# “Smart Contactless Vital Monitoring System Using Mobile Camera And Smart AI Coach”

## Major Project Report

*Submitted in Fulfillment of the*
*Requirements for the Degree of*

## Master OF TECHNOLOGY

## IN

## EMBEDDED SYSTEM

By

**Disha Bhatti**
**(19MECE03)**


**Department of Electronics and Communication Engineering,**
**Institute of Technology,**
**Nirma University,**
**Ahmedabad 382 481**

**May 2021**

# Acknowledgement

It gives me immense pleasure in expressing thanks and profound gratitude to **Dr Ruchi Gajjar**, Assistant Professor, Department of Electronics & Communication Engineering, Institute of Technology, Nirma University, Ahmadabad for her valuable guidance and continual encouragement throughout this work. The appreciation and continual support she has imparted has been a great motivation to me in reaching a higher goal. Her guidance has triggered and nourished the intellectual maturity that I will benefit from, for a long time to come.

I have been fortunate to have a mentor like Dr Nagendra Gajjar, Professor & PG Coordinator (Embedded Systems), Institute of Technology, Nirma University, Ahmadabad for continuous support.

It gives me immense pleasure to thank **Dr Dhaval Pujara**, Hon'ble Head of Department of Electronics & Communication Engineering, Institute of Technology, Nirma University, Ahmedabad for his kind support and providing basic infrastructure and a healthy research environment.

A special thank you is expressed wholeheartedly to **Dr Rajesh N Patel**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout this work.

I would also thank the Institution, all faculty members of Electronics and Communication Engineering, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

It was a pleasure to be associated with Teksun Inc and I would like to thank Mr Darshil Modi and Ms Jasmin Parmar and all my teammates for them valuable guidance and continual encouragement throughout this work.

Disha Bhatti

# ABSTRACT

As an AI intern, I was associated with Teksun Microsys and was part of two projects: 1. Smart Contactless Vital Monitoring System Using Mobile Camera and 2. Smart AI Coach.

Vital signs monitoring is important for clinical diagnostics and in-home health monitoring. It is measurements of the basic functions of the human body. Medical professionals monitored four main vital signs routinely as Temperature, Pulse Rate, Respiration Rate, Blood Pressure. Vital signs such as heart rate (HR), HR variability (HRV), and respiratory rate (RR), are usually measured with non-invasive electrocardiography (ECG) or photoplethysmography (PPG) sensors in clinical examination or with commercial wearable devices in health monitoring. The measurements in both scenarios often employ contact sensors, which may be inconvenient or cause discomfort in long-term monitoring sessions. For example, it is hard to put sensors on young children and ask them to keep them still during the monitoring session. Some pioneering works reveal a possible approach of remote vital signs measurement with contactless sensors. Extracted HR from face videos, based on the small colour change on the face that is consistent with the pulse signal. This technology is called remote PPG (rPPG). We are using this technique in our project. We are using the rPPG signal and extract the values of the RGB signal after applying some signal processing and based on that information we predict the values of HR, RR and SpO2.

**Smart AI Coach:** A fitness trainer leads or assisting a client to reach their physical fitness goal. Trainer determines the clients' fitness level, individual goal, skills, develop training program according to clients' need and also, monitor the progress. To accomplish all tasks there need to trainer and client present physically, which is not possible for both client and trainer. Smart AI Coach overcomes this issue and helps a client to get online training. AI coach useful to a user so that s/he do not need to go the gym and guided at home only to perform an exercise, also, no need to keep count of sets and measure correct posture thus avoid injuries. We trained a CNN model for 21 different exercises, for that we have created our dataset and used it in our model. We faced so many challenges for exercise which has the same posture while doing exercise, its result comes in misclassification. To resolve this misclassification we used the LSTM algorithm and we got good accuracy with accurate classification.

## Company Profile:

Teksun is a global Electronic R&D Services company providing services such as PCB Design, Hardware Development, Firmware Development, Technology Research, Product Engineering Services, New Product Development, Product Re-Engineering, Product Sustenance, and Manufacturing to several industry verticals such as Consumer Electronics, Automotive, Healthcare, Industrial, Automation, Home Automation, IoT, Security & Surveillance, Aerospace & Defense, Cloud and Mobility.

Teksun provides end to end product development services to its clients. Engineers blend innovation with engineering expertise to offer a quality solution that is cost-effective and ready to deliver.

## Group Profile:

I am associated with this company in the AI Department as an AI intern, working under the guidance of Mr Darshil Modi who is an AI developer in the company.

# INDEX

# LIST OF FIGURES

# LIST OF TABLES

# NOMENCLATURE

**Abbreviations**

| | |
|---|---|
| HR | Heart Rate |
| RR | Respiration Rate |
| Hrv | Heart Rate variability |
| SpO2 | Oxygen Saturation |
| ECG | Electro Cardiogram |
| rPPG | Remote photoplethysmography |
| CNN | Convolution Neural Network |
| LSTM | Long short term memory |

# Chapter 1

# Introduction

This report has been divide into two major project

1. Smart Contactless Vital Monitoring System Using Mobile Camera, and
2. Smart AI Coach

This report has been designed in such a way so that reader can easily understand the purpose behind the project and its implementation of it.

## 1.1 Project: Smart Contactless Vital Monitoring System Using Mobile Camera

The concept of the vital sign monitoring system has been derived to observe the basic function of the human body without any delay by self. This system will help to measure all the basic signs simultaneously on a mobile device. In this era of technology, with help of medical equipment, it is easy to measurement vital sign for a patient health examination. But, contactless measurement is quite difficult. Still, using some ML approach and signal processing techniques we can get a good result in this field.

### 1.1.1 Introduction

Vital signs monitor means the most basic functions of the human body. The main vital sign monitored by most professionals as follows:

- Heart Rate
- Respiration Rate
- Oxygen Saturation etc.

Vital monitoring system used for monitoring patient's health condition and help to providing an absolute way for the treatment.

Generally, these vital signs measured by ECG, Pulse Oximetry (measured by physical contact sensor). But the proposed system will provide the contactless vital monitoring system which will be very helpful for the patient like young children or aged people, because, It's difficult to place sensors on young kids and older person as they could not keep still during the monitoring session. We can see the example of the contactless vital sign monitoring system in fig 1.1.


Fig 1.1: depict contactless vital sign measurement[14]

1) Heart Rate

Heart Rate means the no. of times heartbeats per minute. It is measured as beats per minute(BPM). Normal HR for a healthy adult is in the range of 60 to 100 BPM. The heart pushes the blood via veins and veins expand[8]. During this time of flow of blood, we measure the change of skin colour using the Remote photoplethysmography technique, also known as rPPG. We also use some signal processing method to calculate the heart rate using this rPPG signal. Finally, we train the machine learning model and apply it to predict the actual Heart Rate of the person from the face.


Fig 1.2: Heart Rate in BPM

2) Oxygen saturation

Similar as we measure the Heart Rate, we also measure the oxygen level. The only change is the machine learning trained model. At this stage, we apply that model which is trained on the oxygen saturation data.

3) Respiration Rate

In simple word Respiration Rate count is the no. of breath per minute. In other word counting how many times the chest rises in one minute[8]. Normally the RR range is 12 to 16 for an adult person.

We are measuring the heart rate and oxygen saturation vital sign measurement using the following steps,

1. **Real-time Face detection** to detect the person in the camera frame
2. **Applying signal processing techniques** to get the information about the RGB pixel value of the detected face
3. **Extracting rPPG signal**
4. **Featurization using FFT**
5. **Apply SVR regression AI model**

All the stages we will discuss in the detail in the next chapter no. 3.

### 1.1.2 Objective of the project

There are mainly three objectives for this project first we will capture the face using an android device then in real time we will detect the face, in other words, we will find the location

of the face in the frame. The second objective is to extract the rPPG signal by applying signal processing and finally we predict the vital signs.

### 1.1.3 Motivation

Contactless Vital sign measurement is the application which uses the only video from camera and state of the art of AI algorithm to calculate basic vital sign. The scope of the application is as follows,

1 The patient can monitor their vital sign on their own
2 No need for wearable device or sensors to measure vital sign
3 No need for any sensor to contact the body while monitoring
4 All the vital sign can be a monitor at a time using a camera only

## 1.2 Project: Smart AI Coach

If we just turn back and see the history of India, we can see how Indian culture conscious about health. Smart AI coach application is designed to practice exercise without having a personal coach present in front of the user. Although the user will be guided by an AI coach same as the user can get guidance from his/her coach. Hence it's named 'Smart AI Coach'.

### 1.2.1 Introduction

A personal trainer motivates or assists a client in achieving their physical health goals. Trainers assess clients' fitness levels, individual goals, and abilities, develop skills training highly personalized, and monitor their progress. To complete all activities, both the mentor and the client must be physically present, which is not possible for both. As we can see fig 1.3 a fitness coach is guiding a person for doing correct exercise. But not always it is possible the availability of personal guidance. For that, the solution is Smart AI Coach. Which guides the user as if it is a personal tutor of his.


Fig 1.3: Coach Guides Client to do exercise

## What is AI Coach?

AI Coach is a virtual fitness trainer which Guides user to perform the exercise by detecting the exercise, counting the repetition and detecting a faulty moment.

## Why AI Coach is needed?

1. It is useful so that the user does not need to go to the gym to perform an exercise, s/he can be guided at home only.
2. No need to keep count of sets
3. Focus only on exercise
4. Measure correct posture thus avoid injuries

### 1.2.2 Objectives

For any AI project, the most important parameter is its dataset. model behavior and accuracy of the AI model is depending on the dataset. We used AWS cloud service for data collection and data storage. So, our first objective is to Data collection on the AWS platform and the next objective is the data pre-processing in which we clean our data; here doesn't mean clean with soap and water, but data cleaning means removing the unwanted data and noisy data. Now, we can train our model, once training

is completed we can test our model on the validation dataset to check model performance in terms of accuracy.

### 1.2.3 Motivation

Fitness is a major concern for all human being. this project helpful for those people who are not able to take a personal visit. the motivation behind this project is a person can be able to take training and guidance from his/her fitness coach at home. they no need to visit personally in case of any adverse situation.

## 1.4 Organization of the Report

As mentioned earlier this project has been divided into two parts, we follow that sequence respected to project 1 and 2, and hence the next chapter 2 will introduce the literature review of respiration Rate in the project 'Smart Contactless Vital Monitoring System'. In chapter no.3 will discuss the implementation of the vital sign system and next chapter 4 will discuss the experimental result. We will continue with another project 'Smart AI Coach' through the next chapters. Chapter 5 will introduce the literature review for CNN and different machine learning algorithms. Next chapter no 6 and 7 will introduce the implementation of the CNN model and experiment results respectively. Chapter no 8 dedicated to the comparison of the accuracy of different approaches. Finally, we conclude both projects with chapter 9.

# Chapter 2

## Literature Review

This chapter will introduce the working of Respiration Rate from a different research paper, what they have done with it, and what kind of result they achieve and how it can be used for this project. After the literature survey, we can decide that what kind of approach will helpful for our objectives. In this survey three most related paper were chosen and tried to understand their approach and how it will use for us.

Table 2.1: Literature Survey On Respiration Rate

| Sr. No. | Paper | The method used for RR monitoring | Description |
|---|---|---|---|
| 1 | Respiratory Rate Estimation from Face Videos[3] | Using face RR monitoring in corporate HR, HRV | The paper represents the method for remote Respiratory Rate monitoring using face video. HR, HRV, and RR measurements are all included in the proposed system. In the rest of the cases, this approach is efficient and robust[3]. |
| 2 | Multiparameter Respiratory Rate Estimation From the Photoplethysmogram[9] | In the applied algorithm RR estimate from PPG signal. The algorithm combines RIFV, RIIV and RAIV respiratory induced variation. | The Smart Fusion algorithm for online RR estimation from PPG was developed in this paper. The algorithm uses a straightforward mean calculation to combine the components of three respiratory-induced variations (RIFV, RIIV, and RIAV) and excludes estimations deemed inaccurate due to conflict between individual RR |

| | | | |
|---|---|---|---|
| | | | estimations or the detection of artefacts. |
| 3 | Robust respiration detection from remote photoplethysmography[15] | Pulse and respiration induced skin colour variation | In this paper, The close similarity between pulse and respiration induced colour changes of the skin can be observed with a camera in both visible and dark lighting conditions. The proposed approach was thoroughly tested using 52 difficult videos that included both seated adults performing guided breathing and newborns breathing spontaneously in a supine position. The proposed approach outperformed previous PPG-based respiration measurement approaches by a wide margin. |

To count the RR from the face in real-time, we have to apply basic filters and signal processing method. First, we used region-of-interest (ROI) to obtain the pulse signal on the face in the video sequence, i.e., the rPPG signal.

Then temporal filtering is applied to exclude the energy out of the typical HR range. To keep the frequency of the rPPG signal within the HR range, a temporal bandpass filter was used. To improve the signal quality, add an infinite impulse response (IIR) filter with a narrower bandpass to the rPPG signal.

The Photoplethysmography method can be a very useful method for respiration monitoring.

# Chapter 3

## 3.1 Implementation

In this chapter, we will discuss hardware and software requirements for this application. The software we used to develop the application and the hardware is used to deploy the developed application for the real-time use case.

### 3.1.1 Hardware Requirement

Nowadays many applications deployed in the smartphone. we cannot use any live application without a smartphone.

For this project, we deploy our model on the Android Device. It is very helpful because we have a model that can be measure all the vital signs in real-time. But it is worthless if we can not use it in real-time. Android devices used to show all the vital results on the screen in real-time.

Fig 3.1: Android device

### 3.1.2 Software Requirement

The software requirements describe the features and functionality of the target device. Users' expectations of the software product are expressed in requirements. For vital sign measurement product we used mainly two software tool which is, Open CV, ML Kit API Android Studio. We can see the logo of this software tool in fig 3.2.


Fig 3.2: Opencv and ML Kit

### 3.1.3 Open CV-Library

OpenCV is an open-source software library. This library requires more than 2500 optimized algorithm inside it. For object detection, face detection, classify human action, stitch images, tracking object movement, find a matching image in the database, etc. with help of all these optimized algorithms. This system integrates the OpenCV library with an android studio for face detection. All dependency of OpenCV import in an android project and import all modules for further process.

### 3.1.4 MLKit for face Detection

ML Kit is an open-source, Google provided Mobile SDK that is useful in machine learning application to use the custom TensorFlow Lite model in mobile apps, such as Android and iOS. MLKit provides so many APIs for Video and images analysis to label image, barcodes detection, text, faces, and object detection also[1].

## 3.2 Flow diagram of vital sign measurement

The following flow graph will give a deep understanding of the vital sign monitoring system. All the stages will be explained in detail. Using this diagram we will understand the working principle of all the block.

1. **Live Video feed(Camera):**

   Use an Android Phone camera to implement the vital sign monitoring system. This camera capture real-time raw data frame, apply a face-detection algorithm on each frame. This system

deals with the real-time situation in which a camera captures the real-time human face on the camera preview. The algorithm of face detection only works on the raw frame. Hence real-time capturing by the camera is divided into many frames.

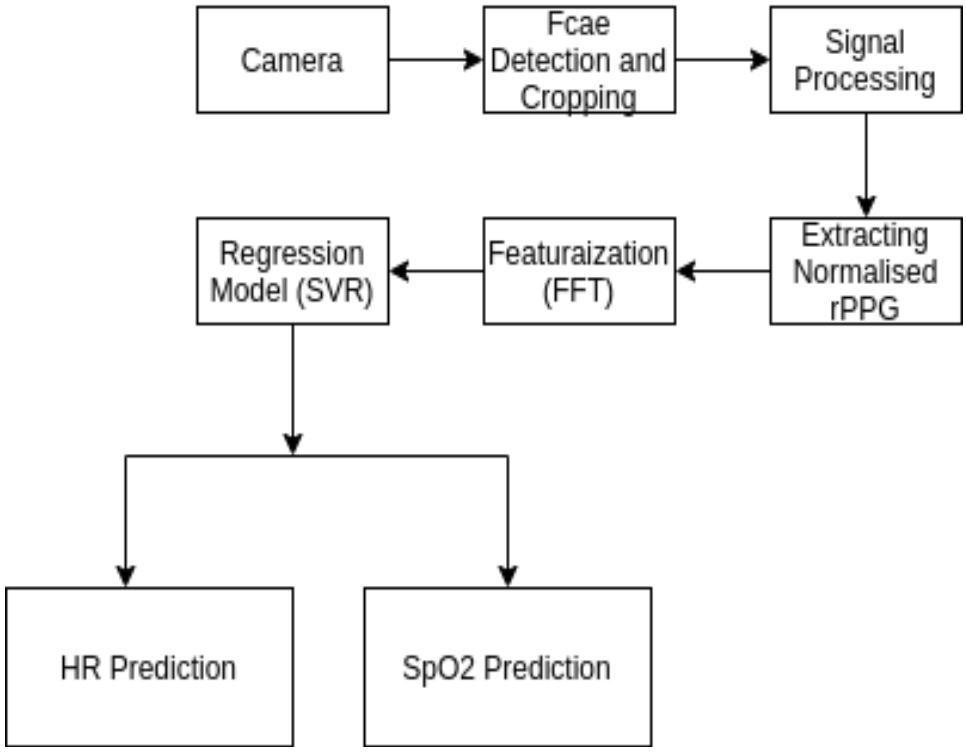



Fig 3.3: Vital sign Monitoring Flow Diagram

**2. Face detection model:**

ML-kit faces detection model used at this stage in which image should be of at least 480*360 pixels of dimensions and to accurately detect faces, the image must contain faces with sufficient pixel data. Also, this product gives a landmark on the detected face as shown in Fig.

Fig 3.4: Facial landmark detection of MLkit model

3. **Face cropping:**

Using the OpenCV library we cropped the face area only and converted it into Mat to separate the YUV signal from an image. Then we separate the YUV channel of this Mat. Face cropping is the essential step to remove the noise for further processing and also less computational power will be used.

4. **Extract rPPG signal:**

` To get rPPG signal from the face we take 25 equal part of the image and average pixel value of each part in the array because all the value inside the array is from 0 to 255.

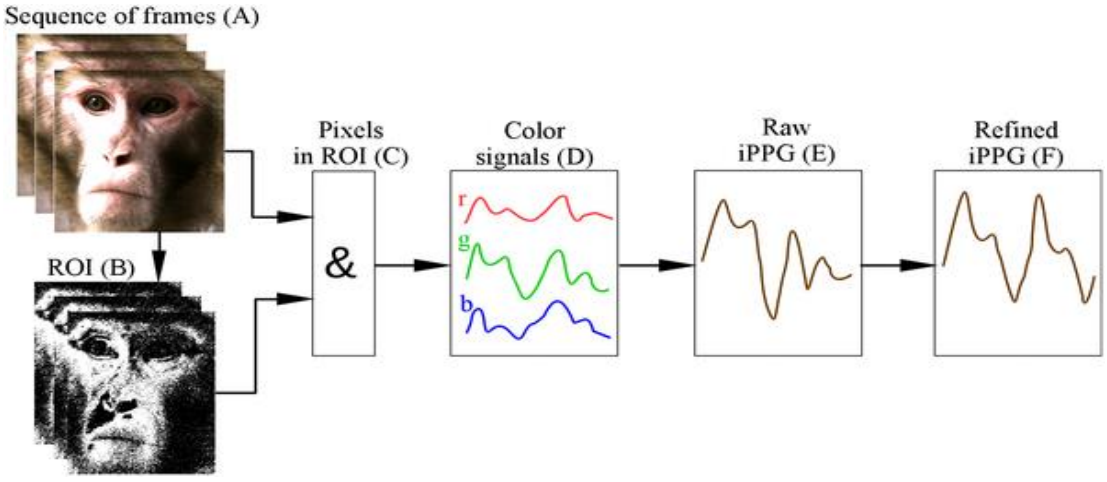



Fig 3.5: Extracting rPPG signal from the image

**5. Featurization :**

We use FFT Method to extract the desired feature from the signal. i.e. how many picks we get in one minute in the signal, that number we consider as the heart rate of a person, and how many times the signal crosses the zero line.

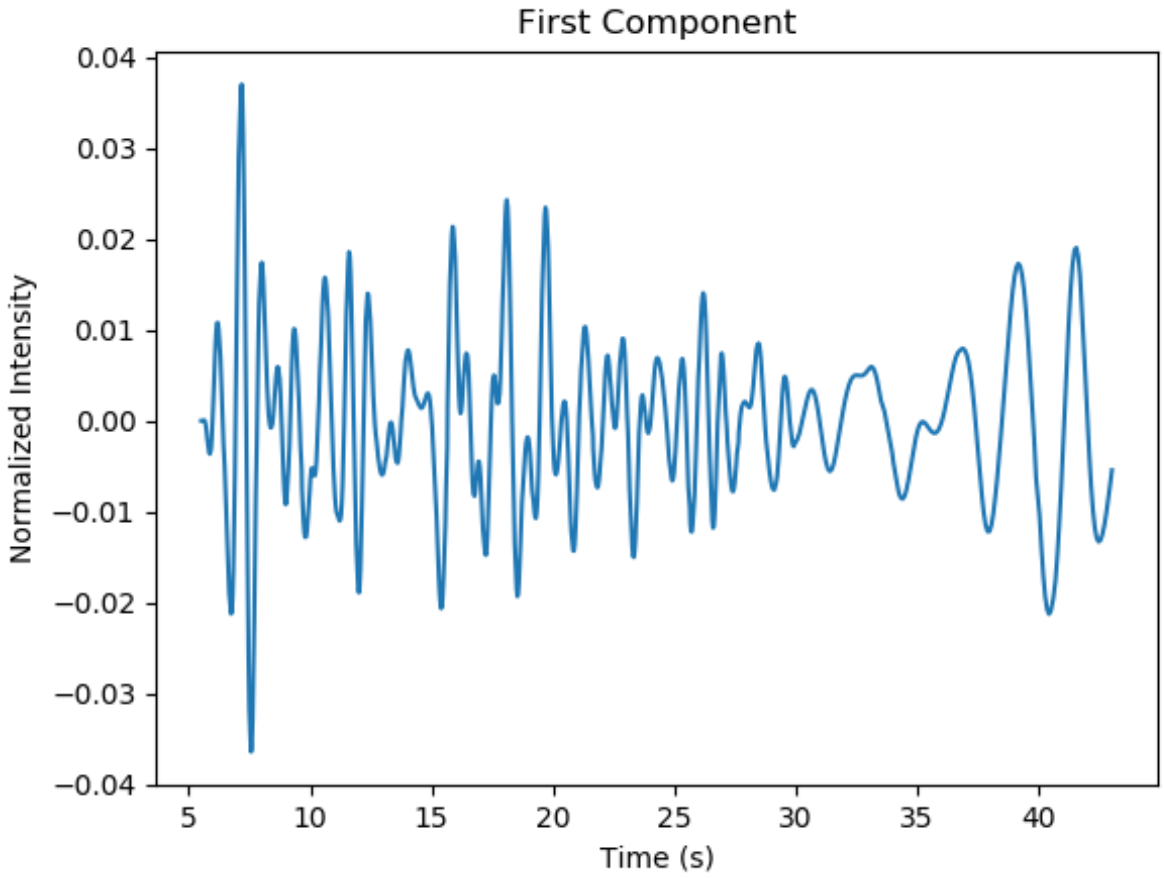



Fig 3.6: Understanding of FFT signal for HR count

**6. Support Vector Regression:**

Support Vector Regression is a supervised learning model for analysing data that is used in classification and regression analysis. In the high dimensional feature space, it performs linearly, using insensitive loss it tries to reduce model complexity. It maintains all the main feature. Using this algorithm, we predict the value of HR and SpO2(oxygen Level)Vital sign.

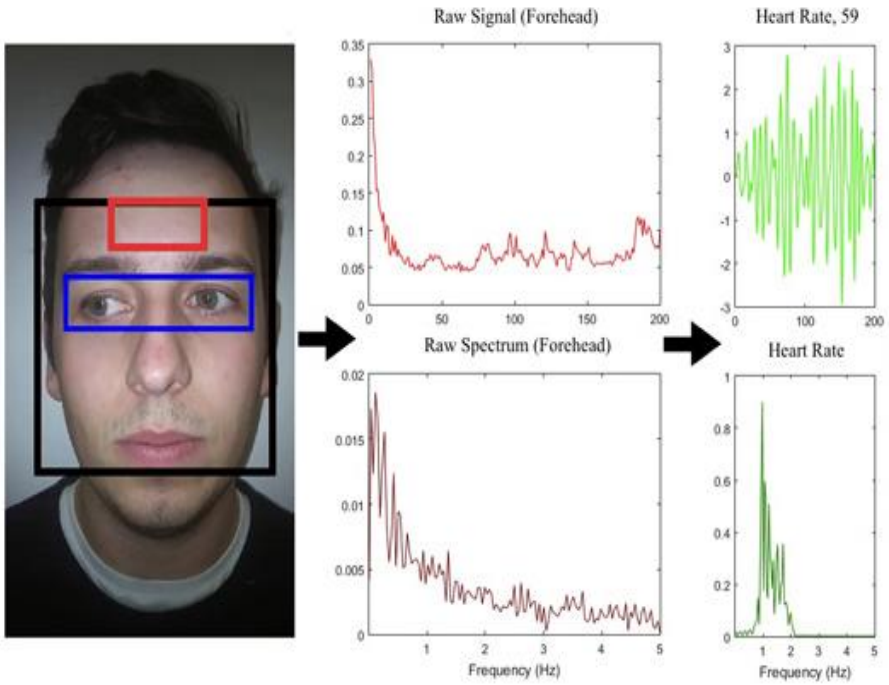



Fig 3.7: prediction of HR using face

Using the rPPG signal and signal processing algorithm we can extract the HR prediction using face only. The user has to be set the position in front of the camera so that model can get a spectrum from the forehead of the user and then using the FFT model can predict the heart rate of that user.

# Chapter 4

## 4.1 Experimental Results

As we discuss in the previous chapter, we used MLKit for face detection and Landmark detection also. we get all the landmark on the exact location of the facial part i.e. eye, nose, lip, and jaws as shown in figure 4.1

Fig 4.1: Facial Landmark Detection

Fig 4.2 is an example of the final result. After deploying the model in the android device application will run and gives the prediction of all the vital signs.

As shown in the figure while the camera starts the capturing frame of the person within a minute the system will give the measurement of all the vital sign from the face only. We convert the AI model into a tflite file, which is supported in the android platform. The tflite file will run in android and the model will predict the output for the person who is facing the camera.

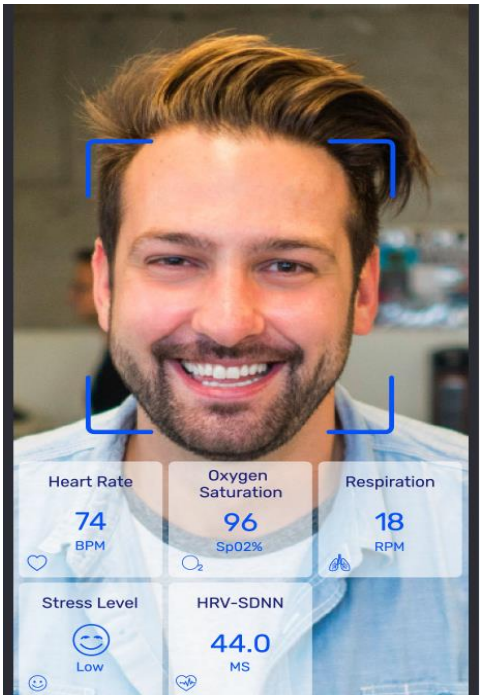



Fig 4.2: Vital Sign Measurement Result

As we know all the vital sign is very important for the healthy life. If a person is not able to go to the professional immediately, he/she can monitor his/her health status and take advice from the professionals.

# Chapter 5

## Literature Review

This chapter will introduce the working of Human Activity Recognition from a different research paper, what they have done with it, and what kind of result they achieve and how it can be used for this project.

Table 5.1: Literature Survey On Human Activity Recognition

| Sr. No. | Paper | The machine learning approach used | Accuracy | Description |
|---|---|---|---|---|
| 1 | Anticipating Human Activities Using Object Affordances for Reactive Robotic Response[10] | KNN, Random Forest, SVC | Very less accuracy, points got jittered after pose detection. | The approach is easy but it is very difficult to measure the points. In the approach, they got minimum and maximum points at a very different level. So they cannot predict the actual position of key points[10]. |
| 2 | Human Activity Recognition from Accelerometer Data Using Convolutional Neural Network [11] | CNN | 92.7% | The method in the paper used is CNN based. The method used in human recognition is using triaxial accelerometer data collection. In CNN based approach the accuracy is 92.7% and in the random forest accuracy is 89.10%.[11] |

| | | | | |
|---|---|---|---|---|
| 3 | Automated daily activity recognition for video surveillance using neural network[12] | Image processing techniques used to classify human movements. | Not mentioned | Image processing techniques are used to classify human movements. Multi-layer feedforward is used to classify the activity models. There are two choices of a feed-forward network. The first is a Single layer and the Second is a multi-layer feed-forward. Only linearly separable patterns can be learned by a single-layer perceptron. Multi-layer networks, on the other hand, are made up of multiple layers of computational units that are normally linked in a feed-forward manner. Each neuron in one layer is connected to the neurons in the next layer through guided connections. A sigmoid function is used as an activation function in networks. |

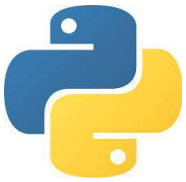
| | | | | |
|---|---|---|---|---|
| 4 | Abnormal human activity recognition using Bayes classifier and Convolutional neural network[13] | CNN | 96% | The technique used is Bayes classifier and convolutional neural network. Activities recognition are walking, running, punching and tripping. The accuracy they got throughout the experiments is 88%, 92%, in Bayes classification and 92%, 96% in CNN. A common application where Kalman filer is used in guidance, navigation and control of vehicles, aircraft. A Kalman filter is widely applied in the application of time series.[13] |
| 5 | A novel feature map for human activity recognition | Feature map, SVM | 94% | In this paper, the pretrained Clarifai net and VGG net-D model are used for tunning of networks. Clarifai net has a kernel size of 7x7 in first convolutional, which contains higher resolution. Both networks have a strong ability to learn the features of data. It has a sliding window approach to detect human of different size. After feature extraction of each candidate, these feature vectors are |

| | | | | |
|---|---|---|---|---|
| | | | | compared with a single network in feature extraction. PCA based approach used to extract the features. Binary classification is used to train the classifier. After that comparison will be held and according to that get the output. |

After the literature review, we have seen that they have used both CNN and Machine learning algorithm. ( i.e. KNN, Random forest, SVM and so on.) but, they got higher accuracy using CNN to compare to the machine learning algorithm. Hence we decide to use the CNN model in our project.

# Chapter 6
# Implementation

## 6.1 Tools and Technology

Python


Jupyter notebook & google Colab


AWS Cloud Service

We have used such technology to accomplish this project i.e. a) LabelImg Tool b) python c) Jupyter notebook and Google Colab d) AWS Cloud Service.

a) AWS Cloud Service:

Amazon Web Services (AWS) is a protected cloud services platform that provides computing power, database stockpiling, content conveyance and other usefulness to help organizations scale and develop. In other words, AWS permits us to do the – Running web and application servers in the cloud to have dynamic websites. We create our dataset in the office, to store the data and retrieve it for future usage we use AWS cloud Services.

b) Python:

For the Preparing AI model, we use python language as python is a very popular language for machine learning and AI. Python is high level, interpreted programming language.

c) Jupyter notebook and Google Colab:

Jupyter notebooks an open-source web application that permits us to make and share python file that contains live code, detail, representation and logical content.

Google Colab permits anyone to write and execute self-assertive python code through the browser and is particularly appropriate to AI, ML, data analysis and education.

d) LabelImg:

This tool is used for annotation of the image with a respective label. It is a very popular tool in ML & AI world for annotating images.

## 6.2 Approaches

In AI there are different architectures, a different model exists. Since the different model has different accuracy we can not decide which one is the best fit for our dataset. We have to try and examine which model is best for our dataset. We tried mainly two approaches, 1) Dark flow, 2) CNN+LSTM. We will discuss it one by one.

1) Dark Flow

Dark Flow is a python implementation of YOLOv2 using TensorFlow. We have used YOLO's pretrained model for object detection and configured the model to work for our 21 classes for exercises. Firstly we annotate the images with a respected label of a particular exercise and train our model. Testing on TensorFlow object detection with "SSD Mobile Net v2 "model and we get output. We observe that we didn't get desired output by this approach as we get misclassification as well as false multi-classification. The result will be shown in the next chapter.

2) CNN+LSTM

This model is created using the tensor flow library and Keras. The model uses a combination of Convolution Neural Network and LSTM. The reason for using CNN is that CNN uses a hierarchical model to create a network and then outputs a fully connected layer in which all neurons are connected and the output is processed. It uses a sliding window that traverses through each part of the image and identifies the most important pixels and then uses that information for classification.

- **LSTM**

LSTM stands for long short term memory. This is used since we are dealing with videos and videos is a combination of frames. So the model needs to know what was the last frame along with the current frame to make a prediction and determine the exercise based on movement. LSTM used for complex problem domains like action recognition, machine translation, speech recognition, and more. LSTMs, on the other hand, use multiplications and additions to make minor changes to the data. information flows through a system known as cell states in LSTMs. LSTMs may selectively recall or forget things in this way. There are three different dependencies on the information at a specific cell state.

The information that was present in the memory after the previous time step was stored in the previous cell state.

The hidden state is the state that is the same as the previous cell's output. The input at the current time step where the new information that is being fed in at that moment.

Another important feature of LSTM is its analogy with conveyor belts!

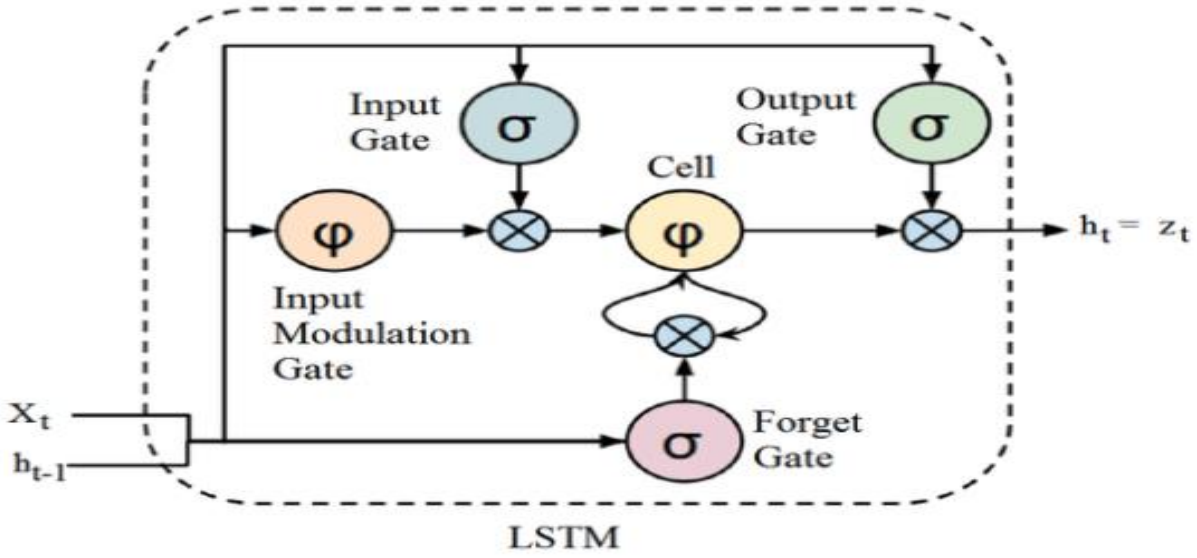



Fig 6.1: Working of LSTM

They are used in industries to move products around for various processes. This mechanism is used by LSTMs to transfer data around.

Information may be added, modified, or removed as it passes through the various layers, much as a product may be moulded, painted, or packed while on a conveyor belt. Our model uses CNN with a bottleneck and LSTM layer to first identify important features in an image and then remember them for further frames to predict accurately for the entire video.

# Chapter 7

## 7.1 Experimental Results

In this chapter, we will discuss the results of two different approaches. We can compare both of the results concerning model accuracy and precise output. First, we discuss the Dark Flow approach and its result.

First, we annotate the images concerning the labels and train the Darknet model with annotating the dataset. Figure 7.1 a-b shows the output of the Dark flow model. This approach is worthless to use in our project as it gives the label of the detected exercise along with a bounding box around the detected object which is incorrect results and misclassified the exercise.

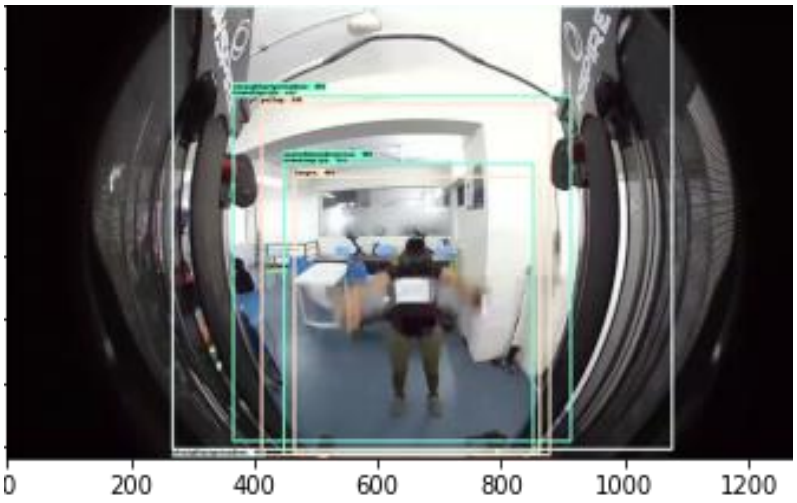



Fig 7.1(a): Object detection result of Dark Flow for exercise 1

In Fig 7.1 (a) and 7.1(b) a person performs two different exercises but the model gives incorrect and more than one bounding box and label. We are not sure why this is happening but one reason would be because Dark Flow is the implementation of YOLOv2 and YOLO is useful for multi-object detection. For this one reason, we do not get correct detection over the image.

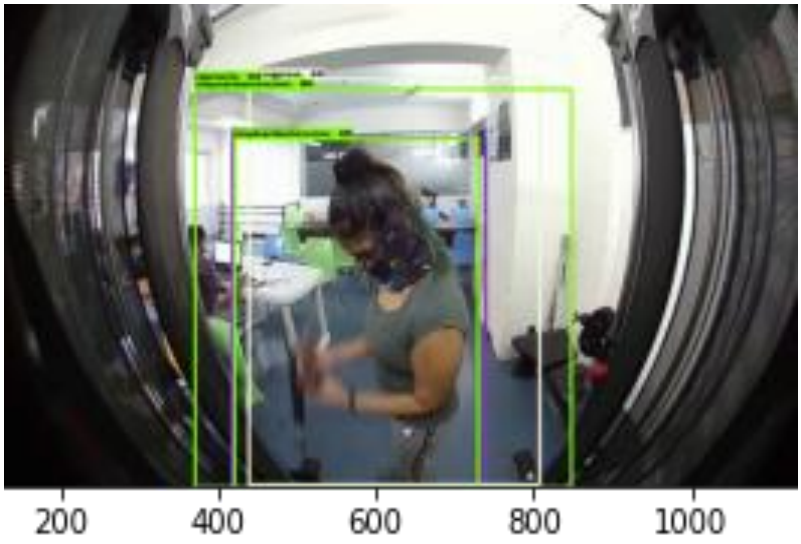



Fig 7.1(b): Object detection result of Dark Flow for exercise 2

Somehow background is also affected. In the validation data if the background changed there is a high chance to get misclassification and less accuracy.

Hence we tried the second approach in which first we convert RGB Image into absolute difference format. The absolute difference is the technology in which it detects the motion between two consecutive frames and captures only edges. This method is very helpful to make an image background-independent.

Fig 7.2(a): RGB image

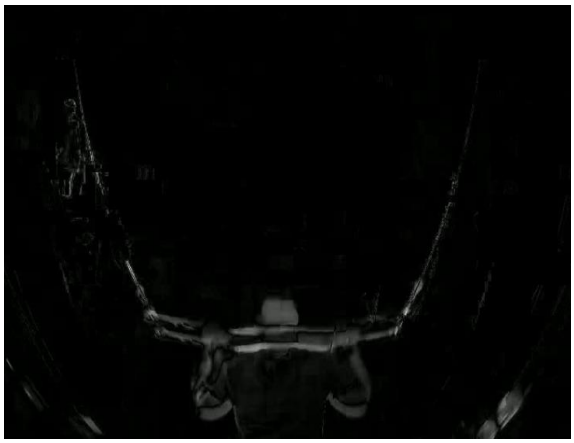


Fig 7.2(b): absolute difference Image

After converting in Abs (absolute difference) format we feed these images in the CNN+LSTM model as input then the model will predict the actual output with good accuracy.


Fig 7.2(c): Exercise detection using CNN+LSTM model

In Fig 7.2(c) the output of the CNN+LSTM model gives a very accurate result. This model predicts only actual exercise, unlike Dark Flow if does not gives a bounding box.

# 7.2 Accuracy comparison

In this section, we will compare the different model accuracy concerning the performance. We have used four approaches to complete the goal of exercise detection. That is DarkFlow, CNN, ResNet and a combination of CNN and LSTM.

CNN and ResNet model get overfitted again and again on training data. If we change the data or add more data for training still model remain overfitted. So we left with CNN and ResNet.`

Table 8.1: Accuracy comparison for a different model

| Model | Accuracy |
|---|---|
| 1. DarkFlow | Misclassified |
| 2. CNN | Overfitted(training acc-95%, test-30%) |
| 3. ResNet | Overfitted(training acc-95%, test-10%) |
| 4. CNN+LSTM | Around 95% |

From the above table, we can see that the Dark Flow model is not worth using anyway as it misclassified all the exercise, and CNN and ResNet model was overfitted hence these two models also are not useful for us. The only approach, CNN+LSTM is useful as it gives about 95% accuracy.

# Chapter 9

# Conclusion and Future scope

## 9.1 Smart Contactless Vital Monitoring System Using Mobile Camera

The ‘Contactless Vital Monitoring System’ uses the camera to detect the person and using signal processing cropping the face from the frame and Analyze the colour changes in the veins due to blood flows using rPPG signal analysis. Then apply FFT for featurization which will give us the information of Heart rate and oxygen saturation measurement, after that, we apply the SVR machine learning algorithm to predict the Heart Rate and Oxygen Level in real-time. To run this application on an end device i.e. android, we convert this model into a tflite file and deploy it in android studio, then we able to get an actual prediction of the value of the vital sign.

## 9.2 Smart AI Coach

Firstly, we convert the RGB video frame to Abs format and fed it to our AI model. Till now we tried four approaches, Dark Flow, CNN, Resnet, and CNN+LSTM. As Dark Flow is used for multi-object detection and it gives misclassification on all exercises, hence it is not so useful for us. So, we left this approach and explore the new one which is CNN+LSTM. This approach gives us the best output with around 95% accuracy.

## 9.3 Future Scope

In the ‘Contactless Vital Monitoring System’ We almost close to the results but the main issue that we are facing of FPS(Frame Per Second). We are using the MLKit model for face detection that is only supported by the CPU. For better performance, we need to run the model on GPU to increase FPS as well.

In the ‘Smart AI Coach’, we have a total of 50 exercises. Till now we have done 20 exercises and now we are going to add the remaining 30 different exercises with different posture and movement.

# References


[1] "Detect faces with ML Kit on Android | Google Developers", *Google Developers*, 2020. [Online]. Available: https://developers.google.com/ml-kit/vision/face-detection/android. [Accessed: 04- Dec- 2020]

[2] [online] Available at: <http://arXiv:1909.03503v1> [Accessed 16 September 2020].

[3] 'Respiratory Rate Estimation from Face Videos' Arxiv.org. 2020. [online] Available at: <https://arxiv.org/pdf/1909.03503.pdf> [Accessed 16 September 2020].

[4] [Online]. Available: https://www.worldbestlearningcenter.com/tips/Android-crop-image.htm. [Accessed: 05- Dec- 2020].

[5] multi-points, A. Res and A. Res, "Android - Crop an image from multipoint", Stack Overflow, 2020. [Online]. Available: https://stackoverflow.com/questions/11579645/android-crop-an-image-from-multipoints. [Accessed: 05- Dec- 2020].

[6] "Pulse oximeter, Blood Oxygen Monitor", PulseOximeter.org, 2020. [Online]. Available: https://www.pulseoximeter.org/. [Accessed: 05- Dec- 2020].

[7] Pathpartnertech.com, 2020. [Online]. Available: https://www.pathpartnertech.com/camera-tuning-understanding-the-image-signal-processor-and-isp-tuning/. [Accessed: 05- Dec- 2020].

[8]"Vital Signs (Body Temperature, Pulse Rate, Respiration Rate, Blood Pressure)", 2020. [Online]. Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/vital-signs-body-temperature-pulse-rate-respiration-rate-blood-pressure#:~:text=Vital%20signs%20are%20measurements%20of,Respiration%20rate%20(rate%20of%20breathing). [Accessed: 05- Dec- 2020].

[9] W. Karlen, S. Raman, J. M. Ansermino and G. A. Dumont, "Multiparameter Respiratory Rate Estimation From the Photoplethysmogram," in IEEE Transactions on Biomedical Engineering, vol. 60, no. 7, pp. 1946-1953, July 2013, doi: 10.1109/TBME.2013.2246160.

[10] H. S. Koppula and A. Saxena, "Anticipating Human Activities Using Object Affordances for Reactive Robotic Response," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 1, pp. 14-29, 1 Jan. 2016, doi: 10.1109/TPAMI.2015.2430335.

[11] Song-Mi Lee, Sang-Min Yoon and Heeryon Cho, "Human activity recognition from accelerometer data using Convolutional Neural Network," 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, Korea (South), 2017, pp. 131-134, doi: 10.1109/BIGCOMP.2017.7881728.

[12] M. Babiker, O. O. Khalifa, K. K. Htike, A. Hassan and M. Zaharadeen, “Automated daily human activity recognition for video surveillance using neural network,” 2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), Putrajaya, Malaysia, 2017, pp. 1-5, doi: 10.1109/ICSIMA.2017.8312024.

[13] C. Liu, J. Ying, F. Han and M. Ruan, "Abnormal Human Activity Recognition using Bayes Classifier and Convolutional Neural Network," 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), Shenzhen, China, 2018, pp. 33-37, doi: 10.1109/SIPROCESS.2018.8600483.

[14] M., B., 2021. *Video Chat Algorithms Can Now Detect Your Heart Rate and Stress Levels*. [online] Dogtown Media. Available at: <https://www.dogtownmedia.com/video-chat-algorithms-can-now-detect-your-heart-rate-and-stress-levels/> [Accessed 14 May 2021].

[15] van Gastel, M., Stuijk, S. and de Haan, G., 2016. Robust respiration detection from remote photoplethysmography. *Biomedical Optics Express*, 7(12), p.4941.