shape of PPG signal. Absorbtivity ratio is an estimation between the local peak and trough, which the dents along signal can distract our measurement. According to the nature property of dicrotic which has the length or dircortic parts them-self compares to the true local maxima and minima. Obviously, the distance of two successive max and min in dicrotic component is much smaller than that of the real local extremes.

Therefore, by sorting all the distance between two consecutive peaks and troughs and looking for the points where the abnormal change occurs, we can remove all the local extremes from the beginning up to this point as considered to be dicrotic parts. At this stage a sliding window is used to segment the signal into small chunks and predict the oxygen saturation in each of them.

(4) AC ( $I_p$ ) and DC( $I_b$ ) component of each channel (wavelength) is obtained using the corresponding standard deviation or (difference of max and min value) and mean of each ppg signal.

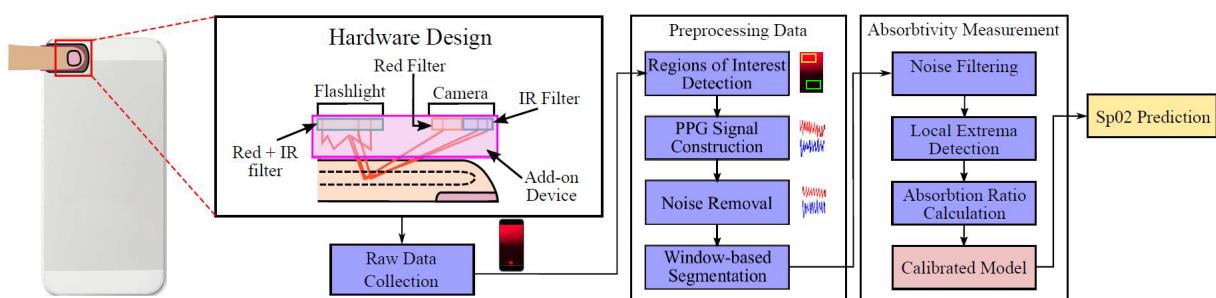


Figure 5.9 Overall system design and architecture of measuring Sp02 with dual filters.

After that, we can receive the absorptivity ratio using

$$R = \frac{AC(Sr)/DC(Sr)}{AC(Sir)/DC(Sir)}$$

(5.12)

(5) The intensity of light can be referred to the number of photon perceived per a unit area. Theoretically, it should be dependent on wavelength, but, due to the Quantum Efficiency which defines the percentage of photons can be successfully converted. Specifically, lights penetrates differently on the depletion layer of Charges Couple Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) according to their wavelength. Therefore, a calibration procedure is needed to compensate the degradation. In practice, the coefficients are non-linear calibrated with ground truth data to suit with different camera models.

- Using a wake lock mechanism to Make sure that device doesn't go to sleep mode until finish O2S processes "Wakelocks are power-managing software mechanisms, which make sure that your Android device doesn't go into deep sleep (which is the state that you should strive for), because a given app needs to use your system resources".
- Open device camera and set orientation to 90 degree.
- Turn on Flash Light of Camera and start record frames for 30 Sec.
- Store system time as a start time for the analyzing process.
- Getting frames data "PPG signal" of Red and Blue Colors from the camera by applying image processing method on each frame to get the RGB intensities and start the heartbeat process.
- Count number of frames in 30 seconds.
- Check if we got a good red intensity to process.
- Calculating the sampling frequency using FFT Algorithm.
- Perform some medical equations to calculate the value of O2S Using value of heart rate.
- Storing the O2S value in the user's History log.

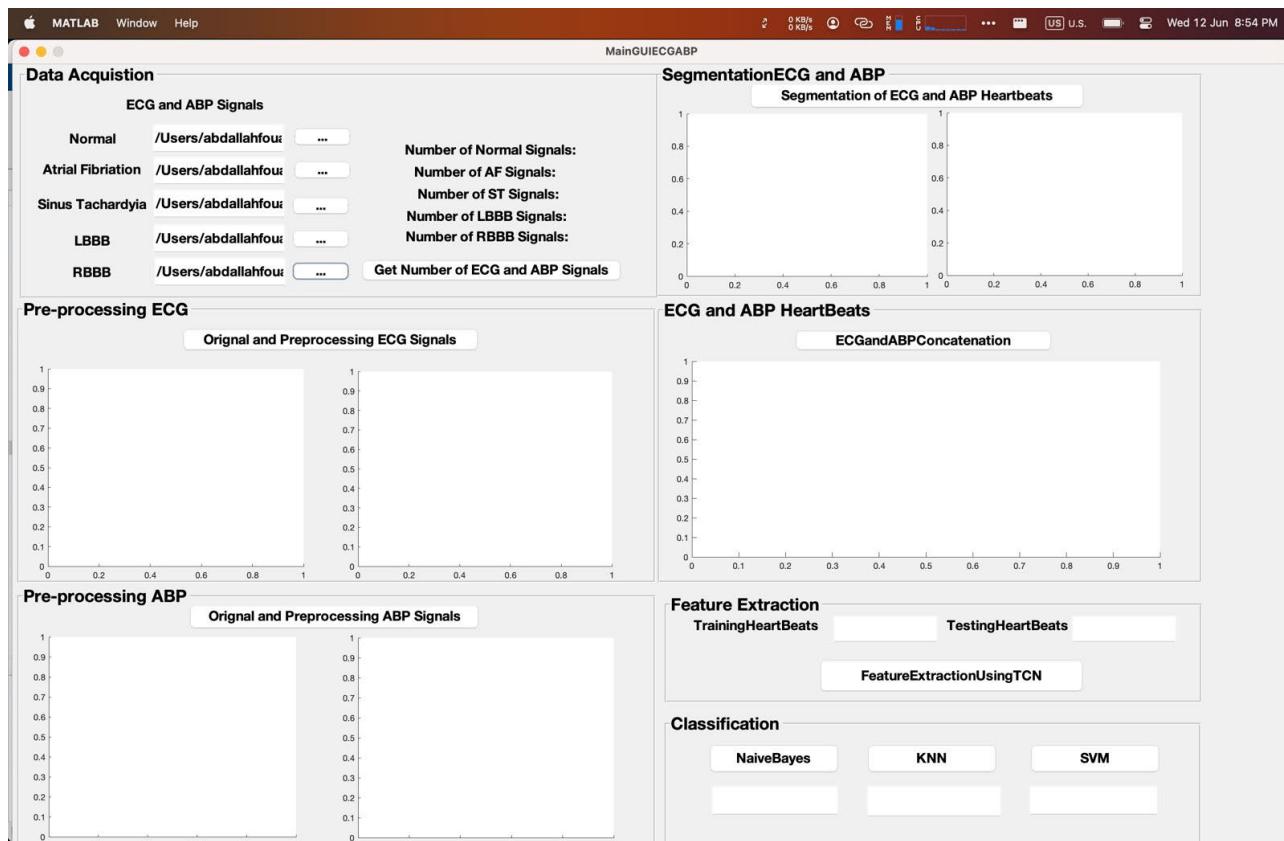
# **Experimental Results**

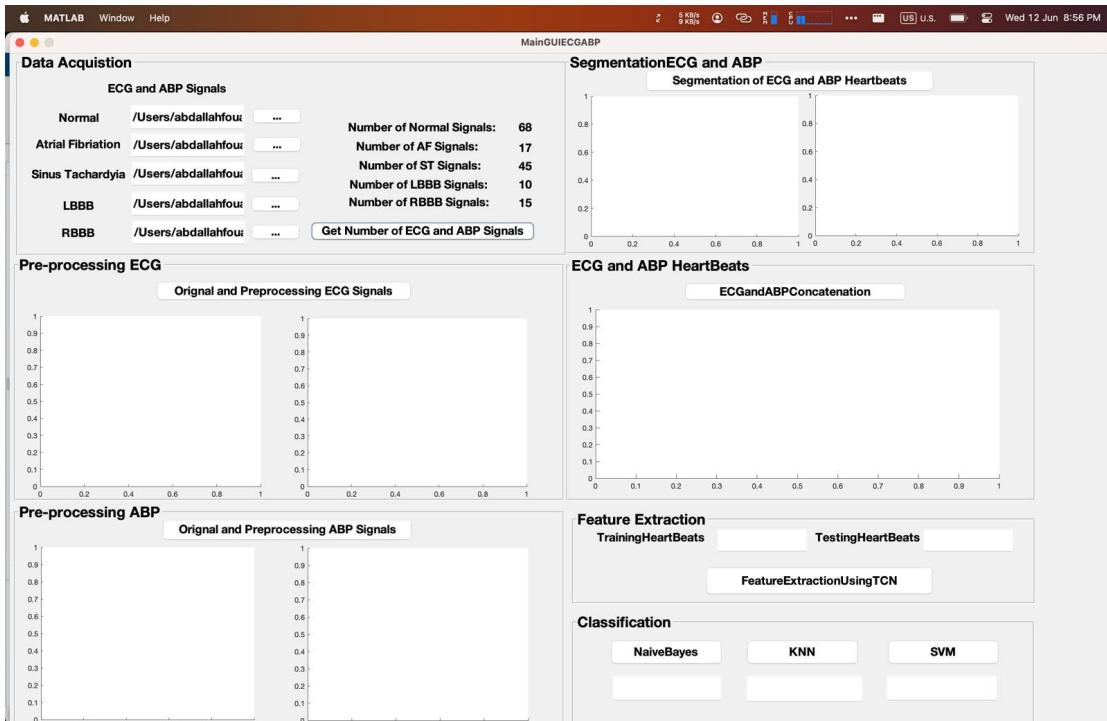
## 6. Experimental Results

We have developed a mobile application interface. This application provides two main features. The first feature is the diagnosis and the second features is a complete check up on the vital signs in the human body. The diagnosis as mentioned before is based on ECG combined with ABP to determine five main heart diseases, and it is also based on the combination of ECG and PCG to determine the normality and the abnormality. The diagnosis is implemented using Matlab and integrated with the mobile application.

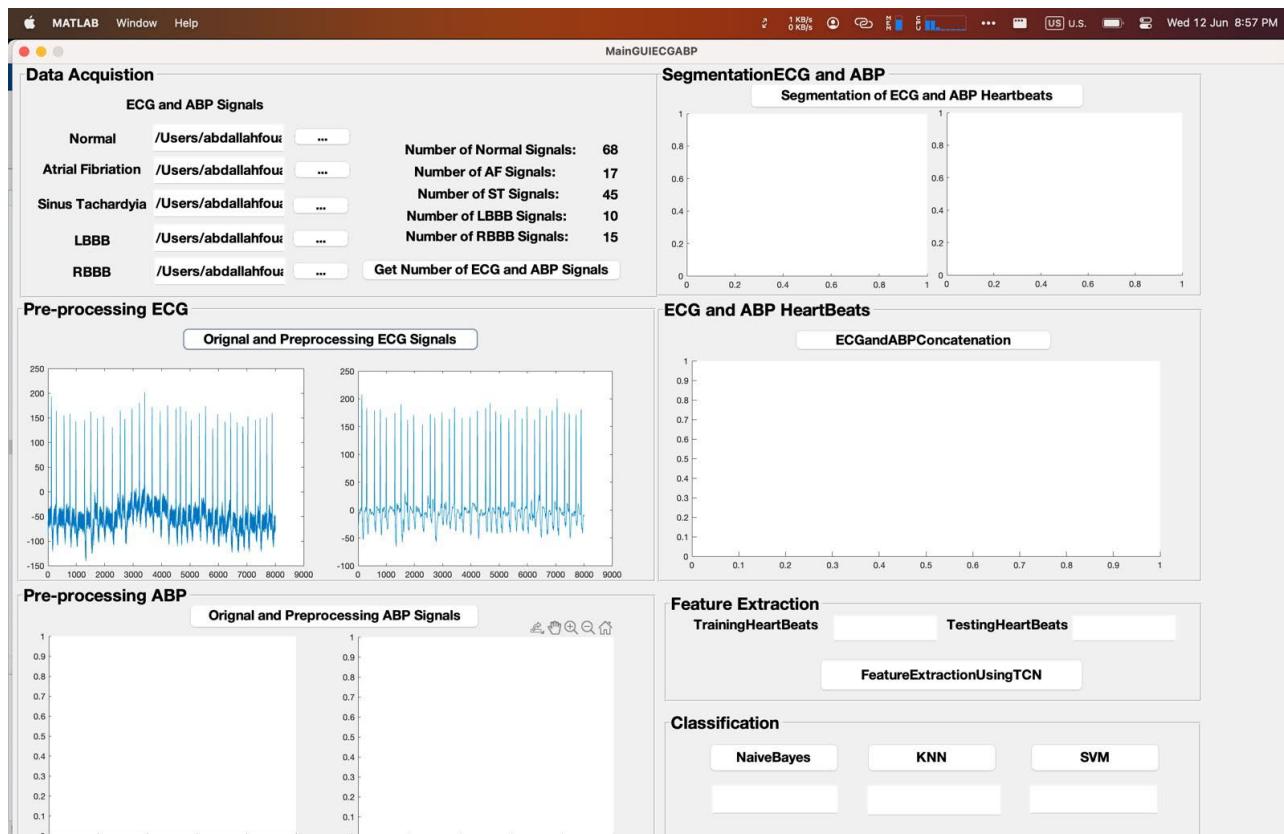
### 6.1 ECG-ABP Methodology

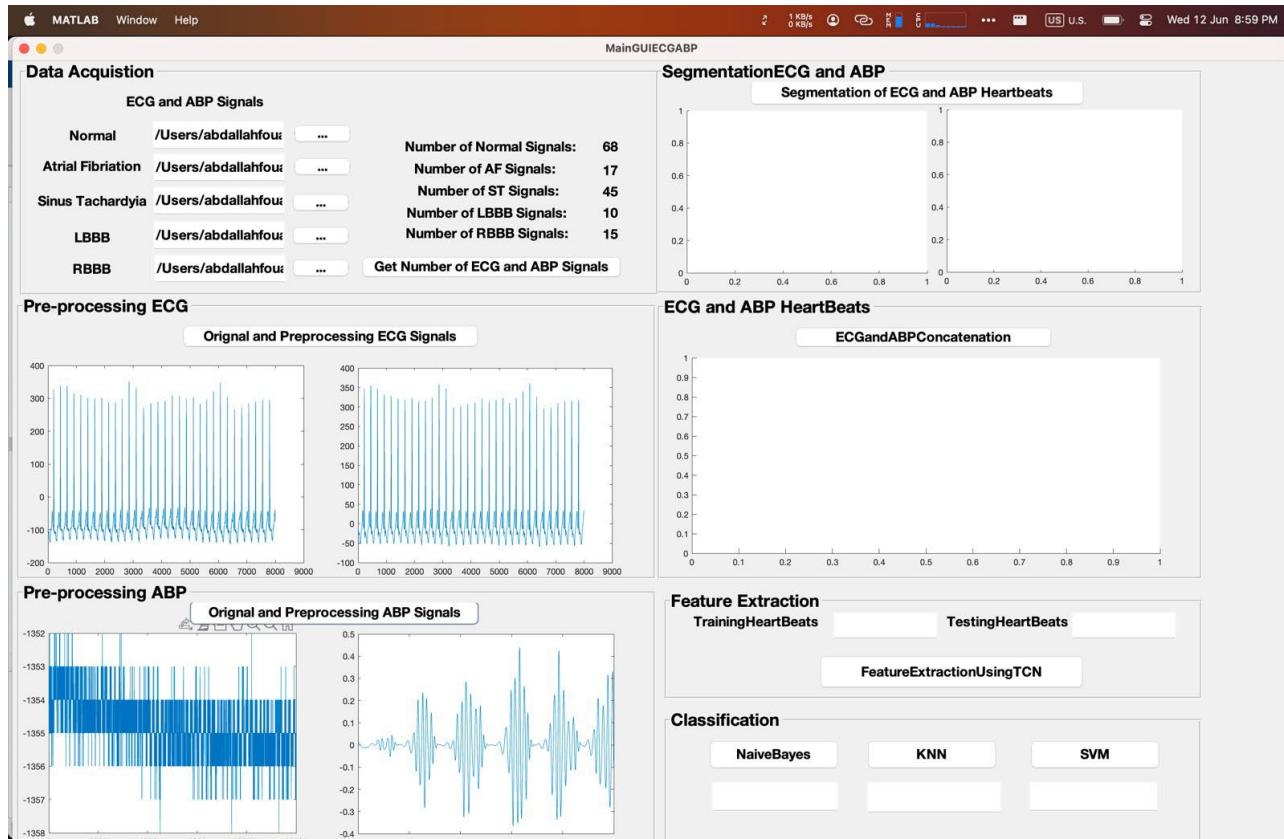
#### 6.1.1 ECG-ABP Signals Upload



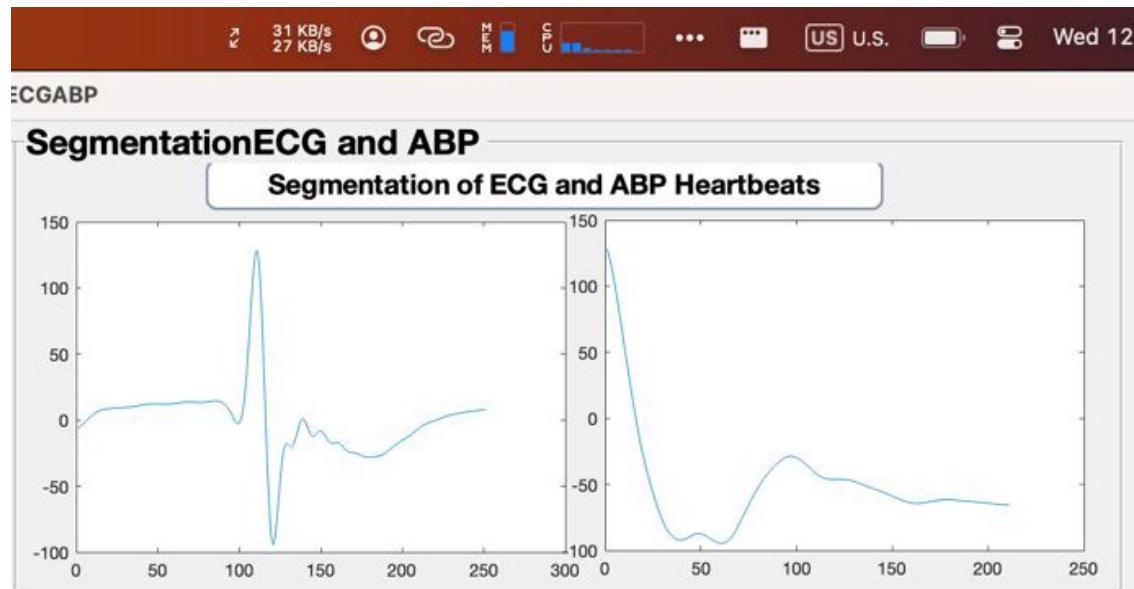


## 6.1.2 ECG-ABP Signals Pre-processing

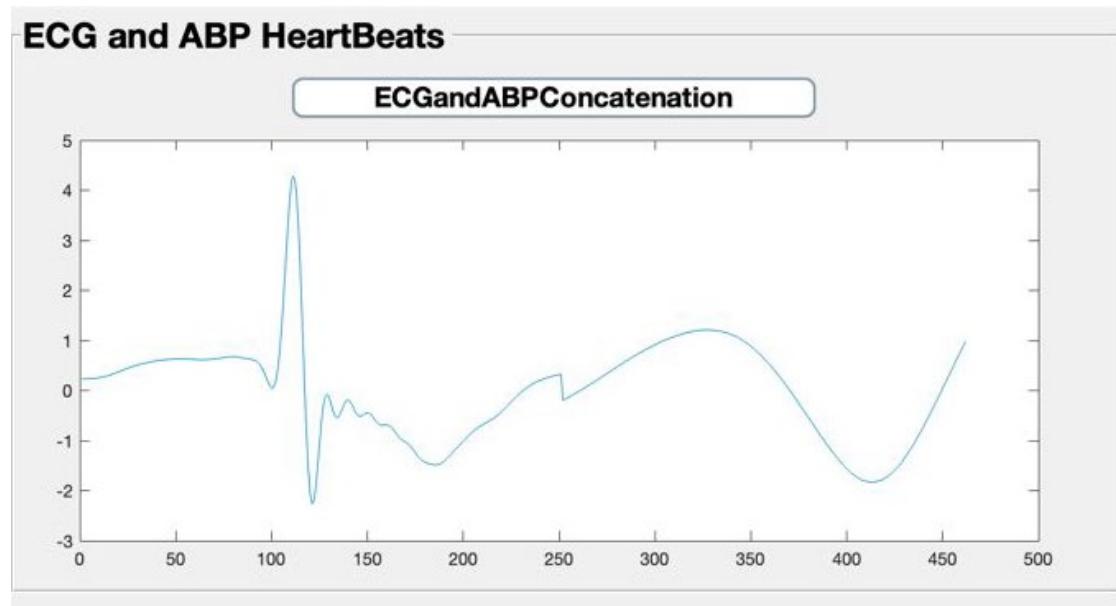




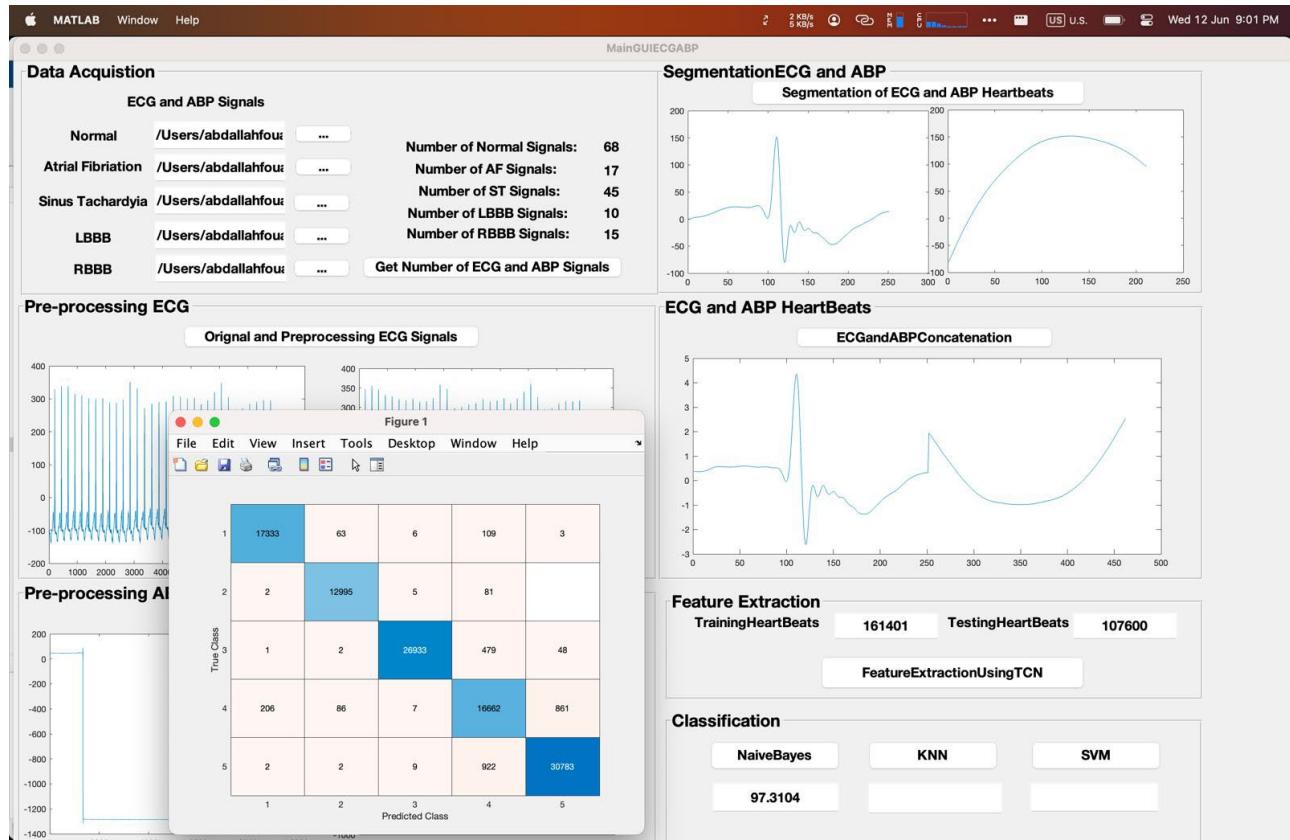
### 6.1.3. ECG-ABP Signals Segmentation

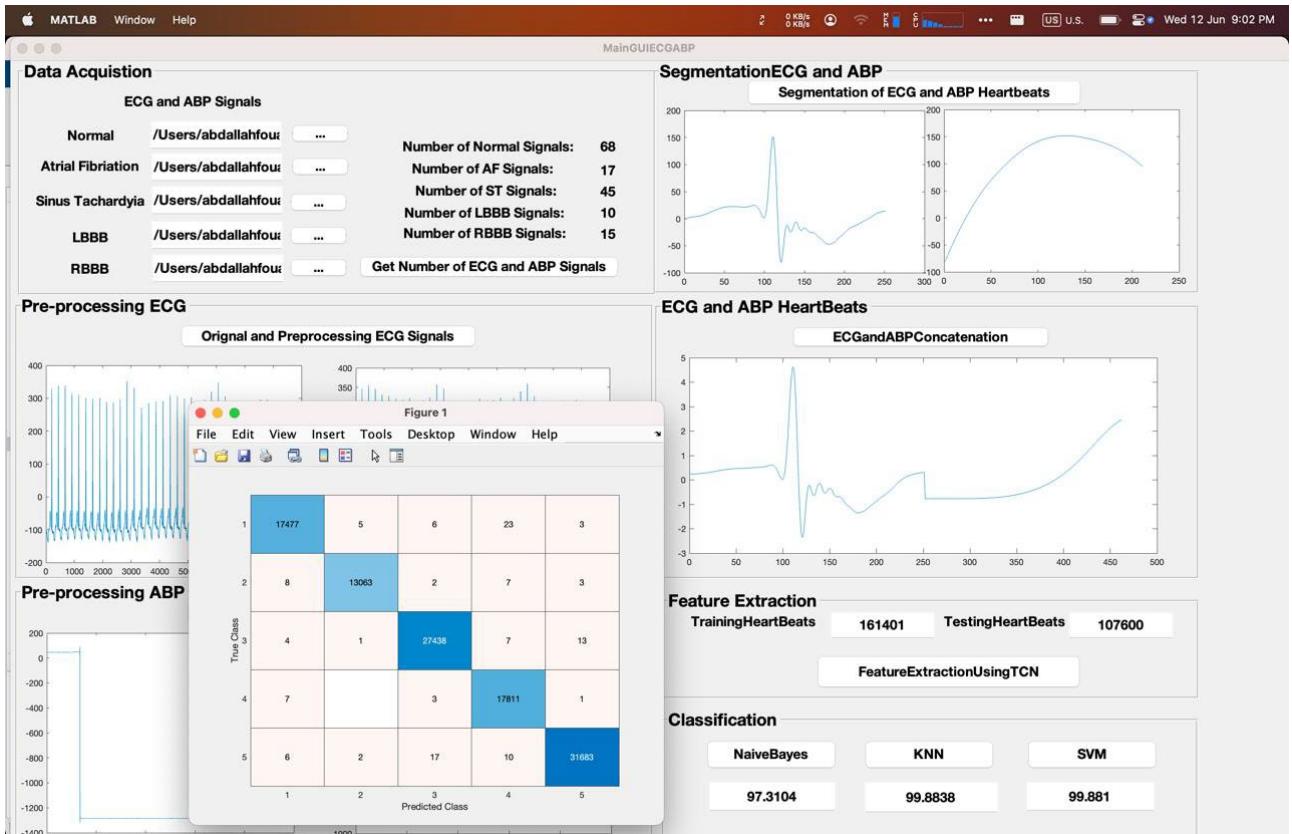
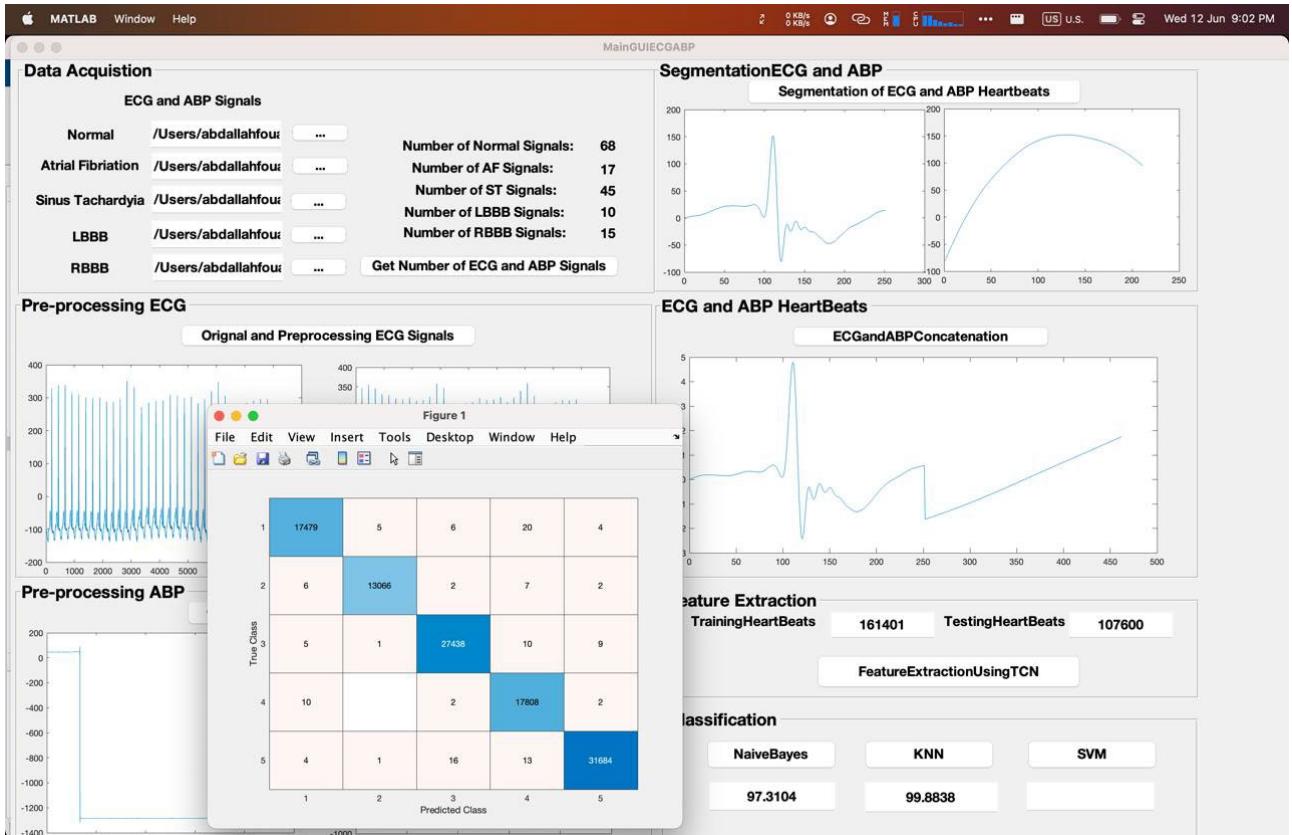


## 6.1.4. ECG-ABP Signals Concatenation



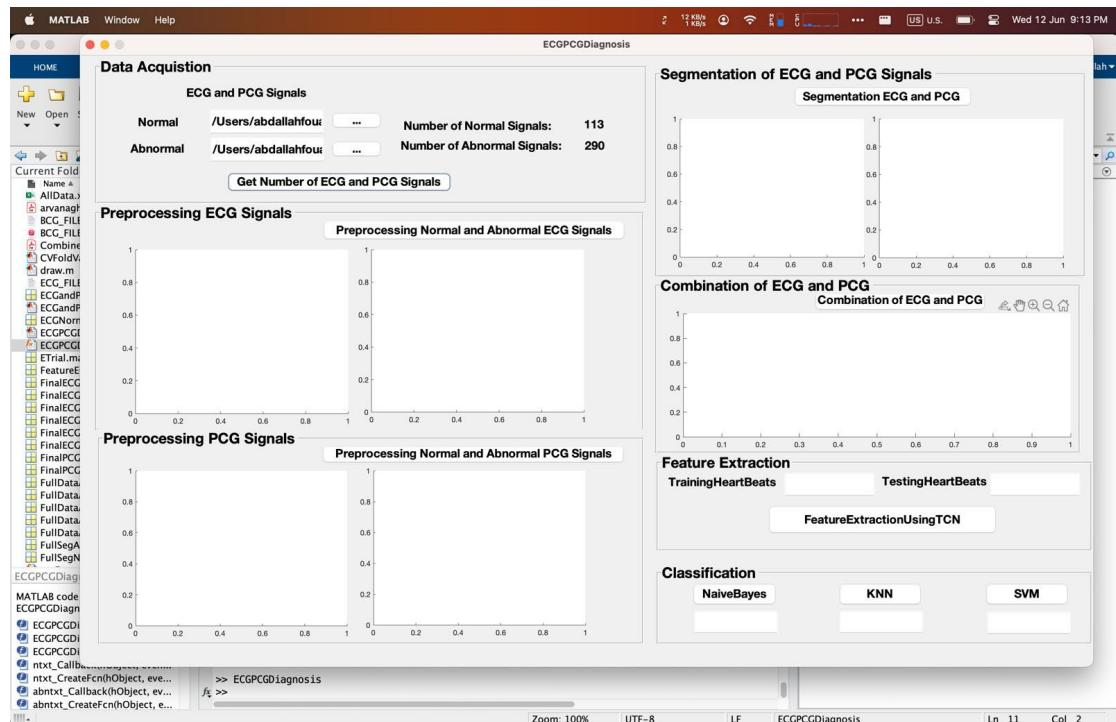
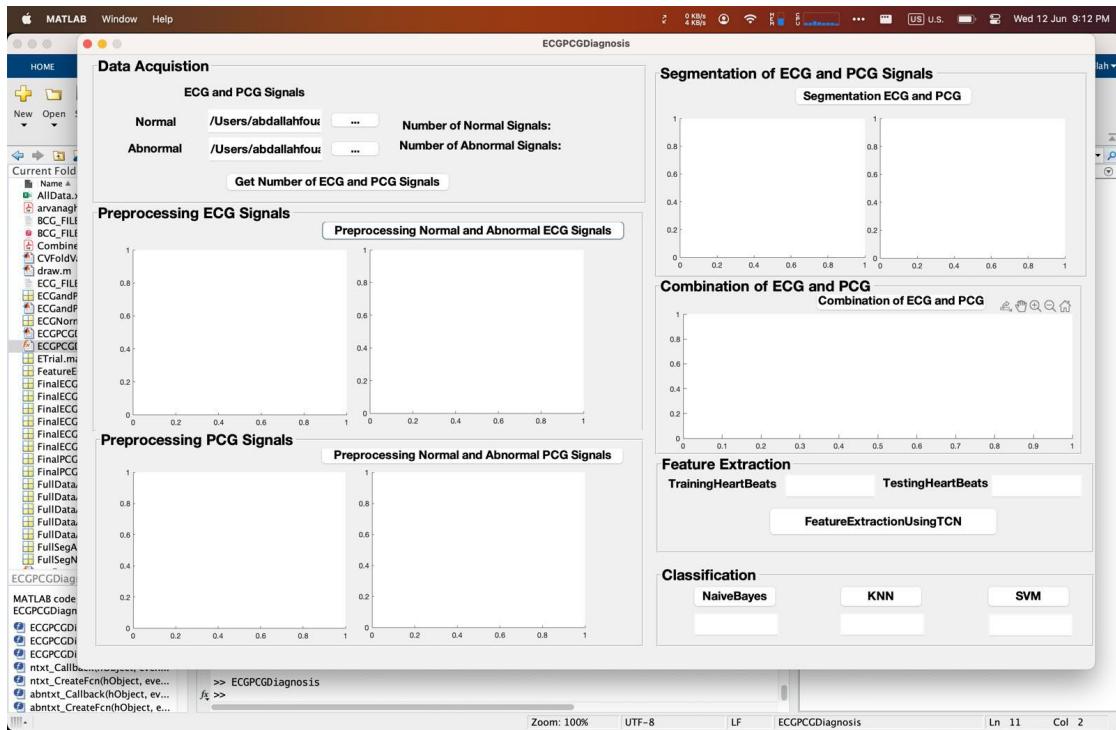
## 6.1.5 ECG-ABP Signals Feature extraction and Classification



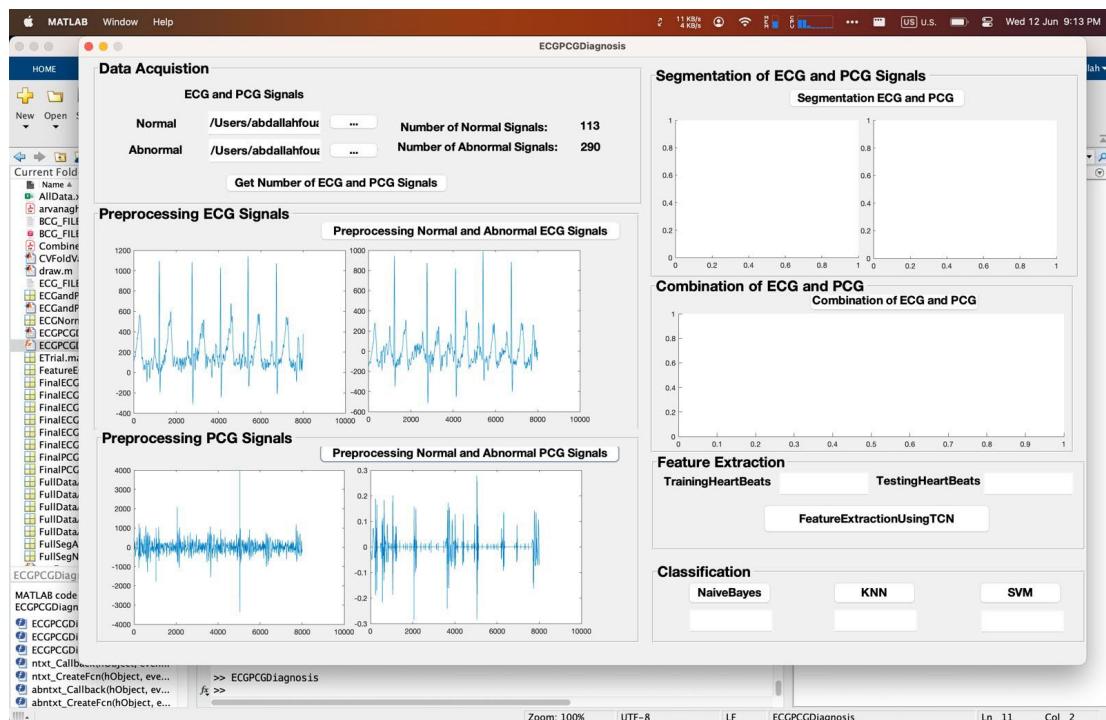
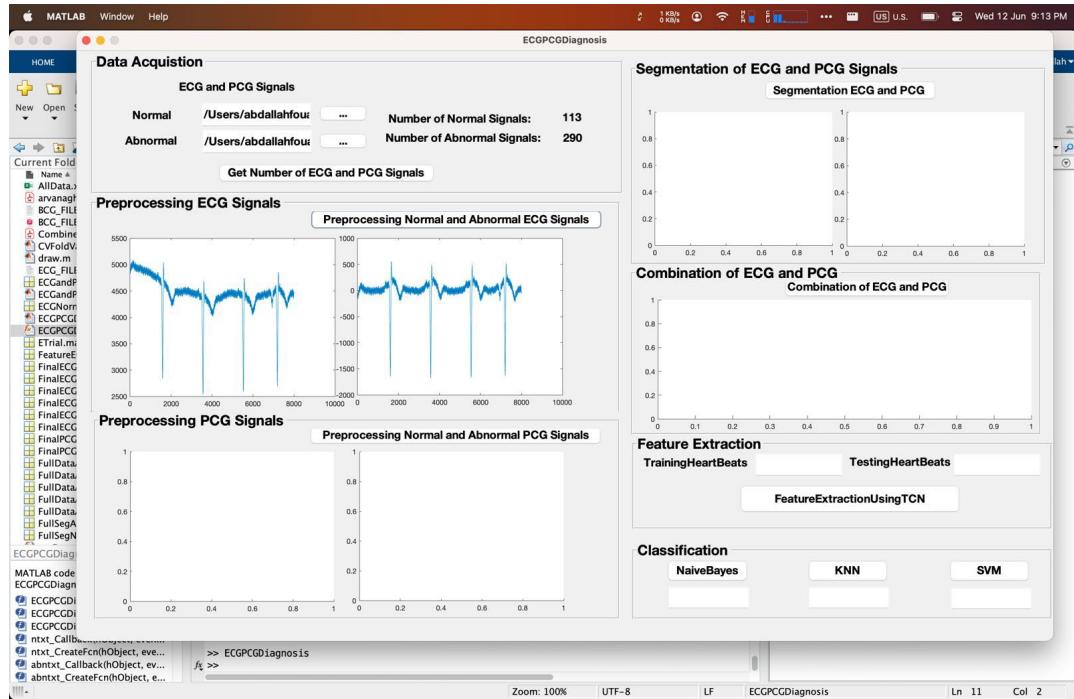


## 6.2 ECG-PCG Methodology

### 6.2.1 ECG-PCG Signals Upload



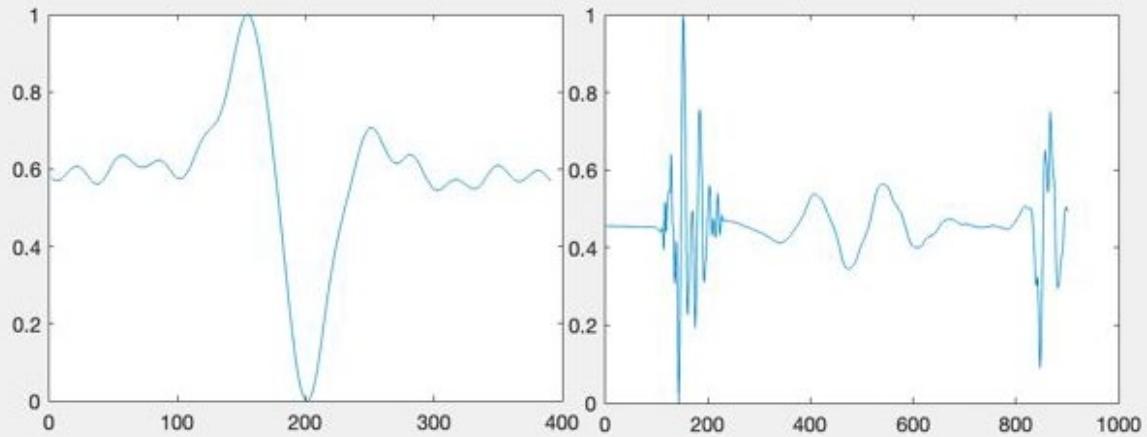
## 6.2.2 ECG-PCG Signals Preprocessing



### 6.2.3 ECG-PCG Signals Segmentation

#### Segmentation of ECG and PCG Signals

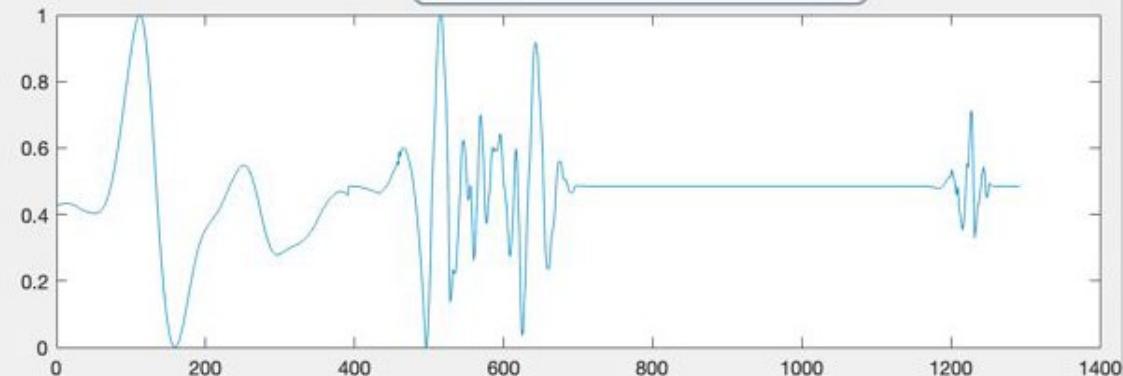
Segmentation ECG and PCG



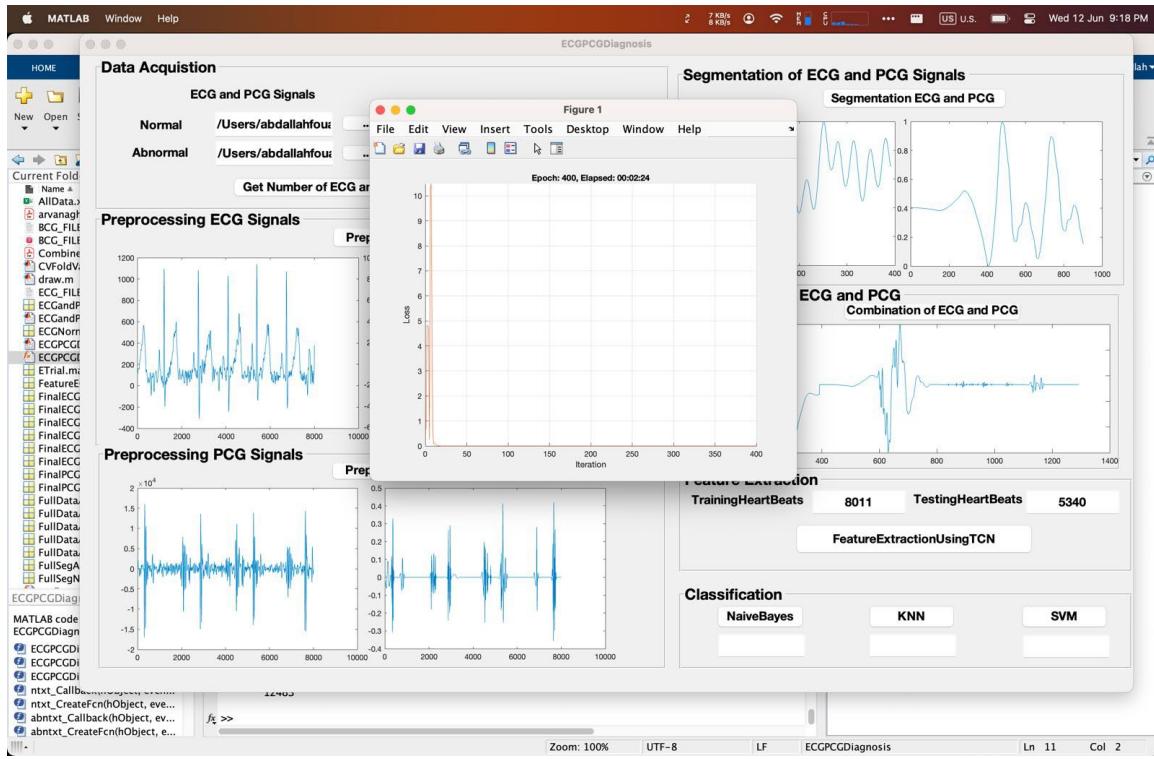
### 6.2.4 ECG-PCG Signals Concatenation

#### Combination of ECG and PCG

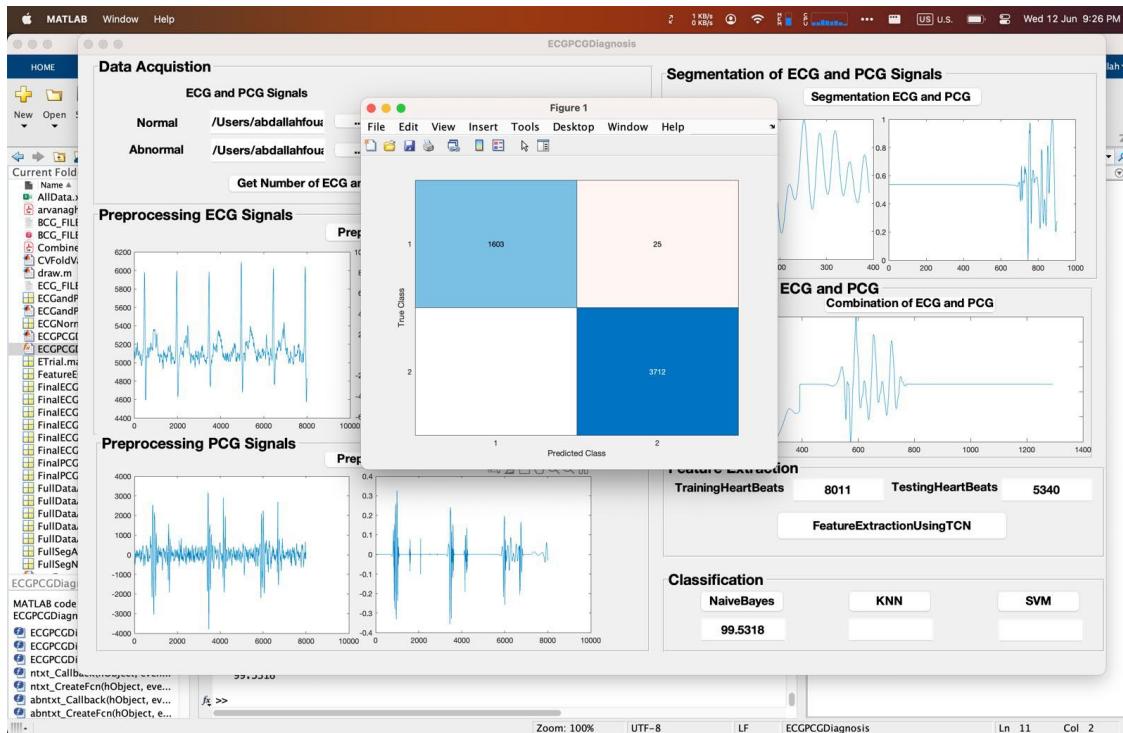
Combination of ECG and PCG

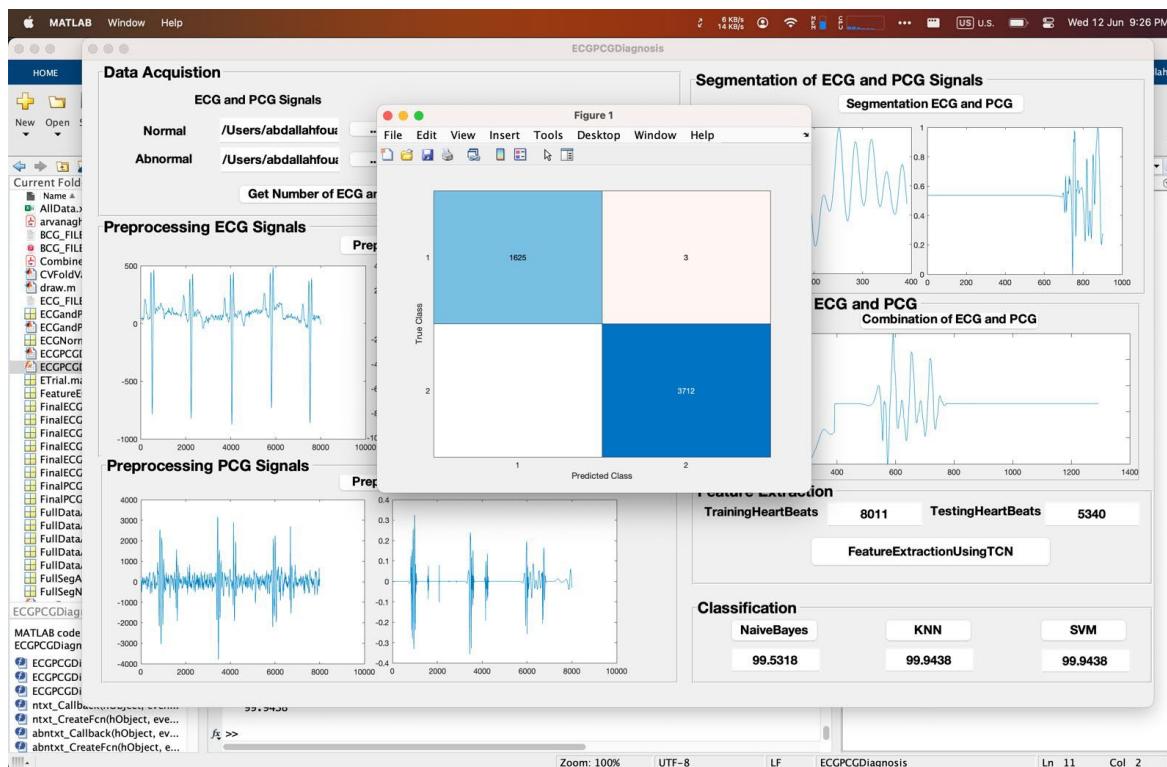
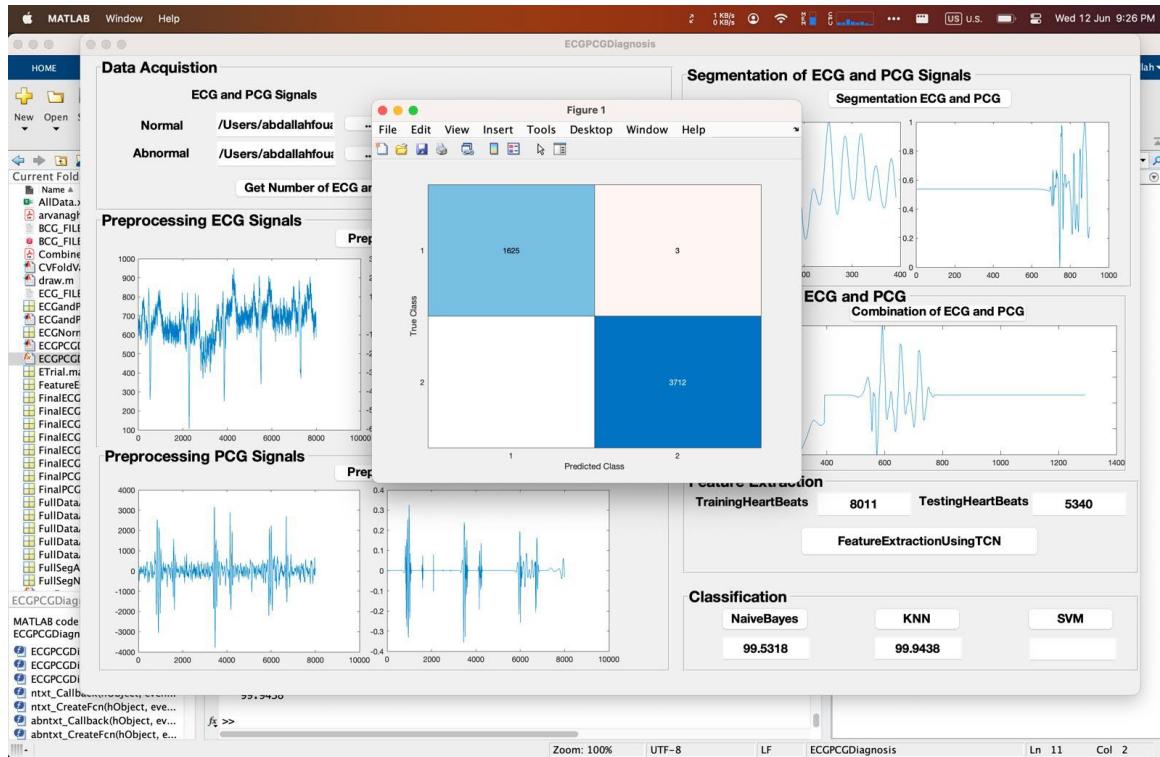


## 6.2.5 ECG-PCG Signals Feature Extraction

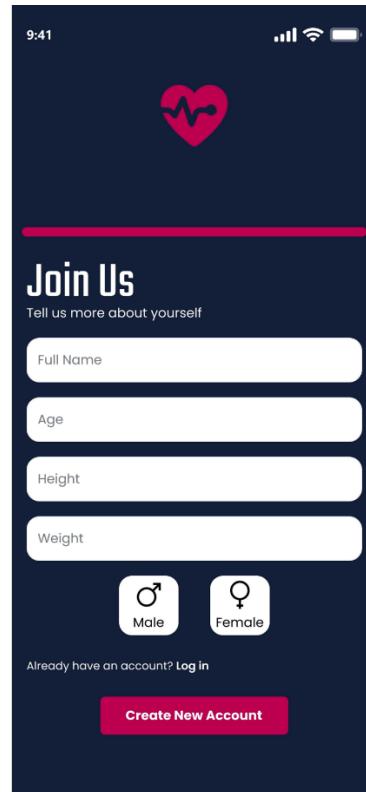
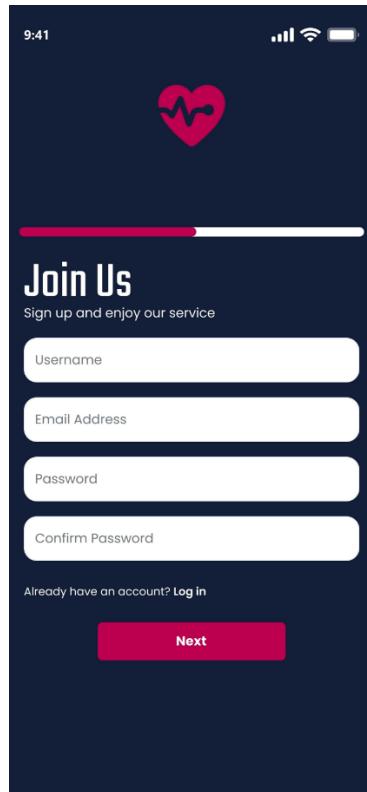
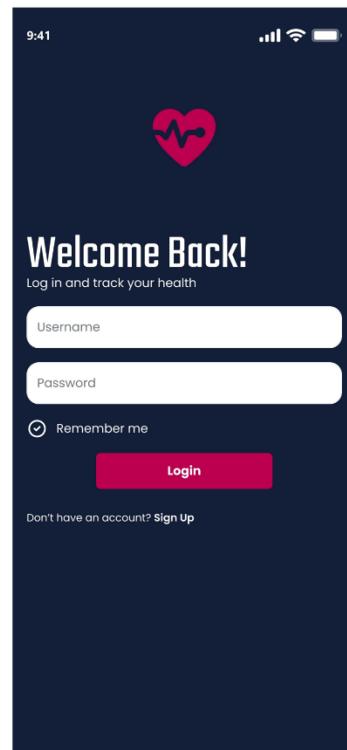
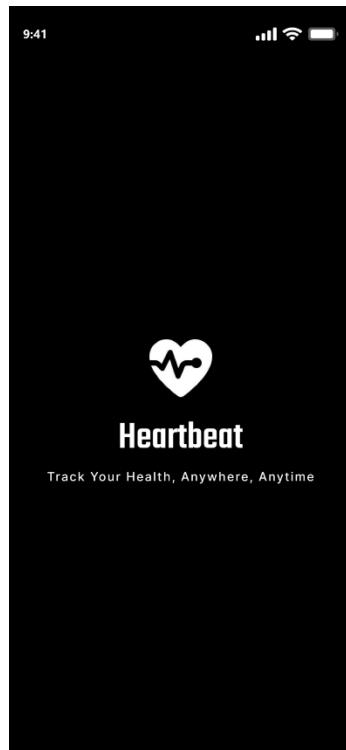


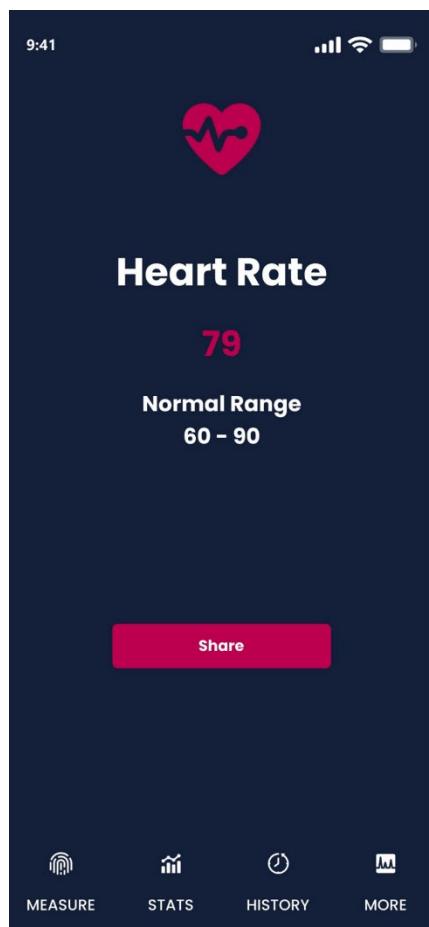
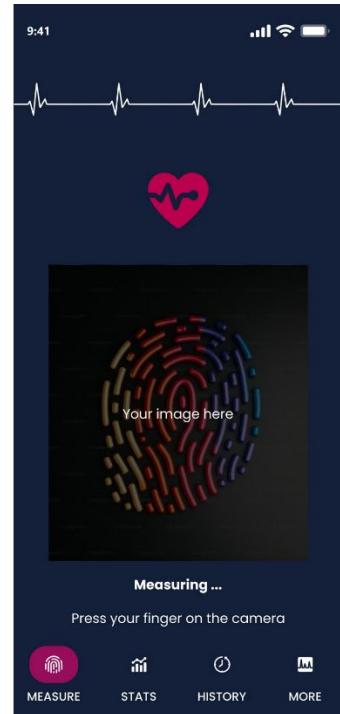
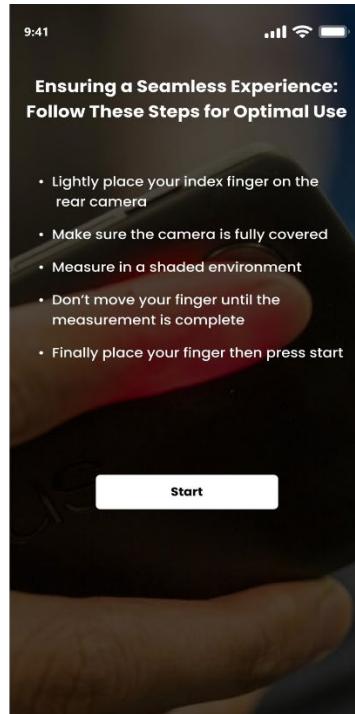
## 6.2.6 ECG-PCG Signals Classification

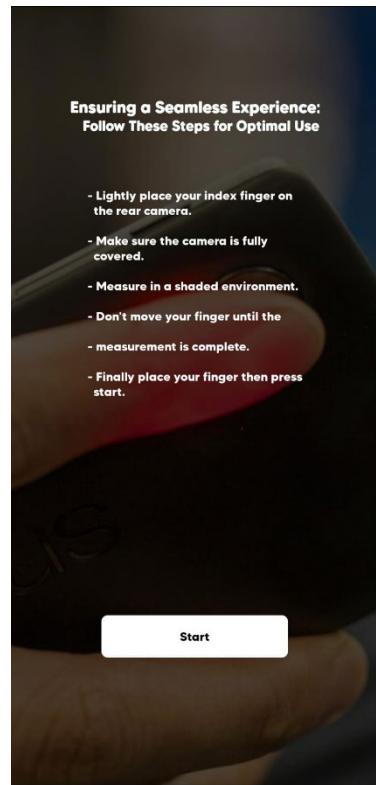
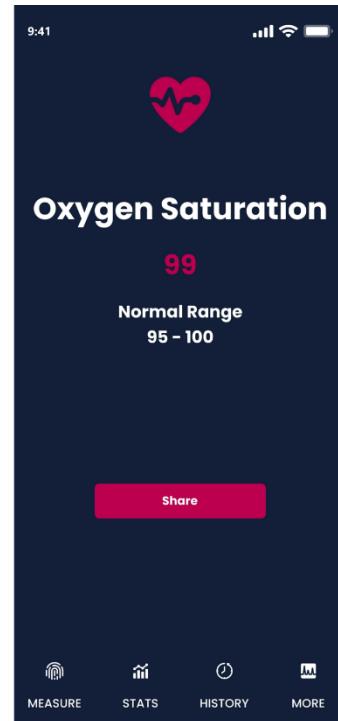
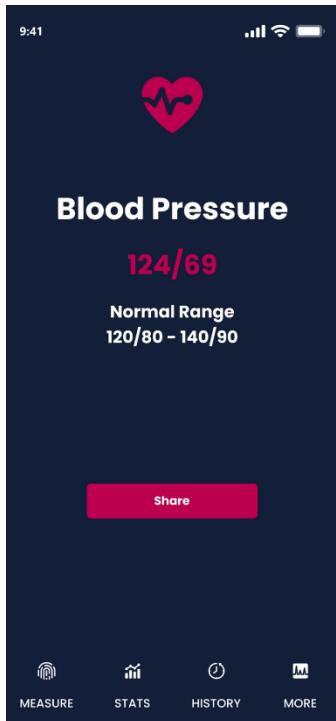


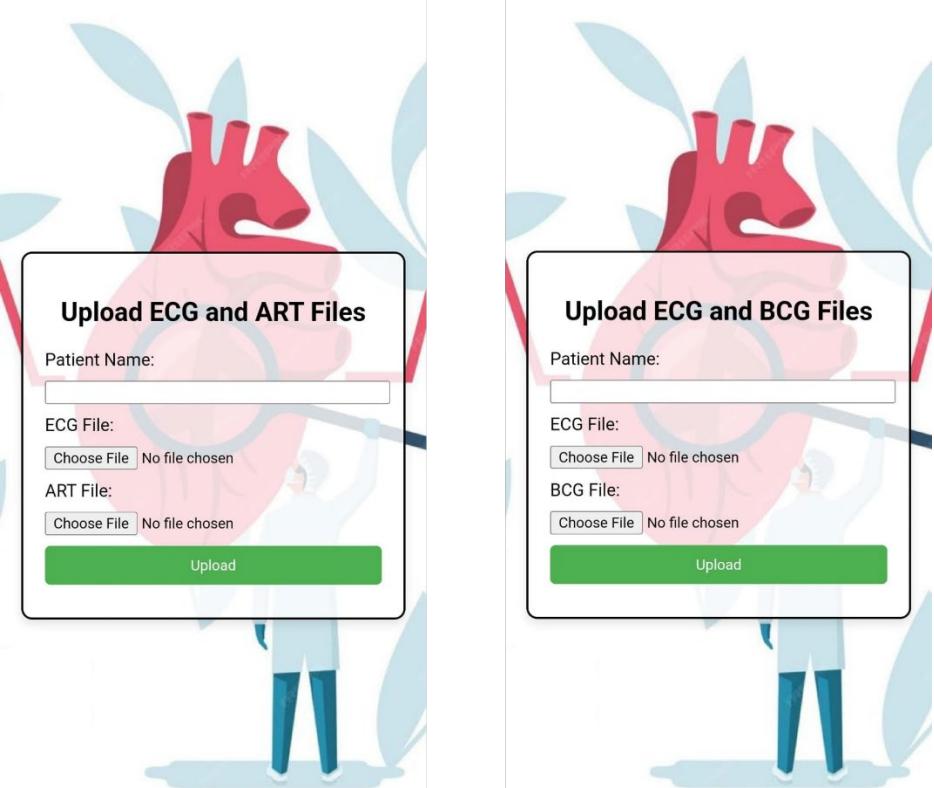


## 6.3 Mobile Application









### Upload ECG and ART Files

Patient Name:

ECG File:  No file chosen

ART File:  No file chosen

**Upload**

### Upload ECG and BCG Files

Patient Name:

ECG File:  No file chosen

BCG File:  No file chosen

**Upload**

## **References:**

1. Alberdi, A. Aztiria, A. Basarab, Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review, *J. Biomed. Inform.* 59 (2016) 49–75.
2. Illustration from Anatomy & Physiology, Connexions Web Site, 2013 <http://cnx.org/content/col11496/1.6/>.
3. M. Merone, P. Soda, M. Sansone, C. Sansone, ECG databases for biometric systems: a systematic review, *Expert Syst. Appl.* 67 (2017) 189–202.
4. R.G. Afkhami, G. Azarnia, M.A. Tinati, Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals, *Pattern Recogn. Lett.* 70 (2016) 45–51.
5. S. Dilmac, M. Korurek, ECG heart beat classification method based on modified ABC algorithm, *Appl. Soft Comput.* 36 (2015) 641–655.
6. S. Banerjee, M. Mitra, Application of Cross Wavelet Transform for ECG pattern analysis and classification, *IEEE Trans. Instrum. Meas.* 63 (2014) 326–333.
7. S. Shadmand, B. Mashoufi, A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization, *Biomed. Signal Process.* 25 (2016) 12–23.
8. J. Mateo, A.M. Torres, A. Aparicio, J.L. Santos, An efficient method for ECG beat classification and correction of ectopic beats, *Comput. Electr. Eng.* 53 (2016) 219–229.
9. E.J.D. Luz, W.R. Schwartz, G. Camara-Chavez, D. Menotti, ECG-based heartbeat classification for arrhythmia detection: a survey, *Comput. Methods Programs Biomed.* 127 (2016) 144–164.
10. K. Soualhi, A. Elloumi Oueslati, N. Ellouze, ECG image representation of normal sinus rythm, 1 St International Conference on Advanced Technologies for Signal and Image Processing (AT SIP 2014) (2014) 225–230.
11. S. Pal, M. Mitra, Increasing the accuracy of ECG based biometric analysis by data modelling, *Measurement* 45 (2012) 1927–1932.
12. J. Herzig, A. Bickel, A. Eitan, N. Intrator, Monitoring cardiac stress using features extracted from S1 heart sounds. *IEEE. Trans. Biomed. Eng.* 62(4) (2015) 1169-1178.
13. Allen J (2007) Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas* 28(3):R1
14. Moraes JL, Rocha MX, Vasconcelos GG, VasconcelosFilho JE, De Albuquerque VHC, Alexandria AR (2018) Advances in photoplethysmography signal analysis for biomedical applications. *Sensors* 18(6):1894.
15. Moço AV, Stuijk S, de Haan G (2018) New insights into the origin of remote PPG signals in visible light and infrared. *Sci Rep* 8(1):1–15
16. Hartmann V, Liu H, Chen F, Qiu Q, Hughes S, Zheng D (2019) Quantitative comparison of photoplethysmographic waveform characteristics: efect of measurement site. *Front Physiol* 10:198.
17. Chan ED, Chan MM, Chan MM (2013) Pulse oximetry: understanding its basic principles facilitates appreciation of its limitations. *Respir Med* 107(6):789–799.

18. Joyner MJ, Casey DP (2015) Regulation of increased blood flow (hyperemia) to muscles during exercise: a hierarchy of competing physiological needs. *Physiol Rev* 95:549–601.
19. Tamura T, Maeda Y, Sekine M, Yoshida M (2014) Wearable photoplethysmographic sensors—past and present. *Electronics* 3(2):282–302.
20. Sun Y, Thakor N (2015) Photoplethysmography revisited: from contact to noncontact, from point to imaging. *IEEE Trans Biomed Eng* 63(3):463–477.
21. Siaron KB, et al. Blood Pressure measurements are site dependent in a cohort of patients with neurological illness. *Sci Rep*. 2020;10(1):3382.
22. Bur A, et al. Factors influencing the accuracy of oscillometric blood pressure measurement in critically ill patients. *Crit Care Med*. 2003;31(3):793–9.
23. Zawadzki MJ, Small AK, Gerin W. Ambulatory blood pressure variability: a conceptual review. *Blood Press Monit*. 2017;22(2):53–8.
24. Viera AJ, Tuttle L, Zeng J. Dollars and discomfort: What will people be willing to give for better blood pressure assessment? *J Clin Hypertens (Greenwich)*. 2016;18(5):422–3.
25. Ruzicka M, Hiremath S. Accuracy-limiting factor of home blood pressure monitors? *Am J Hypertens*. 2017;30(7):661–4 Oxford University Press (OUP).
26. Shriram R, et al. Continuous cuffless blood pressure monitoring based on PTT. In: 2010 International Conference on Bioinformatics and Biomedical Technology: IEEE; 2010. p. 51–5. 13.
27. Luo N, et al. Flexible piezoresistive sensor patch enabling ultralow power cuffless blood pressure measurement. *Adv Funct Mater*. 2016;26(8):1178–87.
28. Urteaga J, et al. “Automated detection of pulse using continuous invasive arterial blood pressure in patients during cardiopulmonary resuscitation”, in 2021 Computing in Cardiology (CinC). Brno: Czech Republic; 2021.
29. Johnson AEW, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035
30. Lee J, Kim J, Shin M. Correlation Analysis between Electrocardiography (ECG) and Photoplethysmogram (PPG) Data for Driver's Drowsiness Detection Using Noise Replacement Method. *Procedia Computer Science*. 2017;116:421–6.
31. Mazaheri, V., & Khodadadi, H. (2020). Heart arrhythmia diagnosis based on the combination of morphological, frequency and nonlinear features of ECG signals and metaheuristic feature selection algorithm. *Expert Systems with Applications*, 161, 113697.
32. Khalil, M., & Adib, A. (2020). An end-to-end multi-level wavelet convolutional neural networks for heart diseases diagnosis. *Neurocomputing*, 417, 187-201.
33. Huang, J. S., Chen, B. Q., Zeng, N. Y., Cao, X. C., & Li, Y. (2020). Accurate classification of ECG arrhythmia using MOWPT enhanced fast compression deep learning networks. *Journal of Ambient Intelligence and Humanized Computing*, 1-18.
34. Xu, X., Jeong, S., & Li, J. (2020). Interpretation of electrocardiogram (ECG) rhythm by combined CNN and BiLSTM. *IEEE Access*, 8, 125380-125388.
35. Shaker, A. M., Tantawi, M., Shedeed, H. A., & Tolba, M. F. (2020). Generalization of convolutional neural networks for ECG classification using generative adversarial networks. *IEEE Access*, 8, 35592-35605.

36. Eltrass, A. S., Tayel, M. B., & Ammar, A. I. (2021). A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-Q non-stationary Gabor transform. *Biomedical signal processing and control*, 65, 102326.
37. Rath, A., Mishra, D., Panda, G., & Satapathy, S. C. (2021). Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomedical Signal Processing and Control*, 68, 102820.
38. Panganiban, E. B., Paglinawan, A. C., Chung, W. Y., & Paa, G. L. S. (2021). ECG diagnostic support system (EDSS): A deep learning neural network based classification system for detecting ECG abnormal rhythms from a low-powered wearable biosensors. *Sensing and Bio-Sensing Research*, 31, 100398.
39. Qin, L., Xie, Y., Liu, X., Yuan, X., & Wang, H. (2021). An end-to-end 12-leading electrocardiogram diagnosis system based on deformable convolutional neural network with good antinoise ability. *IEEE Transactions on Instrumentation and Measurement*, 70, 1-13.
40. Tyagi, A., & Mehra, R. (2021). Intellectual heartbeats classification model for diagnosis of heart disease from ECG signal using hybrid convolutional neural network with GOA. *SN Applied Sciences*, 3(2), 1-14.
41. Dixit, S., & Kala, R. (2021). Early detection of heart diseases using a low-cost compact ECG sensor. *Multimedia Tools and Applications*, 80(21), 32615-32637.
42. Monedero, I. (2022). A novel ECG diagnostic system for the detection of 13 different diseases. *Engineering Applications of Artificial Intelligence*, 107, 104536.
43. Anand, A., Kadian, T., Shetty, M. K., & Gupta, A. (2022). Explainable AI decision model for ECG data of cardiac disorders. *Biomedical Signal Processing and Control*, 75, 103584.
44. Rath, A., Mishra, D., Panda, G., Satapathy, S. C., & Xia, K. (2022). Improved heart disease detection from ECG signal using deep learning based ensemble model. *Sustainable Computing: Informatics and Systems*, 35, 100732.
45. Shin, S., Kang, M., Zhang, G., Jung, J., & Kim, Y. T. (2022). Lightweight Ensemble Network for Detecting Heart Disease Using ECG Signals. *Applied Sciences*, 12(7), 3291.
46. Chauhan S, Wang P, Sing Lim C, Anantharaman V (2008) A computer-aided MFCC-based HMM system for automatic auscultation. *Comput Biol Med* 38: 221–233.
47. Çomak E, Arslan A, Tuřkoglu I (2007) A decision support system based on support vector machines for diagnosis of the heart valve diseases. *Comput Biol Med* 37: 21–27.
48. Wu JB, Zhou S, Wu Z, Wu XM (2012) Research on the method of characteristic extraction and classification of phonocardiogram. In: *Systems and Informatics (ICSAI), 2012 International Conference on*. pp. 1732–1735.
49. Jiang Z, Choi S (2006) A cardiac sound characteristic waveform method for inhome heart disorder monitoring with electric stethoscope. *Expert Systems with Applications* 31: 286–298.
50. Avci E, Turkoglu I (2009) An intelligent diagnosis system based on principle component analysis and anfis for the heart valve diseases. *Expert Systems with Applications* 36: 2873–2878. 7. Avci E (2009) A new intelligent diagnosis system for the heart valve diseases by using genetic-svm classifier. *Expert Systems with Applications* 36: 10618–10626.
51. Abdel-Motaleb I, Akula R (2012) Artificial intelligence algorithm for heart disease diagnosis using phonocardiogram signals. In: *Electro/Information Technology (EIT), 2012 IEEE International Conference on*. pp. 1–6.
52. Boutana D, Benidir M, Barkat B (2011) Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis. *Signal Processing, IET* 5: 527–537.

53. Babaei S, Geranmayeh A (2009) Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals. *Computers in Biology and Medicine* 39: 8–15.
54. Jabbari S, Ghassemian H (2011) Modeling of heart systolic murmurs based on multivariate matching pursuit for diagnosis of valvular disorders. *Computers in Biology and Medicine* 41: 802–811. 12. Syed Z, Leeds D, Curtis D, Nesta F, Levine R, et al. (2007) A Framework for the Analysis of Acoustical Cardiac Signals. *Biomedical Engineering, IEEE Transactions on* 54: 651–662.
55. Gupta CN, Palaniappan R, Swaminathan S, Krishnan SM (2007) Neural network classification of homomorphic segmented heart sounds. *Applied Soft Computing* 7: 286–297.
56. Turkoglu I, Arslan A, Ilkay E (2002) An expert system for diagnosis of the heart valve diseases. *Expert Systems with Applications* 23: 229–236.
57. Simarjot Kaur Randhawa, Mandeep Singh, "Classification of Heart Sound Signals Using Multimodal Features," Second International Symposium on Computer Vision and the internet, Elsevier, Procedia Computer science, vol. 58, 2015, pp. 165-171.
58. Cota Navin Gupta, Ramaswamy Palaniappan, Sreeraman Rajan, Sundaram Swaminathan, S.M. Krishnan, "Segmentation and Classification of heart sound," International Conference: Canadian Conference on Electrical and Computer Engineering, June 2005, DOI: 10.1109/CCECE.2005.1557305, IEEE Xplore.
59. Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* 2015, 521, 7553.
60. Joyanta Kumar Roy & Tanmay Sinha Roy, "A Simple technique for heart sound detection and real-time analysis," Proceedings of ICST 2017 held at Macquarie University Sidney, Sensing Technology (ICST), 2017 Eleventh International Conference, 4-6 Dec. 2017, 10.1109/ICSensT.2017.8304502.
61. S. Barma, B.-W. Chen, W. Ji, F. Jiang, and J.-F. Wang, "Measurement of duration, the energy of instantaneous-frequencies, and splits of subcomponents of the second heart sound," *IEEE Transactions on Instrumentation Measurement*, vol. 64, no. 7, pp. 1958–1967, Jul. 2015.
62. M. Tavel, "Classification of systolic murmurs: Still in search of a consensus," *Am. Heart, J.*, vol. 134, no. 2, pp. 330–336, 1997.
63. Priya Ranjan Muduli, Atindra Kanti Mandal, and Anirban Mukherjee; An Anti-Noise-Folding Algorithm for the Recovery of Biomedical Signals from Noisy Measurements, *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2909-2916, 2017.
64. Tanmay Sinha Roy, Prof. Joyanta Kumar Roy, Dr. Nirupama Mandal, " A Study of Phonocardiography (PCG) Signal Analysis by K-Mean Clustering", In: Mandal, J.K., Roy, J.K. (eds) *Proceedings of International Conference on Computational Intelligence and Computing 2022, Algorithms for Intelligent Systems*. Springer, Singapore. [https://doi.org/10.1007/978-981-16-3368-3\\_16](https://doi.org/10.1007/978-981-16-3368-3_16).
65. Madhusudhan Mishra, Sanmitra Banerjee, Dennis Thomas, Sagnik Dutta, and Anirban Mukherjee; Detection of Third Heart Sound using Variational Mode Decomposition, *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1713-1721, 2018.

66. Madhusudhan Mishra, Hrishikesh Menon, Anirban Mukherjee, "Characterization of S1 and S2 Heart Sounds Using Stacked Autoencoder and Convolutional Neural Network", IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 9, pp. 3211-3220, 2019.
67. El-Segaier M, Lilja O, Lukkarinen S, Sörnmo L, Seppanen R, Pesonen E., Computer-based detection and analysis of heart sound and murmur, Ann Biomed Eng., 2005, Jul;33(7):937-42, <http://www.ncbi.nlm.nih.gov/pubmed/16060534>.
68. Nygaard, H., et al. Assessing the severity of aortic valve stenosis by spectral analysis of cardiac murmurs (spectral vibrocardiography).Part I: Technical aspects. J. Heart Valve Dis. 2(4):454-467, 1993.
69. Anju and Sanjay Kumar, "Detection of Cardiac Murmur." Anju et al., International Journal of Computer Science and Mobile Computing, Vol.3 Issue.7 July2014, pg. 81-87, ISSN 2320– 088X.
70. G. Buchanna, P. Premchand, A. Govardhan, Classification of Epileptic and Non-Epileptic Electroencephalogram (EEG) Signals Using Fractal Analysis and Support Vector Regression, 2022, Vol. 6, No.1, Doi: 10.28991/ESJ-2022-06-01-011.
71. En Zhou Ye, En Hui Ye, Maxime Bouthillier, Run Zhou Ye, Deep Image Translator V2: analysis of multimodal medical images using semantic segmentation maps generated through deep learning, bioRxiv 2021, doi: <https://doi.org/10.1101/2021.10.12.464160>.
72. Sukor, J.A.; Redmond, S.J.; Lovell, N.H. Signal quality measures for pulse oximetry through waveform morphology analysis. Physiol. Meas. 2011, 32, 369–384. [Google Scholar] [CrossRef]
73. Di, U.; Te, A.; De, J. Awareness of Heart Disease Prevention among Patients Attending a Specialist Clinic in Southern Nigeria. Int. J. Prev. Treat. 2012, 1, 40–43. [Google Scholar]
74. Tun, H.M. Photoplethysmography (PPG) Scheming System Based on Finite Impulse Response (FIR) Filter Design in Biomedical Applications. Int. J. Electr. Electron. Eng. Telecommun. 2021, 10, 272–282. [Google Scholar] [CrossRef]
75. Ram, M.R.; Madhav, K.V.; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. IEEE Trans. Instrum. Meas. 2011, 61, 1445–1457. [Google Scholar] [CrossRef]
76. Luke, A.; Shaji, S.; Menon, K.U. Motion artifact removal and feature extraction from PPG signals using efficient signal processing algorithms. In Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 19–22 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 624–630. [Google Scholar] [CrossRef]
77. Otsuka, T.; Kawada, T.; Katsumata, M.; Ibuki, C. Utility of second derivative of the finger photoplethysmogram for the estimation of the risk of coronary heart disease in the general population. Circ. J. 2006, 70, 304–310. [Google Scholar] [CrossRef] [Green Version]
78. Shintomi, A.; Izumi, S.; Yoshimoto, M.; Kawaguchi, H. Effectiveness of the heartbeat interval error and compensation method on heart rate variability analysis. Healthc. Technol. Lett. 2022, 9, 9–15. [Google Scholar] [CrossRef]

79. Moraes, J.L.; Rocha, M.X.; Vasconcelos, G.G.; Vasconcelos Filho, J.E.; De Albuquerque, V.H.; Alexandria, A.R. Advances in photoplethysmography signal analysis for biomedical applications. *Sensors* **2018**, *18*, 1894. [Google Scholar] [CrossRef] [Green Version]
80. Hwang, S.; Seo, J.; Jebelli, H.; Lee, S. Feasibility analysis of heart rate monitoring of construction workers using a photoplethysmography (PPG) sensor embedded in a wristband-type activity tracker. *Autom. Constr.* **2016**, *71*, 372–381. [Google Scholar] [CrossRef]
81. Moshawrab, M.; Adda, M.; Bouzouane, A.; Ibrahim, H.; Raad, A. Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review. *Sensors* **2023**, *23*, 828
82. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1. [Google Scholar] [CrossRef] [Green Version]
83. Kumar, P.S.; Harikumar, R. Performance Comparison of EM, MEM, CTM, PCA, ICA, entropy and MI for photoplethysmography signals. *Biomed. Pharmacol. J.* **2015**, *8*, 413–418. [Google Scholar] [CrossRef] [Green Version]
84. Almarshad, M.A.; Islam, S.; Al-Ahmadi, S.; BaHammam, A.S. Diagnostic Features and Potential Applications of PPG Signal in Healthcare: A Systematic Review. *Healthcare* **2022**, *10*, 547. [Google Scholar] [CrossRef]
85. Yousefi, M.R.; Khezri, M.; Bagheri, R.; Jafari, R. Automatic detection of premature ventricular contraction based on photoplethysmography using chaotic features and high order statistics. In Proceedings of the 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, Italy, 11–13 June 2018; pp. 1–5. [Google Scholar] [CrossRef]
86. Ukil, A.; Bandyopadhyay, S.; Puri, C.; Pal, A.; Mandana, K. Cardiac condition monitoring through photoplethysmogram signal denoising using wearables: Can we detect coronary artery disease with higher performance efficacy? In Proceedings of the Computing in Cardiology Conference (CinC), Vancouver, BC, Canada, 11–14 September 2016; pp. 281–284. [Google Scholar] [CrossRef]
87. Almanifi, O.R.A.; Khairuddin, I.M.; Razman, M.A.M.; Musa, R.M.; Majeed, A.P.A. Human activity recognition based on wrist PPG via the ensemble method. *ICT Express* **2022**, *8*, 513–517. [Google Scholar] [CrossRef]
88. Paradkar, N.; Chowdhury, S.R. Coronary artery disease detection using photoplethysmography. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 100–103. [Google Scholar] [CrossRef]
89. Neha; Kanawade, R.; Tewary, S.; Sardana, H.K. Neha; Kanawade, R.; Tewary, S.; Sardana, H.K. Photoplethysmography based arrhythmia detection and classification. In Proceedings of the 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 7–8 March 2019; pp. 944–948. [Google Scholar] [CrossRef]
90. Prabhakar, S.K.; Rajaguru, H.; Lee, S.W. Metaheuristic-based dimensionality reduction and classification analysis of PPG signals for interpreting cardiovascular disease. *IEEE Access* **2019**, *7*, 165181–165206. [Google Scholar] [CrossRef]

91. Sadad, T.; Bukhari, S.A.C.; Munir, A.; Ghani, A.; El-Sherbeeny, A.M.; Rauf, H.T. Detection of Cardiovascular Disease Based on PPG Signals Using Machine Learning with Cloud Computing. *Comput. Intell. Neurosci.* 2022, 2022, 1672677. [Google Scholar] [CrossRef]
92. Chakir, F., Jilbab, A., Nacir, C., & Hammouch, A. (2020). Recognition of cardiac abnormalities from synchronized ECG and PCG signals. *Physical and Engineering Sciences in Medicine*, 43, 673-677.
93. Arvanaghi, R., Daneshvar, S., Seyedarabi, H., & Goshvarpour, A. (2017). Fusion of ECG and ABP signals based on wavelet transform for cardiac arrhythmias classification. *Computer methods and programs in biomedicine*, 151, 71-78.
94. Lea, C., Vidal, R., Reiter, A., & Hager, G. D. (2016, October). Temporal convolutional networks: A unified approach to action segmentation. In European Conference on Computer Vision (pp. 47-54). Springer, Cham.
95. Lara-Benítez, P., Carranza-García, M., Luna-Romera, J. M., & Riquelme, J. C. (2020). Temporal convolutional networks applied to energy-related time series forecasting. *Applied Sciences*, 10(7), 2322. <https://doi.org/10.3390/app10072322>