Team ID: PNT2022TMID14098

Team members: Balaji V M, Dharunkumar S, Gokulavanan P, Karthick S

REAL-TIME COMMUNICATION POWERED BY AI FOR SPECIALLY ABLED

Project Report Format

1. **INTRODUCTION**

a. Project Overview

In our society, we have people with disabilities. The technology is developing day by day but no significant developments are undertaken for the betterment of these people. Communications between deaf-mute and a normal person has always been a challenging task. It is very difficult for mute people to convey their message to normal people. Since normal people are not trained on hand sign language. In emergency times conveying their message is very difficult. The human hand has remained a popular choice to convey information in situations where other forms like speech cannot be used. Voice Conversion System with Hand Gesture Recognition and translation will be very useful to have a proper conversation between a normal person and an impaired person in any language.

The project aims to develop a system that converts the sign language into a human hearing voice in the desired language to convey a message to normal people, as well as convert speech into understandable sign language for the deaf and dumb. We are making use of a convolution neural network to create a model that is trained on different hand gestures. An app is built which uses this model. This app enables deaf and dumb people to convey their information using signs which get converted to human-understandable language and speech is given as output.

b. Purpose

The main purpose is to communication between deaf and dumb people and normal people so that they can express their feeling easily.

2. LITERATURE SURVEY

2.1 Existing problem

1.From the input RGB image, the hand is separated and morphological operations are performed to identify the region of interest. The features of the gesture are then extracted and compared to a database of features of standard gestures. Finally, based on the comparison the output is generated.

2.The methodology used is similar to [1] except that, instead of bare hands, the system requires the user to wear gloves to extract hand gesture

3.The image is converted into grayscale and the edges of the fingers are detected using Canny edge detection. Then using the detected finger tips the gesture is recognized.

4.The RGB image is converted into a binary image. Certain coordinates are mapped to the binary image. Using a pattern matching algorithm the coordinates are then compared to the coordinates in a database. Based on the comparison, the gesture is identified.

2.2References

[1] Sood Anchal, and Anju Mishra, "AAWAAZ: A communication system for deaf and dumb," 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). IEEE, 2016.

[2] Shraddha R. Ghorpade, Surendra K. Waghamare, “Full Duplex Communication System for Deaf & Dumb People,” International Journal of Emerging Technology and Advanced Engineering (IJETAE), Volume 5, Issue 5, May 2015, ISSN 2250-2459.

[3] Er. Aditi Kalsh, Dr N.S. Garewal, “Sign Language Recognition System,” International Journal of Computational Engineering Research (IJCER), Volume 03, Issue 6, June 2013.

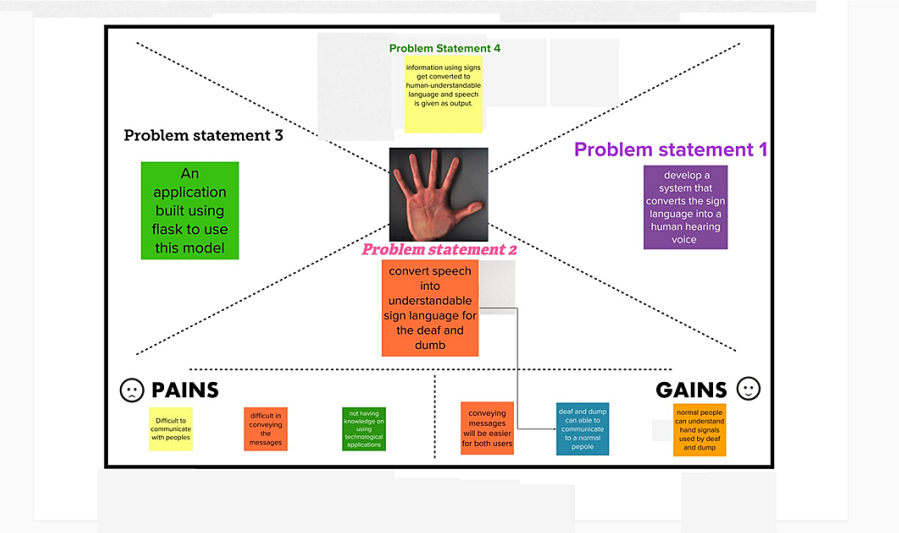
[4] Sawant Pramada, Deshpande Saylee, NalePranita, NerkarSamiksha, Mrs.Archana S. Vaidya “Intelligent Sign Language Recognition Using Image Processing,” IOSR Journal of Engineering (IOSRJEN), Volume 3, Issue 2, Feb. 2013, PP 45-51.

2.3 Problem Statement Definition

Problem Statement (PS)	I am	I'm trying to	But	Because	Which makes me feel
1. Deaf and Dumb peoples can't communicate to normal people	Person with Hearing impairment	Convey my message to a normal people	They were not able to understand our gestures	They were not aware of the hand gestures used by us	Very difficult to convey and communicate with the normal people
2. Normal people not able to communicate with PwD	Person who lives along with a people of PwD	Understand the messages conveyed by the PwD(dumb and Deaf).	I can't able to understand the communicati on They were made to me	I don't know the meaning of the hand gestures they use	Feels useless when I am not able to understand and not able to help them.

3. IDEATION & PROPOSED SOLUTION

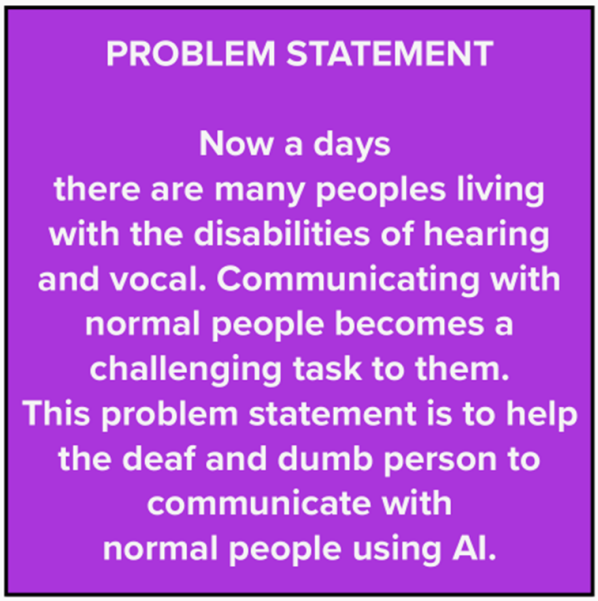
3.1 Empathy Map Canvas



3.2 Ideation&Brainstorming

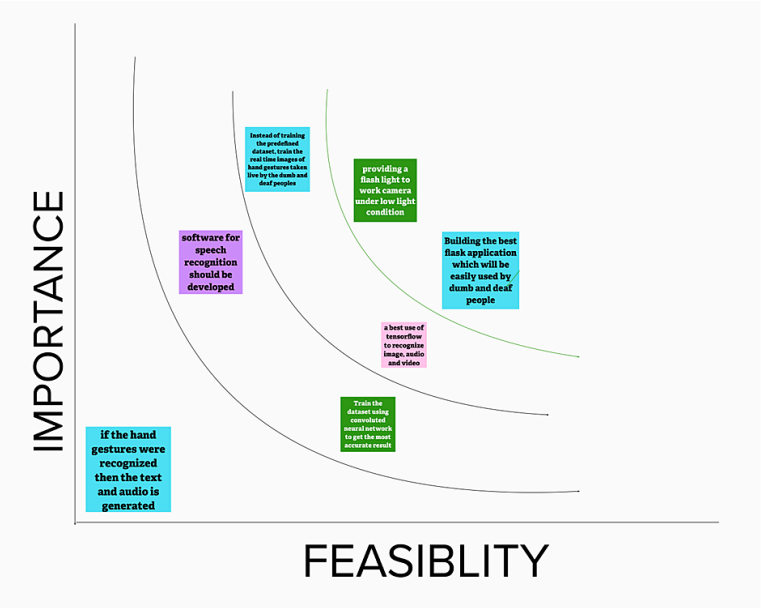
Step-1: Team Gathering, Collaboration and Select the Problem Statement

Our team were discussed about the problem statement and defined it



Step-2: Brainstorm, Idea Listing and Grouping

3. IDEA PRIORITIZATION



3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	**To Develop a model which is very useful to communicate to normal people by using hand signal and gestures.**
2.	Idea / Solution description	● **Using CNN model of image recognition to identify the accurate hand gestures** ● **A quick result of voice and text after the gestures get identified**
3.	Novelty / Uniqueness	**Image to sound detection is the uniqueness of this project. After analysing the hand signals the gesture get identified and provides a sound**
4.	Social Impact / Customer Satisfaction	● Disabled people experience a great deal of difficulty with day-to-day activities ● Normal people who not able to communicate with disabled peoples can now easily get communicate with them. It will be the great impact and provide a satisfaction
5.	Business Model (Revenue Model)	1. A person who needs this model can afford at low price and this provide a income. 2. This advanced technology make life easier and will get great demand in market of technology.
6.	Scalability of the Solution	Depth, Width and Resolution were scaled in CNN. This model will be working under even low light conditions. Even a Child with basic knowledge can use this app.

3.4 Problem solution Fit

1.CUSTOMER SEGMENTS(CS)	6.CUSTOMER CONSTRAINTS (CC)	5.AVAILABLE SOLUTIONS(AS)
There were two customers: • Deaf and Dumb People who can't convey the message properly. • The normal People who are trying to communicate them were customers.	• Specially Abled Person use their hand signals to get communicate with other. • Normal people will face difficulty in understanding the sign language.	• CNN to identify the hand gestures. • AI to communicate with gesture and voice • Flask to develop application.
2. JOBS TO BE DONE	**9. PROBLEM ROOT CAUSE(RC)**	**7. BEHAVIOUR(BE)**
• Create an efficient app to convert hand gestures to voice and text. • Develop cnn model to recognize the voice and text. • Developing and Training the Dataset is major task	• The Communication barrier is root cause. • Problem of conveying message properly to the normal people. • The proper expression of the feel was not expressed	• Searching the medium to express the feelings. • Searching a device to get translate.
3. TRIGGERS(TM)	**4. YOUR SOLUTION(SL)**	**8.CHANNELS OF BEHAVIOUR (CH)**
• The ability of the customers to communicate efficiently at serious and necessary situations.	• This application help in communication between the normal people and dumb and deaf people	The Application developed by us is the main channel of the behavior. Online translation is also and Channel of Behaviour.

4.REQUIREMENT ANALYSIS

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User communication	User Must Know the Sign Language
FR-2	User Communication	The user has to communicate in front of the Camera

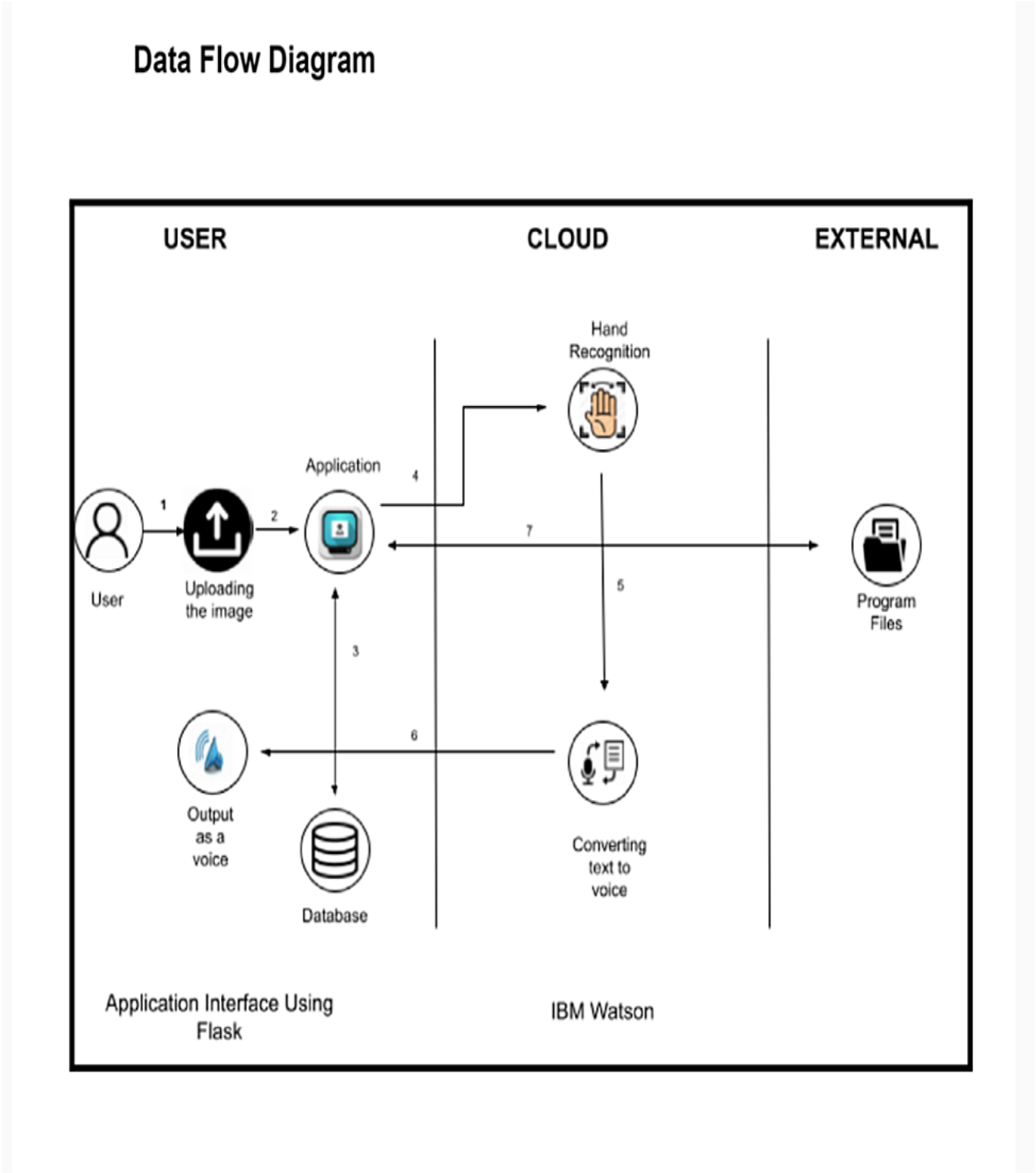
Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

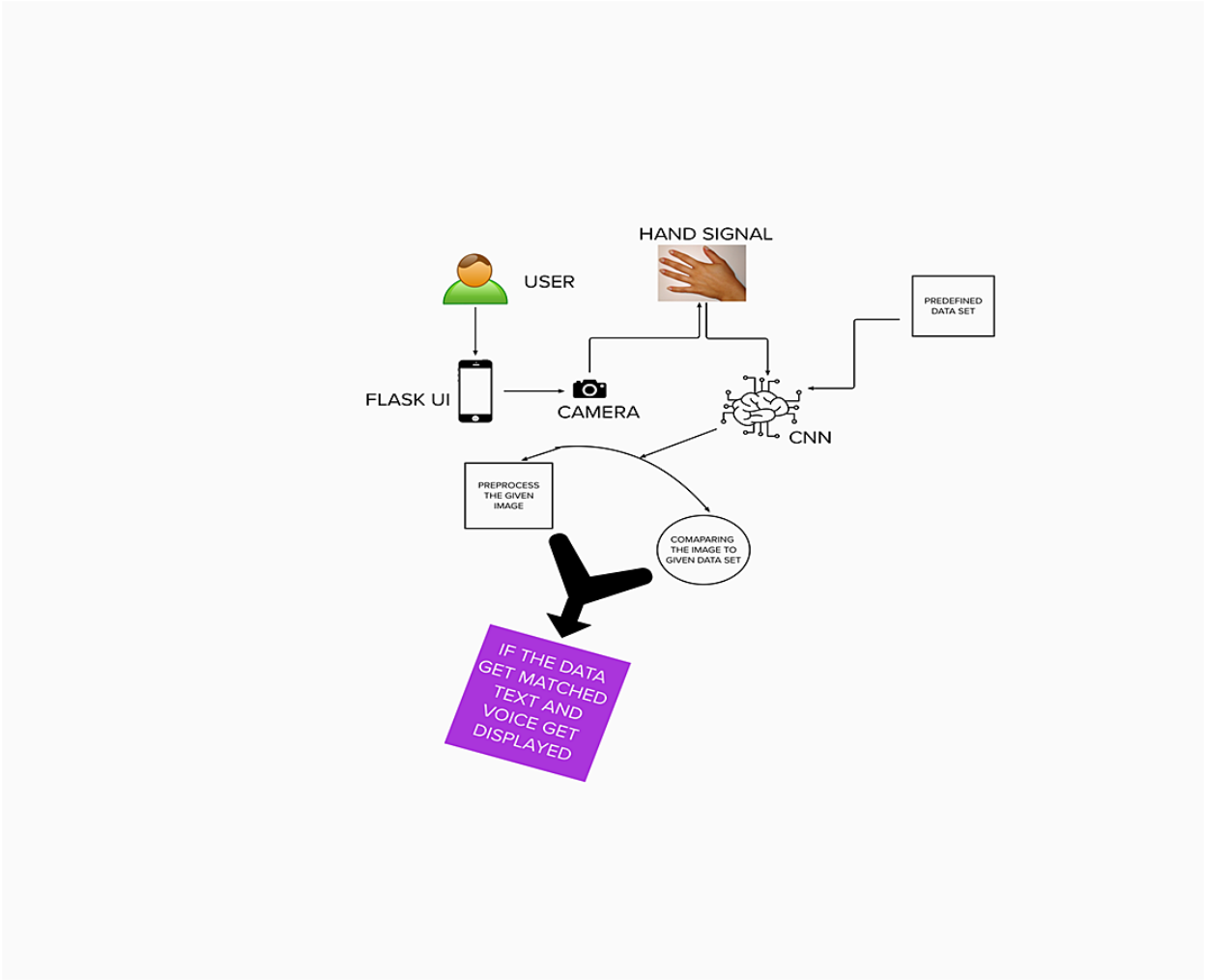
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The camera captures all expressions including facial expressions and hand gestures which can be easily used by all age groups.
NFR-2	**Reliability**	The system is very liable, it can last for long amounts of time if well maintained.
NFR-3	**Performance**	The cost-effective nature of the system makes it extremely liable and thus, efficient.
NFR-4	**Availability**	The solution fits all the sign languages when we train the model for all the sign languages. So, it is used by all the countries with different languages.
NFR-5	**Scalability**	The system gives output rapidly. It also predicts quickly when it gets so many inputs at a time. It predicts different types of sign language at a time.

5.**PROJECT DESIGN**

5.1 Data Flow Diagram



5.2.Solution & Technical Architecture



6. **PROJECT PLANNING AND SCHEDULING**

6.1.Sprint Planning & Estimation

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	Not Required		High	-
Sprint-1	Login	USN-2	Not Required	1	High	-
Sprint-1	Main page	USN-4	As a User, **I can enter the web page once clicked, which provides be the Guidelines** to use the app	1	Medium	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**
Sprint-2	Convert Sign	USN-6	As a User, **I can click the button Convert sign** , which directs me towards the Main screen	4	Medium	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-2	**Camera(Hand movement detection)**	USN-7	As a User, **I can show my hand sign towards the camera which converts them into text manner.**	8	High	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**
Sprint-3	Voice mode	USN-8	Once the text is obtained, As a User **I can click on the voice mode which provides the text in the form of speech.**	3	High	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**
Sprint-1	**Provide the necessary functionalities required to use the app.**		As an Executive, I can **provide the Specifications of Camera required, and other factors** that are required for smooth functioning of the app.	1	Low	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**
Sprint-4	**Check the performance of the app**		As an Executive, I can **check the usage and queries obtained from the end users.**	1	Medium	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**
Sprint-4	**Receive queries based on the usage**		As an Admin, I can **take the queries from the customer care and perform the testing phase again , loading the other signs in the dataset,** in order to make the customers to use the app effectively.	2	High	**BALAJI V M DHARUNKUMAR S GOKULAVANAN P KARTHICK S**

6.2 Sprint Delivery schedule

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	6	6 Days	3 NOV 2022	8 NOV 2022		29 Oct 2022
Sprint-2	12	6 Days	8 Nov 2022	13 Nov 2022		13 Nov 2022
Sprint-3	3	3Days	13 Nov 2022	15 Nov 2022		15 Nov 2022
Sprint-4	4	4 Days	17 Nov 2022	20 Nov 2022		20 Nov 2022

Velocity:
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

7. CODING & SOLUTIONING

7.1 DETECTION OF HAND SIGNALS CLEARLY :

```
def detect(jpeg):
img = resize(jpeg, (64, 64, 3))
copy = img.copy()
copy = copy[150:150 + 200, 50:50 + 200]
cv2.imwrite('image.jpg', copy)
copy_img = image.load_img('image.jpg')
x = image.img_to_array(copy_img)
x = np.expand_dims(x, axis=0)
prediction = np.argmax(model.predict(x), axis=1)
pred = vals[prediction[0]]
print("it indicates : ", pred)
```

```
    return pred
```

7.2 GETTING THE RESULT FROM HTML PAGE

```
<!DOCTYPE html>
<html>
<head>
  <title>html page</title>
</head>
<body>
  <h1>video streaming</h1>
  <img id="video" src="{{ url_for('video_feed') }}">
</body>
</html>
```

7.3 READING LIVE STREAM FRAME USING PYTHON CODE WITH CLEAR PIXELS:

```
import cv2

class VideoCamera():
    def __init__(self):
        # Open a camera
        self.cap = cv2.VideoCapture(0)



    def __del__(self):
        self.cap.release()

    def get_frame(self):
        ret, frame = self.cap.read()

        if ret:
            ret, jpeg = cv2.imencode('.jpg', frame)
            return jpeg.tobytes()

        else:
            return None
```

```
AND BELOW CODE IS TO DISPLAY


def gen():
    global video_camera
    global global_frame

    if video_camera == None:
        video_camera = VideoCamera()

    while True:
        frame = video_camera.get_frame()

        if frame != None:
            global_frame = frame
            yield (b'--frame\r\n'
                   b'Content-Type: image/jpeg\r\n\r\n' + frame + b'\r\n\r\n')
        else:
            yield (b'--frame\r\n'
                   b'Content-Type: image/jpeg\r\n\r\n' + global_frame +
b'\r\n\r\n')
```

8. **TESTING**

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issuesof the project at the time of the release to User Acceptance Testing (UAT

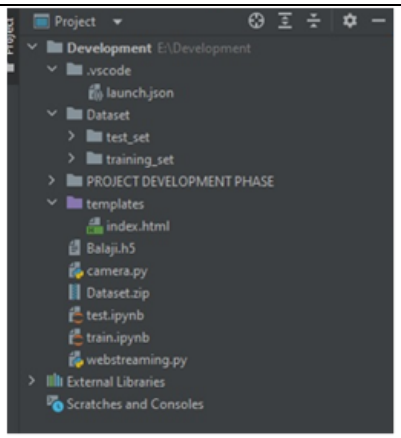
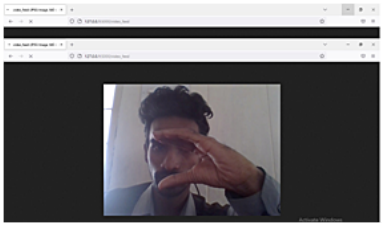
2. Test Case Analysis

This report shows the numberof test cases that have passed, failed,and untested

Section	Total Cases	Not Tested	Fail	Pass
Camera detection	1	0	0	1
Train the model and saving	7	0	0	7

Frame capturing and output	2	0	2	0

9. **PERFORMANCE TESTING**

S.No.	Parameter	Values	Screenshot
1.	Project structure	PYTHON FILE HTML FILE FLASK APP LOADED MODEL	
2.	OUTPUT	Camera enable and letter display	

10.ADVANTAGES AND DISADVANTAGES

ADVANTAGES:

i.) Main advantage is normal people can easil communicate to PwD.

ii.) PwD people can easily express their feelings to everyone

DISADVANTAGES

i.)Lack of knowledge on using the application

ii.) Working under low light camera may be a disadvantages.

11. **CONCLUSION:**

To Develop a model which is very useful to communicate to normal people by using hand signal and gestures Disabled people experience a great deal of difficulty with day-to-day activities Normal people who not able to communicate with disabled peoples can now easily get communicate with them. It will be the great impact and provide a satisfaction.. A person who needs this model can afford at low price and this provide a income. This advanced technology make life easier and will get great demand in market of technology.

12. **FUTURE SCOPE**

a.) A new module can be developed that working under low light condition.

b.) CNN algorithm can be tuned more to get an accurate result that desired.

12. APPENDIX:

Source Code:

#webstreaming.py

```
import numpy as np
import cv2
import os
from keras.models import load_model
from flask import Flask, render_template, Response, jsonify, request
from camera import VideoCamera
from keras.preprocessing import image
```

```
global graph
global writer
from skimage.transform import resize

writer = None
model = load_model('Balaji.h5')
vals = ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I']
app = Flask(__name__)
print("[info] accessing video stream...")
vs = cv2.VideoCapture(0)


def detect(jpeg):
    img = resize(jpeg, (64, 64, 3))
    copy = img.copy()
    copy = copy[150:150 + 200, 50:50 + 200]
    cv2.imwrite('image.jpg', copy)
    copy_img = image.load_img('image.jpg')
    x = image.img_to_array(copy_img)
    x = np.expand_dims(x, axis=0)
    prediction = np.argmax(model.predict(x), axis=1)
    pred = vals[prediction[0]]
    print("it indicates : ", pred)
    return pred


video_camera = None
global_frame = None


@app.route('/')
def index():
    return render_template('index.html')


def gen():
    global video_camera
    global global_frame

    if video_camera == None:
        video_camera = VideoCamera()

    while True:
        frame = video_camera.get_frame()

        if frame != None:
            global_frame = frame
            yield (b'--frame\r\n'
                   b'Content-Type: image/jpeg\r\n\r\n' + frame + b'\r\n\r\n')
        else:
```

```
            yield (b'--frame\r\n'
                   b'Content-Type: image/jpeg\r\n\r\n' + global_frame +
b'\r\n\r\n')
            img = resize(frame, (64, 64))
            x = image.img_to_array(img)
            x = np.expand_dims(x, axis=0)
            prediction = np.argmax(model.predict(x), axis=1)
            pred = vals[prediction[0]]
            print("it indicates : ", pred)


@app.route('/video_feed')
def video_feed():
    return Response(gen(), mimetype='multipart/x-mixed-replace; boundary=frame')


if __name__ == '__main__':
    app.run(host='0.0.0.0', debug=True)
```

#camera.py

```
import cv2

class VideoCamera():
    def __init__(self):
        # Open a camera
        self.cap = cv2.VideoCapture(0)



    def __del__(self):
        self.cap.release()

    def get_frame(self):
        ret, frame = self.cap.read()

        if ret:
            ret, jpeg = cv2.imencode('.jpg', frame)
            return jpeg.tobytes()

        else:
            return None
```

#index.html

```
<!DOCTYPE html>
<html>
<head>
  <title>html page</title>
</head>
<body>
  <h1>video streaming</h1>
  <img id="video" src="{{ url_for('video_feed') }}">
</body>
</html>
```

#train.ipynb

#%%
from keras.preprocessing.image import ImageDataGenerator
train_datagen = ImageDataGenerator(rescale = 1./225,
shear_range=0.2,zoom_range=0.2,horizontal_flip=True)
test_datagen = ImageDataGenerator(rescale = 1./225)
#%%
x_train =
train_datagen.flow_from_directory('Dataset/training_set',target_size=(64,64),
batch_size=300,class_mode='categorical', color_mode ="grayscale")
#%%
x_test =
train_datagen.flow_from_directory('Dataset/test_set',target_size=(64,64),
batch_size=300,class_mode='categorical', color_mode ="grayscale")
#%%
from keras.models import Sequential

from keras.layers import Dense
from keras.layers import Convolution2D
from keras.layers import MaxPooling2D
from keras.layers import Dropout
from keras.layers import Flatten
#%%
model=Sequential()
#%%
model.add(Convolution2D(32,(3,3), input_shape=(64,64,1), activation = 'relu'))
#%%
model.add(MaxPooling2D(pool_size=(2,2)))
#%%
model.add(Flatten())
#%%
model.add(Dense(units=512,activation='relu'))
model.add(Dense(units=9,activation='softmax'))

#%%
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
#%%
model.fit_generator(x_train, steps_per_epoch=24, epochs=10,validation_data=x_test,validation_steps=40)
#%%
model.save('Balaji.h5')

#test.ipynb

```
from keras.models import load_model
import numpy as np
import cv2

model=load_model('Balaji.h5')

from skimage.transform import resize
def detect(frame):
    img = resize(frame,(64,64,1))
```

```
    img = np.expand_dims(img,axis=0)
    if(np.max(img)>1):
        img = img/255.0
    prediction = model.predict(img)
    print(prediction)
    predictions = np.argmax(model.predict(img), axis=1)
    print(predictions[0])


frame=cv2.imread(r"E:\Development\Dataset\test_set\B\1.png")
data= detect(frame)
```

Github link:

click here to navigate to github
or else copy paste this
https://github.com/IBM-EPBL/IBM-Project-39209-1660400483

DEMO LINK:

click here for the video