HAND GESTURE MAGIC

Capstone Project Report

Final Evaluation

Submitted by:

(401603009)	Deepanshu Udhwani
(401603010)	Devak Ohri
(401603019)	Mansi
(401633001)	Kali Nigam

BE Third Year, COEM

CPGM No: ___3___

Under the Mentorship of

Dr. Manju Khurana,

Assistant Professor

Computer Science and Engineering Department,

TIET, Patiala

December, 2019

ABSTRACT

The project is introducing an application using Hand Gesture Recognition computer vision. A camera records a live stream of video from which an interface is used to take a snapshot. For each type of count hand gestures, the system is trained at least once. After that, it gives the system a test gesture and it recognizes it.

Our idea is to develop an application that would translate hand gestures and vocal inputs to texts and prevent the differently-abled community from their daily life struggles, help them communicate without any hesitation and trouble.

The application is built upon the concepts of Hand Gesture Recognition in computer vision. All that the user will have to do is, perform the desired gestures in front of the camera and the camera will record the live stream of video from which an interface is used to take a snapshot. For each type of count hand gestures, the system is trained at least once. After that, it gives the system a test gesture and it recognizes it.

Various stages in which the project will move forward includes study about various concepts of Machine Learning, Data Analytics, Deep Learning followed by rigorous study about various machine learning models related to our project and their behavior, pre-processing of data including extraction of features required to build a machine learning model and applying various testing procedures to figure out the best or the most accurate model that can deliver the desired and correct output.

DECLARATION

We hereby declare that the design principles and working prototype model of the project entitled "Hand Gesture Magic" is an authentic record of our own work carried out in the Computer Science and Engineering Department, TIET, Patiala, under the guidance of Dr. Manju Khurana during 6th semester (2019).

Date:

Roll No.	Name	Signature
401603009	Deepanshu Udhwani	
401603010	Devak Ohri	
401603019	Mansi	
401633001	Kali Nigam	

Counter Signed By:

Faculty Mentor:

Dr. Manju Khurana,
Assistant Professor,
CSED,
TIET, Patiala

ACKNOWLEDGEMENT

We would like to express our thanks to our mentor Dr. Manju Khurana. She has been of great help in our venture, and an indispensable resource of technical knowledge. She is truly an amazing mentor to have.

We are also thankful to Dr. Maninder Singh Head, Computer Science and Engineering Department, entire faculty and staff of Computer Science and Engineering Department, and also our friends who devoted their valuable time and helped us in all possible ways towards successful completion of this project. We thank all those who have contributed either directly or indirectly towards this project.

Lastly, we would also like to thank our families for their unyielding love and encouragement. They always wanted the best for us and we admire their determination and sacrifice.

Date:

Roll No.	Name	Signature
401603009	Deepanshu Udhwani	
401603010	Devak Ohri	
401603019	Mansi	
401633001	Kali Nigam	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 - INTRODUCTION

1.1. PROJECT OVERVIEW

Communication brings people together, closer to each other. It bridges the gap between individuals and groups through flow of information and understanding between them. It is the most basic requirement for a healthy life. But this is basic requirement is not too easy to fulfill, especially if you are a differently-abled person. Expressing even a single thought becomes a tough challenge for those who are deaf and mute. This is mainly because how they communicate or rather try to communicate is by using sign language, which is generally not understood by the majority of the common audience. Now to be able to express themselves, they either take the help of a human interpreter or carry pen-paper with them wherever they go. This, however, causes them to feel dependent and lowers their confidence.

This project is our way to help them regain their confidence. Our application requires high computing power and hence can be installed in schools and particular areas where they can easily access. It can be provided to them as a handy, pocket-friendly interpreter.

1.1.1. Technical terminology

A significant challenge with data-driven algorithmic composition is: what data to use? What parts or features to be used? This is where the Machine Learning and Data Analytics approach comes into being. Data Analytics is the science of examining raw data with the purpose of finding patterns. Whereas, Machine Learning focuses on the development of computer programs that can access data and use it to learn for themselves.

1.1.2. Problem Statement

For a population of about 1.3 million deaf and mute people, there are only 250 trained interpreters available in our country. This poor proportion causes the majority of these people to face all sorts of communication-related problems. Talking to people, getting an education, visiting a doctor, everything's a struggle for them. Our app would bridge the day to day hurdles

faced by our audience while communicating with each other. It would be a stepping stone for the people related to this field as there are no other major sources available to them.

1.1.3. Goal

Our goal is to provide the differently-abled, a tool with which they'll be able to communicate with everyone all by themselves. The growth in education sector when considered for deaf and mute is very low, our app aims to increase the education rate. The doctors face many problems while interacting with these people so our app will be useful for them as well

1.1.4. Solution

- The solution according to our project involves numerous steps:
- To study the existing models, related projects and analyze them to deduce the feasibility and working of our own project.
- To gather and preprocess the input dataset and apply various data mining operations them to extract useful data.
- To optimize the chosen model to generate output with the desired accuracy.
- To make the outcome available over an android application.

1.2. NEED OF THE PRODUCT

Most of the differently-abled people face a lot of problems in communicating with others. It is difficult for them to express what they feel and also to show what they have understood.

With our app, it would not only help these people but also for those who haven't bought mobile phones yet (we observed it doing our surveys) as for them, hearing and speaking disability have left them with no other reason to buy a phone so the electronics segment of the market would

experience a certain rise too as now we are providing them with an easier and a comfortable platform to communicate through.

Our product will definitely be a source of aid for the differently-abled people as they will be provided with a new medium for communication.

There is a big need to rise in the education sector in our country as well, our app will help in providing aid to the cause and can be used as a mode of education for the customer segment we have our focus on. It is observed that the doctors and government running organizations have many difficulties too when it comes on the ground level in dealing with the deaf & mute, so for them as well. Our one-click away app will help them too.

These are basically where we aim to tap on. Other people willing to learn and communicate with their dear ones via our application are also most welcome.

1.3. RESEARCH GAPS

In terms of the research approach, we see at least two recurring questions. First, what kind of representations should we use for each gesture as including each and every possible with reference to American sign language is difficult to achieve.

Second, how can we achieve accuracy in displaying the corresponding text with minimal possible latency?

1.4. PROBLEM DEFINITION AND SCOPE

Problem:

Differently-abled people, particularly the ones that are deaf or/and mute, face a number of difficulties in their daily lives especially while communicating with a normal person.

Trying to express what they feel or what they want is a never-ending and exhausting process. Not everyone understands sign language so either they have to carry a pen-paper or tag along with an interpreter wherever they go to express themselves.

Not just communication, getting a decent education is a difficult task because not many schools or colleges offer classrooms with interpreters or teachers capable of using sign-languages.

It's also strenuous for doctors to deal with patients that are deaf or/and mute.

Scope:

India has approximately only 250 qualified Indian Sign Language (ISL) interpreters for a population of over one crore deaf and mute people! This is a large community and there is a need for quality interpreters.

The project has a very wide future scope of further development and expansion. This disproportionality can be successfully eliminated by our app. Our app would prove to be a stepping stone and benefit our deaf and mute people in many ways!

1.5. ASSUMPTIONS AND CONSTRAINTS

S NO.	ASSUMPTIONS
1	We assume that the system used is equipped with a high-quality camera that detects the hand gesture easily and as accurately as possible.
2	One assumption about the product is that it will always be used on computers with high computing power and that has enough performance. If the system does not have enough hardware resources available for the application, for example, the users might have allocated them with other applications, there may be scenarios where the application does not work as intended or even at all.

Table 1.1 – Assumptions

S NO.	CONSTRAINTS
1	It is very important for the hand gesture to be detected precisely and understood by the system.
2	It necessary for the hand gesture made to be present in the existing database as well, then only the conversion to corresponding text can take place.

Table 1.2- Constraints

1.6. APPROVED OBJECTIVES

The objectives of our project are described below:

i. To study the existing models, related projects and analyze them to deduce the feasibility and working of our own project.

ii. To gather and preprocess the input gestures and apply various data mining operations them to extract useful data.

iii. To predict the gesture of the input action and then generate text of the same g using our Machine Learning model.

iv. To optimize the chosen model to generate texts corresponding to the gestures with the desired accuracy.

1.7 METHODOLOGY

• The user opens the application and provides a video feed.

• Frames are taken into consideration.

• Relevant features are extracted and classified.

• The intended gesture is recognized and converted into text.

• The text is then displayed on the screen.

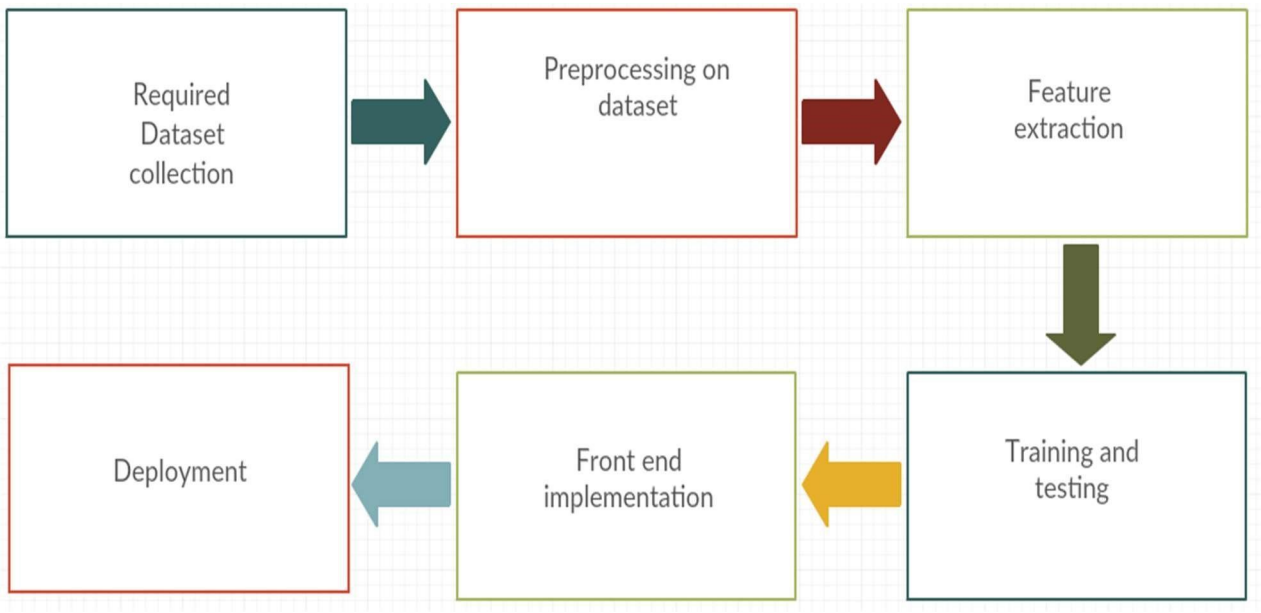


Figure 1.1: Methodology

1.8 PROJECT OUTCOMES & DELIVERABLES

This project aims at successfully interpreting the desired or projected hand gesture into its letter equivalent, thus eliminating the need for anyone to learn sign language and also reducing the chance of miscommunications.

The system can be conveniently used by differently-abled, doctors and in educational institutions dealing with differently-abled students.

Our goal is to create the best in class services along with the pocket-friendly application on the PC itself. The project is aimed at creating a medium for communication between people who are both deaf and mute. This app would bring rescue to doctors as well and also help in increasing the education rate of our country.

1.9 NOVELTY OF WORK

The proposed solution is quite unique as none of our competitors provide services with the ease that we aim to provide them with. Bringing an application easily available on the play store which is also open to all shouts for uniqueness and an increase in development itself is our key strategy.

Currently, in the market, there is no other app that provides the same solutions for the Indian masses. They still rely on paper-based or human-based solutions. Our application will provide them a sense of independence and confidence.

Our application is feasible, accessible, and easy to understand and use by any and every person.

CHAPTER 2 - REQUIREMENT ANALYSIS

2.1. LITERATURE SURVEY

2.1.1. Theory Associated with the Problem Area

As described above, the detection and recognition of gestures are related to the location of a hand in a still image or image sequence, i.e. moving images. The tracking of the hand in the scene that is relevant to the application such as sign language can be followed in the case of moving images. The underlying concept of hand detection is that human eyes can detect objects which machines cannot with that much accuracy as that of human.

We use a variety of factors to determine hand detection including variation in the image plane and pose, skin tone, and other structural components, lighting, and background. The model will then be trained to recognize gestures that are imputed to it. The gestures are then looked up in the existing database and the corresponding text is then conveyed on the screen. The tracking of the hand in the scene that is relevant to the application such as sign language can be followed in the case of moving images.

This project's scope is to build a classification system for real-time gestures that can automatically detect gestures in natural lighting conditions. A real-time gesture-based system is developed to identify gestures to accomplish this objective.

This system will function as one of Artificial Intelligence's futuristic and user interface computer vision. Its method of creating hand gesture recognition is based on various parameters. The main priority of the system, however, is to fulfill requirements of being simple, easy and user-friendly without using any specialized hardware.

2.1.2. Existing Systems & Solutions

Research has been limited to small scale systems able to recognizing a minimal subset of a full sign language. Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz[1]. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and PowerGlove[1]. A popular approach for Hand Gesture Recognition is of Hidden Markov Models, which have also been used for speech recognition [2].

Chan Wah Ng, Surendra Ranganath presented a hand gesture recognition system, they used image furrier descriptor as their prime feature and classified with the help of the RBF network[3].

Their system's overall performance was 90.9%. Claudia Nölker and Helge Ritter presented a hand gesture recognition modal based on recognition of fingertips, in their approach, they find full identification of all finger joint angles and based on that a 3D modal of hand is prepared and using a neural network.

Application running on the Android Platform but only in the Brazilian language is also available on the Android Play Store. We aim to introduce our app for English & further languages would be introduced later.

2.1.3. Research Findings for Existing Literature

S. No.	Project Title	Tools/ Technology	Learnings
1	A Review on Indian Sign Language Recognition[4]	ASLR	Learnt about Neural Networks, SVM, CNN
2	Vision-Based Sign Language Translation Device	Real Time Vision Based System	Image Thresholding
3	Hand Gesture Recognition for Human-Machine Interaction	Image Comparison Techniques	Hausdorf Distance

Table 2.1- Research Findings

2.1.4. The Problem That Has Been Identified

As discussed above, the problems faced by the people or the customer segment are limited but there are no optimal solutions and better options available right now.

It has been observed that the differently-abled people face difficulties and obstacles while communicating with others so our application bridges the gap between the two. The growth in education sector when considered for deaf and mute is very low, our app would prove to be a stepping stone in increasing the education rate. The doctors face many problems while interacting with these people so our app will be useful for them as well.

2.1.5. Survey of Tools and Technologies Used

The following methods or technologies were observed during the literature survey:

i. Data Analytics - Data Analytics is the science of examining raw data with the purpose of finding patterns and drawing conclusions about that information by applying an algorithmic or mechanical process to derive insights. Since our data is in the form of hand gestures, it is the most crucial step to convert these files to machine-understandable form such that finding similarities and patterning the data becomes easy.

ii. Machine Learning - Machine Learning focuses on the development of computer programs that can access data and use it to learn for themselves. Once the data is extracted from raw files, it is fed into the best-suited machine learning model giving the desired accuracy. Devising methods of testing and increasing our model accuracy also remains an objective of this project. Tools Used for the same include "JupyterNotebook" and we may even use Google's Cloud Computing services during model building and training.

iii. Deep Learning - Deep

iv. Learning is a part of broader family of Machine Learning methods based on learning data representations and can be supervised, semi-supervised or unsupervised. Deep learning models are used in Neural Networks and have been applied to the fields including Natural language processing, Audio recognition, speech recognition, computer vision, etc.

2.2 STANDARDS

IEEE 802.11 - It is a set of media access control (MAC) and physical layer (PHY) specifications for implementing wireless local area network (WLAN) computer communication, They are the world's most widely used wireless computer networking standards, used in most home and office networks to allow laptops, printers, and smartphones to talk to each other and access the Internet without connecting wires.

IEEE 610.4-1990 IEEE 610.4-1990-It is a specification IEEE Standard for Pattern Recognition Terminology which was established by IEEE.

2.3. SYSTEM REQUIREMENT SOFTWARE

2.3.1. INTRODUCTION

2.3.1.1. Purpose

This document is intended to provide a detailed overview of our software product, its parameters and its objectives. This document specifies in detail various functional and non-functional requirements and also specifies which software feature satisfies these requirements. It also describes different constraints and standards that apply to this domain's software. It includes in the development of this software description of all software / hardware and third-party dependencies.

2.3.1.2. Intended Audience and Reading Suggestions Majority of our users and customers can be classifies as:-

- Differently-abled people.
- MHRD programs for welfare of deaf and mute.
- Schools for Deaf and Mute.
- Employers dealing with Deaf and Mute Employees.

- Hospitals and Clinics.

These are basically where we aim to tap on. Other people willing to learn and communicate with their dear ones via our application are also encouraged. Not declining the fact that still there is a big need to reach the masses and creating awareness about an all-time available app that would act as an interpreter for them, easily carried anywhere any time just in their pockets in their cell phones, would definitely help them to discover this as a great solution. On an initial note, it would be difficult to accept this but once our product is released in the market and people start to believe in us and in our vision, they won't be disappointed.

2.3.1.3. Scope

With the launch of our application, we'll be the trendsetters in this field. It will also help in creating opportunities for growth in the electronics sector and the literacy rate of our nation as well. With the belief in our vision, we hope that the market is affected in a positive manner only. The purpose of this document is to provide a detailed overview of our software product, its parameters and its objectives. Our app would bridge our audience's day-to-day hurdles while communicating with each other. For people related to this field, it would be a stepping stone as there are no other major sources available for them to develop and grow shoulder to shoulder with other human beings (with no disabilities) and it will also help to increase the literacy rate and the economy.

2.3.2. Overall Description

2.3.2.1. Product Perspective

A machine learning model will act as the generating core of the system which will translate the gesture to its corresponding text. The camera will be used to input the gesture made by the user. Using feature extraction algorithms implemented by CNN, the system will detect the gesture and look into its database for the corresponding text. The text will then be displayed on the screen.
The final product will be able to do the following:

1. Get gesture as input from the user.
2. Feature extraction and classification to detect the gesture.
3. Compare the gesture with the existing database.

4. Generate the corresponding output.
5. Input audio from the user and get the corresponding text.

2.3.2.2. Product Features

Some of the features of the product are as follows:

1. Real-time conversion of visual data into text.
2. Create a user-friendly platform for communication.
3. Real-time conversion of audio into text.

2.3.3. External Interface required

2.3.3.1. User Interfaces

PC Device

2.3.3.2. Hardware Interfaces

Hardware	Minimum requirements
Any Computing Device	2 GHz minimum, multi-core processor
Memory (RAM)	At least 4GB, preferably higher
disk space	At least 1 GB
Camera	2 MegaPixel or Higher

Table 2.2- Hardware Requirements

2.3.3.3 Software Interfaces

Software	Minimum Requirements
Operating System	Windows
IIS(Internet Information Services)	Version 8 or Higher
Tensor Flow Framework(v 4.6.1 or Higher)	Mobiles And Web or Higher
OpenCV(Open Computer Vision)[5]	Version 2.10 or Higher
Google Play Services	Google Play services 14.1.94 or Higher

Table 2.3- Software Requirements

2.3.4 Other Non-Functional Requirements

2.3.4.1 Performance Requirements

Model selection and optimization is crucial since it will directly affect the response time. Better response time and accuracy detection will improve the efficiency of the product.

2.3.4.2 Safety Requirements

The user interface should be safe, secure, friendly and easy to use so that people not having much knowledge about sign language and technology are also able to use the product easily.

2.3.4.3 Security Requirements

Payments in case of downloading options should be made securely.

2.4 COST ANALYSIS

(i) The first novelty of the proposed method is that the hand gesture recognition is based on the result of finger recognition. Therefore, the recognition is accomplished by a simple and efficient rule classifier instead of the sophisticated but complicated classifiers such as SVM and CRF.

(ii) Some previous works need users to wear a data glove to acquire hand gesture data. However, the special sensors of data glove are expensive and hinder its wide application in real life. In the work, the authors use the TOF camera, that is, Kinect sensor, to capture the depth of the environment and a special tape worn across the wrist to detect hand region. Our approach only uses a normal camera to capture the visual information of the hand gesture meanwhile does not need the help of the special tape to detect hand regions.

(iii) The third advantage of the proposed method is that it is highly efficient and fit for real-time applications.

(iv) Hand Gesture Recognition System in form of a software that is compatible either on firstly Windows/Android and then followed by iOs.

2.5. RISK ANALYSIS

There are a few risk factors included in the building and completion of this project like: the device will use a large amount of processing power and a will need some noticeable amount of time to respond. Another risk factor we have considered is that the model might produce some unrequired text in some cases until it attains the desired accuracy. The fact that we will be needing a high computational system for model building and training purposes is another risk factor.

CHAPTER 3 METHODOLOGY ADOPTED

3.1. INVESTIGATION TECHNIQUES

Our project is based on Experimental investigation since we are trying to design a model that will take input a hand gesture and will detect it and match it in the existing database. Further if matched then the corresponding text will be displayed on the screen. Initially, the android studio platform is used in the formation of the application further than we use various Machine Learning algorithms and the LSTM Deep Learning model to test and train our model.

3.2. PROPOSED SOLUTION

An application that will bring aid to the cause is the appropriate solution. Explaining how the application would work.

A new user will first sign up through an email-id and set a password. After the initial sign-up, the user can log in to their accounts.

He/She can then allow permission to access the camera and microphone, so as allowing the application to access them. Now when the user clicks on the Generate gesture, the camera will be opened and then the action made would be detected.

Gesture captured will be processed and matched from the existing database, after the processing only the text corresponding to the gesture will be displayed.

If the gesture made doesn't exist in the database, the user will be asked to try again.

3.3. WORK BREAKDOWN STRUCTURE

Since we believe in teamwork, every task is performed by the group. So segregating each person's individual role is quite difficult. Still, the main front end development of the application & training of the model is managed by Mansi & Kali. They are also working hard in doing literature surveys and research work required in the field. The back end development of the app including testing the model made, surveys, contacting people related to this field & researchers willing to provide us with helpful information, these tasks are seen by Devak & Deepanshu.

3.4. TOOLS AND TECHNOLOGIES USED

- Data Analytics - Data Analytics is the science of examining raw data with the purpose of finding patterns and drawing conclusions about that information by applying an algorithmic or mechanical process to derive insights. Tools Used for the same include "Jupyter-Notebook", "Pycharm" and various Python libraries like Libros, pyAudioAnalysis, etc.
- Keras as the open-source neural network library written in python. It was developed as part of the research effort of project ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System)Google Camera is a camera application developed by Google for Android.
- The recognition is accomplished by a simple and efficient rule classifier instead of the sophisticated but complicated classifiers such as SVM and CRF.
- Amazon Web Services (AWS) is a subsidiary of Amazon that provides on-demand cloud computing platforms to individuals, companies, and governments, on a metered pay-as-you-go basis. In aggregate, these cloud computing web services provide a set of primitive, abstract technical infrastructure and distributed computing building blocks and tools. One of these services is Amazon Elastic Compute Cloud, which allows users to have at their disposal a virtual cluster of computers, available all the time, through the Internet.

- Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object and tools for working with these arrays. It is the fundamental package for scientific computing with Python.

  Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data.

CHAPTER 4 - DESIGN SPECIFICATIONS

4.1. SYSTEM ARCHITECTURE

4.1.1. ARCHITECTURE

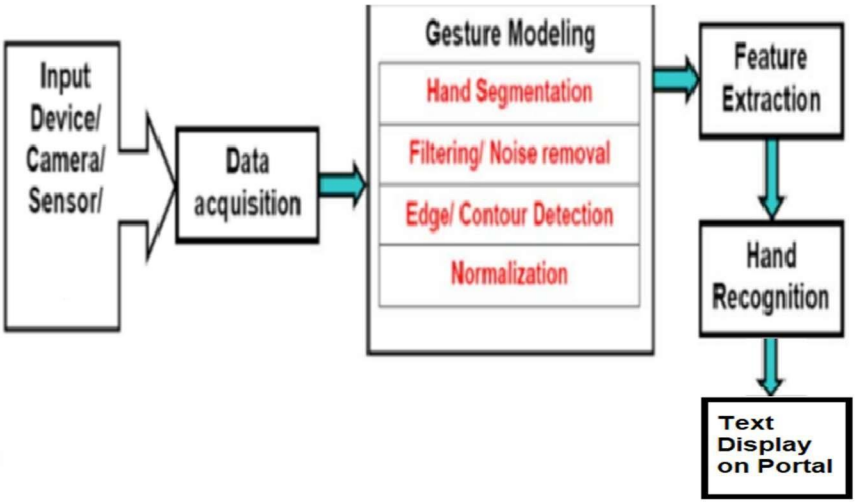


Figure 4.1- Architecture

4.1.2. TIER ARCHITECTURE

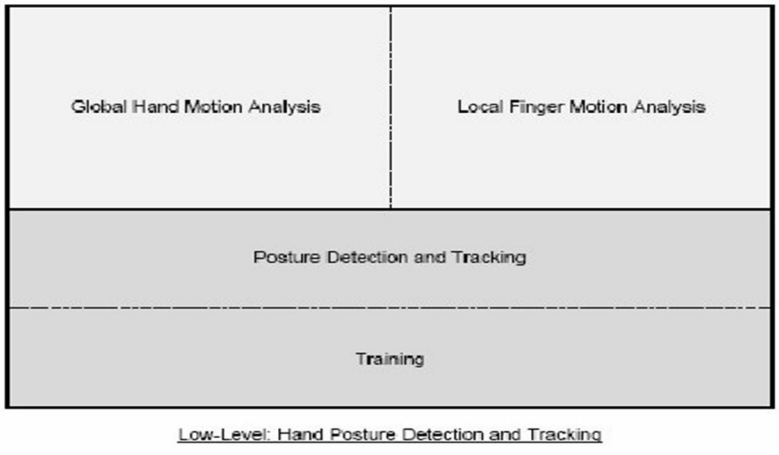


Figure 4.2- Three tier Architecture

4.1.3. MVC ARCHITECTURE

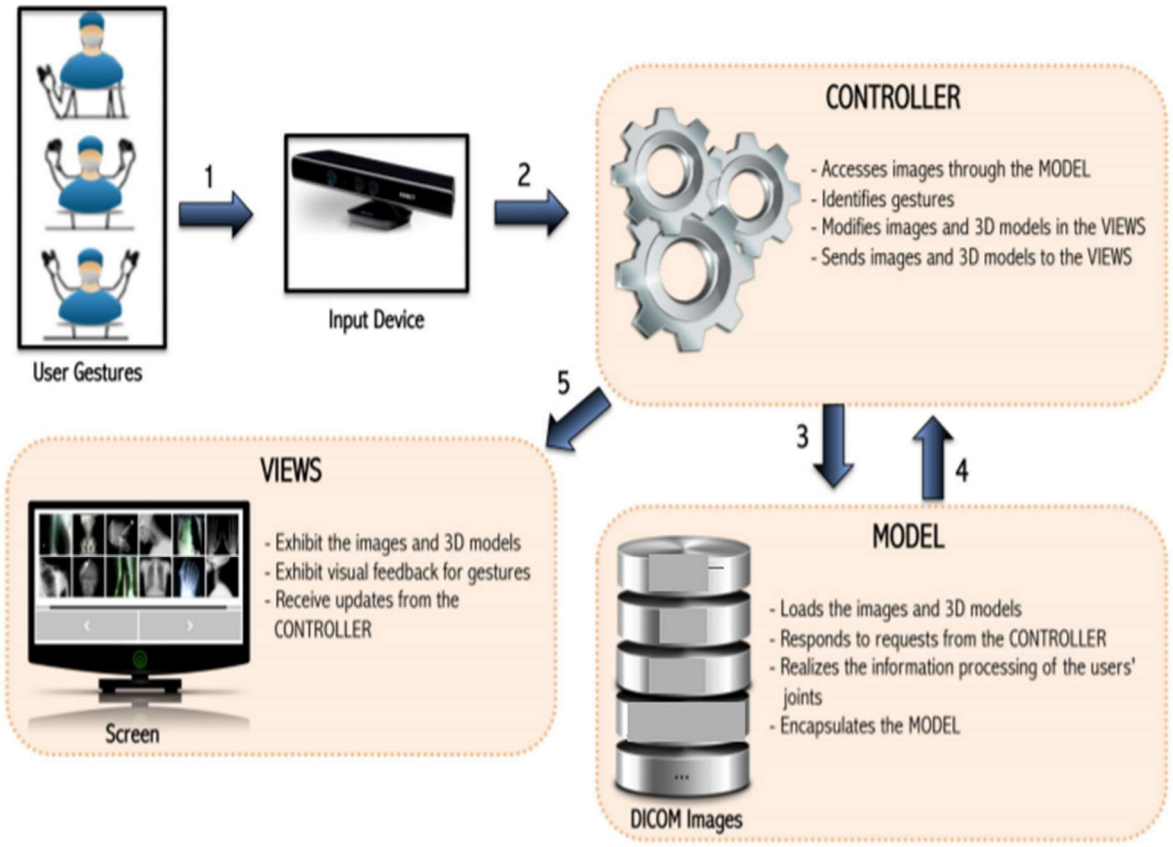


Figure 4.3- MVC Architecture

4.2. DESIGN LEVEL DIAGRAMS

4.2.1. E-R DIAGRAM

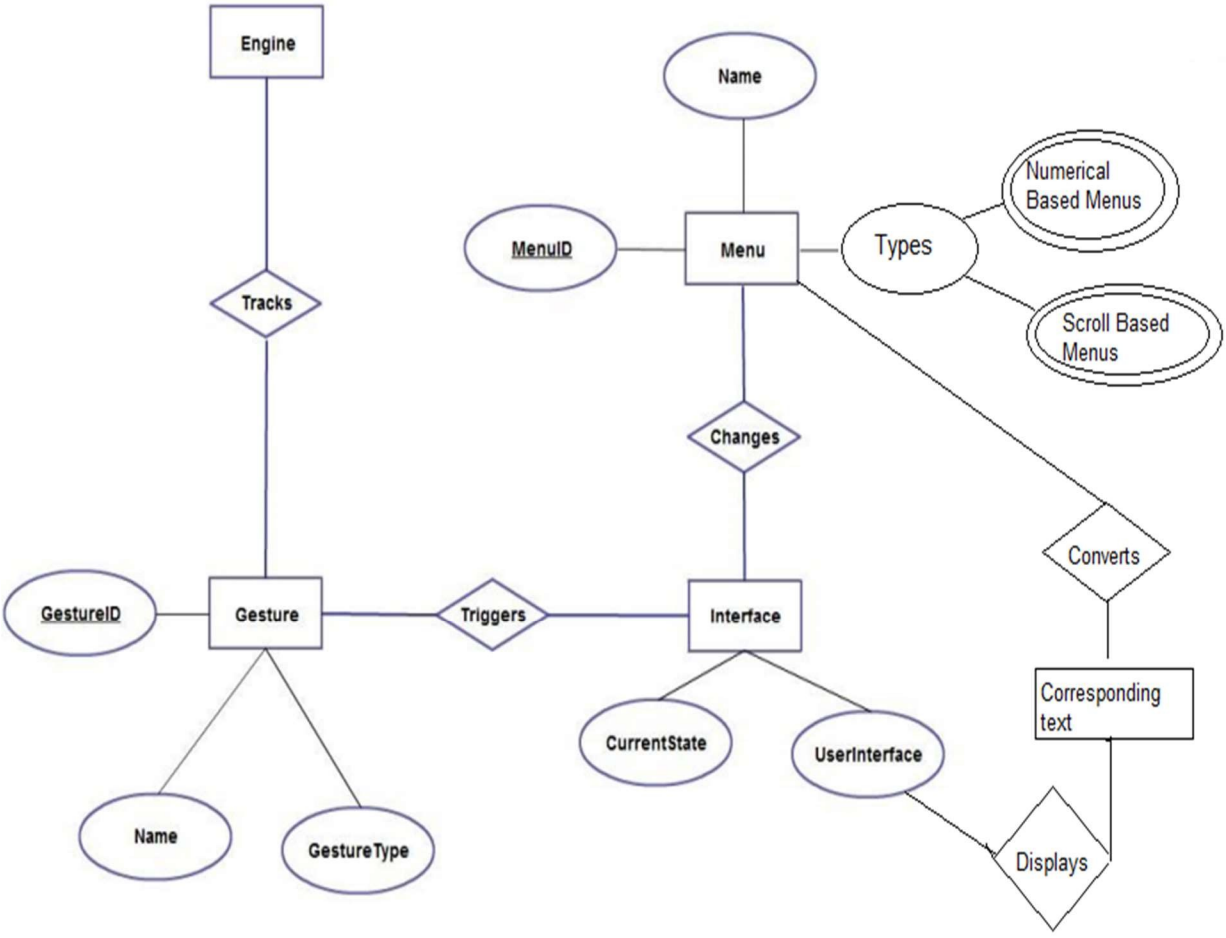
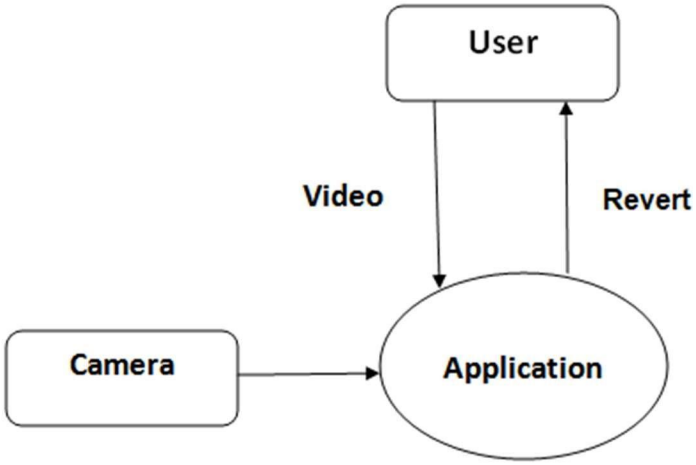
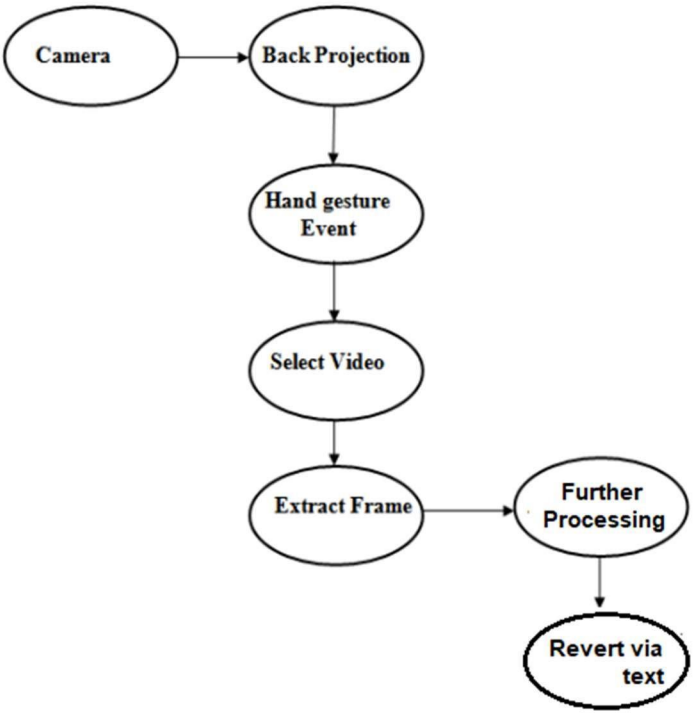


Figure 4.4 – E-R Diagram

4.2.2. DATA FLOW DIAGRAM



Level 0



Level 1

Figure 4.5 – Data Flow Diagram

4.2.3. CLASS DIAGRAM

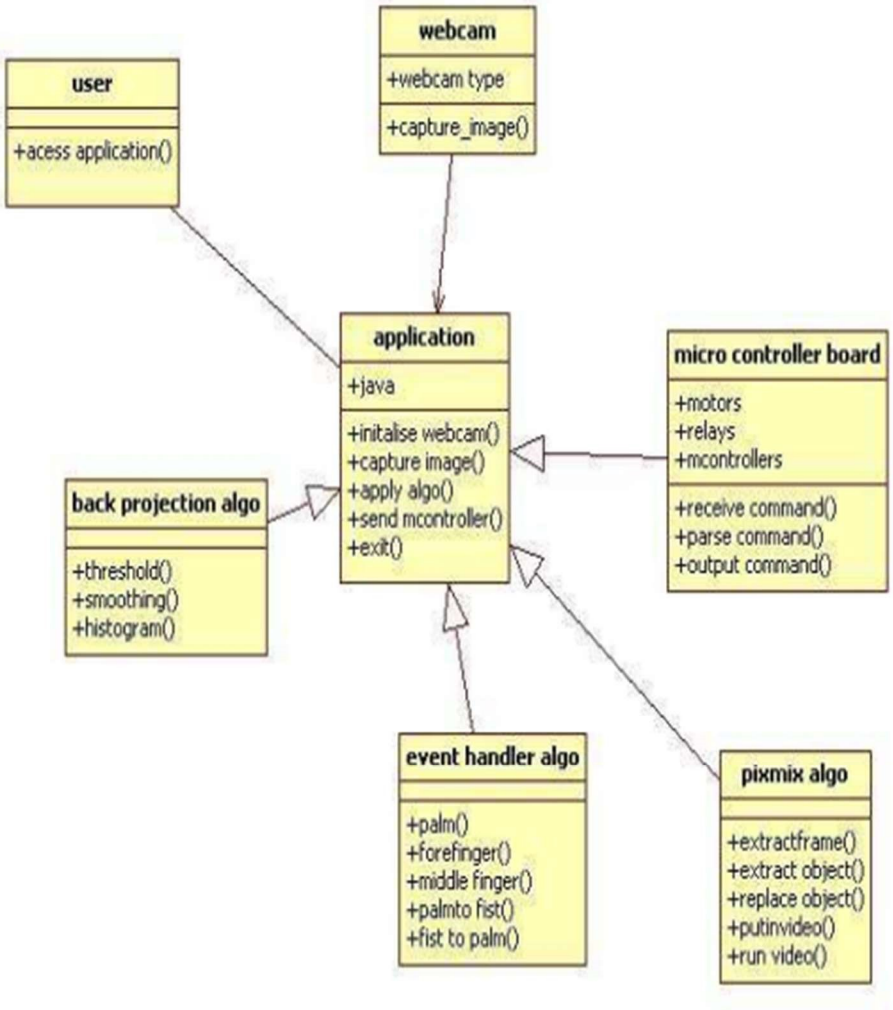


Figure 4.6 – Class Diagram

4.3. USER INTERFACE DIAGRAM

4.3.1. ACTIVITY DIAGRAM

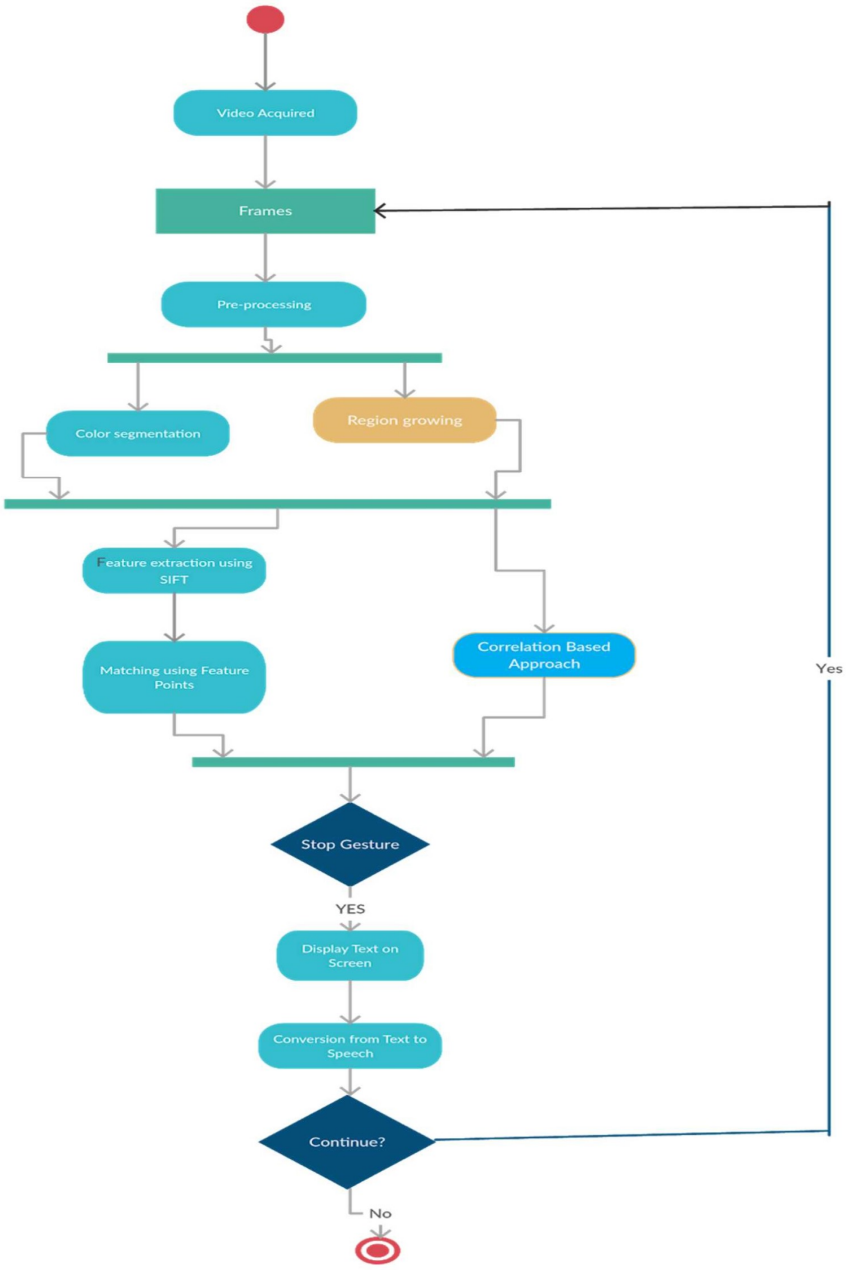


Figure 4.7- Activity Diagram

CHAPTER 5- IMPLEMENTATION AND EXPERIMENTAL SETUP

5.1 Experimental Setup

We found out about the various sign languages and the datasets we could derive from them. We searched and analyzed various data preprocessing techniques and training of machine learning models to detect the gestures and output of the corresponding text. We searched for various datasets that can be used for model training and construction. Then we researched the data preprocessing techniques and various Machine Learning models that could be implemented in our project. The project has been depicted via data flow diagrams, use cases, activity diagrams, system architecture, and a few other data representation techniques earlier.

5.2 Experimental Analysis

5.2.1 Data

Initially, we faced difficulties in gathering a huge database with respect to the language we selected as our benchmark i.e. American Sign Language but somehow we managed to gather the one according to the selected language. The transition from single pictures to segmenting videos into multiple video frames was quick. The segmented video frames were taken under consideration for preprocessing and image classification techniques.

We applied a slider bar that would adjust the HSV values and segment the hand color. In the last, we came up with something even better and finally with the idea that gestures already added as per the ASL could also be combined with an option available to the user that he/she could add gestures as per his/her own suitability.

Now with the approach of the user to add images as per his suitability, we felt the need to increase the database of the same by flipping the images. An instant folder would be created having 1200 pictures of 50*50 pixels each and then adding the flipped ones also, it would result in 2400 sample images available for a single gesture.

5.2.2 Performance Parameter

(Accuracy & QoS parameters)

We used Keras which runs on the top of Tensor Flow using python as our writing language, after initial training in Phase 1, we observed that elimination of noise has to be implemented and there's a big room for increasing accuracy. The letters were getting displayed by the end of our major time span but after adding support for the user to add gesture and name of gesture as per his wish, accuracy and lagging became our major concerns.

We adopted different algorithms in the next phases, in which some worked and some just displayed errors, finally, the approach of increasing the database of a single gesture by flipping images in the existing database we had, also taking the background noise elimination into consideration, we managed to achieve approximately 90 percent of accuracy.

5.3 Working of the project

5.3.1 Procedural Workflow

When we talk about the procedural workflow, the initial stepping stone is creating a gesture. We will first set our hand histogram. A histogram is a graphical display of data using bars and different heights. In a histogram each bar group numbers into range. Initially, after giving the command for setting the hand histogram, the camera will open up and 50 small squares(5x10) would appear on the screen.

The hand has to be put on these boxes covering them up to the full. Sidewise we can see that the skin of the hand is first segmented and on the thresholded image, it is clearly visible that the only the area with the presence of the user's hand (skin color segmented) is shown in white patches and the rest is black. With the continuous collection of the gesture image, it captures 1200 frames for the same in minimal time. As soon as we get an optimal display for the hand segmented now we would save it in our database.

The new gesture created is stored in the database, and then using computer vision techniques, we flip and sharp them, creating another set of 1200 copies for the same. So in total, we have 2400 frames available for training and testing our model.

Since we are ready with our database we need to set up the environment for testing our model. We installed Keras and imported libraries and modules. We then loaded the image dataset from MNIST. Preprocessing input data for Keras did take a while. We then defined the model architecture and compiled our model.

The Fitness of the model was checked on the training data. We then retrain the model every time we add or remove any gesture to our database. The latter part comes to evaluate the model on test data. In testing, we get the confusion matrix, f scores, precision and recall for the predictions by the model.

The predictions made by the model are displayed as text on the screen as per the user makes any gesture using the camera. We also run python library pickle to serialize the objects and make sure that our model doesn't miss any frames rendered or made by the user i.e performs object serialization. The windows subsystem's libraries were used to load text to voice modules and add it as a silver lining for only dumb people to use. We also kept high accuracy and minimal lag possibility under salient features while training and testing of our model. That's the holistic approach of the processing workflow of our model in real-time.

5.3.2 Algorithmic Approaches Used

We used SQLite as the system software for managing our database. SQLite is a relational database management system which is not a client server-based engine, rather it is embedded into the end program and henced helped us in storing our database gathered on the system. We used OpenCV (Open source computer vision) for computer vision related applications. It is a library of programming functions mainly aimed at real-time computer vision. In simple language, it is a library used for Image Processing techniques.
We used NumPy for calculations related to matrices made and multidimensional arrays. It is a library for the Python programming language, adding support for large matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

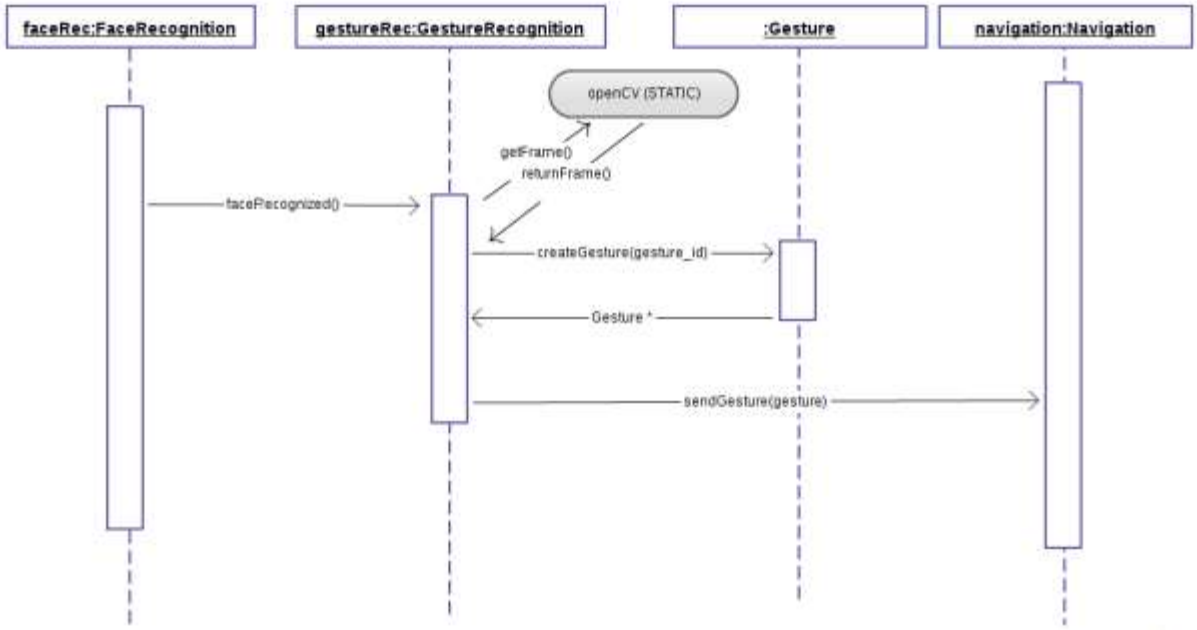
TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. We used various machine learning platforms and

algorithms in which we preferred Keras as the open-source neural network library written in python.

We chose Python as the language to work upon the existing software used. Keras, in turn, was a very great choice as it could easily run on top of various other platforms.

It helped us in fast experimentation and testing of deep neural networks. It definitely is user-friendly and focuses on being modular and easily extensible.

5.3.3 Project Deployment



5.3.4 System Screenshots

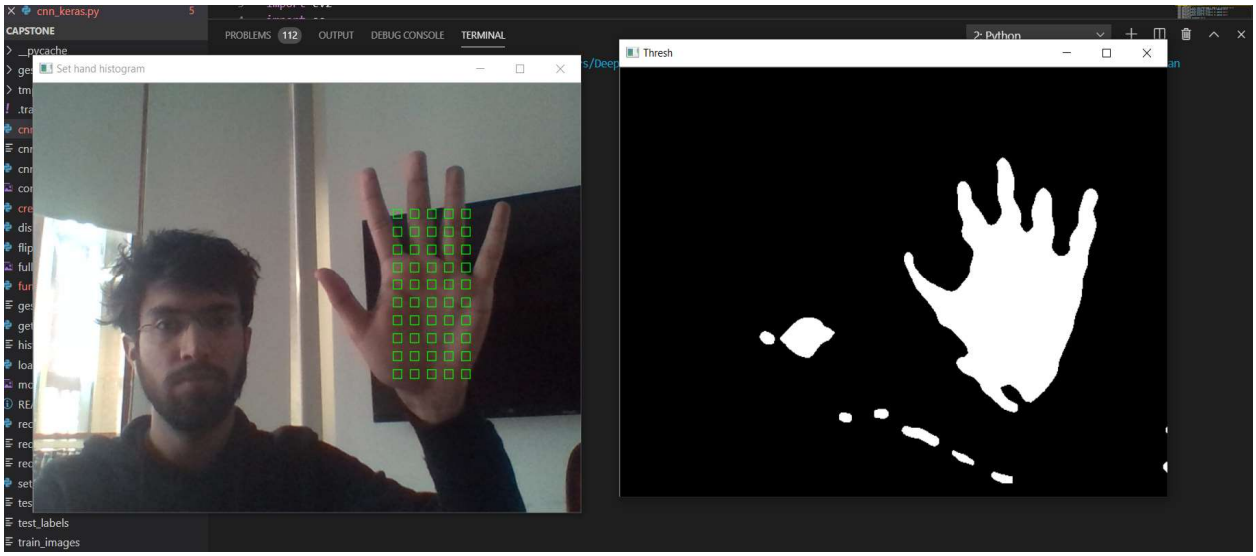
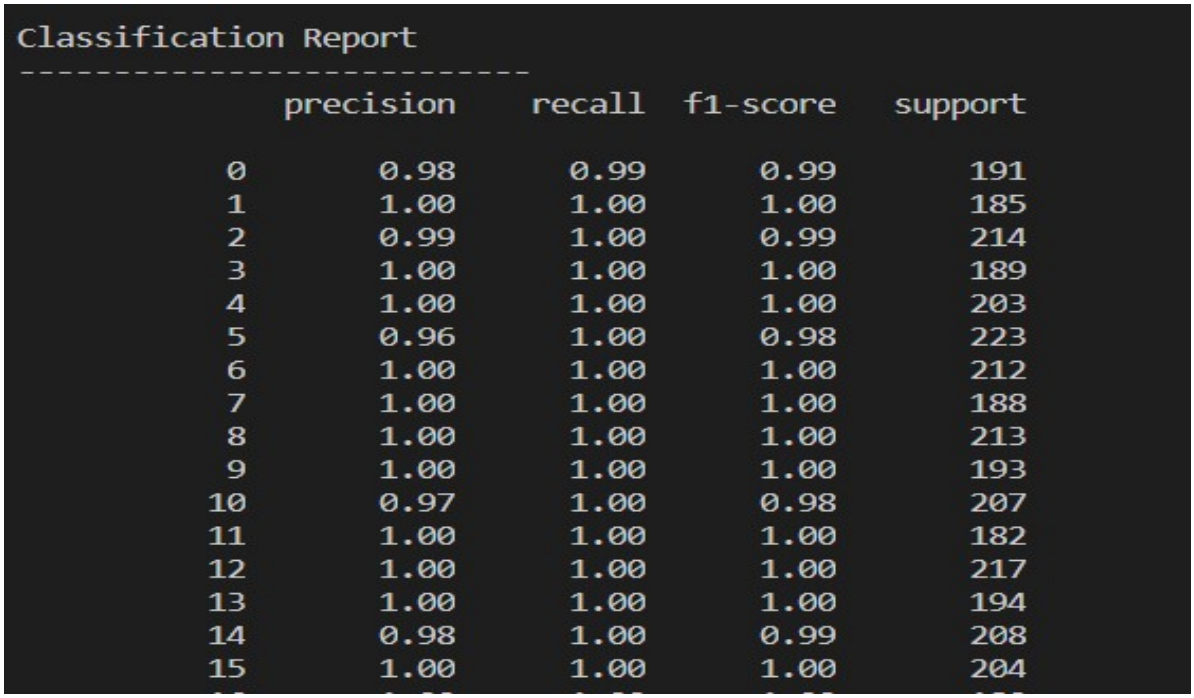


Fig 5.1. Hand Histogram Capture

accuracy			0.98	9000
macro avg	0.96	0.98	0.97	9000
weighted avg	0.96	0.98	0.97	9000

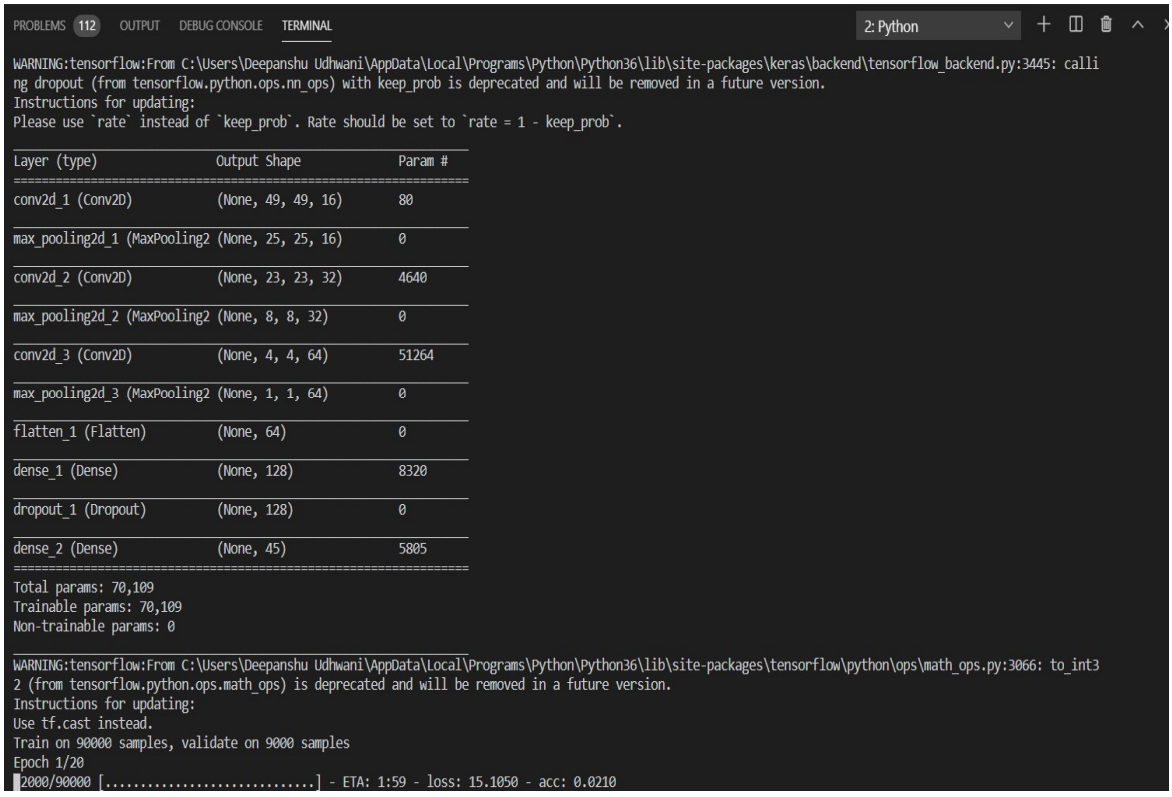
Fig 5.2. Accuracy Matrix



Classification Report

	precision	recall	f1-score	support
0	0.98	0.99	0.99	191
1	1.00	1.00	1.00	185
2	0.99	1.00	0.99	214
3	1.00	1.00	1.00	189
4	1.00	1.00	1.00	203
5	0.96	1.00	0.98	223
6	1.00	1.00	1.00	212
7	1.00	1.00	1.00	188
8	1.00	1.00	1.00	213
9	1.00	1.00	1.00	193
10	0.97	1.00	0.98	207
11	1.00	1.00	1.00	182
12	1.00	1.00	1.00	217
13	1.00	1.00	1.00	194
14	0.98	1.00	0.99	208
15	1.00	1.00	1.00	204

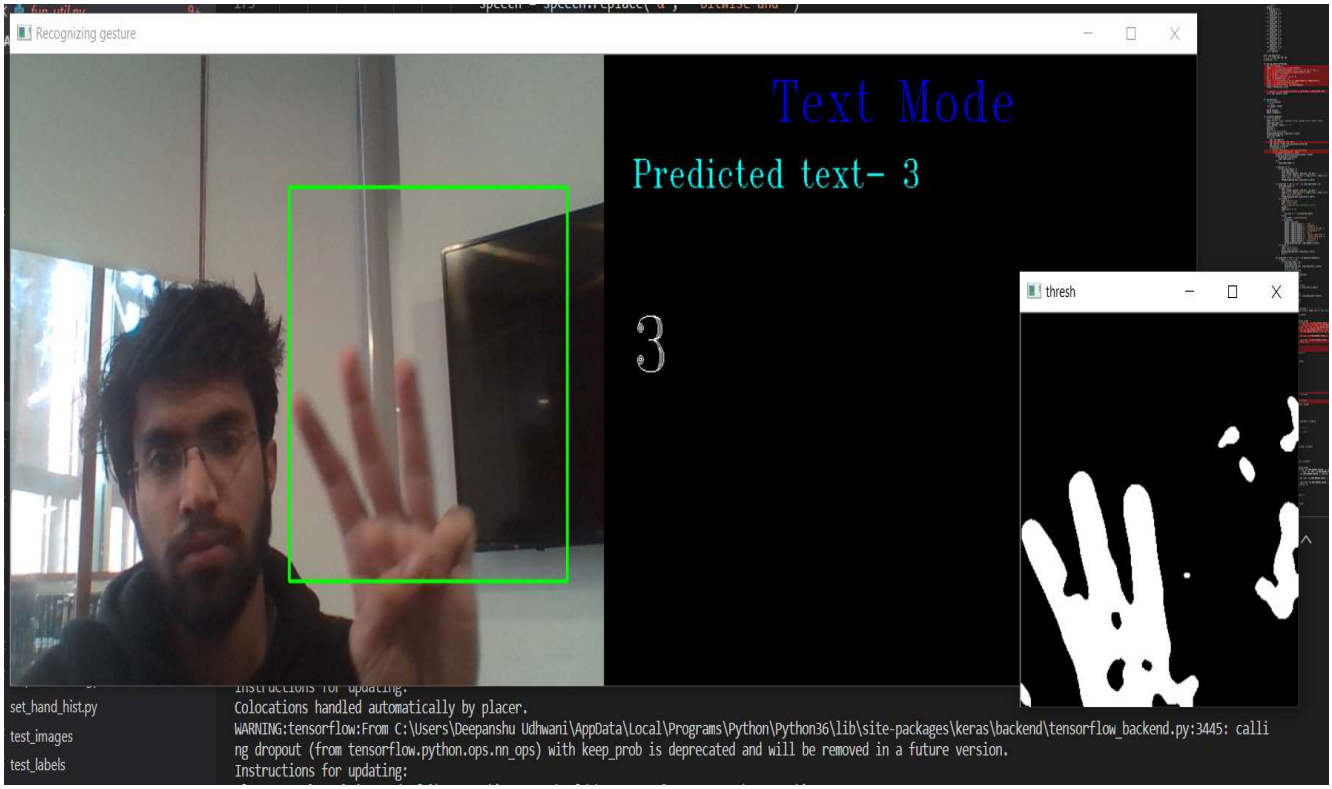
Fig 5.3.a. Model trained using Keras



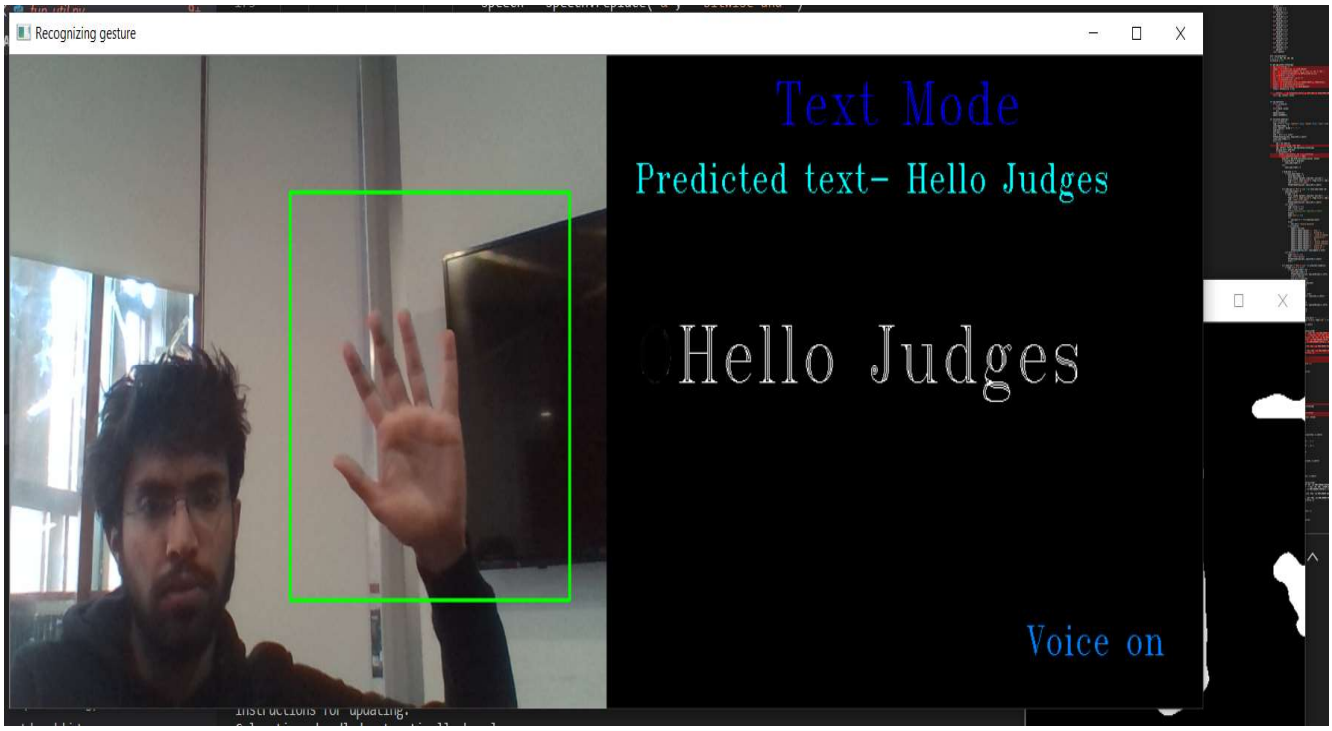
```
PROBLEMS 112  OUTPUT  DEBUG CONSOLE  TERMINAL                    2: Python

WARNING:tensorflow:From C:\Users\Deepanshu Udhwani\AppData\Local\Programs\Python\Python36\lib\site-packages\keras\backend\tensorflow_backend.py:3445: calli
ng dropout (from tensorflow.python.ops.nn_ops) with keep_prob is deprecated and will be removed in a future version.
Instructions for updating:
Please use `rate` instead of `keep_prob`. Rate should be set to `rate = 1 - keep_prob`.
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 49, 49, 16)        80
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 25, 25, 16)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 23, 23, 32)        4640
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 8, 8, 32)          0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 4, 4, 64)          51264
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 1, 1, 64)          0
_________________________________________________________________
flatten_1 (Flatten)          (None, 64)                0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               8320
_________________________________________________________________
dropout_1 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 45)                5805
=================================================================
Total params: 70,109
Trainable params: 70,109
Non-trainable params: 0
_________________________________________________________________
WARNING:tensorflow:From C:\Users\Deepanshu Udhwani\AppData\Local\Programs\Python\Python36\lib\site-packages\tensorflow\python\ops\math_ops.py:3066: to_int3
2 (from tensorflow.python.ops.math_ops) is deprecated and will be removed in a future version.
Instructions for updating:
Use tf.cast instead.
Train on 90000 samples, validate on 9000 samples
Epoch 1/20
 2000/90000 [..............................] - ETA: 1:59 - loss: 15.1050 - acc: 0.0210
```

Fig 5.3.b. Model trained using Keras



5.4.a. Sample text of gesture showing “3”



5.4.b. Sample text of gesture showing “Hello judges”

5.4 Testing Process

Testing generates the report for the tracking and execution of the entire testing of the project. The testing of software is essential to assure the quality of a product before it is released. Testing can be used to find the dead code of modules in project development.

Two types of software testing:

- Black Box Testing
- White Box Testing

5.4.1 Test Plan

Test Plan is a document describing software testing activities and scope. It is the basis for formally testing any software/product. The development of this project was done in versions that included the updating of some functionality in every version update as an agile model of software development was followed. Firstly, we analyzed the product about its target audience, use cases, how will it work and what exact functionalities will it be providing. Full documentation was prepared before starting the actual implementation of the project. Every objective to be delivered was worked upon by our team precisely.

5.4.1.1 Features to be tested

- Capture of required frame from camera
- Recognition of gesture from given frame
- Translation of given gesture to corresponding text.
- Pipelining of Data

5.4.1.2 Test Strategy

Test strategy gives the software testers a clear picture of the project at any instance. It deals with the review and approval of test documents prepared for the purpose of full proof testing. Every version after development has been checked for defects and all of the modifications have been done

simultaneously. It helped us to list all the estimated risks. All of the open-source software rules and licenses were followed in the development and taken care of.

5.4.1.3 Test Techniques

- Unit Test - Test the smallest piece of verifiable software in the application.
- API Testing - Test the API's created for the application.
- Integration Test - Individual software module combinations are tested as a group.
- System Test - Conducted on a complete, integrated system to evaluate the system's compliance with its specified requirements.
- Install/uninstall Testing - Focuses on what customers will need to do to install /uninstall and set up/remove the new software successfully

5.4.2 Test Cases

The test cases include firstly when we were trying to segment the hand only and do the skin segmentation for capturing frames and creating our own database. After the complete creation of the database and training of the model, the final test cases included the correct gesture recognition. the ability to extract correct information and match it with the database and then render it on the screen on runtime with minimal latency and also keeping the light changes under consideration.

5.4.3 Test Results

The letters corresponding to the gestures in the database are displayed on the screen as the output. There is also an audio module loaded using the windows subsystem itself, that would convert the text to audio. We can easily form words from letters using spaces. Spaces can be generated by taking the hand away from the area where the gesture has to be made and is expected to be detected by the model. The accuracy has always been our concern and it also varies with the illumination of the surrounding environment. These are the test results performed by us.

5.5 Results and Discussions

The letters corresponding to the gestures in the database are displayed on the screen as the output.

The user could add gesture and with having farthest approximations of every gesture stored in the database, texts corresponding to the gesture to be displayed became achievable.

The gesture has to be made within the green box area displayed on the screen as it is the segmented one and would be mapped with the ASL further. We have also used windows subsystem's libraries to convert text displayed on the screen to audio, so the text is also displayed and read aloud. Not only alphabets, but words could also be easily formed now.

The user has to take his/her hand outside from the segmented box area to generate a space and let the system understand that the previous gestures made are linked and have to be displayed together.

5.6 Inferences Drawn

- The real-time running project will be a source of communication for the deaf and mute.
- The system requires high computing power to run with minimum latency but is easily available too.
- Effective strategies and precautions to be taken under consideration for the achievement of high accuracy and illumination concerns
- With more development, this project could easily be deployed on the android platform.

5.7 Validation of Objectives

Our objective of displaying the gesture as per made by the user is fulfilled. The letters can be separated by spaces to form words and are available to the user to use them in sentences. The text is also converted into audio at the runtime with minimum possible latency and high accuracy. Hence all our objectives are achieved and validated.

CHAPTER 6 – CONCLUSION AND FUTURE DIRECTIONS

6.1 CONCLUSIONS

In the previous submission, we had very well concluded that the existing application fulfills these requirements in Brazilian and we focused on languages that we commonly speak. We conducted surveys to conclude the need for our model to be implemented and found that there is a big fat room for our project to gloom as a remarkable product in the outside world as well. We gathered the database, trained and tested our machine learning model well using multiple iterations and also applying different techniques. The corresponding letters to the gestures in the database are being displayed on the screen. The letters can also be converted into words and displayed on the screen. The text to audio feature is also added for convenience. We have for now achieved 80 percent accuracy as per after eliminating the noise from the background in a suitable environment.

6.2 ENVIRONMENTAL, ECONOMIC AND SOCIETAL BENEFITS

Our project will help in bringing the society even closer. Through the medium of our project, the differently-abled people can also communicate easily. These people are way underrated in terms of being productive just because they aren't able to express and communicate with others. This project will help them overcome these hurdles and bring happiness home. It is also a beneficial source for doctors and medical hospitals to use for regular check-ups and an easier mode to communicate. This is just like a pebble of good deed thrown in the societal pond, which will definitely create more positive ripples in the pond itself.

6.3 FUTURE WORK

We are focusing on adopting different methodologies to create awareness among deaf and mute, to make them realize its need and what they're lacking from the modern world. We have achieved the formation of words from the letters and words that are displayed on the screen.

Since we want our product to be released for the betterment of the specially-abled in the world, we're looking for transporting the system onto android software.

CHAPTER 7 - PROJECT METRICS

7.1. CHALLENGES FACED

While doing this project, we faced several challenges. Our initial challenge was a lack of understanding of various problems faced by deaf and mute people. We visited a nearby school for deaf and mute in Patiala to get a better insight into the problems of the deaf and the mute. We then had to find the various technologies that could be developed to help such people. We decided to use OpenCV to classify the images detected by the camera. We had a few troubles segmenting the frames captured by each video. It was difficult to capture a perfect frame for the hand as we had several issues including noise and other components that had to be considered.

7.2. RELEVANT SUBJECTS

Subject Code	Subject Name	Description
UCS503	Software Engineering	In this course, we learnt about Software Development Life Cycle (SDLC). We implemented the concepts throughout the duration of this project.
UCS742	Deep Learning	Implemented the concepts of deep learning training model using Deep Learning concepts such as Models including CNN and KNN.
UCS632	Computer Vision	We used various libraries and concepts of Computer Vision for Image Classification.

7.3. INTERDISCIPLINARY KNOWLEDGE SHARING

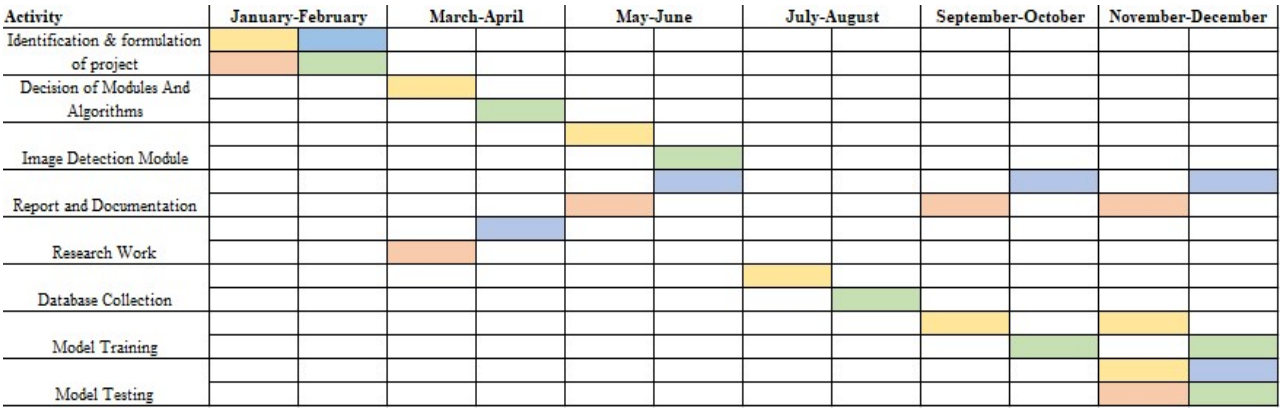
While working on this project, we only used computer science-related concepts. Nevertheless, we have had the opportunity to make use of theories belonging to different subfields within Computer Science. We applied Software Engineering principles throughout this project's period. We identified criteria, planned them, created the softer, carried out different types of testing and reported our research. To help us design and plan the project, we used the principles related to the Image Classification area. We needed to learn the best practices in this area to support the deaf and mute people from this campaign.

7.4. PEER ASSESSMENT MATRIX

Student		Deepanshu	Devak	Mansi	Kali
Evaluation by	Deepanshu	-	5	5	5
	Devak	5	-	5	5
	Mansi	5	5	-	5
	Kali	5	5	5	-

7.5. ROLE PLAYING AND WORK SCHEDULE

Since we believe in teamwork, every task is performed by the group. So segregating each person's individual role is quite difficult. Still the application & training of the model is managed by Mansi & Kali. They are also working hard in doing literature surveys and research work required in the field. The back end development of the app including testing the model made, surveys, contacting people related to this field & researchers willing to provide us with helpful information, these tasks are seen by Devak & Deepanshu.



Colour	Member
	Deepanshu
	Mansi
	Kali
	Devak

7.6. STUDENT OUTCOMES (SO) DESCRIPTION AND PERFORMANCE INDICATORS (PI) (A-K MAPPING)

SO	Description	Outcome
A1	Applying mathematical concepts to obtain analytical and numerical solutions.	Mathematical concepts were required in processing of the input images.
A2	Applying basic principles of science towards solving engineering problems.	Use models like CNN and KNN for image classification.
A3	Applying engineering techniques for solving computing problems.	We trained the dataset with camera clicked inputs to provide functioning in max devices .
B1	Identify the constraints, assumptions and models for the problems.	Assumptions and constraints were feasible with time.
B2	Use appropriate methods, tools and techniques for data collection.	We used ASL Database and trained the model to recognize corresponding text to the gesture.
C1	Design software system to address desired needs in different problem domains.	We used Python to develop the given project.
C2	Can understand scope and constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability.	It will help social growth of differently-abled people.
D1	Fulfill assigned responsibility in multidisciplinary	We collectively did the entire

	teams.	work together.
D2	Can play different roles as a team player.	Members played role of team members as well as team leaders during the course of the project.
E1	Identify engineering problems.	Identified that problems in investigation can be solved by engineering a product.
E2	Develop appropriate models to formulate solutions.	To make the system efficient enough to recognize the hand gestures and generate the corresponding text.
F1	Showcase professional responsibility while interacting with peers and professional communities.	We achieved a higher level of professionalism during interaction between members of the team and learnt to present our projects more professionally.
F2	Able to evaluate the ethical dimensions of a problem.	It will help the differently-abled community and their related ones to communicate better.
G1	Produce a variety of documents such as laboratory or project reports using appropriate formats.	Throughout the project, we prepared and submitted documents as per the format prescribed by the coordinators.
G2	Deliver well-organized and effective oral presentation.	Took an opportunity to explain our work to various panels.

H1	Aware of environmental and societal impact of engineering solutions.	The project has no negative effect on the environment. It has a positive effect on the societal growth.

7.7 BRIEF ANALYTICAL ASSESSMENT

Q1. What sources of information did your team explore to arrive at the list of possible Project Problems?

Ans: We explored the topics over the internet. At that time, we came to know about the Smart India hackathon. So finalized the topic from the problem statements of SIH, and then searched for some innovation that can be done in the project. Finally added one safety feature in our project, and came up with the final topic -> Hand Gesture Detection for Deaf and Dumb, we called it Hand gesture Magic.

Q2. What analytical, computational and/or experimental methods did your project team use to obtain solutions to the problems in the project?

Ans: We visited one of the nearby Deaf and mute schools in Patiala itself, where we came to know about the daily life that they live, what expectations do they have from society, in what ways they can and are capable of returning back to the community. Surfed about systems similar to this which are already in use but in Brazillian Language.

Q3. Did the project demand demonstration of knowledge of fundamentals, scientific and/or engineering principles? If yes, how did you apply?

Ans: Yes. The project demanded demonstration of knowledge of the following fundamentals, scientific and/or engineering principles:

- Image Processing - Image segmentation techniques to extract gestures from an image.
- Software Engineering - The agile method of development has been followed.

- Keras - The experimentation of the artificial deep neural networks has been performed well here.
- Database Management System- Created MySQL database, for storing and updating user data.

Q4. How did your team shares responsibility and communicate the information of schedule with others in the team to coordinate design and manufacturing dependencies? .
Ans: Teamwork has been one of the major focuses of our course curriculum. We used to meet in the library, discuss the ideas, and distributed work to each other. We maintained the coordination through WhatsApp and Gmail.

Q5. What resources did you use to learn new materials not taught in class for the course of the project?
Ans: The Major resource was our Mentor. For further knowledge, we surfed over the internet and went through some research papers related to the technology to be implemented.

Q6. Does the project make you appreciate the need to solve problems in real life using engineering and could the project development make you proficient with software development tools and environments?
Ans: Yes, our project made us appreciate the need to solve problems in real life using engineering. As we chose a real-life problem and gave a proper solution that can be implemented in real life for the betterment of the population of India.

APPENDIX A: REFERENCES

[1] Lee Ch. and Xu Yan., "Online, Interactive Learning of Gestures for Human/Robot Interfaces", The Robotics Institute Carnegie Mellon University Pittsburgh, Pennsylvania 15213, USA,1996.

[2] Yang Tie. , Xu Yan. , "Hidden Markov Model for Gesture Recognition", CMU-RI-TR-94 10 The Robotics Institute Carnegie Mellon University Pittsburgh, Pennsylvania 15213(May 1994).

[3] Chan Wah Ng, Ranganath S., "Real-time gesture recognition system and application", Volume 20, Issued 13-14, December 2002

[4] Hu Peng, "Application Research on Face Detection Technology based on Open CV in Mobile Augmented Reality", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 2 (2015).

[5] Pavlovic V I., Sharma R., T. S. Huang, "Visual Interpretation of Hand Gestures for HumanComputer Interaction: A Review," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.19, pp.677-695, 1997.

APPENDIX B: PLAGIARISM CHECK