Software Engineering Department

ORT Braude College

Capstone Project Phase A – 61998

Camera measurement of physiological vital signs monitoring system

23-2-D-7

Shimon Rubin | shimonr111@gmail.com

Lior Guzovsky | liorguzovsky@gmail.com

Supervisors:

Dr. Dan Lemberg | lembergdan@braude.ac.il

Mrs. Elena Kramer | elenak@braude.ac.il

Table Of Contents

Abstract

Measurement of heart rate is necessary to assess the status of the patient's health. Performing non-contact heart rate measurement can provide many benefits in various fields such as medicine, hospitals, and sports. In addition, it can be mainly significant in the field of remote medicine. For it to be a worthy replacement for the existing technologies that include contact equipment for heart rate measurement, there needs to be great accuracy for the heart rate measurement results, which is something that is not easy to achieve given the fact that it is a measure that changes in a very short period. In our project, we suggest a camera-based heart rate monitoring system that detects objects using image processing methods such as the Haar cascade classifier. We use the Fast Fourier Transform (FFT) for analyzing frequency domain heart rate signals. Our non-contact method using a camera provides a better option for heart rate monitoring than existing techniques that assess changes in pixel brightness on the forehead.

Keywords:

Image processing, Real-time, Heart rate, Video, Fast Fourier Transform (FFT), Haar Cascade classifier.

1 Introduction

In today's fast-paced world, the demand for quality healthcare remains paramount, especially in nursing homes where the well-being of our beloved family members becomes a top priority.

From a personal story of a family member who was in a nursing home, we have noticed that the therapist-patient ratio during the night shifts is relatively small and therefore dangerous for the patients. This experience highlighted the need for improved patient-therapist ratios, particularly during the night shifts, when resources are limited, and monitoring becomes a challenging task.
Inspired by this personal story, the idea for our system to measure heart rate from video emerged and was approved as our final project.

Traditional methods rely on the use of pulse oximeters or other contact-based devices to obtain vital signs. However, such devices often require direct physical contact with the patient, which can be intrusive and uncomfortable, especially for elderly individuals in nursing homes. In addition, the traditional methods demand more therapists which is a problem we won’t be able to solve.

In recent years, the field of non-contact heart rate monitoring through video has gained significant attention, and particularly in the post-COVID era. The use of such systems (contactless monitoring systems) was introduced in airports where automated thermal screening systems have been deployed to monitor body temperature without physical contact.

As a solution, we propose to develop a system that will monitor the heart rate of the patients and present them to the medical team, to enable life-saving treatment if necessary. The system will process video coming from live streaming and derive heart rate data. Then, it will locate the region of interest (ROI), more especially the face and forehead, which offer indicators of the underlying cardiovascular activity, by using a Haar cascade classifier which is a very fast machine learning algorithm. From the ROI we have found, the green channel information is then extracted from the video data, and the Fast Fourier Transform (FFT) is used to examine the frequency spectrum of those frames. This enables us to precisely extract the crucial information and identify conspicuous peaks that correspond to the heart rate.

In the subsequent sections of this article, we will explore deeper into the complexities of contactless heart rate monitoring from video in nursing homes.
we will explore the existing related work in the field. Then present the background information necessary for a thorough understanding of the challenges and opportunities associated with contactless monitoring. Building upon the knowledge gained from related work and background analysis, we will describe the expected achievements of our system. The subsequent section will detail the research and engineering process we have made to develop the system including the architectural design and data flow to provide an overview of the system's structure and functionality. Finally, we will discuss the evaluation and verification plan.

2 Background and Related work

The field of monitoring heart rate from video data has gained considerable attention in recent years due to its potential for non-contact vital sign measurement. In the medical domain, accurate and timely monitoring of heart rate is crucial for diagnosing and managing various cardiovascular conditions.

This section is dedicated for the description of the medical background, PPG and RPPG theory, mathematical background, RGB model, ICA algorithm, Bandpass filter, GrabCut algorithm and related work we found relevant during our research to solve the problem we have described.

2.1 Medical background – cardiovascular system

The cardiovascular system is responsible for transporting blood throughout the body and is composed of the heart, blood vessels, and blood.

Hemoglobin is a vital component of the cardiovascular system. It is an iron-containing protein found in red blood cells that responsible for the transport of oxygen from the lungs to the body's tissues. When hemoglobin binds with oxygen in the lungs, it forms oxyhemoglobin, which provides a bright red color to oxygenated blood. This oxygen-rich blood is then carried by arteries to various organs and tissues, where oxygen is released for cellular respiration.

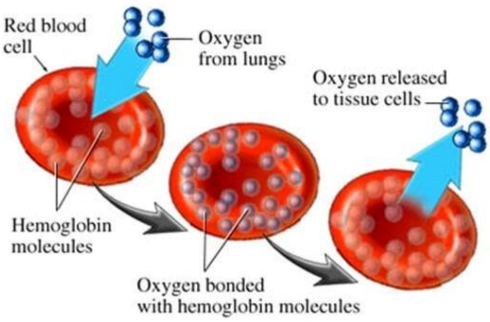


Figure 1

Blood circulation involves the heart pumping oxygenated blood from the lungs into arteries, which carry it throughout the body. Deoxygenated blood returns to the heart through veins and is pumped back to the lungs for re-oxygenation.

When The heart beats by contracting and relaxing its muscles in a rhythmic pattern, creating pressure that forces blood out of the heart and into arteries during systole and allows blood to flow back into the heart during diastole.

Blood flows through arteries, arterioles, capillaries, venules, and veins in a continuous loop throughout the body.

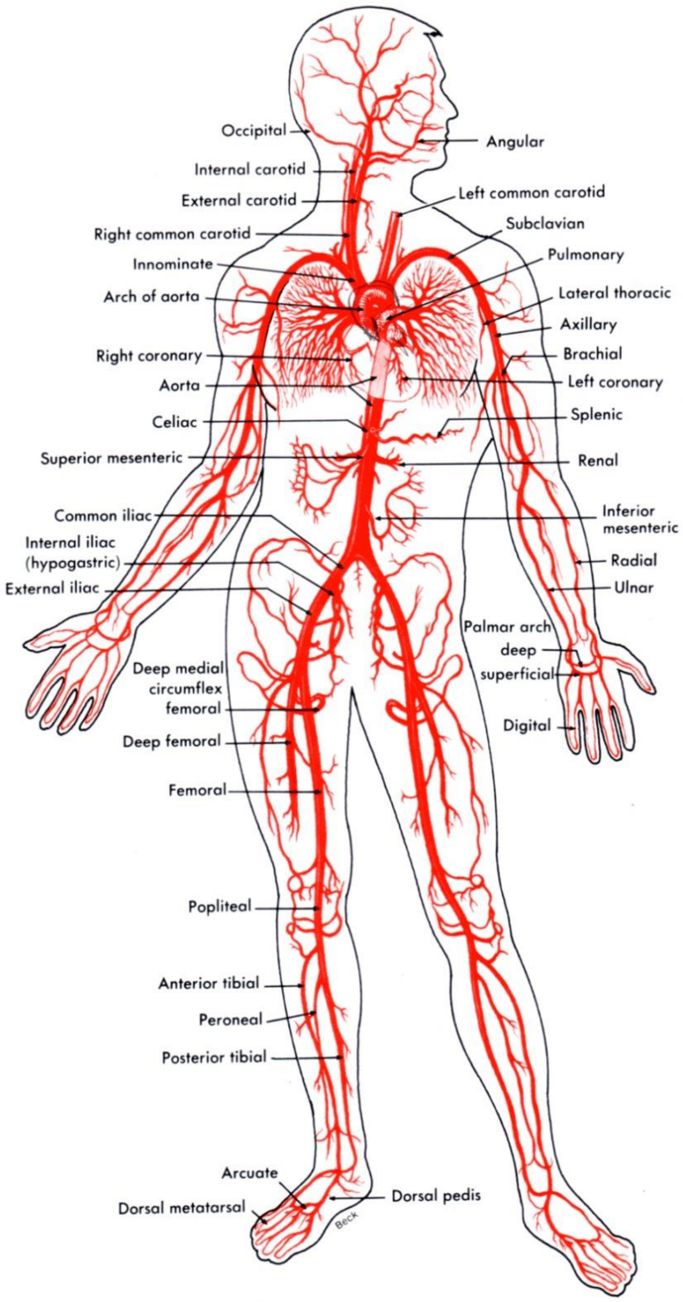
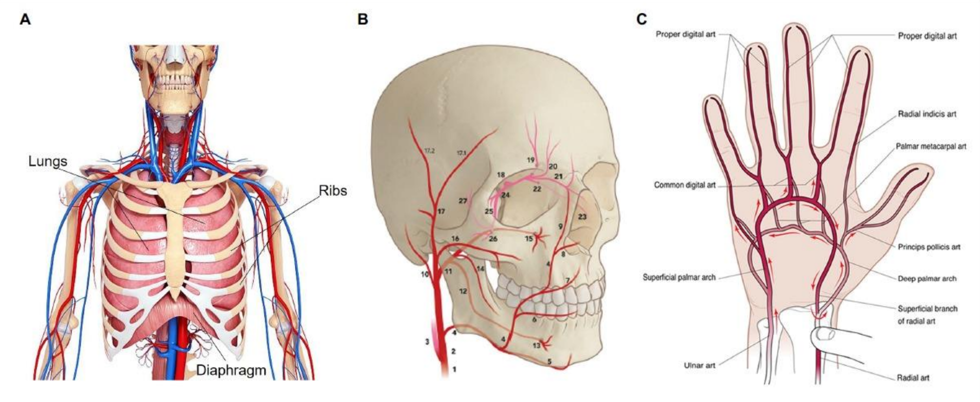


Figure 2

The major arteries that supply blood to the face and head are located on either side of the neck: carotid arteries (internal and external) and vertebral arteries. In particular, the forehead receives blood primarily from branches of both internal carotid arteries, including the supraorbital artery [Fig 1 – B].



Anatomy of the respiratory system (A), face (B; adapted from von Arx et al., 2018), and hand (C) vascularization (from Habib et al., 2012). In (B), numbers indicate the main arteries that carry the blood to the forehead and cheeks, specifically: common carotid artery (1), external carotid artery (2), internal carotid artery (3), facial artery (4), transverse facial artery (16), superficial temporal artery (17) and frontal branch (17.1), ophthalmic artery (18).

Figure 3

Normal heart rate parameters vary depending on age, sex, and physical activity level but typically range between 60-100 beats per minute at rest which are corresponding to a frequency range of 1-1.67 Hz.

2.2 PPG and RPPG

2.2.1 PPG

Photoplethysmography, sometimes known as PPG, is a non-invasive optical technique that assesses changes in blood volume in peripheral tissues.
It involves shining a light into the skin and detecting changes in light absorption created by the continuous nature of blood flow.

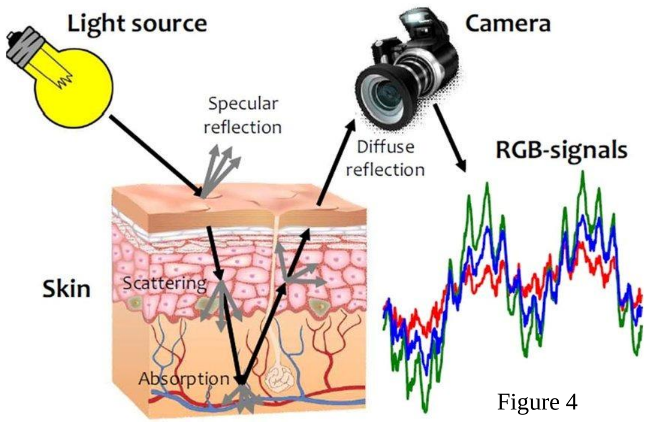


Figure 4

In PPG, light passes onto the skin, generally at the fingertip or earlobe from a light source. The blood arteries below the tissues partially absorb some of the light. The amount of blood in the vessels affects how much light is absorbed. Blood flow during heartbeats changes the blood vessels volume on a regular basis, changing how much light is absorbed in the process.

The PPG waveform, an electrical signal, is created from these intensity variations. The PPG waveform symbolizes a pulsatile blood flow element and can reveal information regarding the heart rate.

PPG is used in many different types of medical research and monitoring. To determine the heart rate and evaluate its variability, the waveform is typically studied for heart rate monitoring. Pulse oximeters[1] also use PPG to measure peripheral perfusion and measure blood oxygen saturation levels.

Improvements in technology have ended up in the development of wearable devices with PPG sensors, enabling non-invasive continuous monitoring of vital signs.

2.2.2 RPPG

RPPG, or Remote Photoplethysmography, is a new method that builds on the fundamentals of photoplethysmography (PPG) to monitor physiological parameters without having to encounter the subject directly. RPPG uses video-based photography to capture small color changes on the subject's face or other exposed areas, unlike traditional PPG, which necessitates direct physical contact between the light source and the skin.

In RPPG, a camera collects a video of the subject while focusing on specific areas like the forehead or cheeks. The recorded video offers details about

[1] The pulse oximeter, or Pulse Ox, is an electronic device that measures the saturation of oxygen carried in your red blood cells.

changes in blood volume beneath the skin's surface that affect skin color. The continuous blood flow caused by the cardiac cycle is connected to these variations in color.

The continuous signals can be extracted from the video frames using RPPG algorithms, which can then produce a remote photoplethysmogram. The remote photoplethysmogram shows distinctions in blood volume over time similarly to a traditional PPG. The subject's heart rate, heart rate variability, blood flow characteristics, possibly other cardiovascular parameters, can all be learned from this information.

Due to several variables that may affect the precision and dependability of the remote measurements, RPPG does, however, also present difficulties. These elements include of subject mobility, alterations in skin tone and changed ambient lighting conditions. As a result, current research concentrates on creating reliable algorithms and strategies to deal with these constraints and raise the precision of RPPG-based measurements.

2.3 FFT - Fast Fourier Transform

A fast Fourier transform (FFT) is a mathematical algorithm used to transform a time-domain signal into its frequency-domain representation. A fast Fourier transform can be used in various types of signal processing. It may be useful in reading things like sound waves, or for any image-processing technologies. A fast Fourier transform can be used to solve various types of equations or show various types of frequency activity in useful ways.

The Fast Fourier Transform is an optimized implementation of the Fourier transform algorithm. It significantly reduces the computational complexity of the transform, making it feasible to compute the frequency-domain representation of signals in real-time or near real-time applications. For heart rate detection, the frequency domain of a PPG signal can be inspected for peaks in the region corresponding to a reasonable heart rate.

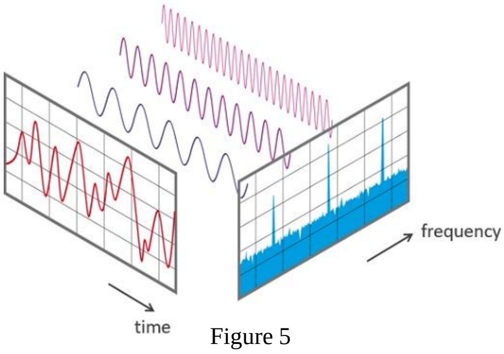


Figure 5

2.4 RGB Model

A typical color model used in digital imaging and displays is RGB (Red-Green-Blue). Red, green, and blue light are combined in various intensities to represent colors in the RGB model. Each color component can have values between 0 and 255, which represent brightness or intensity. A wide spectrum of colors can be produced by adjusting the red, green, and blue component intensities.

Due to the optical properties of human tissues and the absorption spectrum of hemoglobin, photoplethysmography (PPG) chooses to detect heart rate using the green channel. PPG makes use of the green channel to find a balance between reducing interference from other tissue elements and raising the sensitivity to changes in blood volume. Both oxygenated and deoxygenated hemoglobin partially absorbs green light, making it possible to record a distinct rapid signal. When using the green channel, we can notice that more hemoglobin-rich blood flows into the veins close to the skin's surface when the heart beats and the blood volume rises. As a result, there is an increase in the absorption of green light and a corresponding decrease in pixel intensity. Heart rate can be determined by tracking these changes over time.

Although PPG can also be performed using red and near-infrared channels, they are more sensitive to interference from things like ambient light and tissue variance. Green light is commonly used for PPG measurements of heart rate and other cardiovascular parameters because it provides a far better compromise between signal quality and interference.

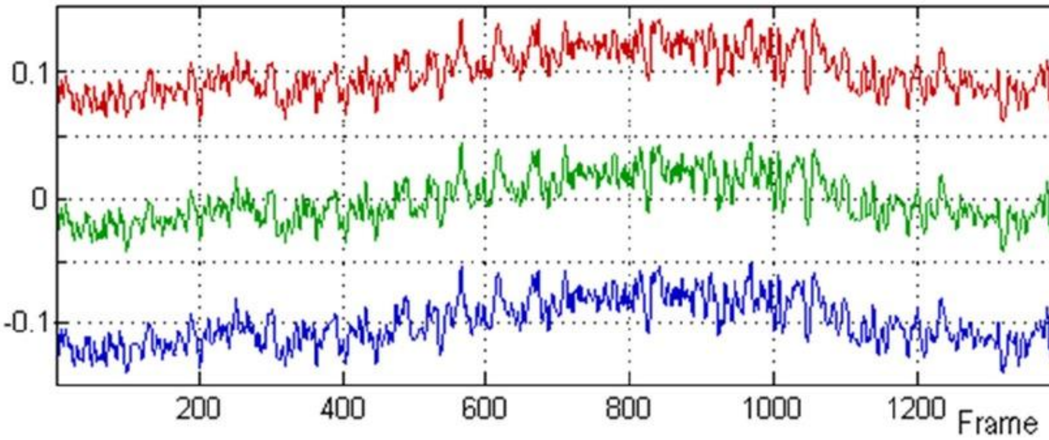


Figure 6

2.5 ICA Algorithm

Independent Component Analysis is referred to as ICA. It is a method for computing the independent parts of a multidimensional signal. ICA is used to separate the independent source signals from the observed mixed color signals. The observed color changes in the three channels (red, green, and blue) are assumed to be the result of three source signals. According to ICA, these source signals are combined linearly to produce the observed mixed signals. These distinct source signals can be extracted using ICA to produce more precise signals that can be used to determine heart rate from video data.

2.6 Bandpass filter

A bandpass filter is a sort of electronic filter that reduces (decreases) signals outside of a specific frequency range while allowing signals within that range to pass through. A bandpass filter can be used to isolate frequency peaks in the power spectrum in the range of 0.75 to 4 Hz, which correspond to physiological heart rate ranges of 45 to 240 bpm. The accuracy of heart rate measurements from video data can be increased by applying a bandpass filter to remove noise and other undesired frequencies from the signal.

2.7 GrabCut algorithm

GrabCut is an image segmentation method created by Rother et al. that is used to distinguish between an image's foreground and background pixels. By creating a graph model to describe the image and figuring out the smallest cut for the graph to generate two sets of nodes, which represent the foreground and background pixels of the image, it is possible to achieve the algorithm's goal of iteratively decreasing an energy cost function. GrabCut is used during the stage of ROI selection in the context of monitoring heart rate from video of a person's face. GrabCut is used to change the image after a region of interest (ROI) on the person's forehead has been chosen where blood vessels are visible and there is little motion artifact.

Within a few iterations, GrabCut is used to change the bounding box around the face of the subject and remove background pixels. This helps remove the background pixel variations that was present when using a simple, rigid bounding box ROI, which may improve the accuracy of heart rate measurements from video data.

2.8 Haar Cascade classifier

A machine learning-based object detection system called the Haar cascade classifier is used to identify objects in images or video streams. Viola and Jones first mentioned it in their study from 2001.

The research article "Rapid Object Detection using a Boosted Cascade of Simple Features" by Viola and Jones makes use of edge or line detection characteristics. The algorithm is trained on both many positive photos with faces in them and many negative images without any faces in them.

When using Haar cascade classifier the characteristics of the image make it simple to identify the image's borders or lines as well as regions where there is a sharp change in the pixel brightness. Below you can see the filters used by the Haar cascade classifier:

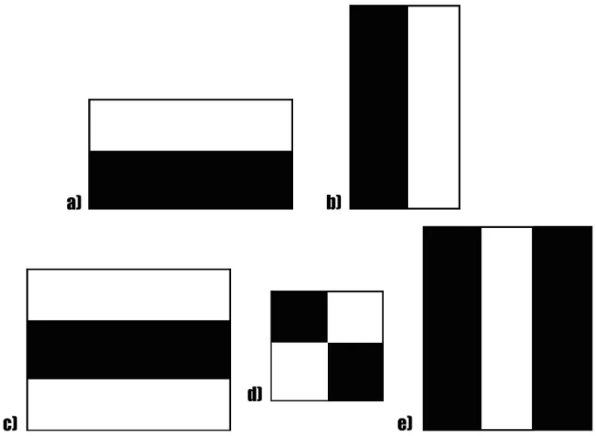


Figure 7

In the above filters, the goal is to calculate the total of all the image pixels located in the Haar feature's darker and lighter areas, respectively. then identify what makes them different. The Haar value will be closer to 1 if there is an edge in the image separating light pixels on the left from dark pixels on the right. In other words, if the Haar value is closer to 1, we declare that an edge has been found. Since the Haar number is distant from 1, there is no edge in the previous scenario.

For example, the left picture consists of arbitrary pixel values and the right one is the Haar filter. After calculating the result, we get the outcome on the right.

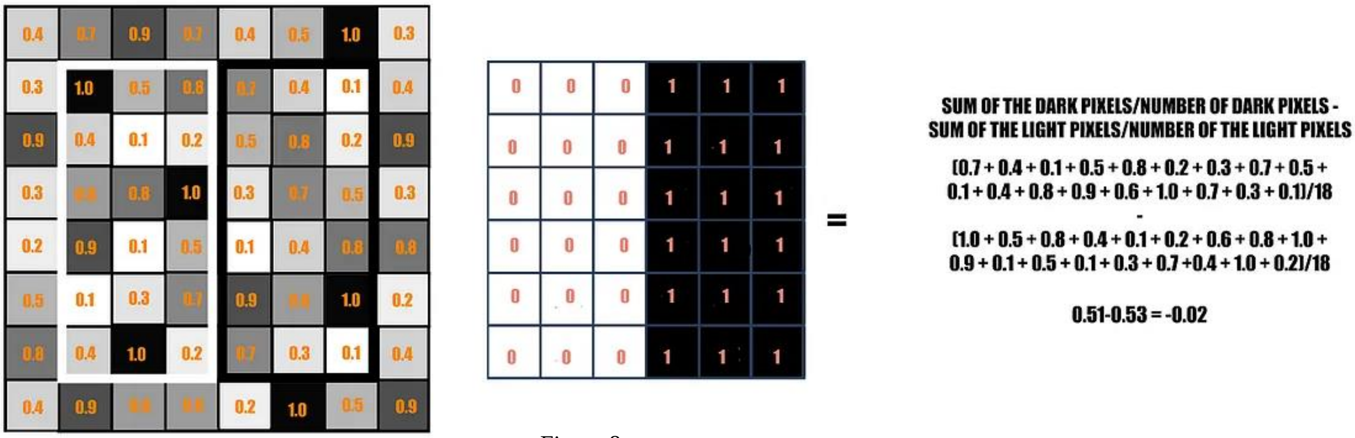


Figure 8

This is just an instance of a specific Haar characteristic that divides a vertical edge. There are currently more Haar characteristics as well, which will recognize any additional image structures as well as edges in other directions. The Haar feature must go across the entire image to find an edge anywhere in it.

Traversing the Haar features on an image would need numerous mathematical calculations. Even with a high-performance system, this would be an intensive operation.

The authors developed a different solution known as **The Integral Image** to address this problem and carry out the same function. Each pixel in an integral image is calculated from the original image so that it equals the total of all the pixels to its left and to its upper right.

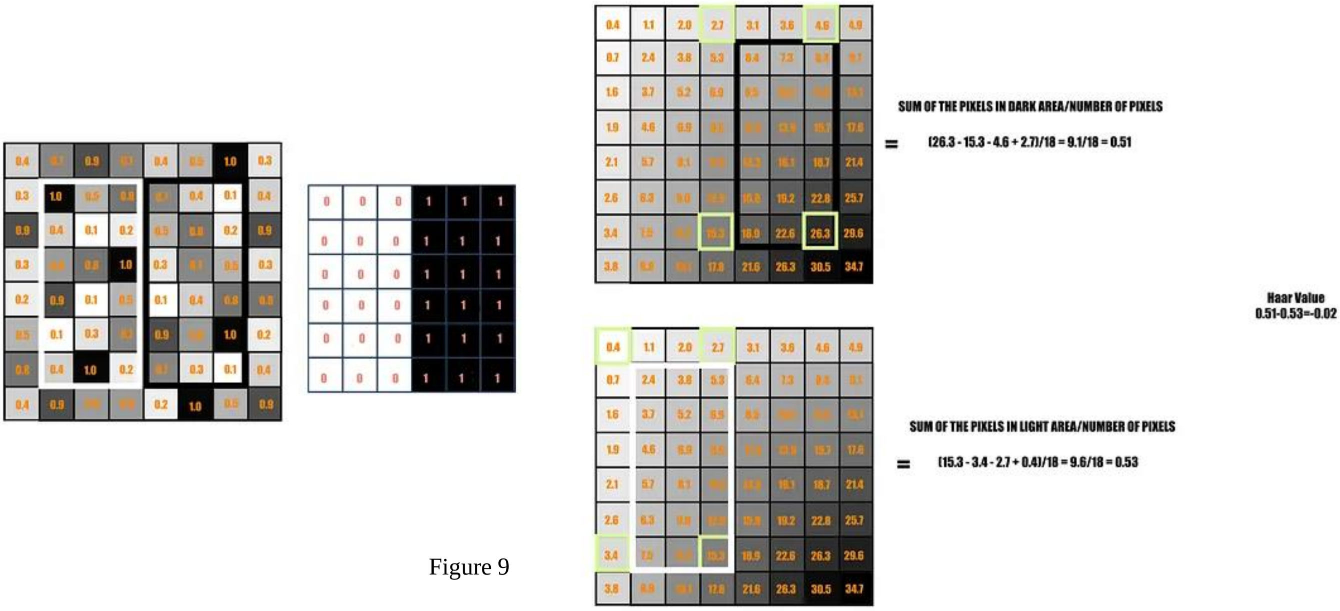


Figure 9

In contrast to the 18 additions previously, the Integral Image simply requires 4 constant value adds each time for any feature size. As a result, the number of adds does not depend on the number of pixels enclosed anymore, continually decreasing the time complexity of each addition.

2.9 AdaBoost

The focus of the previous chapter was on the features and representation of the image utilized in the Haar Cascade research. Let's look at some of the implementation details right now.

There is a collection of features that would accurately depict specific facial features, such as the lips, the bridge connecting the two eyes, or the brows. Most of these features won't fit the facial features well or won't be relevant to them because they are too random to be of any use. In this situation, they need a feature selection technique to pick a small subset of features from a large set that would not only pick features that performed better than the others but would also remove the irrelevant ones. They used a technique known as AdaBoost. Their final collection of features was reduced using this method from 180,000 to a total of 6000 features.

We created *weak learners* by applying each of these 180,000 features to the photos separately. Some of them, however, produced low error rates because they distinguished between Positive and Negative images more effectively than the others. These weak learners are created so that they would incorrectly classify a minimum number of photos.

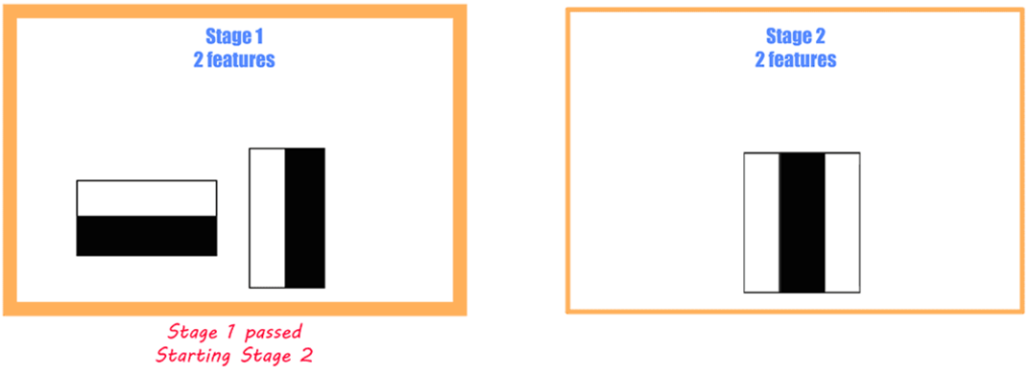
2.10 Attentional Cascade

The next step is the cascading section is to determine if a facial feature is present or absent. The subset of all 6000 features will again be run on the training photos. A standard window size of 24x24 was chosen for the feature detection to be running in at this point.

Another method called The Attentional Cascade was suggested to make this easier to understand. Not all features must be used on every window, according to the rationale behind this. We can infer that the face features are absent if a feature fails on a specific window. Therefore, we can continue to the following windows where facial traits might be present.

- The images have features in stages added to them. Beginning stages include simpler features compared to later stages, which have features that are sophisticated enough to reveal the finer details on the face. The window itself will be removed from the remaining process and replaced with the next window if the initial step is unable to detect anything on it. As the irrelevant windows won't be processed in the bulk of the steps, a significant amount of processing time will be saved.

- Only when the first stage's features are found in the image would the second stage of processing begin. This is how the process proceeds: if one stage succeeds, the window is moved on to the next; if it fails, the window is discarded.



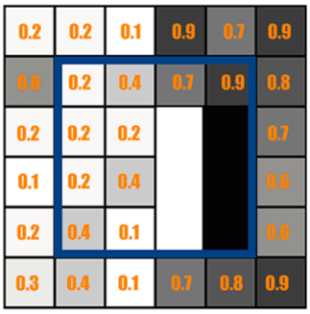


Figure 10

2.11 Related work

There is a growing amount of related work and research because of the significant interest in the subject of heart rate measurement using video data. Numerous studies have investigated different methods and strategies for precisely obtaining heart rate data from video recordings, allowing for non-contact. These studies addressed a wide range of topics, including machine learning, signal processing, computer vision, and healthcare.

For estimating heart rate from video data, researchers developed new technique, proposed innovative algorithms, and analyzed the efficacy of different approaches. In this section, we provide an overview of the body of research on heart rate measurement from video, highlighting significant contributions and methodology.

In the paper “Measuring Heart Rate from Video” [1] by Isabel Bush, Stanford Computer Science faculty, we have encountered the first method for extracting heart rate from video.

There are multiple steps involved in the process of extracting heart rate from video data using non-contact techniques based on photoplethysmography (PPG). The person's face is first videotaped using a standard color camera to give a clean view of the forehead and cheeks. The input for the associated analysis is this video.

The person's forehead is the next area to be chosen as a region of interest (ROI), as blood vessels are visible and motion artifacts are at a minimum there. This ROI serves as the basis for further research. The ROI is then tracked throughout time, considering any movements or position changes, to ensure accurate measurement. For this, tracking algorithms like template matching or optical flow are used. GrabCut algorithm is used to separate the foreground (in this case, the person's face) from the background in an image or video frame. It is used in the context of measuring heart rate from video of a person's face because it can help to exclude background pixels and hair that may interfere with accurate heart rate measurements.

Signals are taken from each video frame within the chosen ROI when the ROI has been successfully monitored using ICA method (Independent Component Analysis) to extract independent source signals from observed mixed color signals. For heart rate estimation, these signals show changes in pixel brightness over time and offer useful data. These signals do, however, frequently include noise and undesirable deviations. Therefore, to improve the quality of the recovered signals and eliminate undesired noise, signal processing techniques including bandpass filtering and Fourier analysis are used.

The heart rate is determined after signal processing by examining the frequency components of the processed signals. Usually, techniques like peak detection or spectrum evaluation are used for this. The heart rate can be precisely calculated in the frequency domain by locating peaks or dominating frequencies.

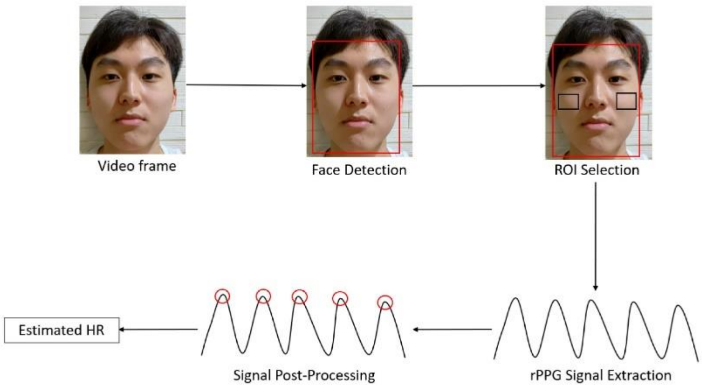
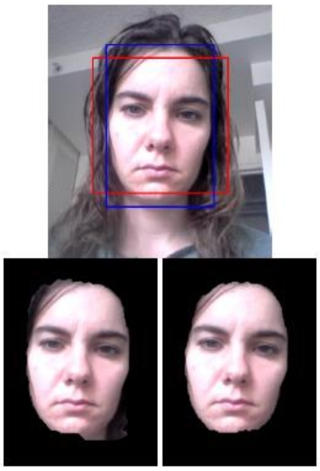


Figure 11

In the paper “Beat-to-Beat Cardiac Pulse Rate Measurement from Video” [2] by Brian L. Hill, University of California, Los Angeles, Xin Liu, University of Washington and Daniel McDuff, Microsoft, a new method for extracting heart rate from video was introduced. The authors proposed a neural architecture and model for camera-based vital sign measurement that effectively captures both spatial and temporal information, enabling the recovery of cardio-pulmonary signals from video recordings.

Several steps are involved in the procedure, all of which help to accurately extract the heart rate. The process begins with a camera taking a clear video of the subject's face in ideal lighting. The subject's face is then precisely located inside the video frames using face identification algorithms, which combine both conventional computer vision approaches and cutting-edge deep learning techniques (basically they used landmark points on the face such as corners of the eyes, nose, and mouth to find the ROI). A region of interest (ROI) is carefully chosen around the forehead area once the face has been detected because this is where small variations in skin color caused by changes in blood flow are most noticeable.

The remote photoplethysmography (RPPG) signal, which shows variations in skin color over time, is extracted using the ROI as the basis. Various signal processing techniques, including peak identification, normalization, and filtering, are used to improve the RPPG signal's quality. These techniques extract relevant cardiac pulse rate-related characteristics. The subject's heart rate is finally calculated over time using these derived features.



The original bounding box in red and adjusted bounding box that was input into the GrabCut segmentation algorithm in blue (top) as well as the first two iterations of the GrabCut implementation (bottom).

Figure 12

3 Expected Achievements

3.1 Outcome

The outcome of this project is the creation of a reliable and precise system for estimating pulse rate through video analysis. The project intends to non-invasively estimate pulse rates from live video or recorded video streams by utilizing computer vision techniques, signal processing algorithms, and innovative methodologies. The system will manage different skin tones and demographics, minimize noise and artifacts, and deliver real-time or almost real-time pulse rate estimation while overcoming environmental factor problems. The result will be a approachable solution that can help with a variety of applications, including remote health monitoring, fitness tracking, and wellness evaluations.

3.2 Criteria for success

To ensure precise and effective pulse rate estimation from live video or recorded video streams, the following main objectives will serve as the project's success criteria:

3.2.1 Accurate pulse estimation

Based on the video input, the algorithm should estimate the pulse rate accurately with approximated error of 4-6 bpm. By comparing the predicted pulse rates with reference readings received from other trustworthy sources, such as medical equipment or manual counting, the accuracy can be assessed.

3.2.2 Robustness and generalizability

The algorithm must be reliable enough for use in a range of conditions with various subjects, skin tones, lighting, and camera configurations. It must be able to recognize changes in facial expression and record precise pulse rates under diverse conditions.

3.2.3 User-friendly interface

A success criterion may include creating an intuitive and user-friendly interface that enables people to quickly comprehend and interpret the estimated pulse rate if the project contains a user interface or visualization component. Clear and useful feedback should be provided through the interface. Intuitive system is one of our main goals.

3.2.4 Performance across various populations

The system needs to be tested on a wide range of people from all ages, genders, and nationalities. For a wide range of people, it should be able to calculate pulse rates accurately with the approximated error described above.

3.2.5 Real-time processing

Real-time or almost real-time processing could be a success criterion if the project seeks to work with live video streams. The algorithm should effectively analyze the video frames and deliver estimated pulse rates in matter of seconds.

3.3 Special components and engineering/research obstacles

In our project, which involves estimating pulse rate from video, there are several special components and engineering/research obstacles that one may anticipate.

3.3.1 Non-invasive measurement

The fact that video-based pulse estimation provides a non-invasive way to measure pulse rate is one of its main benefits. Video-based approaches use computer vision algorithms to extract physiological information from video records, in contrast to conventional techniques that demand physical contact with sensors or devices. It is more practical and user-friendly to monitor pulse rate in a variety of situations due to its non-invasive nature.

3.3.2 Noise reduction

Videos may have a variety of distortions and noise that make it difficult to determine the pulse with precision. Examples of artifacts include image artifacts, motion artifacts, and variations in illumination. To increase the precision and dependability of pulse rate estimation from video, it is essential to develop efficient noise reduction strategies and artifact handling approaches.

3.3.3 Real-time processing and efficiency

The project must achieve efficient processing if its goal is to deliver real-time pulse rate estimation. The algorithm should be created to process video frames in real-time, making use of parallel processing, efficient algorithms, or hardware acceleration approaches as needed. Real-time video-based pulse estimation has significant engineering problems in terms of ensuring minimal processing requirements and effective resource use. After conducting the research, we chose a technique that minimizes the time complexity.

Overall, overcoming part of these difficulties is essential to creating a reliable and accurate video-based pulse rate estimation system. Additionally, advanced computer vision and signal processing methods must be used to extract useful physiological data from video recordings.

4 Research / Engineering Process

In this project we were faced with a lot of information about our topic, so we had to carefully filter and select the articles and sources that would be most relevant to our idea. After that, reading and understanding these sources was no less challenging, since it is a very complex and new topic for us. After researching the issues in depth, we debated which approaches to choose at each stage of implementing the system. Also, although this is the book phase of the project, we already have a forward-looking view regarding the implementation of this system in part B of the project, therefore this also complicates the considerations of choosing the different approaches.

4.1 Process

4.1.1 Medical Research

A thorough research procedure including several fields of study is necessary to comprehend the medical basis for heart rate and the circulatory system. To begin with, a thorough understanding of cardiovascular physiology and anatomy is essential to understanding the complex operations of the heart and blood vessels. This includes researching subjects including the cardiac cycle, the anatomy of the heart, and the mechanics of blood flow.

An in-depth knowledge of the medical field is necessary to develop a remote heart rate monitoring system. Understanding hemoglobin and its function in determining heart rate in the face is crucial. Photoplethysmography, a non-invasive technology for heart rate monitoring, relies heavily on hemoglobin, the protein in red blood cells that carries oxygen. Accurate measurements may be taken from areas with strong blood flow, like the forehead or cheeks, by understanding the concepts underlying this technique related to its medical background.

There are various difficulties while researching medical information as a non-medical professional. It can be difficult to comprehend and interpret the complicated vocabulary and technical language used in scientific articles when conducting research. It is tricky to separate reliable sources from personal experiences since it is difficult to evaluate the quality and reliability of medical information without experience.

To sum up, the medical basis for heart rate and the circulatory system must be understood, and this requires extensive research in many different academic areas. This include studying cardiovascular anatomy, physiology, and the function of hemoglobin in influencing heart rate. Understanding ideas like photoplethysmography and its dependency on hemoglobin is necessary for creating a remote heart rate monitoring system. However, there are issues when conducting medical research as a non-medical professional, such as understanding scientific papers, understanding technical language, and assessing the reliability of sources.

4.1.2 Different image processing techniques

A thorough research effort is necessary for understanding the various image processing methods used by a heart rate monitoring system. First and foremost, it is essential to examine the literature and academic research on image processing algorithms for heart rate detection. This involves looking into several techniques, including photoplethysmography (PPG), color-based analysis, and motion estimation methods.

We came across several methods for obtaining pulse information via video footage throughout our research. These methods included approaches like GrabCut, which isolates the relevant object from the background while segmenting the face. To extract color signals from the video, we also looked at independent component analysis (ICA) approaches. We delved deep into these many subjects during our study process, investigating its principles, benefits, and disadvantages.

As we dug more into our research, we struggled to decide which technique would be best for our system. We used a spiral research methodology, which means that after looking into a single subject, like a facial recognition system, we kept looking into new techniques. Through this iterative process, we were able to evaluate more options and acquire information from multiple points of view.

We realized the importance of quick response times while working on a pulse detection system from video. Finding methods with high accuracy and minimal processing cost became essential. To meet this requirement, we looked for techniques that balance computational effectiveness with accurate outcomes. Our main goal was to select the methods that would best allow our system to operate accurately, efficiently, and optimally.

We carefully considered several aspects while narrowing down the options for pulse detection in video, considering the limitations. These restrictions mainly concerned the complexity of the algorithms' running times, their resource efficiency, and the trade-off between accuracy and computational complexity. To make sure that the techniques we chose would satisfy the pulse detection system's real-time needs, it was crucial to achieve a balance between these factors.

4.1.3 Hardware – Quality and Resolution

The development of a system that measures heart rate from video depends significantly on the clarity and resolution of the video frames that were recorded.

For reliable heart rate monitoring, high-quality video frames with sharp resolution offer sharper and more distinct details, such as changes in facial color. To successfully extract the pulse signal, the development procedure involves developing algorithms that can assess these aspects efficiently. Images that are clearer and sharper boost the accuracy of the measures.

It can be difficult to obtain accurate heart rate data from low-quality or pixelated video frames because they can introduce noise and artifacts. To reduce the effects of noise and enhance signal quality, noise reduction techniques are incorporated into the development process. The use of sophisticated image processing techniques can assist remove unnecessary noise, resulting in heart rate values that are more reliable and accurate.

Higher resolution cameras can capture finer details, which can help in the precise detection of tiny variations in skin color related with the heart rate. This is important for reliable heart rate monitoring from video. To achieve sufficient detail, choosing cameras with resolutions of at least 720p (1280x720 pixels) or greater is typically advised.

4.2 Product

4.2.1 Projects pipeline

Our system extracts and analyzes tiny variations in facial color from video footage using effective image processing and computer vision techniques, allowing precise real-time heart rate measurement of the subject. We will give an overview of our system's steps and a thorough explanation of how each stage contributes to the precise measurement of heart rate from video footage in the parts that follow.

4.2.1.1 Receiving video input

The system should be able to receive both live and recorded video streams as input.

4.2.1.2 Locating the eyes using the Haar cascade classifier

An efficient machine learning approach for identifying objects, such as face traits like eyes, is the Haar cascade classifier. The goal is to precisely identify the eyes in the video frames.

4.2.1.3 Locating the Region of Interest (ROI)

Once the eyes have been located, the algorithm should move on to choose the forehead as the ROI because it frequently offers a good location for detecting the pulse (related to the hemoglobin we explained earlier).

4.2.1.4 Address the green color channel

Convert the ROI image into a suitable color space, such as RGB (Red, Green, Blue). Split the ROI image into its component red, green, and blue color channels. Then, while working with the picture channels, access the green channel by choosing the proper index.

4.2.1.5 Perform FFT

Fast Fourier Transform, sometimes known as FFT, is a mathematical procedure that converts time-domain data into the frequency domain. The system expects to examine the signal's frequency components by applying the FFT to the extracted data. In addition, it contributes to the noise reduction in our input data.

4.2.1.6 Locating peaks in the necessary frequency range

The approach should locate peaks in a particular frequency band, often between 0.75 Hz and 4 Hz, within the converted frequency domain using bandpass filter. These peaks most likely represent heart rate indicators.

4.2.1.7 Counting peaks within a set time range

An estimate of the pulse rate can be obtained by the algorithm by counting the peaks within the given time and frequency ranges (time range of 15-30 seconds). It is assumed that the number of peaks reflects the number of heartbeats that took place within the specified period.

4.2.1.8 Calculation of the pulse

The algorithm should determine the patient heart rate based on the recognized peaks.

4.2.1.9 Result

The estimated pulse rate should eventually be displayed or visually represented in our GUI.

4.2.2 Requirements

Functional:

Table 1

1	The system should be able to process video input from a camera or recorded video files
2	The system should accurately detect and track the subject's face throughout the video
3	The system should apply appropriate pre-processing techniques to enhance image quality
4	The system should apply appropriate pre-processing techniques to reduce noise
5	The system should extract color signals associated with the pulse from the facial ROI
6	The system should use signal processing methods to evaluate the retrieved pulse signals and determine the heart rate
7	The system should provide real-time heart rate measurements
8	The system should provide continuous monitoring
9	The system should provide immediate feedback
10	The system should provide an alert if the heart rate is dangerous

Non-functional:

1	Easy to use graphical interface for the user
2	Provide accurate heart rate measurements within an acceptable margin of error
3	Consistently deliver reliable results under varying conditions, such as different lighting conditions and subject appearances
4	Process video frames efficiently
5	Provide heart rate measurements in minimal delay
6	Handling variations in skin tone
7	Compatible with different camera devices
8	Steady framerate (20fps at least)
9	Dangerous heart rate will be colored in red

Table 2

4.2.3 Architecture overview

Our architecture consists of several key components:

1. Heart rate monitoring application
2. Video input center which provides input data to the monitoring application
3. Registration DB for user authentication in the login phase and for patients information.

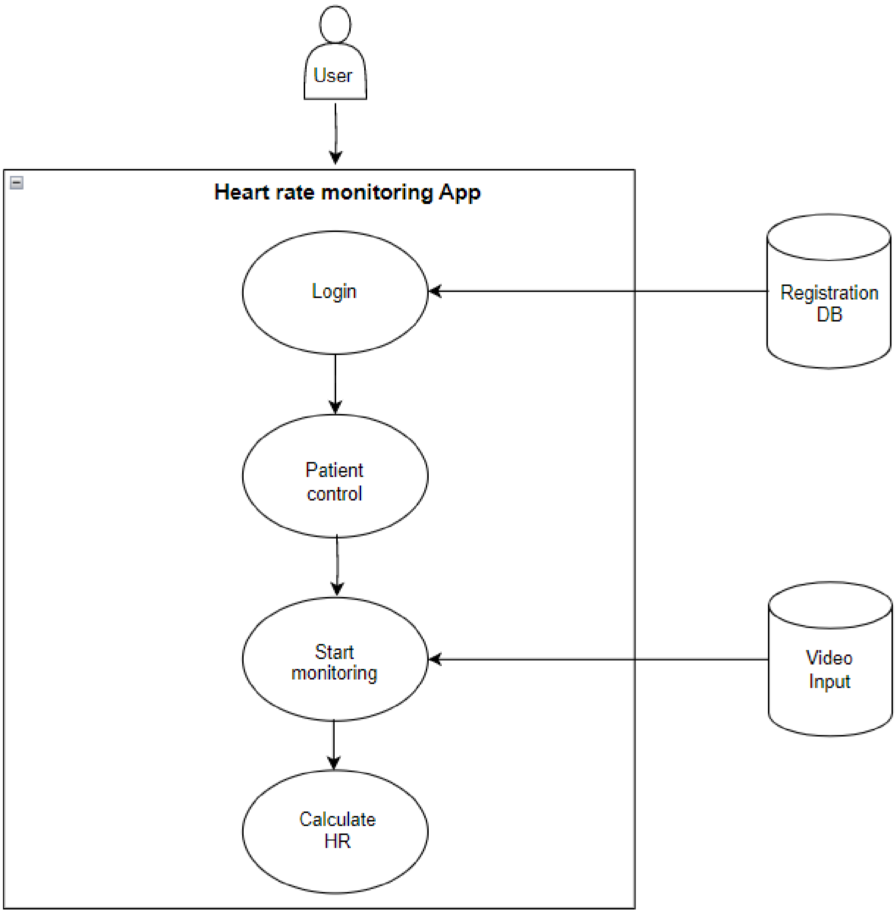


Figure 13

4.2.3.1 Heart rate monitoring Application

A heart rate monitoring application built using video analysis follows a similar architectural pattern known as the Model View Controller (MVC) architecture, which consists of three distinct layers. The application's structure can be visualized in the following figure.

Our application will be using the MVC architecture in the following method:

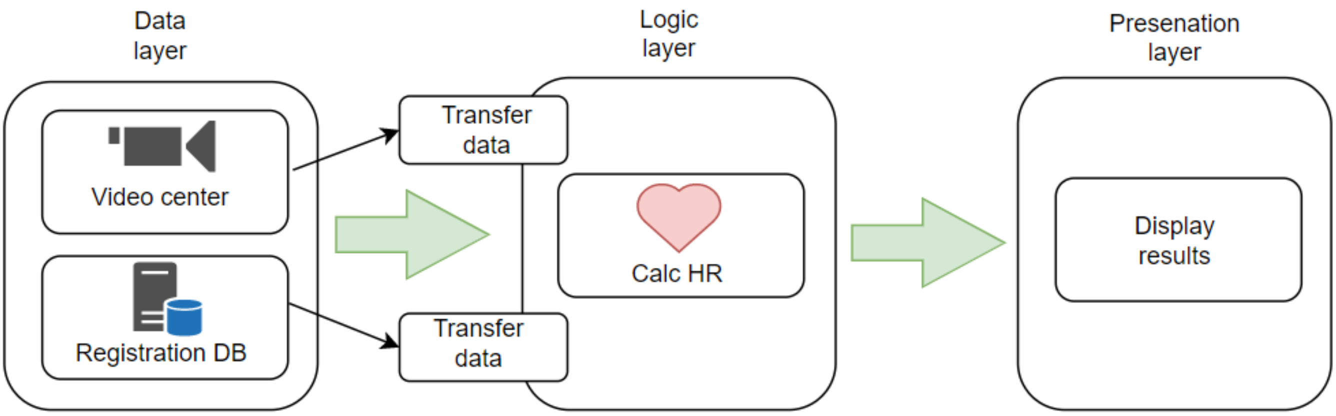


Figure 14

1. *Model*

   Our video center is responsible for sending the video input towards the logic layer where the data is analyzed. The registration DB contains both authorized users and patients.

2. *Controller*

   The controller layer will be responsible to analyze the data in real time, finding the ROI, extracting the signal, calculating the peaks and sum the result. In addition, this layer is responsible for receiving the input video stream from the model layer.

3. *View*

   The view layer is the interface layer where the user navigates between different panels and show different patients. This layer will display the calculated HR from the controller layer and will present visualizations of the HR analysis.

4.2.4 GUI Panels

The login page of the heart rate monitoring system created for nursing home staff members offers a safe and convenient way to access the system. Employees can validate their identification by entering their specific credentials, such as username and password. Only authorized workers can access the program and read sensitive information related to the residents' heart rate data, thanks to the login screen. It acts as the primary entry point to the system and supports data security, confidentiality, and privacy in nursing homes.

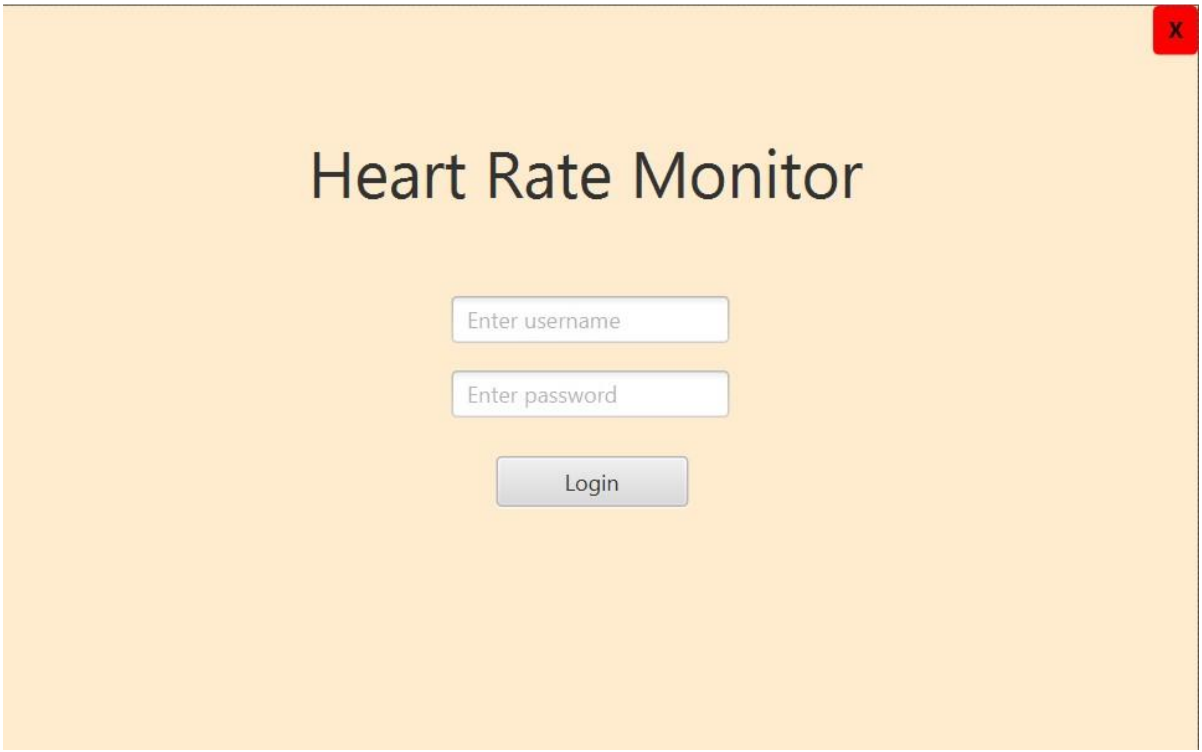


Figure 15

The nursing home personnel can manage patient-related tasks effectively thanks to the menu of the heart rate monitoring application. The menu has three essential selections:

1. The staff can easily register new patients in the system by selecting "Sign up new patient". Each new patient can have their information entered and a profile created by selecting this option, guaranteeing that they are included in the heart rate monitoring system.

2. Another dashboard called the "Patient Control Panel" shows a list of all patients. A tiny video feed displaying each patient's live facial footage is displayed next to their name. The panel also displays the associated heart rate numbers, giving a brief overview of each patient's present heart rate state.

3. The "Logout" button provides an easy way for the user to safely leave the system. The user can preserve their privacy and data by logging out of their account by clicking this button.

Nursing home employees may easily navigate the heart rate monitoring program, add new patients, retrieve patient information, and safely logout as needed through this user-friendly menu.

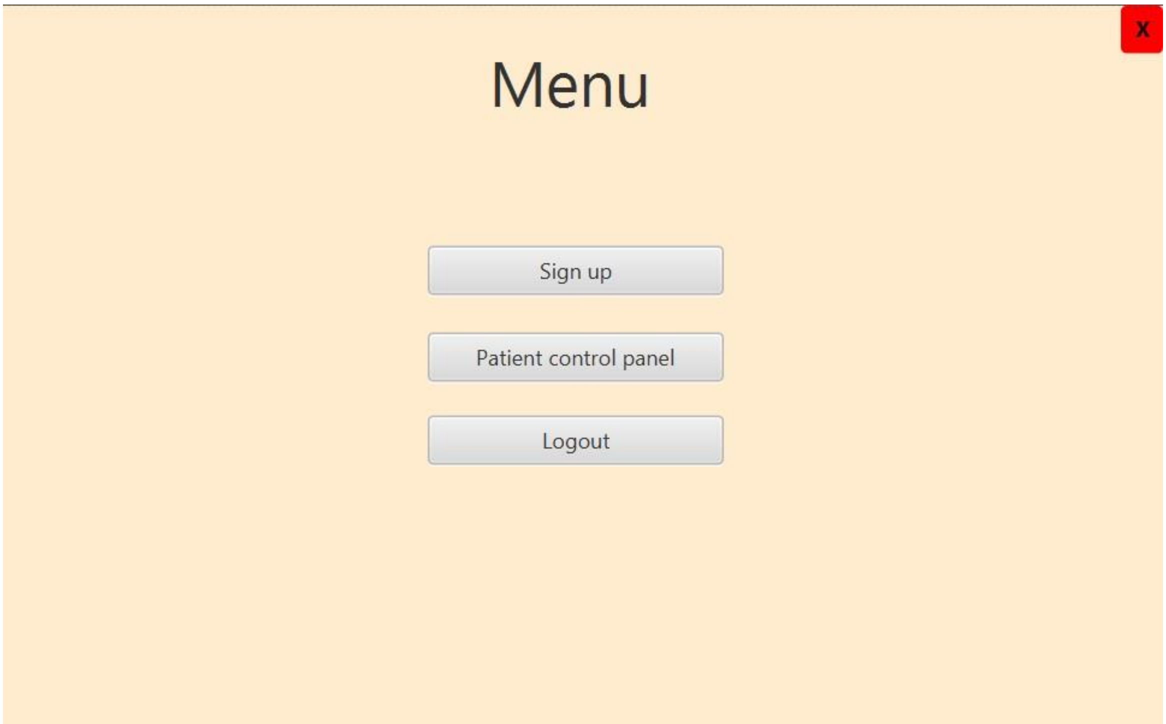


Figure 16

Figure 17

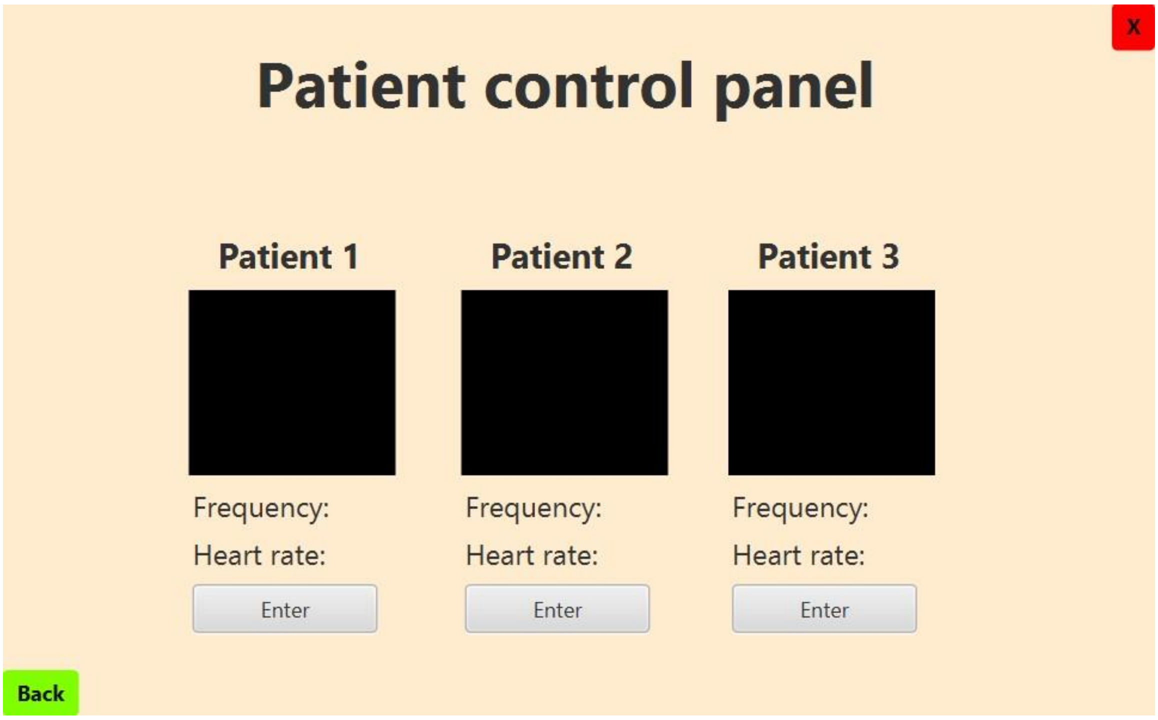


Figure 18

The heart rate monitoring system opens a special window suited to that patient when a specific patient is selected from the control panel and the "Enter" button is pressed. Nursing home workers can access several crucial tools in this window to track the patient's heart rate. They are given the choice of uploading a previously recorded video or turning on the webcam to enter live video. This provides reliable heart rate analysis in real-time monitoring of the patient's facial region. The patient's current heart rate is also shown in the window, giving quick feedback on their cardiovascular health. The window also displays graphs that show the Fast Fourier Transform (FFT) and signal of the RGB (Red, Green, Blue) components to further assist in analysis. This makes it easier to understand the heart rate data and its underlying properties.

Figure 19

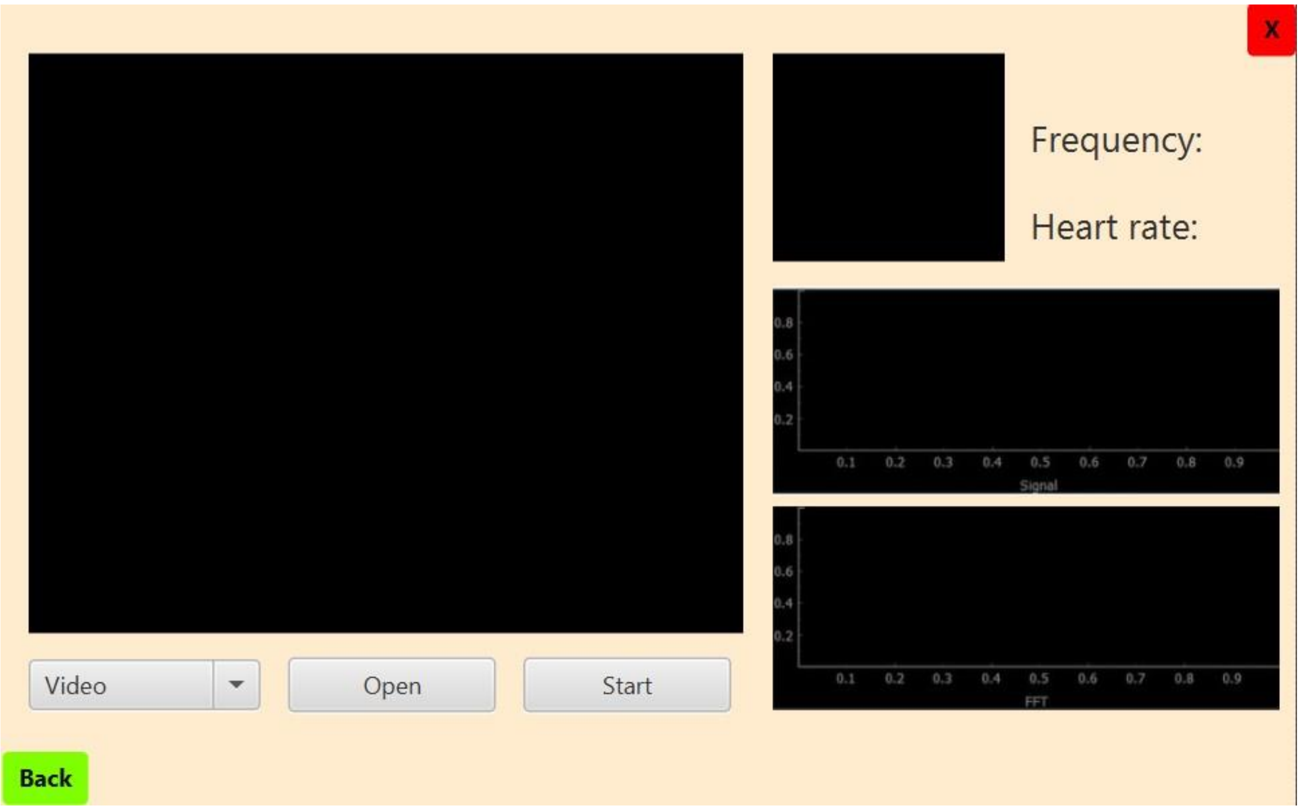


Figure 20

4.2.5 Databases

To manage user accounts and patient information, the heart rate monitoring system makes use of two different databases. Usernames and related passwords for system access are kept in the "User Accounts" database. It manages user rights inside the program and makes sure that authentication is secure. The "Patient Details" database, on the other hand, includes vital details about specific patients, including their first and last names and distinctive patient IDs. This database makes it possible to handle patient records effectively and to quickly retrieve patient-specific information as needed. Together, these databases offer a structured and organized method for managing users and storing patient information, ensuring the heart rate monitoring system runs smoothly and effectively.

{} reg_auth.json > [ ] users

```json
{
    "users": [
      {
        "username": "drcohen",
        "password": "123456"
      },
      {
        "username": "moti12",
        "password": "654789"
      },
      {
        "username": "dan123",
        "password": "dan@1234"
      },
      {
        "username": "doc123",
        "password": "098123"
      }
    ]
  }
```

Figure 21

{} patients.json > [ ] patients > {} 3

```json

{
    "patients": [
      {
        "firstName": "Shimon",
        "lastName": "Rubin",
        "id": "12345"
      },
      {
        "firstName": "Lior",
        "lastName": "Guzovsky",
        "id": "67890"
      },
      {
        "firstName": "Dan",
        "lastName": "Lemberg",
        "id": "54321"
      },
      {
        "firstName": "Elena",
        "lastName": "Kremer",
        "id": "98765"
      }
    ]
  }
```

Figure 22

4.2.6 Diagrams

4.2.6.1 Use Case

The following use case diagram shows the users interaction with the system.

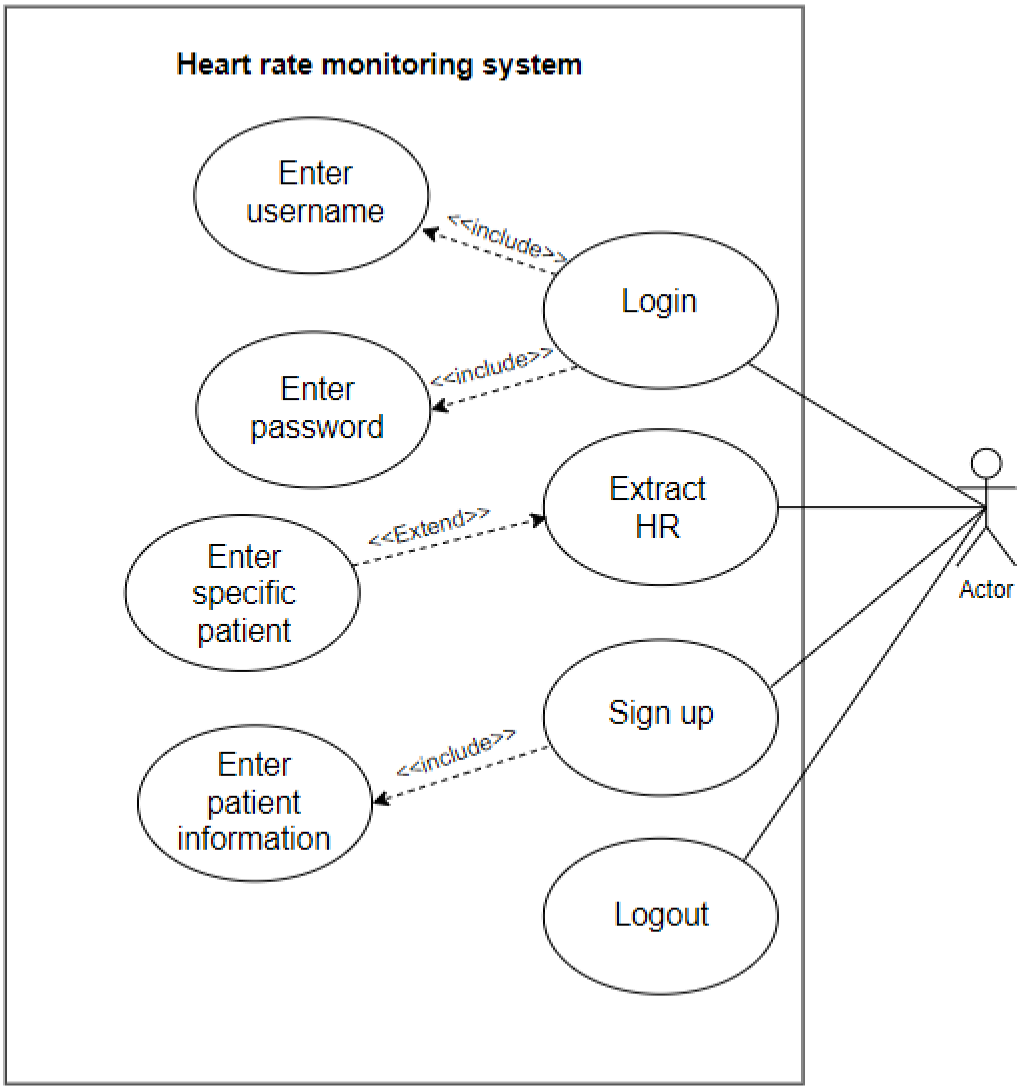


Figure 23

4.2.6.2 Activity diagram

The following activity diagram shows the process of extracting heart rate from a chosen video input by the user, to the point it is shown to the user.

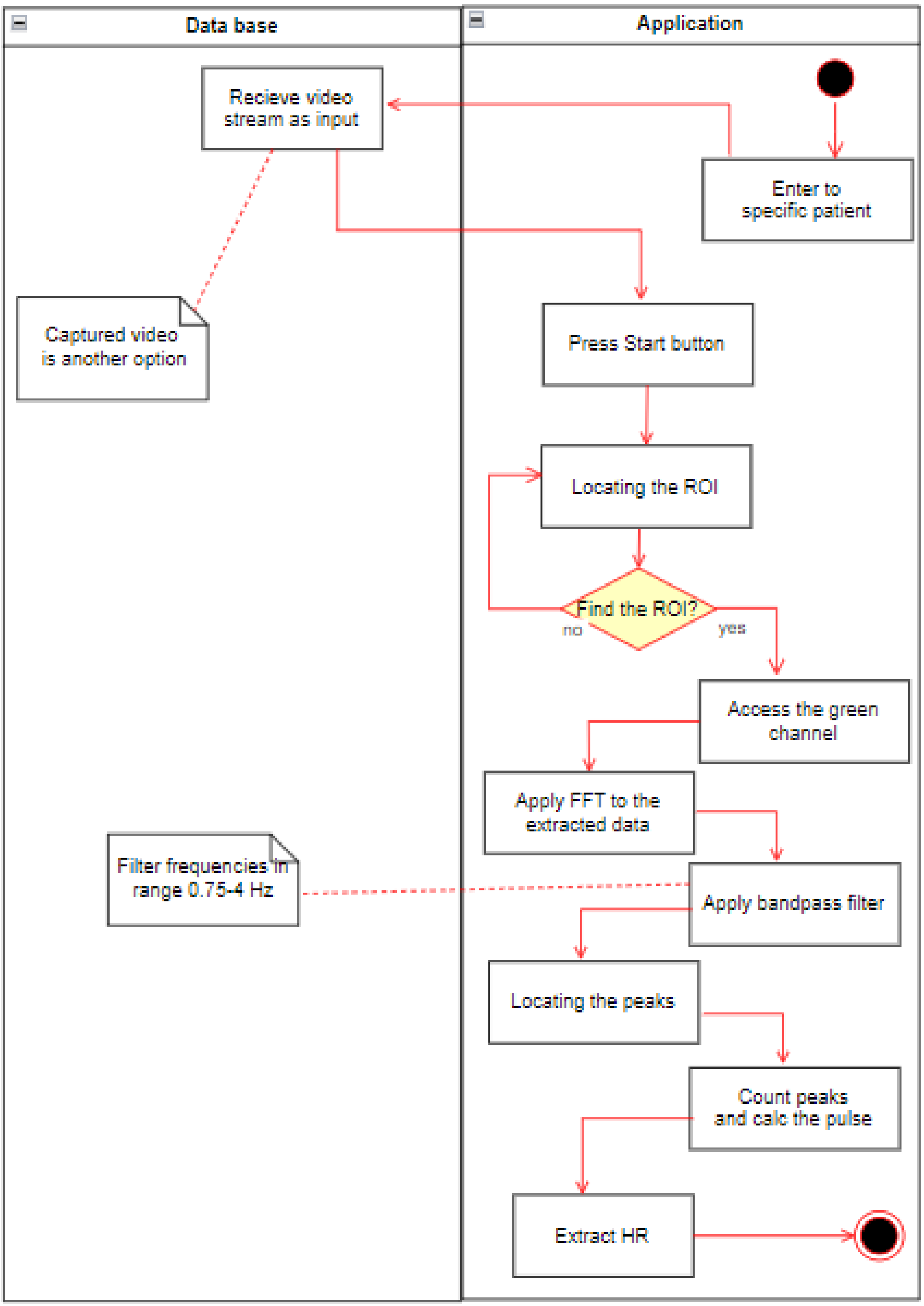


Figure 24

4.2.6.3 Pseudo code for extracting heart rate

```
1. ReadVideoInput():
    - Read and receive live or recorded video streams as input.
    - Continuously analyze each frame of the video.
2. LocateEyes ():
    - Apply Haar cascade classifier to identify the eyes in each video frame.
    - Store the location of the eyes.
3. LocateROI():
    - Locate final ROI (forehead) using eye coordinates from the previous stage.
4. ExtractGreenColorChannel():
    - Convert the ROI image to RGB.
    - Split the image into red, green, and blue color channels.
    - Extract the green channel.
5. PerformFFT():
    - Apply Fast Fourier Transform to the extracted green channel data.
    - Convert the time-domain data into the frequency domain.
6. LocatePeaksInRange():
    - Apply a bandpass filter to identify peaks in the desired frequency range (0.75-4
Hz).
7. CountPeaks():
    - Count the number of peaks within the specified time range (15-30 seconds).
    - Estimate the pulse rate based on the number of peaks.
8. CalculatePulse():
    - Calculate the patient's heart rate based on the recognized peaks.
9. DisplayResult():
    - Display the estimated pulse rate to the user.
10. Repeat steps 2-9 for each frame of the video.
11. End processing when all frames have been analyzed.
```

4.2.6.4 Class diagram

The following class diagram shows the connections and interactions among the key components involved in the process of determining pulse rate from video input.

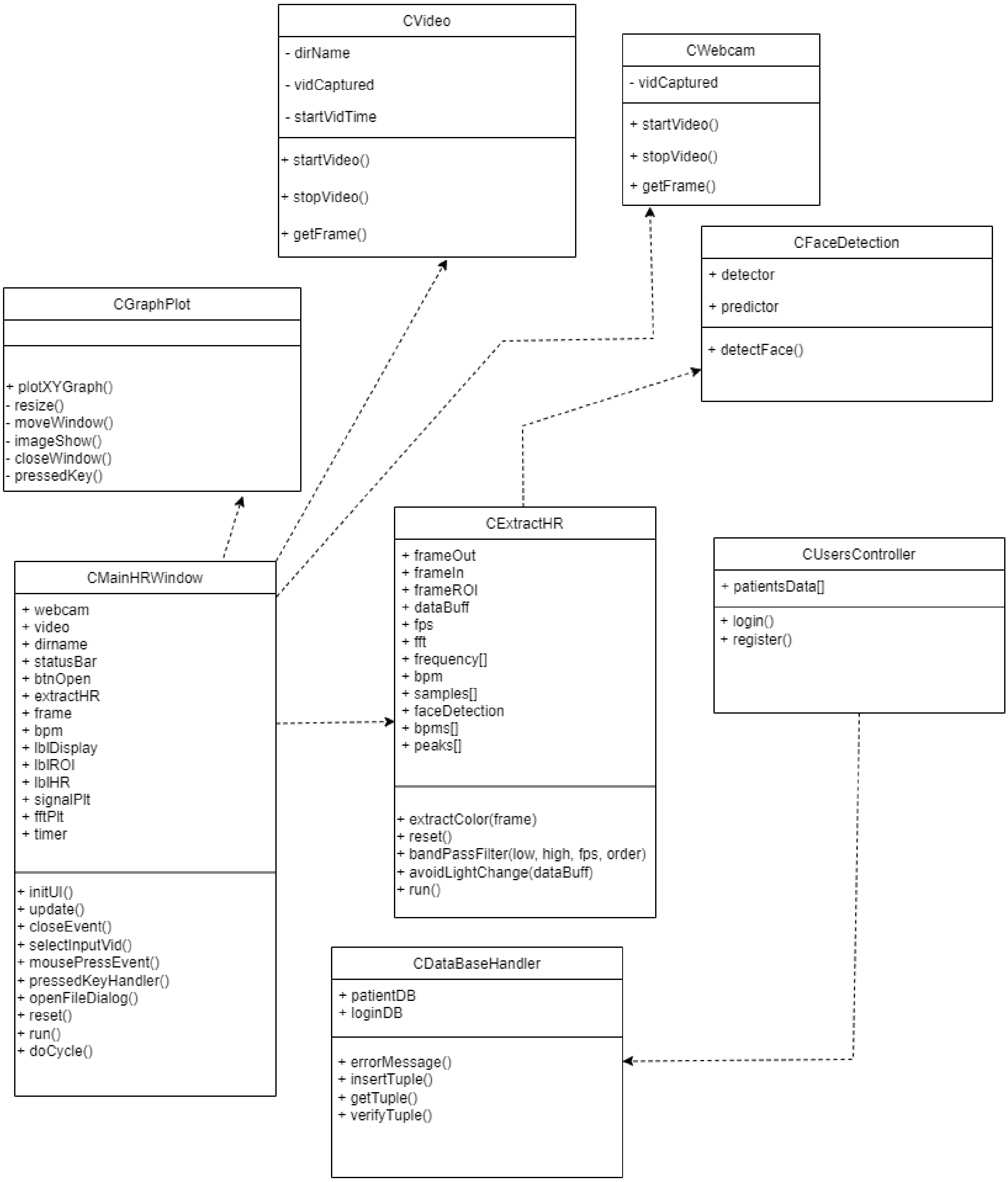


Figure 25

5 Evaluation / Verification plan

5.1 Evaluation

We will be evaluating our heart rate monitoring system in a few ways. First, we will be evaluating our face detection method, which means we will try to assess the system's ability to make a correct ROI segmentation as it is the first phase to detect heart rate from video. We will gather a wide group of videos featuring a range of people, lighting scenarios, and potential problems and test them on this detection method.

Secondly, the most important part is to verify the result obtained by our system in comparison to the real heart rate measured using by another reliable device such as pulse-oximeter.

The heart rate monitoring system's performance is evaluated in the last round of testing using live streaming videos after the face detection technique has been successfully assessed and verified against a pulse oximeter. This phase attempts to verify the system's capability and dependability in precisely recognizing heart rates from dynamic and constantly changing video streams.

5.2 Verification plan

This testing plan describes how to test the Heart Rate Monitoring Application and the related database, that stores patient and user data. To ensure effective and iterative delivery of high-quality software, the testing process aligns to an agile development methodology. The plan intends to confirm the heart rate monitoring application's performance and guarantee the accuracy and dependability of the database. The test number, module, function that is being tested, and the anticipated outcome are all listed in the table below as an overview of the test cases.

Test	Module	Tested Function	Expected Result
1	Login	User login	Only authorized users can login to the Application
2	Registration	User registration	Registered user cannot be re-registered
3	Face detection	Face detection accuracy	System accurately detects and locates faces
4	Face detection	Low lightning	System accurately detects and locates faces
5	Face detection	Different skin tones	System accurately detects and locates faces
6	ROI segmentation	ROI segmentation accuracy	The ROI is accurately segmented by the system
7	ROI segmentation	Head movement	The ROI is accurately segmented by the system during head movements
8	ROI segmentation	Non-relevant facial exclusion	System excludes non-relevant facial features
9	ROI segmentation	ROI not found	Error handling message
10	Heart rate extraction	Heart rate measurement	System accurately extracts and displays the heart rate (compared to the pulse-oximeter results, in range 60-100 bpm)
11	Heart rate extraction	Heart rate alert	When the user's heart rate exceeds the limit, an alert is given
12	Database	Patient data created	User information is successfully updated
13	Database	Patient data not created	User information is not re-updated, handling error message

Table 3

6 References

1. Bush, I. (2016). Measuring Heart Rate from Video. Stanford Computer Science.
2. Hill, B. L., Liu, X., McDuff, D. (2021). Beat-to-Beat Cardiac Pulse Rate Measurement From Video. University of California, Los Angeles.
3. Lee, Y. C., Syakura, A., Khalil, M. A., Wu, C. H., Ding, Y. F., Wang, C. W. (September 2020). A real-time camera-based adaptive breathing monitoring system.
4. Molinaro, N., Schena, E., Silvestri, S., Bonotti, F., Aguzzi, D., Viola, E., Buccolini, F., Massaroni, C. (February 2022). Contactless Vital Signs Monitoring From Videos Recorded With Digital Cameras: An Overview.
5. Viola, P., Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. Mitsubishi Electric Research Labs, Compaq CRL.
6. Chen, X., Cheng, J., Song, R., Liu, Y., Ward, R., Wang, Z. J. (October 2019). Video-Based Heart Rate Measurement: Recent Advances and Future Prospects.