REAL-TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY ABLED

AN IBM PROJECT REPORT

SUBMITTED BY

THE STUDENTS OF

K RAMAKRISHNAN COLLEGE OF ENGINEERING

CARLINE IMAKULATE V - 811519104019

ABIRAMI R - 811519104005

ANGELINE JOY ALEX - 811519104009

GOWTHAMY RM – 811519104037

PROJECT REPORT

1. INTRODUCTION

1.1 Overview

People get to know one another by sharing their ideas, thoughts, and experiences with those around them. There are numerous ways to accomplish this, the best of which is the gift of "Speech." Everyone can very convincingly transfer their thoughts and understand each other through speech. It will be unjust if we overlook those who are denied this priceless gift: the deaf and dumb. In such cases, the human hand has remained the preferred method of communication.

1.2 Purpose

The project's purpose is to create a system that translates sign language into a human-understandable language so that ordinary people may understand it.

2. LITERATURE SURVEY

2.1 Existing problem

Some of the existing solutions for solving this problem are:

Technology: One of the easiest ways to communicate is through technology such as a smart phone or laptop. A deaf person can type out what they want to say and a person who is blind or has low vision can use a screen reader to read the text out loud. A blind person can also use voice recognition software to convert what they are saying in to text so that a person who is Deaf can then read it.

Interpreter :If a sign language interpreter is available, this facilitates easy communication if the person who is deaf is fluent in sign language. The deaf person and person who is blind can communicate with each other via the interpreter. The deaf person can use sign language and the interpreter can speak what has been said to the person who is blind and then translate anything spoken by the blind person into sign language for the deaf person.

Just Speaking: Depending on the deaf person's level of hearing loss, they may be able to communicate with a blind person who is using speech. For example, a deaf person may have enough Residual hearing (with or without the use of an assistive hearing device such as a hearing aid) to be able to decipher the speech of the person who is blind or has low vision. However, this is often not the most effective form of communication, as it is very dependent on the individual circumstances of both people and their environment (for example, some places mayhave too much background noise).

2.2 References

1. Environment Setup:

https://www.youtube.com/watch?v=5mDYijMfSzs

2. Sign Languages Dataset:

https://drive.google.com/file/d/1ITbDvhLwyTTkuUYfNjOKhcIZh7hDgi64/view?usp=sharing

3. Keras Image Processing Doc: https://keras.io/api/preprocessing/image/

4. Keras Image Dataset From Directory Doc:

https://keras.io/api/preprocessing/image/#imagedatasetfromdir

ectory-function

5. CNN using Tensor flow: https://www.youtube.com/watch?v=umGJ30-15_A

6. OpenCV Basics of Processing Image:

https://www.youtube.com/watch?v=mjKd1Tzl70I

2.3 Problem Statement

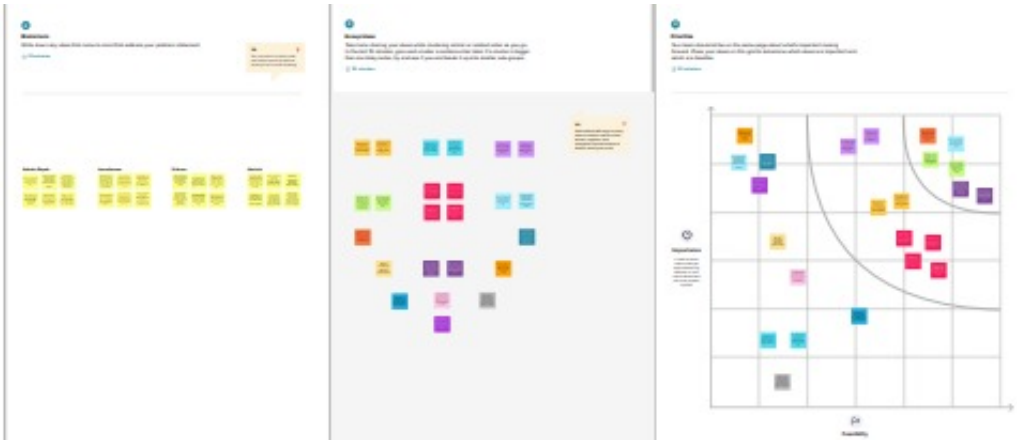
Definition In our society, we have people with disabilities. The technology is developing day by day but no significant developments are undertaken for the betterment of these people. Communications between deaf-mute and a normal person has always been a challenging task. It is very difficult for mute 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.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

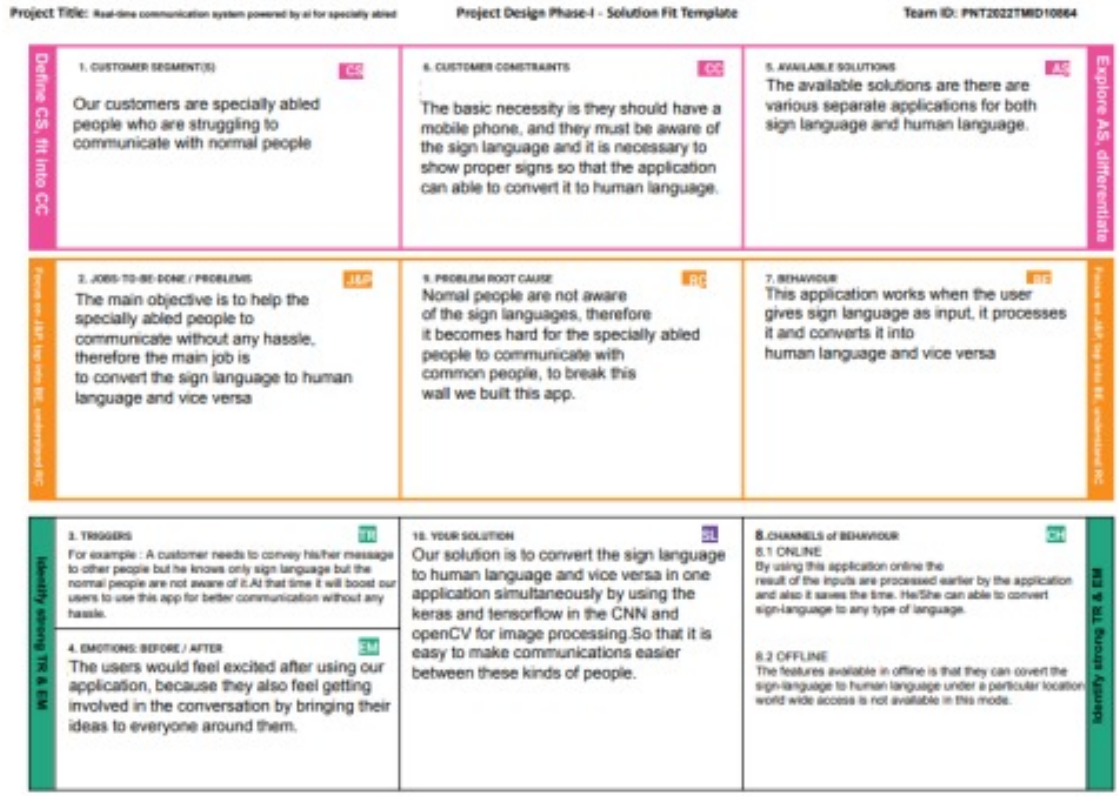


3.3 Proposed Solution

Proposed Solution:

S. No.	Parameter	Description
1.	**Problem Statement (Problem to be solved)**	People with disabilities exist in our society. It has always been difficult to communicate with a deaf-mute and a typical person. Mute persons find it incredibly challenging to communicate with non-mute people. due to the fact that most people are not trained in handsign language. It is exceedingly difficult for them to communicate during times of emergency. In circumstances where other modes of communication, like speech, cannot be employed, the humanhand has continued to be a preferred choice. A proper discussion between a normal person and an impaired person inany language will be made possible by a voice conversion system with hand gesture recognition and translation.
2.	**Idea / Solution description**	We use technology that converts sign language into a human hearing voice in thetarget language to communicate with everyday people and translate speech intoacceptable sign language for the deaf and dumb.
3.	**Novelty / Uniqueness**	To accomplish this, we create a model usinga convolution neural network that is trained on several hand motions. An app is developed using this model. Deaf or stupid people who use this application can communicate using signs that are converted into language that is understandable to people.
4.	**Social Impact / Customer Satisfaction**	The deaf and dumb people as well as the ordinary people can interact without any hassle.
5.	**Business Model (Revenue Model)**	This model is for Deaf-mute person as wellas an ordinary man, communication is the basic necessity as this is an application we can brand this and attract many investors.
6.	**Scalability of the Solution**	In the future we can also help the people who got impaired in the middle who doesn't know the sign language can also learn from this app and communicate without any hassle.

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

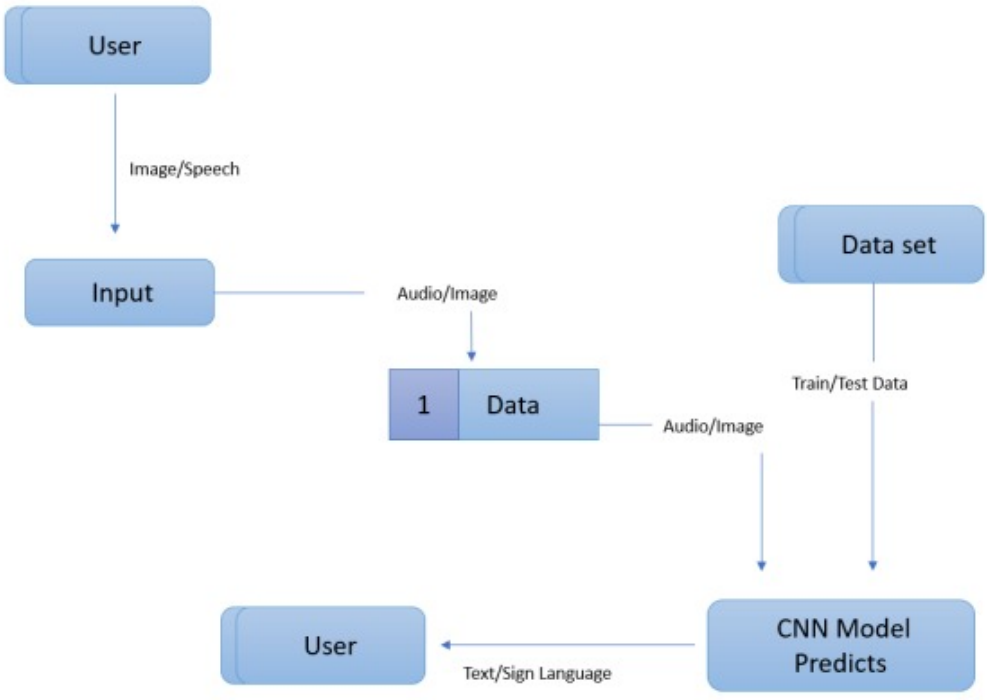
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Gesture recognition	Understanding pointing gestures. Understanding nodding gestures. Understanding focused eye contact.
FR-2	Speech recognition	Understand independent speech. Understand dependent speech. Understand Natural Language.
FR-3	Track	This feature provides the ability to track the movement of the hand, this function is between the user and the GRE.
FR-4	Camera	The phone should have a camera to capture the images of sign language to convert it into text.
FR-5	Microphone	The phone should have a microphone to capture the speech of user to convert it to sign language.

4.2 NON FUNCTIONAL REQUIREMENTS

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The usability is made at ease so that the specially abled can easily access.
NFR-2	**Security**	The system has high security standards and the images gets deleted once they leave the app to maintain privacy.
NFR-3	**Reliability**	This system is made to understand all gestures as well as the languages spoken so it can be dependable without a second thought.
NFR-4	**Performance**	The performance of the system is set to be consistent.
NFR-5	**Availability**	This system is available both online as well as offline, it is made for an access for all the people across the globe.
NFR-6	**Scalability**	Can be used world wide if we feed all the languages in the world, or it can also be customised according to the region.

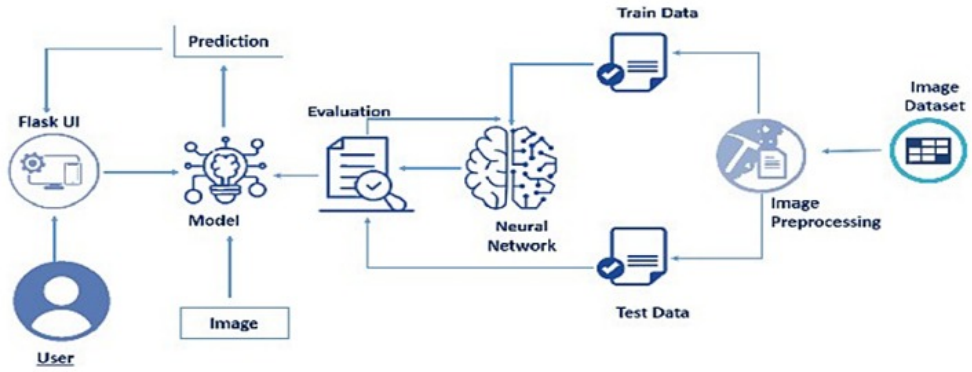
5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS

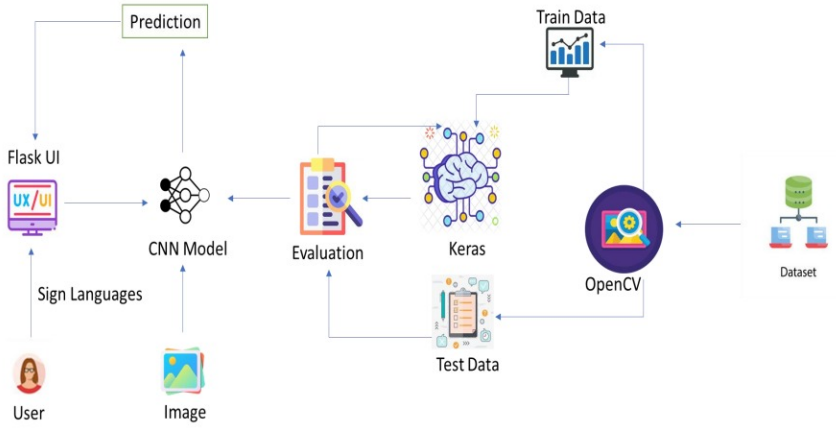


5.2 Solution & Technical Architecture

TECHNICAL ARCHITECTURE



SOLUTION ARCHITECTURE



5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Deaf and mute person	Gesture recognition	USN-1	As the user does the gestures, the system captures it.	I am able to operate the system easily	High	Sprint-1
		USN-2	The gestures which are done by the user, is instantly been stored inside the database.	I am able to see the gestures getting recorded	High	Sprint-1
		USN-3	These stored gestures are processed and finally converted into desired output.	I can instantly see the gestures being converted into speech.	High	Sprint-2
		USN-4	The gestures are converted into text which is understood by a common man.	I can easily communicate with an ordinary man	High	Sprint-1
Ordinary person	Speech recognition	USN-1	The speaker speaks the desired message needed to be conveyed to the specially abled.	I am able to operate the system easily	High	Sprint-1
		USN-2	The system records the speech.	I am able to see the speech getting recorded	High	Sprint-1
		USN-3	The spoken words are stored in the database and converted into desired output.	I can instantly see the speech being converted into gestures.	High	Sprint-2
		USN-4	At last the speech is been converted into gestures which is understood by the specially abled.	I can easily communicate with a specially abled.	High	Sprint-1

6. PROJECT PLANNING AND SCHEDULING

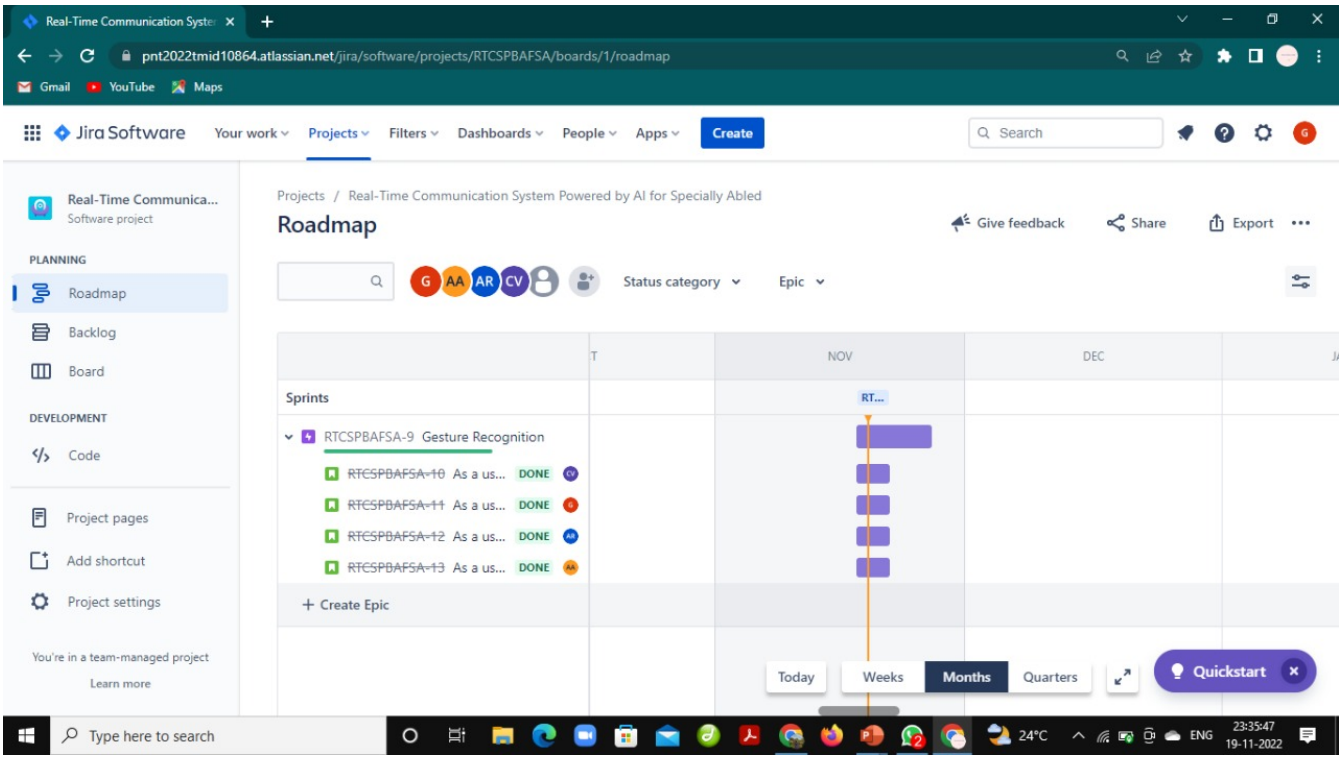
6.1 SPRINT PLANNING AND ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Gesture Recognition	USN-1	As a user, I can login the app and the gesture recognizer automatically captures the gesture done by me.	3	Medium	Carline Imakulate V
Sprint-1	Speech Recognition	USN-2	As a user, the speech recognizer records the speech and converts it into gestures.	2	Low	Gowthamy R M
Sprint-2	Track	USN-3	As a user this feature provides the ability to track the movement of the hand.	3	Medium	Abirami R
Sprint-2	Camera	USN-4	As a user, I can easily use this app without any difficulties as the camera easily captures the gesture.	1	Low	Angeline Joy Alex
Sprint-3	Microphone	USN-5	Speech recognition uses microphone for capturing the audio .	1	Low	Abirami R
Sprint-3	Predicting	USN-6	Using AI algorithm, prediction are made using the paramater provided.	5	Medium	Angeline Joy Alex
Sprint-4	Importing the libraries	USN-7	We have to implement necessary libraries in packages.	5	High	Gowthamy R M
Sprint-4	Result	USN-8	Quality of the nutrition is analyzed.	5	High	Carline Imakulate V

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022

6.3 REPORTS FROM JIRA



7 CODING AND SOLUTION

7.1 Web Application

```
<html lang="en">

<head>
  <meta charset="utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0, shrink-to-fit=no">
  <title>SmartBridge_WebApp_VideoTemplate</title>
  <link rel="stylesheet" href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css">
  <link rel="stylesheet" href="https://use.fontawesome.com/releases/v5.12.0/css/all.css">
  <link rel="stylesheet" href="assets/css/Banner-Heading-Image.css">
  <link rel="stylesheet" href="assets/css/Navbar-Centered-Brand.css">
  <link rel="stylesheet" href="assets/css/styles.css">
</head>

<body style="background: rgb(39,43,48);">
  <nav class="navbar navbar-light navbar-expand-md py-3" style="background: #212529;">
    <div class="container">
      <div></div><a class="navbar-brand d-flex align-items-center" href="#"><span
          class="bs-icon-sm bs-icon-rounded bs-icon-primary d-flex justify-content-center align-items-center me-2 bs-icon"><i
            class="fas fa-flask"></i></span><span style="color: rgb(255,255,255);">Real-Time Communication
          System Powered By AI&amp;nbsp;For Specially Abled</span></a>
      <div></div>
    </div>
  </nav>
  <section>
    <div class="d-flex flex-column justify-content-center align-items-center">
      <div class="d-flex flex-column justify-content-center align-items-center" id="div-video-feed"
        style="width: 640px;height: 480px;margin: 10px;min-height: 480px;min-width: 640px;border-radius: 10px;border: 4px dashed rgb(255,255,255) ;">
        <img src="{{ url_for('video_feed') }}" style="width: 100%;height: 100%;color: rgb(255,255,255);text-align: center;font-size: 20px;"
          alt="Camera Access Not Provided!">
      </div>
    </div>
    <div class="d-flex flex-column justify-content-center align-items-center" style="margin-bottom: 10px;"><button
        class="btn btn-info" type="button" data-bs-target="#modal-1" data-bs-toggle="modal">Quick Reference
        -<strong> ASL Alphabets</strong></button></div>
  </section>
  <section>
    <div class="container">
      <div class="accordion text-white" role="tablist" id="accordion-1">
        <div class="accordion-item" style="background: rgb(33,37,41);">
          <h2 class="accordion-header" role="tab"><button class="accordion-button" data-bs-toggle="collapse"
              data-bs-target="#accordion-1 .item-1" aria-expanded="true"
              aria-controls="accordion-1 .item-1"
              style="background: rgb(39,43,48);color: rgb(255,255,255);">About The Project</button></h2>
          <div class="accordion-collapse collapse show item-1" role="tabpanel" data-bs-parent="#accordion-1">
            <div class="accordion-body">
              <p class="mb-0">Artificial Intelligence has made it possible to handle our daily activities
                in new and simpler ways. With the ability to automate tasks that normally require human
                intelligence, such as speech and voice recognition, visual perception, predictive text
                functionality, decision-making, and a variety of other tasks, AI can assist people with
                disabilities by significantly improving their ability to get around and participate in
                daily activities.<br><br>Currently, Sign Recognition is available <strong>only for
                  alphabets A-I</strong> and not for J-Z, since J-Z alphabets also require Gesture
                Recognition for them to be able to be predicted correctly to a certain degree of
                accuracy.</p>
            </div>
          </div>
        </div>
```

```
            <div class="accordion-item" style="background: rgb(33,37,41);">
              <h2 class="accordion-header" role="tab"><button class="accordion-button collapsed"
                  data-bs-toggle="collapse" data-bs-target="#accordion-1 .item-2" aria-expanded="false"
                  aria-controls="accordion-1 .item-2"
                  style="background: rgb(39,43,48);color: rgb(231,241,255);">Developed By</button></h2>
              <div class="accordion-collapse collapse item-2" role="tabpanel" data-bs-parent="#accordion-1">
                <div class="accordion-body">
                  <p class="mb-0">Students at K.Ramakrishnan College Of Engineering
                    <br><br>1. <strong>R.Abirami</strong> 811519104005<br>2.
                    <strong>Angeline Joy Alex</strong> 811519104009<br>3. <strong>V.Carline Imakulate</strong> 811519104019<br>4.
<strong>R.M.Gowthamy</strong> 811519104037
                  </p>
                </div>
              </div>
            </div>
          </div>
        </div>
      </section>
      <div class="modal fade" role="dialog" tabindex="-1" id="modal-1">
        <div class="modal-dialog" role="document">
          <div class="modal-content">
            <div class="modal-header">
              <h4 class="modal-title">American Sign Language - Alphabets</h4><button type="button"
                class="btn-close" data-bs-dismiss="modal" aria-label="Close"></button>
            </div>
            <div class="modal-body"><img src="{{ url_for('static', filename='img/ASL_Alphabets.png') }}" width="100%"></div>
            <div class="modal-footer"><button class="btn btn-secondary" type="button"
                data-bs-dismiss="modal">Close</button></div>
          </div>
        </div>
      </div>
      <script src="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js"></script>
    </body>
    </html>
```

7.2 Camera

```
from flask import Flask, Response, render_template
import cv2
import numpy as np
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import os

class Video(object):
        def __init__(self):
                self.video = cv2.VideoCapture(0)
                self.roi_start = (50, 150)
                self.roi_end = (250, 350)
                self.model = load_model('savedModel.h5') # Execute IBM Trained Model
                self.index=['A','B','C','D','E','F','G','H','I']
                self.y = None
        def __del__(self):
                k = cv2.waitKey(1)

                self.video.release()
        def get_frame(self):
                ret,frame = self.video.read()
```

```
                frame = cv2.resize(frame,(640,480))
                copy = frame.copy()
                copy = copy[150:150+200,50:50+200]
                # prediction starts
                cv2.imwrite('image.jpg',copy)
                copy_img = image.load_img('image.jpg', target_size=(64,64,3))
                x = image.img_to_array(copy_img)
                x = np.expand_dims(x, axis=0)
                pred = np.argmax(self.model.predict(x), axis=1)
                self.y = pred[0]
                cv2.putText(frame,'The Predicted Alphabet is: '+str(self.index[self.y]),(100,50),cv2.FONT_HERSHEY_SIMPLEX,1,(0,0,0),3)
                ret,jpg = cv2.imencode('.jpg', frame)
                return jpg.tobytes()

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

def gen(object):
        while True:
                frame = object.get_frame()
                yield(b'--frame\r\n'b'Content-Type: image/jpeg\r\n\r\n' + frame +          b'\r\n\r\n')

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


if __name__ == '__main__':
        app.run()
```

8. TESTING

8.1 The American Sign Language Comprehension Test (ASL-CT)

It is a 30-item multiple-choice test that measures ASL receptive skills and is administered through a website. This article describes the development and psychometric properties of the test based on a sample of 80 college students including deaf native signers, hearing native signers, deaf non-native signers, and hearing ASL students. The results revealed that the ASL-CT has good internal reliability ($\alpha = 0.834$). Discriminant validity was established by demonstrating that deaf native signers performed significantly better than deaf non-native signers and hearing native signers. Concurrent validity was established by demonstrating that test results positively correlated with another measure of ASL ability ($r = .715$) and that hearing ASL students' performance positively correlated with the level of ASL courses they were taking ($r = .726$). Researchers can use the ASL-CT to characterize an individual's ASL comprehension skills, to establish a minimal skill level as an inclusion criterion for a study, to group study participants by ASL skill (e.g., proficient vs. nonproficient), or to provide a measure of ASL skill as a dependent variable.

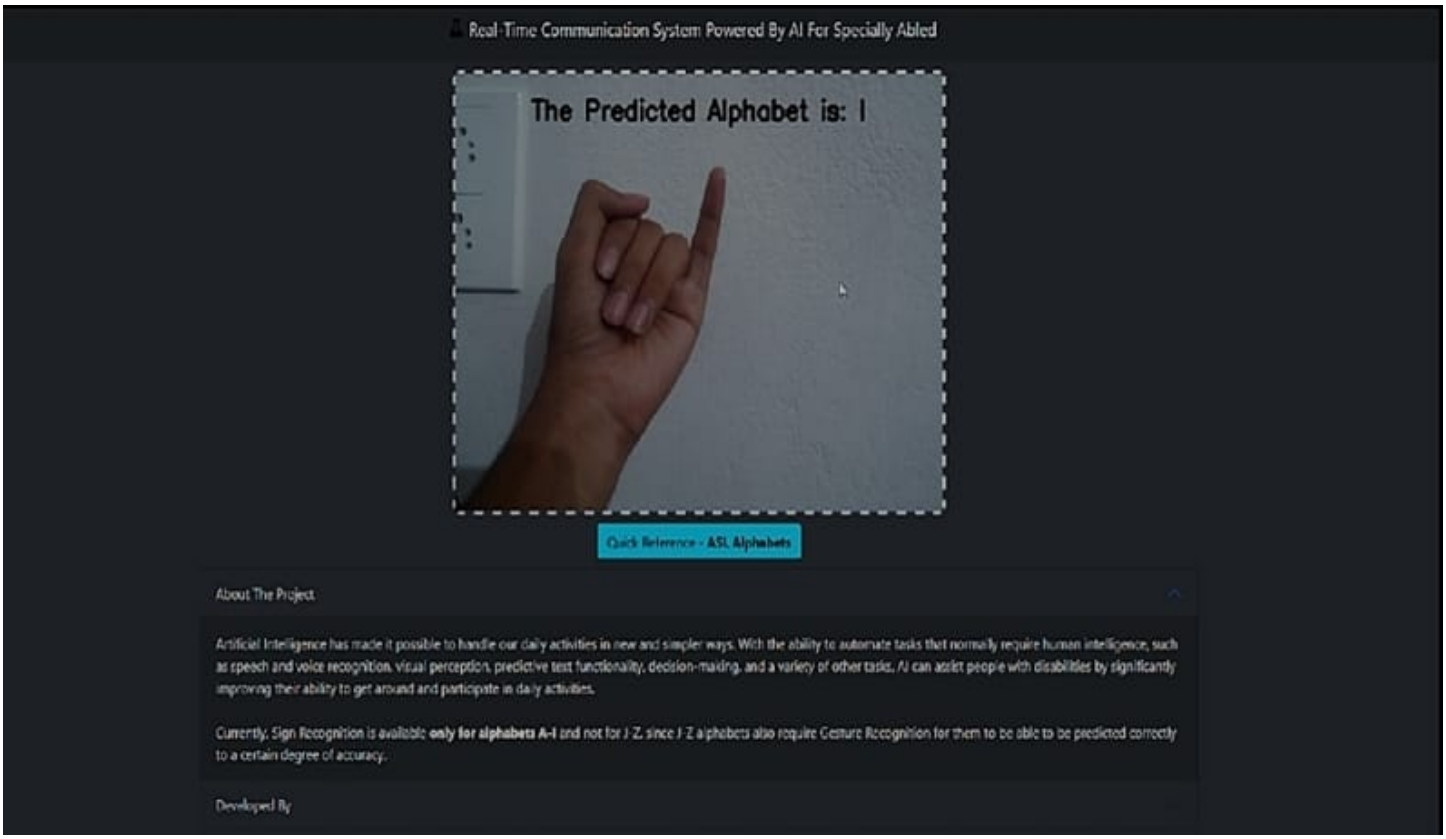
8.2 USER ACCEPTANCE TESTING

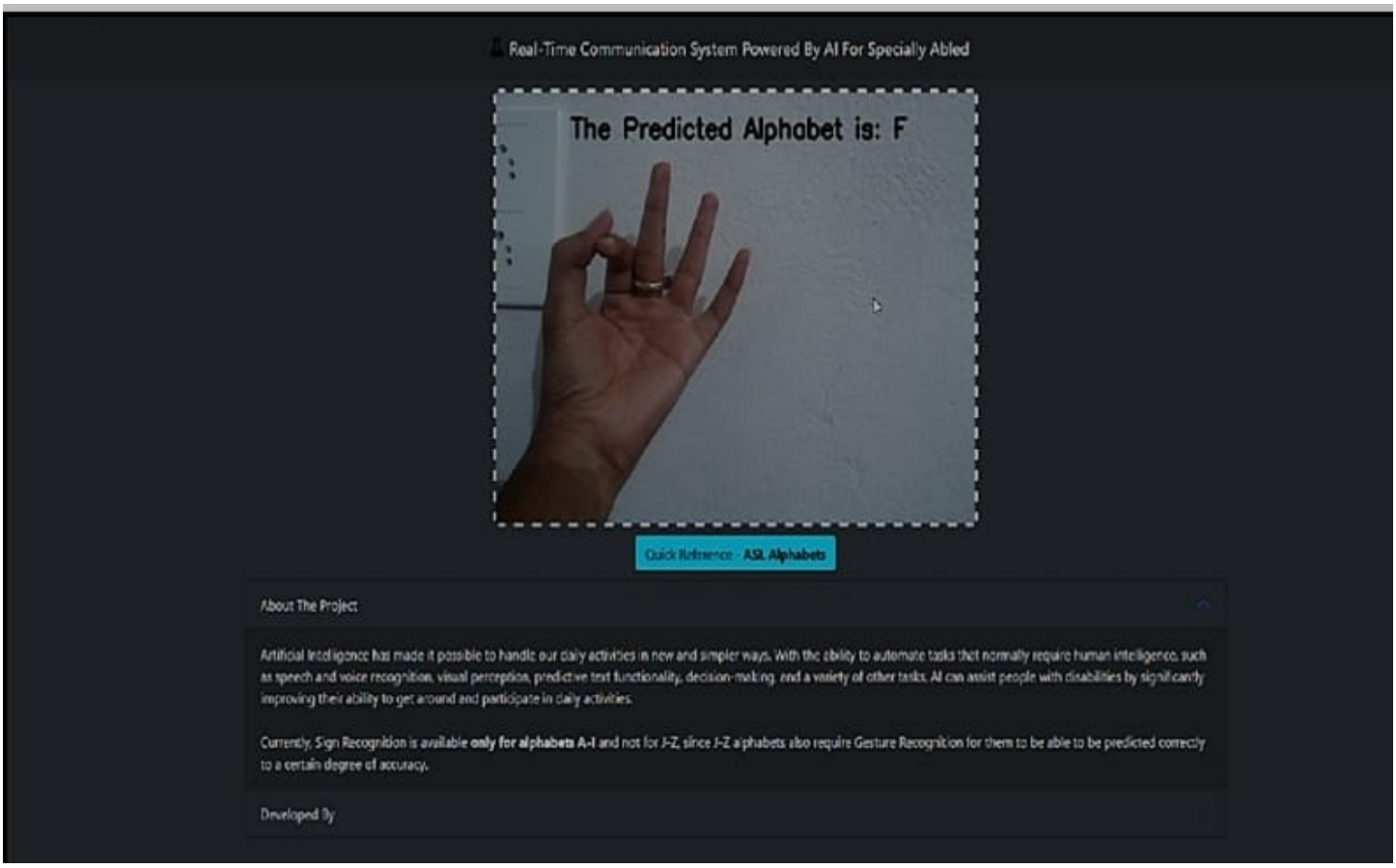
Multiple assistive technologies can help the deaf and mute such as speech-to-text, speech-to visual and sign language. In this study, an offline assistive mobile communication application is developed for the deaf, hard of hearing, mute, and person without disabilities. This can be used as a tool to break the barrier of communication between the people without disabilities and the deaf/mute. Both American Sign Language (ASL) and Filipino Sign Language (FSL) were applied in the developed system that will help the intended users communicate in their everyday activities. The developed system called “BridgeApp” were designed according to the gathered user need of a deaf community based on their specific everyday activity. A thorough system testing was conducted in order to make sure that the mobile app will be ready to use and will aid them in their day to day communication needs. The user acceptance testing results have proven that the system in this study can effectively assist in clear communication between both deaf and mute, as well as to those hearing and speaking individuals

9. RESULTS

9.1 Performance Metrics

The proposed procedure was implemented and tested with set of images. The set of 15750 images of Alphabets from “A” to “I” are used for training database and a set of 2250 images of Alphabets from “A” to “I” are used for testing database. Once the gesture is recognize the equivalent Alphabet is shown on the screen.Some sample images of the output are provided below:





10.ADVANTAGES & DISADVANTAGES

Advantages:

1. It is feasible to develop a mobile application to close the communication gap between the general population and people who are deaf or dumb.

2. The user can select which sign language to read by adding the dataset as new sign language standards are created.

Disadvantages:

1. The current model is limited to the letters A through I.
2. Alphabets from J cannot be recognised in the absence of gesture recognition because they require user input in the form of a gesture.
3. The accuracy isn't great because there aren't many or high-quality photographs in the dataset, but that can be fixed by changing the dataset.

11.CONCLUSION

The use of sign language can help hearing and deaf individuals communicate more effectively. The technology strives to close the communication gap between the deaf community and the rest of society because it supports two-way conversation. The suggested method converts spoken languages into human-understandable English alphabets. With the help of this technology, the model receives hand gestures, recognises them, and then shows the corresponding Alphabet on the screen. This initiative allows deaf-mute people to perform sign language with their hands, which will later be translated into alphabets.

12.FUTURE SCOPE

For persons with particular needs, such as the deaf and dumb, having technology that can translate hand sign language to its appropriate alphabet is a game changer. The web programme may easily be developed to detect letters other than "I," numbers, and other symbols with the addition of gesture recognition. Gesture recognition can also be used to control software and hardware interfaces.

13. GITHUB AND PROJECT DDEMO LINK:

GITHUB: https://github.com/IBM-EPBL/IBM-Project-8356-165891590

DEMO: https://youtu.be/F9Fb-_J8WL8