REAL- TIME COMMUNICATION SYSTEM POWERED BY AI SPECIALLY ABLED

IBM NALAIYA THIRAN PROJECT REPORT

submitted by

KARTHIKEYAN.D (620819104043)
KAVIYARASAN.A (620819104049)
MANOJKUMAR.D (620819104057)
MUGILAN.M (620819104062)

TEAM ID:PNT2022TMID30994

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

GNANAMANI COLLEGE OF TECHNOLOGY, NAMAKKAL – 637018

ANNA UNIVERSITY : : CHENNAI 600 025 NOVEMBER 2022

1. INTRODUCTION

1.1 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.

1.2 Purpose

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.

2. LITERATURE SURVEY

2.1 Existing Problem

1. Based Real Time Communication for Physically and Speech Disabled People (Ong Chin Ann, Marlene Valeriu Lu – 2019)

Communication is a social process of exchanging information from one entity to another in verbaland non-verbal form. It defines our existence and it is an important instrument that connects people together. It comes naturally as a raw skill embedded in most people at birth and we acquired the ways of communication through cognitive learning. Communication is the basis, which drives the process of development in all the fields (Manohar, 2008) and it is the very core of our civilisation.The ability to communicate allows us to express emotion, feelings, convey our thoughts and ideas as well as to relate our experiences. It plays an important role in the dis- semination of information and sharing of knowledge especially in the academic arena. Research has found that human startedto learn how to communicate with each other since they are born not only through spoken and written languages but also body gesture, posture, facial expression and eye contacts (Busso, et al., 2004; Cohen, Grag & Huang, 2000).Communication skill might come as a natural ability in majority of people. However, there are some people inflicted with some form of physical defects which affect their ability tocommunicate. One of the more severe disabilities is known as “cerebral palsy”, a congenital disorder at birth which causes abnormality in their Motor system. It affects their muscle movement and coordination, learning and Speech abilities. Their malfunctioned motor system causes an uncontrollable and involuntary movement. They are unable to control their oralfacial muscles, thus affects their ability to perform facial expression appropriately.From the limitation of the existing tools reviewed (Novita, 2006; Macsolvers, 2009; Standup, 2006; Universiteit van Amsterdam, 2008; Crestwood, 2009; Sci-enceDaily, 2008), there is still a pressing need for more effective and efficient tools to alleviate this problem. One the possible methods are to implement a facial expression recognition system to predict or determine the emotional state of a disabled person through his expression projected on his face. biometrics information system can be employed as a means to detect and classify the physiological aspect of a person in real time. Franco and Treves (2001) further support the notion that facial expression can be used for human computer interaction and usability enhancement.Based on the problem statements deliberated above, we propose an improved real-time communication system using machine learning and computer vision. The aim is to create a communication channel between the specially abled

and the society, so they can express there feelings, thoughts and understand other people's feelings and thoughts through real time communication and facial expressions.

2. Systematic review of computer vision semantic analysis in medical (Antonio Victor Alencar Lundgren, Byron Leite Dantas Bezzerra – 2021)

Medical diagnosing techniques have fascinated us for a long time. It has been common for us to use them in our daily life and implement these technologies. Machine learning and especially computer vision contribute a lot in medical science, which make different difficult tasks easy for doctors and more tolerable for patients. They are widely useful in early detection of disease, and hence are a valuable tool to save human life. Cardio graphic techniques are a must for old age and infant safety.

These include:

• **Retinoscopy -** They although primitive in approach are a must once in a life time and retinoscopy have made yet successful to measure activities of rod and cone receptors in our eyes. Retina has three distinct areas for colors - erythrolabe, chlorolabe and cyanolabe…which are analogical to pixel fixation and identification algorithms on machine learning.

• **Tumor detection -** Cancer is spreading in the world affecting billions of lives both in terms of life and money… machine learning diagnosing systems apply their identification systems to further develop accurate detection in terms of size, location, quality of such tissues which are suspected to become malignant uncontrolled group of fast dividing cells.

• **CT scan - CT scan –** A very common term for cancer patients which uses electromagnetic radiations under manually operated controlled computer vision gratings which are so accurate that it can measure a pigment called c-125 in blood.

3. A survey on Facial Emotion Recognition Techniques (Felipe Zago Canal, Tobias Rossi Muller, Gustavo Gino Scotton – 2022)

Facial expressions recognition is an ability to recognize people by their facial characteristic and differentiate it with one another. Human is born with the ability to recognize other people easily by identifying their facial features such as shape, appearance, skin texture and skin complexion. Other than that, humans also have the ability to express, interpret and differentiate facial expressions. The regular recur-ring ones are happiness, anger,

disgust, fear, surprise and sad (Ekman & Friesen, 1978). The six facial emotions stated above are important and play a major role in expressing emotion as well as recognising facialexpression (Busso, et al, 2004). In real life, inter personal human interaction are performed not only using speech or spoken language, but also nonverbal cues for example hand gesture, body gesture, facial expression and tone of the voice. All these cues are sometimes being used for expressing feeling and give feedback (Busso, et al, 2004; Cohen, et Al., 2000). We can see how human interact with each other using non-verbal cues every day. For example a child cries in front of his mother because he is not happy or dissatisfied with something. Other people might interpret it differently thinking that the child might be in pain. Facial expression interaction is relevant mainly for community social life, teacher and student interaction, credibility in difference contexts, medicine and so on. Besides, facial expression recognition is useful for designing new interactive devices which offers the possibility of new ways for human computer interaction - HCI (Franco & Treves, 2001). Cohen, et al. (2000) conducted survey on their users and noticed that they have been through traditionally HCI consists of the keyboard, mouse, joystick, trackballs, data gloves and touch screen monitors. Facial Expression Recognition System (FER) has been a topic for research since Ekman and Friesen (1978) who pioneered this research and worked from the psychology perspective. In the past 20 years, many researchers have tried to adopt their idea and make improvement, innovation and modification on facial expression recognition by introducing different techniques, mainly concentrated on the improvement in term of accuracy, efficiency, mobility, and speed (Kotsia & Pitas, 2007). With all the enhancements on techniques for facial detection and recognition, the development of the facial expression recognition has also improved (Zhan & Zhou, 2007). The most active researches in computer vision and pattern recognition is face recognition in forensic identification, access control, user interface design (Wang, Plataniotis & Venetsanopoulos, 2005), emotion analysis, interactive video, indexing and retrieval of image and video database, image understanding and synthetic face animation (Zhan & Zhou, 2007).

2.2 References

1.Upendran, S., and Thamizharasi, A., "American Sign Language interpreter system for deaf and dumb individuals", In the Proceedings of the International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp. 1477-1481, 2014.

2. Rajamohan, A., Hemavathy, R., andDhanalakshmi, M., "Deaf-Mute Communication Interpreter", International Journal of Scientific Engineering and Technology, Vol.2, No.5, pp.336-341, 2013.

3. Verma, P., Shimi S. L. and Priyadarshani, R., "Design of Communication Interpreter for Deaf and Dumb Person", Vol.4, no.1, 2013

4. Van Rossum, G., "Python Programming Languag", In Proceedings of USENIX Annual Technical Conference, Vol.41, No.36,2007.

5. Banzi, M., and Shiloh, M.,"Getting Started with Arduino: The Open Source Electronics Prototyping Platform", Maker Media, Inc., 2014.

6. Potdar, K., and Kinnerkar, R., "A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data", International Journal of Science and Research, Vol.5, No.9, pp. 1550-1553, 2016.

7. Desai, P., "Python Programming for Arduino",Packt Publishing Ltd, 2015.

8. Havalagi, P., and Nivedita, S.U., "The Amazing Digital Gloves That Give Voice To The Voiceless", International Journal of Advances in Engineering & Technology, Vol.6, No.1, pp.471-480, 2013

9. D'Ausilio, A., "Arduino: A low-cost multipurpose lab equipment", Behavior research methods, Vol. 44, No,.2, pp.305-313, 2012.

10. Lotti, F.,Tiezzi, P., Vassura, G.,Biagiotti, L., and Melchiorri, C., "UBH 3: and anthropomorphic hand with simplified endo-skeletal structure and soft continuous fingerpads", In Proceedings IEEE International Conference on Robotics and Automation, 2004 (ICRA'04), Vol.5, pp. 4736-474, IEEE, 2004.

11. Goyal, L., andGoyal. V., "Automatic Translation of English Text to Indian Sign Language Synthetic Animations", In Proceedings of the 13th International Conference on Natural Language Processing, p.144, 2016.

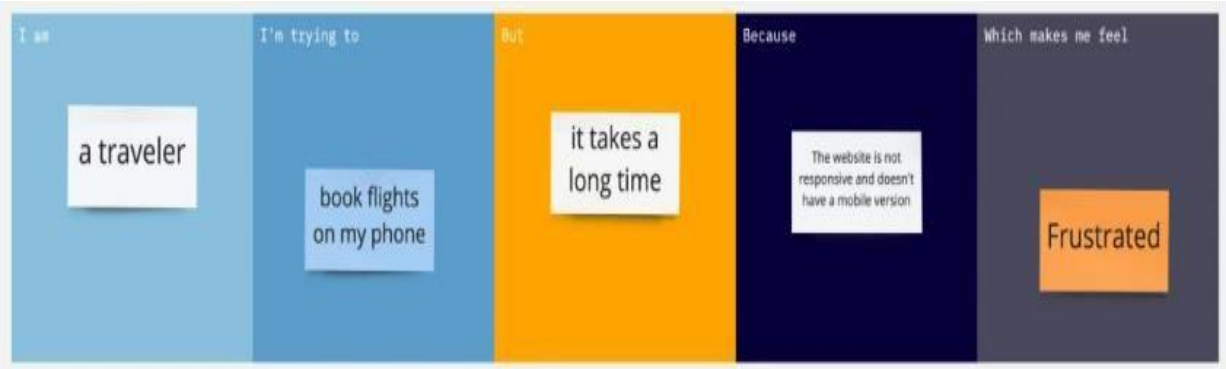
2.3 Problem Statement Definition

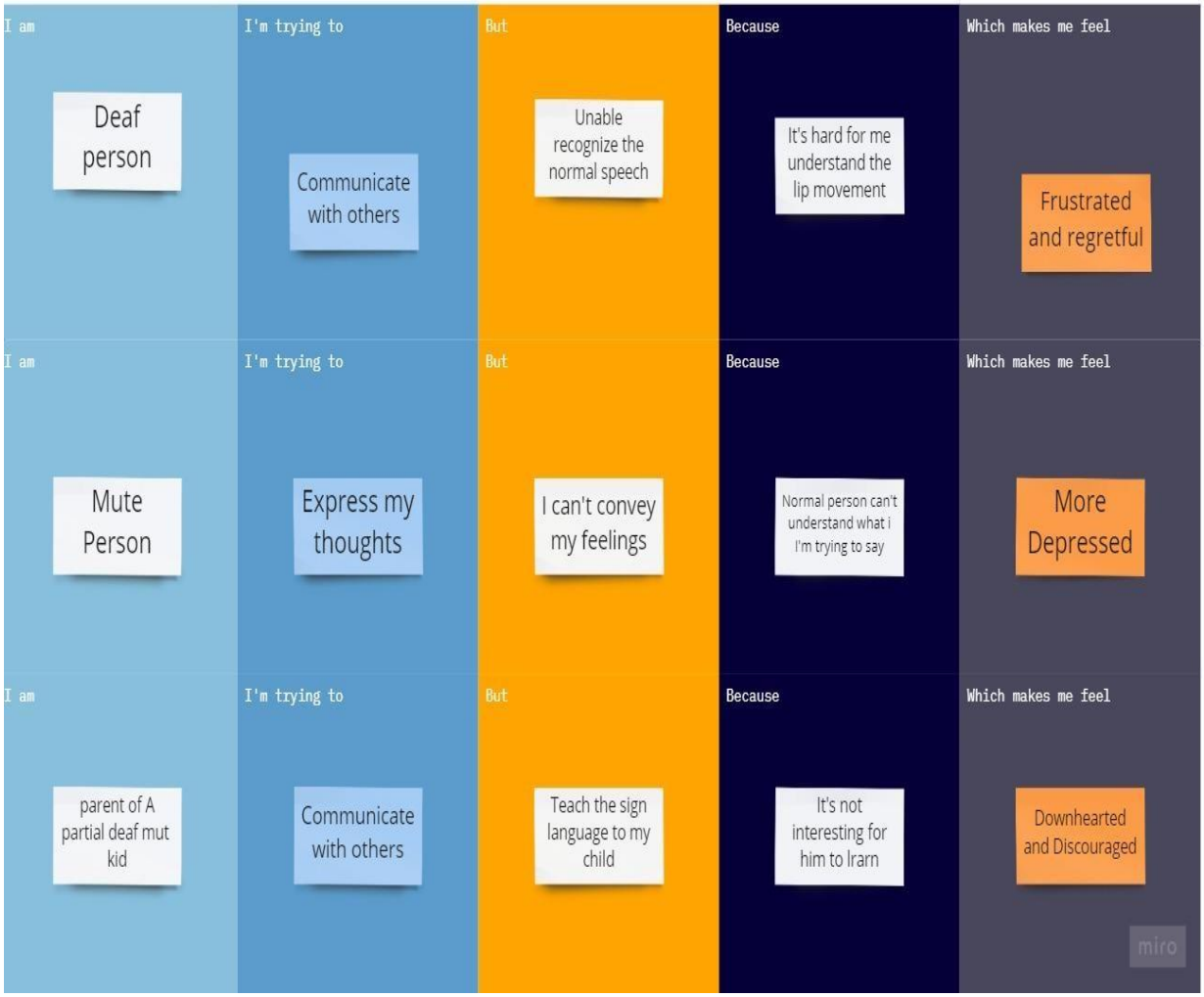
Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love.

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

I am	Describe customer with 3-4 key characteristics - *who are they?*	Describe the customer and their attributes here
I'm trying to	*List their outcome or "job" the care about - what are they trying to achieve?*	List the thing they are trying to achieve here
but	Describe what problems or barriers stand in the way – *what bothers them most?*	Describe the problems or barriers that get in the way here
because	Enter the "root cause" of why the problem or barrier exists – *what needs to be solved?*	Describe the reason the problems or barriers exist
which makes me feel	Describe the emotions from the customer's point of view – *how does it impact them emotionally?*	Describe the emotions the result from experiencing the problems or barriers

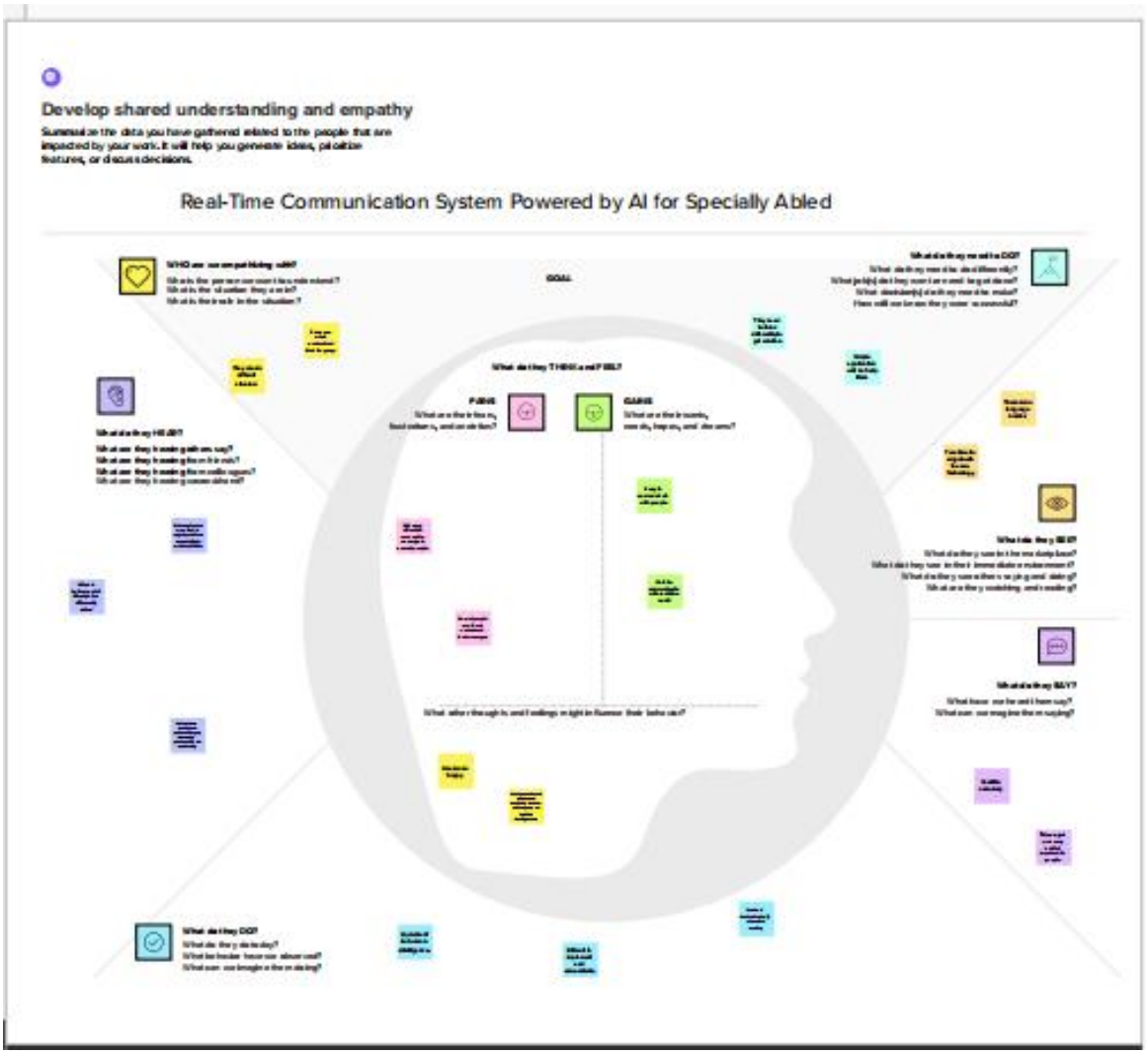
Example:


I am
a traveler
I'm trying to
book flights on my phone
But
it takes a long time
Because
The website is not responsive and doesn't have a mobile version
Which makes me feel
Frustrated


I am
Deaf person
I'm trying to
Communicate with others
But
Unable recognize the normal speech
Because
It's hard for me understand the lip movement
Which makes me feel
Frustrated and regretful
I am
Mute Person
I'm trying to
Express my thoughts
But
I can't convey my feelings
Because
Normal person can't understand what i I'm trying to say
Which makes me feel
More Depressed
I am
parent of A partial deaf mut kid
I'm trying to
Communicate with others
But
Teach the sign language to my child
Because
It's not interesting for him to lrarn
Which makes me feel
Downhearted and Discouraged
miro

3. IDEATION & PROPOSED SOLUTION

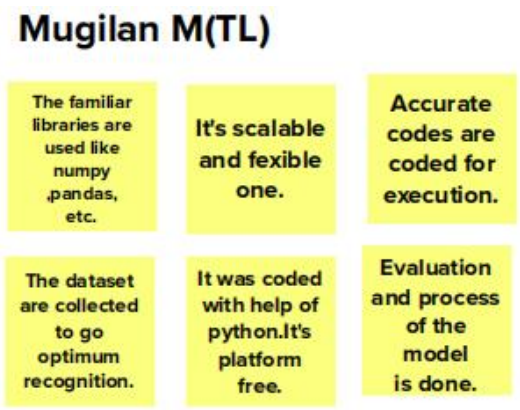
3.1 Empathy Map Canvas

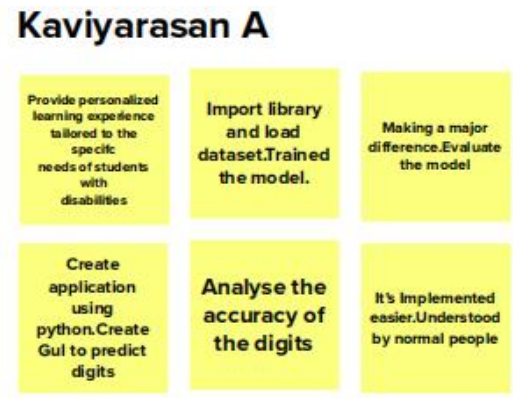


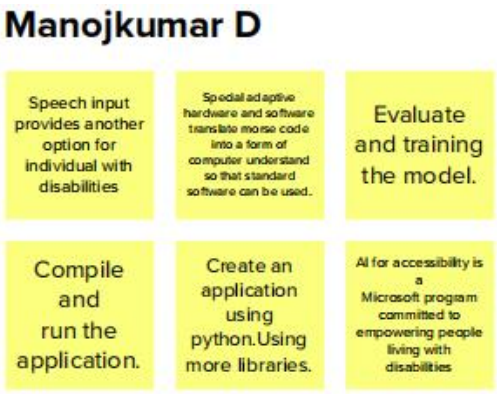
3.2 Ideation & Brainstorming

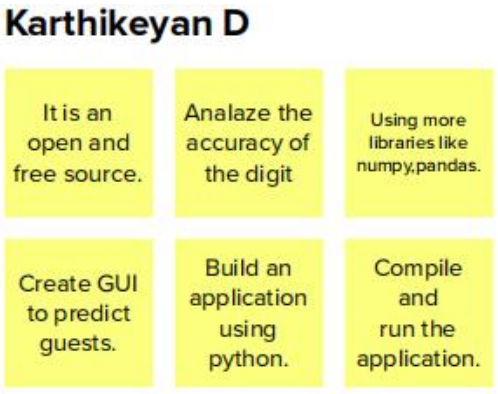
Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Write down any ideas that come to mind that address your problem statement.









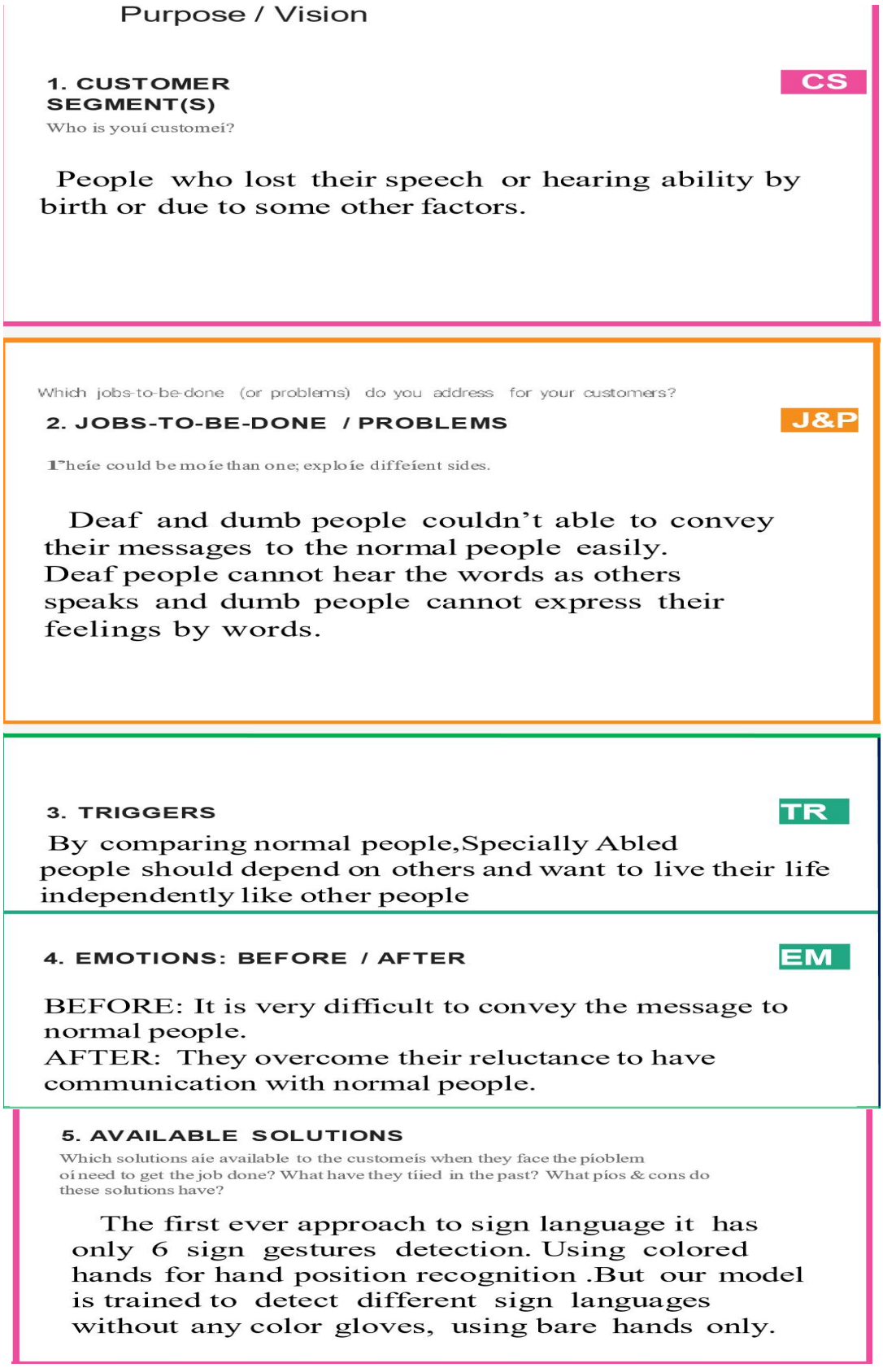
Ideation

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine whichideas are important and which are feasible.

3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	It is 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.
2.	Idea / Solution description	A real time ML based system is built for the real time sign language detection with a Tensor Flow object
3.	Novelty / Uniqueness	This model using SSD ML algorithm recognizing the signs as words instead of old traditional translators, that are very slow and take too much since every alphabet as to be recognized to form the whole statement in old methods
4.	Social Impact / Customer Satisfaction	It drastically reduce communication difference gap between normal people and specially abled people with the help of AI. So they can live their life independently.
5.	Business Model (Revenue Model)	We use freemium business revenue model for making revenue. In our device, we give most of the basic features for free of charge but they have to pay if they need more advanced feature
6.	Scalability of the Solution	The model which is Tensor Flow model that has been used can be replaced with another model as well. The same system can be implemented for different sign languages by substituting the dataset

3.4 Problem Solution Fit



CC

6. CUSTOMER

What constiaints pievent youi customeis fiom taking action oi limit theii choices of solutions? i.e. spending powei, budget, no cash, netwoik connection, available devices.

Difficult accessibility, not user friendly, need more technical knowledge to handle, cost and etc.There are so many choice of solutions availablebut due to these some constraints, choice of solutions were limited.

BE

7. BEHAVIOUR

What does youi customei do to addiess the pioblem and get the job done?
i.e. diiectly ielated: find the iight solai panel installei, calculate usage and benefits;
indiiectly associated: customeis spend fiee time on volunteeiing woik (i.e. Gieenpeace)

In our device, there's an option called problem detection display in which our customer can able to see the type of problem occurs & solution will be displayed.

CH

8. CHANNELS of BEHAVIOUR

8.1 ONLINE
What kind of actions do customeis take online? Extiact online channels fiom 7

Advertise on online with influencers to test the product and promote it also on blog channels

8.2 OÏLINE
What kind of actions do customeis take offline? Extiact offline channels fiom 7and use them foi customei development.

On offline, we have our product experience stores where our customer can experience the product in real

RC

9. PROBLEM ROOT CAUSE

What is the ieal ieason that this pioblem exists? What is the back stoiy behind the need to do this job?
i.e. customeis have to do it because of the change in iegulations.

In Previously developed solution, they have to use coloured hand gloves for hand position recognition. Also, the old method uses traditional translators which take too much of time to process.

SL

10. YOUR SOLUTION

If you aie woiking on an existing business, wiite down youi cuiient solution fiist,fill in the canvas, and check how much it fits ieality.
If you aie woiking on a new business pioposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customei limitations, solves a pioblem and matches customei behavioui.

Using SSD ML algorithm recognizing the signs as words instead of old traditional translators, that are very slow and take too much since every alphabet as to be recognized to form the whole statement in old methods.

4. REQUIREMENT ANALYSIS

4.1 Functional Requirement

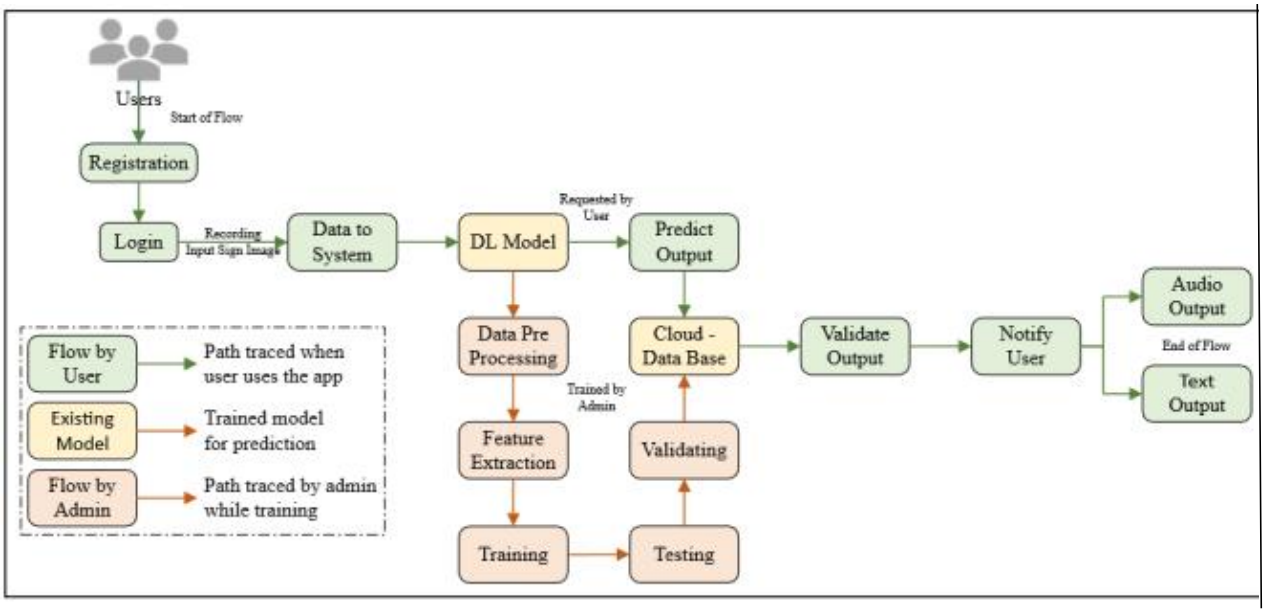
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Web UI/ E-Mail ID. Authentication via OTP.
FR-2	User Confirmation	Confirmation via mail.
FR-3	System	• Desktop/ Mobile with good resolution camera. • Provides system access to capture images/ video and other relevant data.
FR-4	Text conversion	Converts the Sign language into a text using Convolutional Neural Network (CNN) Model.
FR-5	Sentence Translation	To create sentence(s) by recognizing the signs and pauses in the input video stream.

4.2 Non-Functional Reqirement

NFR No.	Non-Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
NFR-1	Usability	Deaf-mute people should be able to use the system with ease. The same applies for normal people who get the system's output. The system should have good UI.
NFR-2	Security	Even though the use-case of the system doesn't need any security feature, it must be ensured that the privacy of user data be maintained and handled appropriately.
NFR-3	Reliability	The translation of sign languages should be reliable. The accuracy of the system should be tested extensively to make sure that it is up to the mark.
NFR-4	Performance	The processing should be done in considerable time so that the conversation can go on without waiting for the system's output.
NFR-5	Availability	The system should be universally accessible. Since sign language is almost same everywhere, the system can be used across the globe.
NFR-6	Scalability	The system should be Scalable to accommodate new features and functionalities and to cater wider range of people in future.

5. PROJECT DESIGN

5.1 Data Flow Diagrams



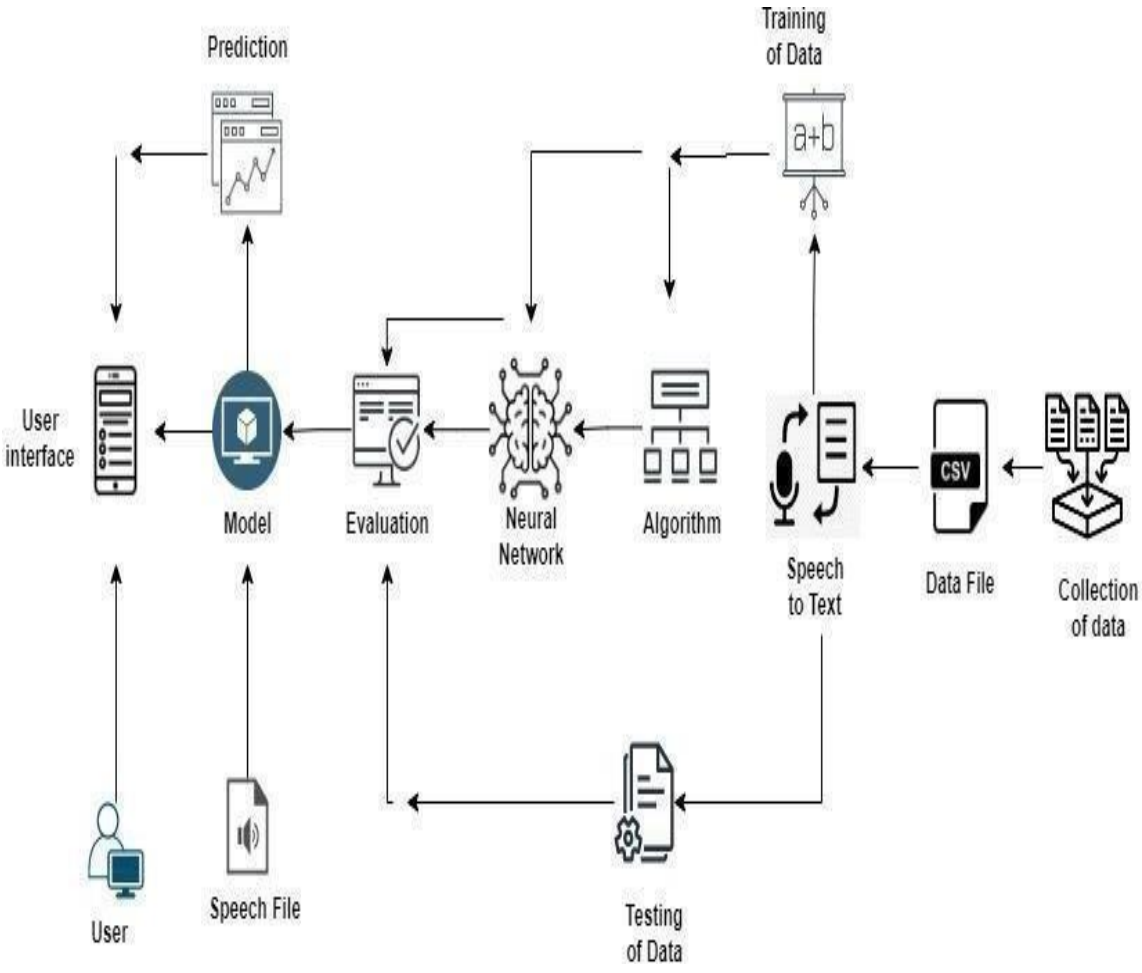
5.2 Solution &Technical Architecture

Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered

Example - Solution Architecture Diagram



Technical Architecture

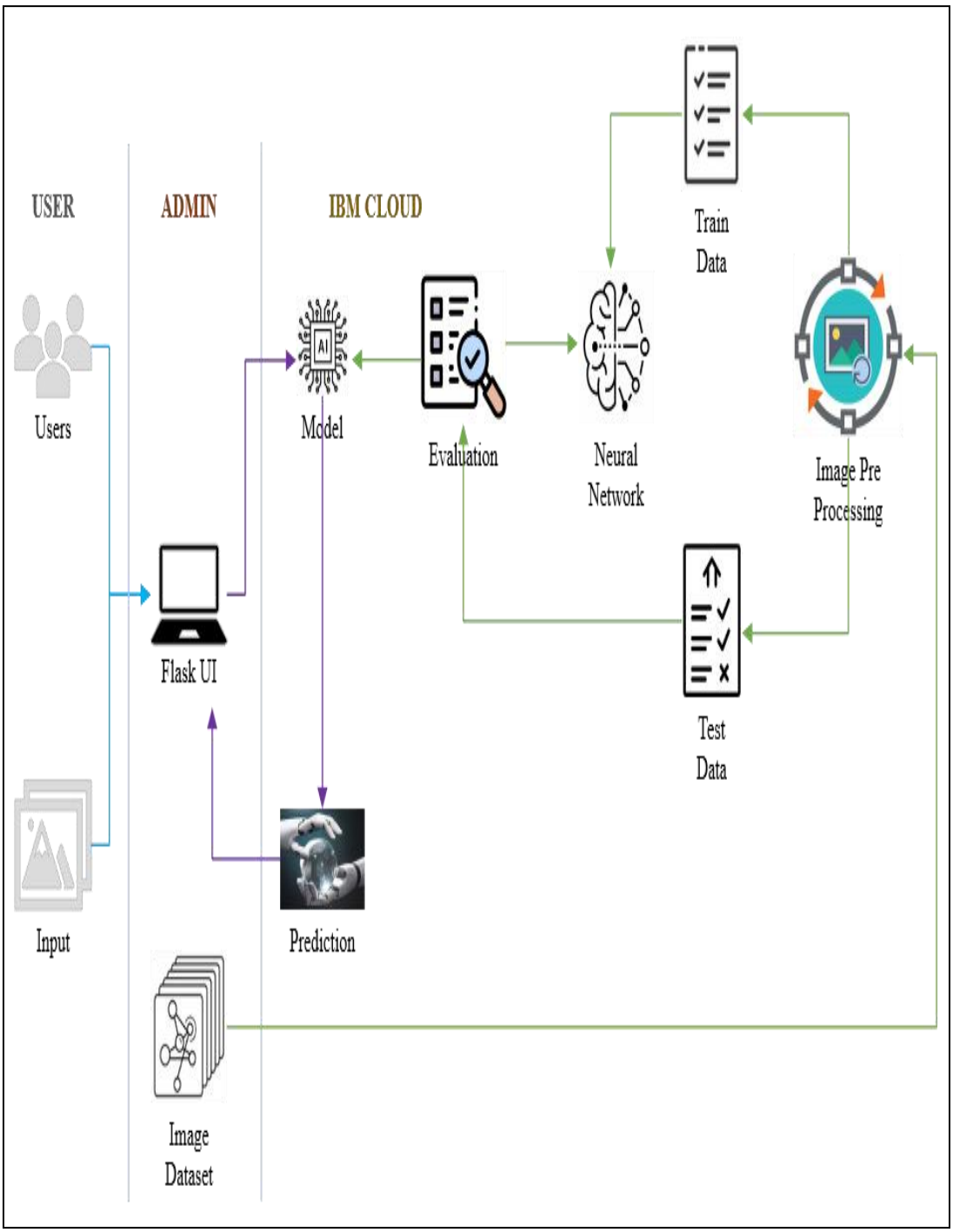


Table -1 Components & Technologies

Si. No.	Components	Description	Technology
1.	User Interface	The user interface is the point of human computer interaction and communication in device.	Python flask, HTML, CSS/JavaScript.
2.	Flash UI	Flash's user interface components let you interact with the users that use your site and gather information.	Using the cloud, it can be executed.
3.	Models	Support Vector Machine (SVM) is subsequently applied to classify our gesture image dataset.	Machine Learning.
4.	Image	Image processing is used to extract signs from the image using neural network.	ANN, CNN, Open CV.
5.	Evaluate data	Aims to estimate the generalization accuracy of a model on future (unseen/out-of-sample) data.	NLP.
6.	Unstructured data	Unstructured data is a conglomeration of many varied types of data that are stored in their native formats.	Natural Language Processing (NLP).
7.	Structured data	Typically categorized as quantitative data is highly organized and easily decipherable by machine learning algorithms.	Machine language and artificial intelligence tools.
8.	File Storage	File storage requirements to store the trained model in order to use it whenever it is needed.	IBM Block Storage or Cloud object.
9.	ML service	Provides a full range of tools and services so that you can build, train, and deploy Machine Learning models.	Python, IBM Watson.

10.	IBM Cloud	IBM Watson Studio empowers data scientists, developers and analysts to build, run and manage AI models, and optimize decisions anywhere on IBM Cloud Pak for Data	IBM Cloud and Watson Studio service
11.	Dataset	First prototype of this system used a dataset of 24 static signs from the Panamanian Manual Alphabet.	AI technology.

Table -2 Application Characteristics

Si. No.	Characteristics	Description	Technology
1.	Open-Source Frameworks	Helps you implement best practices for data automation, model tracking, performance monitoring, and model retraining.	TensorFlow.
2.	Security Implementations	It operates the largest national network of professional monitoring centres and offers a six-month, money-back guarantee to customers.	ADT type of coding.
3.	Scalable Architecture	Three-tier architecture is a well-established software application architecture that organizes applications into three logical and physical computing tiers: the data tier, the presentation tier and the application tier.	3 – Tier Architecture.
4.	Availability	The system will be made ubiquitous so that it is available everywhere.	Web Application.
5.	Performance	The model will be fine-tuned to strike a balance between accuracy vs performance.	Optimization of code and trained model.

5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint - 1
	Authentication	USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	Low	Sprint - 1
	Login	USN-3	As a user, I can log into the application by entering email & password.	I amable to get into the Dashboard	High	Sprint - 2
	Dashboard	USN-4	One place to explore all available features.	I can access my dashboard	High	Sprint - 2
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint - 1
	Authentication	USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm.	Low	Sprint - 1
	Login	USN-3	As a user, I can log into the application by entering email & password	I amable to get into the Dashboard	Low	Sprint - 2
	Dashboard	USN-4	One place to explore all available features	I can access my dashboard	Low	Sprint - 2
	Upload image	USN-5	As a user, I can upload the sign language image for translating into text format	I can be able to see the appropriate text for the sign language	High	Sprint - 3
Administrator	Manage	USN-6	Do-it-yourself service for delivering Everything.	Set of predefined requirements that must be met to mark a user story complete	High	Sprint - 4

6. PROJECT PLANNING & SCHEDULING

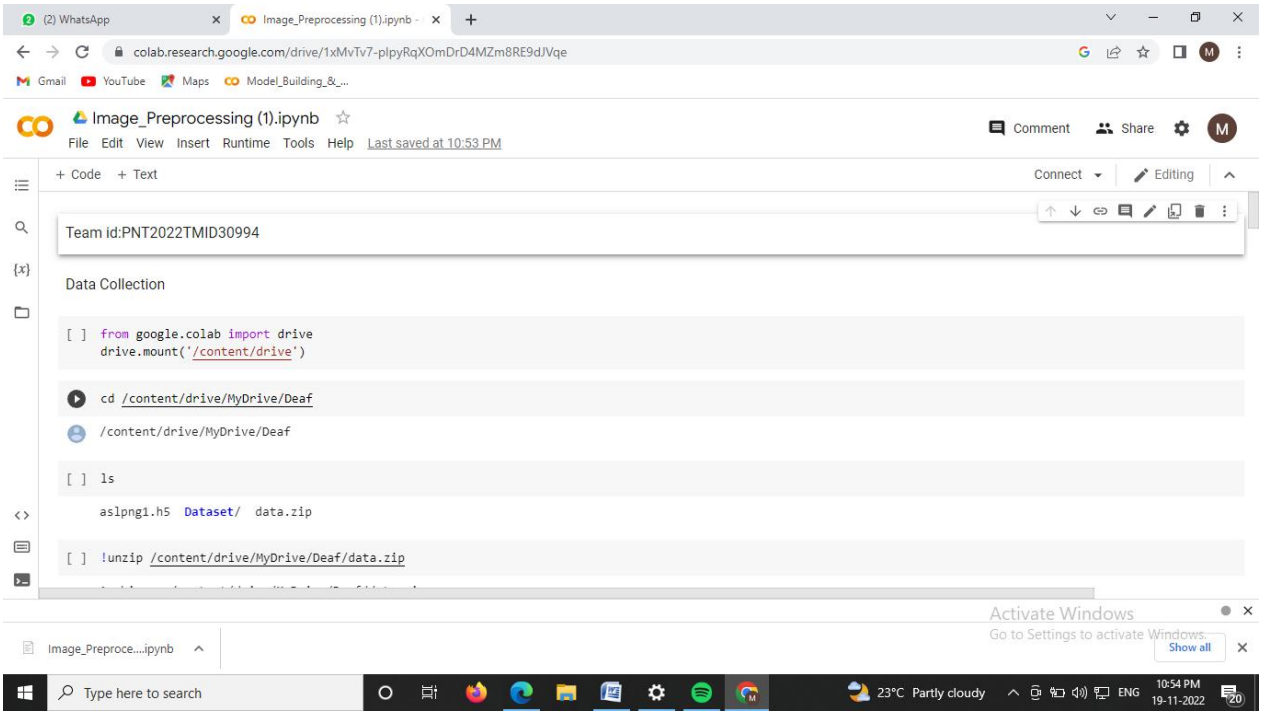
6.1 Sprint Plannig Estimation

SPRINT-1

PHASE: Data Collection and Image Processing

TASKS: Create Train and Test Folders ,Dataset, Import Image Data Generator Library And Configure it ,Apply Image Data generator Functioanality To Train and Test Set.

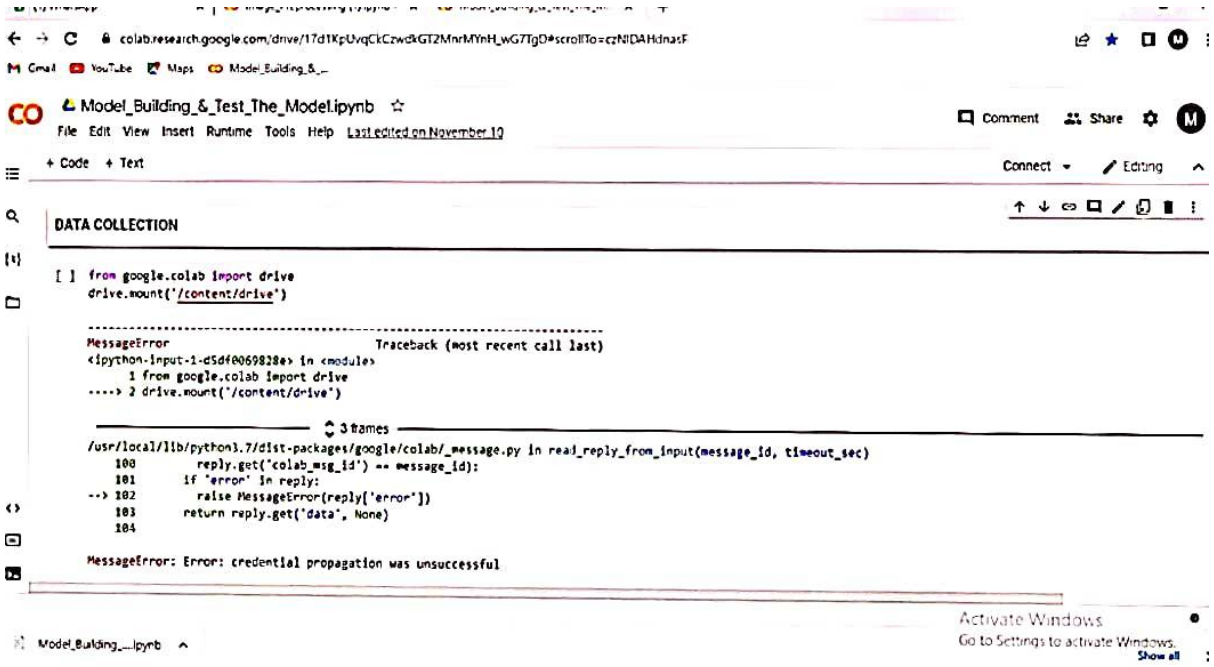
Story points: 16



SPRINT-2:

PHASE: Building and Test the Model

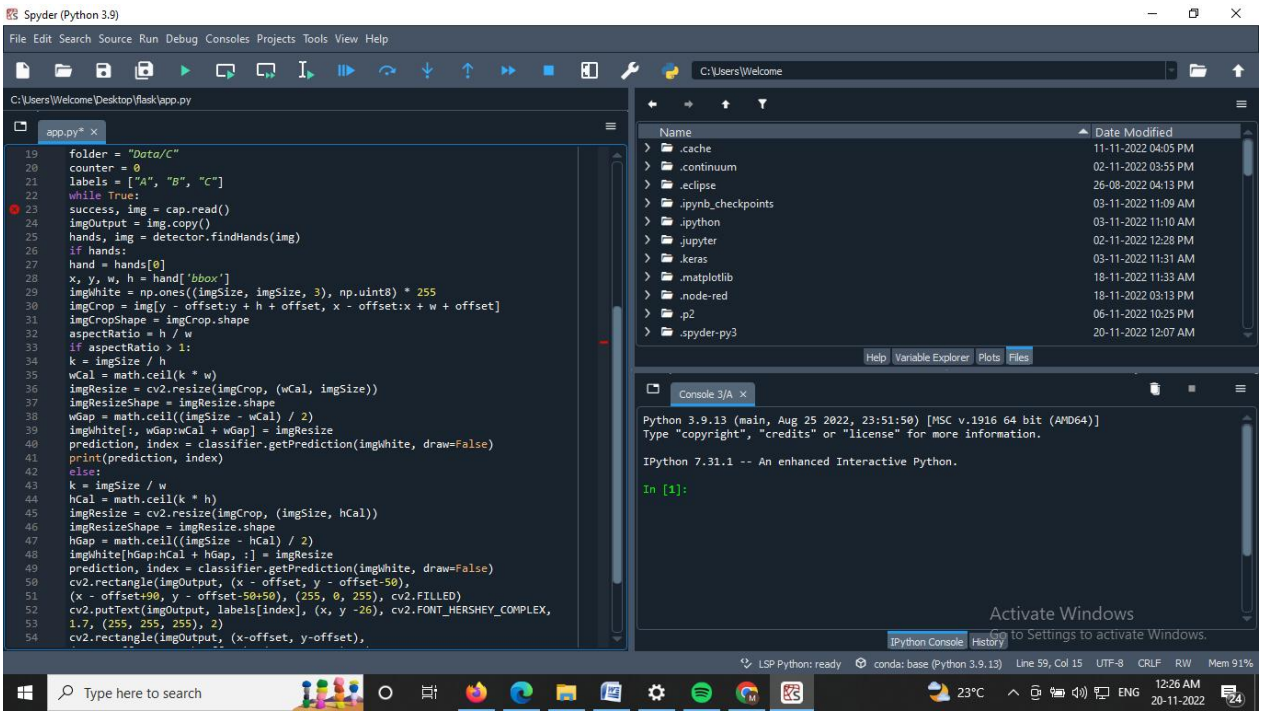
TASKS: Import the Required Model Building Libraries , Initialize The Model, Add The Convolution Layer, Add the Pooling Layer ,Add the Flatten Layer,Adding The Dense Layer , Compile The Model, Fit and Save The Model.



SPRINT-3

PHASE: Application Building

TASKS: Build a Flask Application, Flask App References ,Building flask Application –Part-3,Build The HTML Page ,Output.



SPRINT-4

PHASE: Train CNN Model on IBM

TASKS: Register For IBM Cloud ,Train The Image Classification Model

6.2 SPRINT DELIVERY SCHEDULE

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	8	6 Days	24 october 2022	29 october 2022	0	01 November 2022
Sprint-2	5	6 Days	31 october 2022	31 october 2022		
Sprint-3	4	6 Days	07 November 2022	07 November 2022		
Sprint-4	5	6 Days	14 November 2022	19 November 2022		

7. CODING & SOLUTIONING

7.1 Feature 1

- Create a image processing modules, Model Building.
- All The Models are Integrated Using models (.py)file.

7.2 Feature 2

There are seven HTML files.

- “index.html” is the main file (Homepage).
- “ouput.html” page shows the prediction and its description.
- “app.py” is used to connect frontend and backend files.

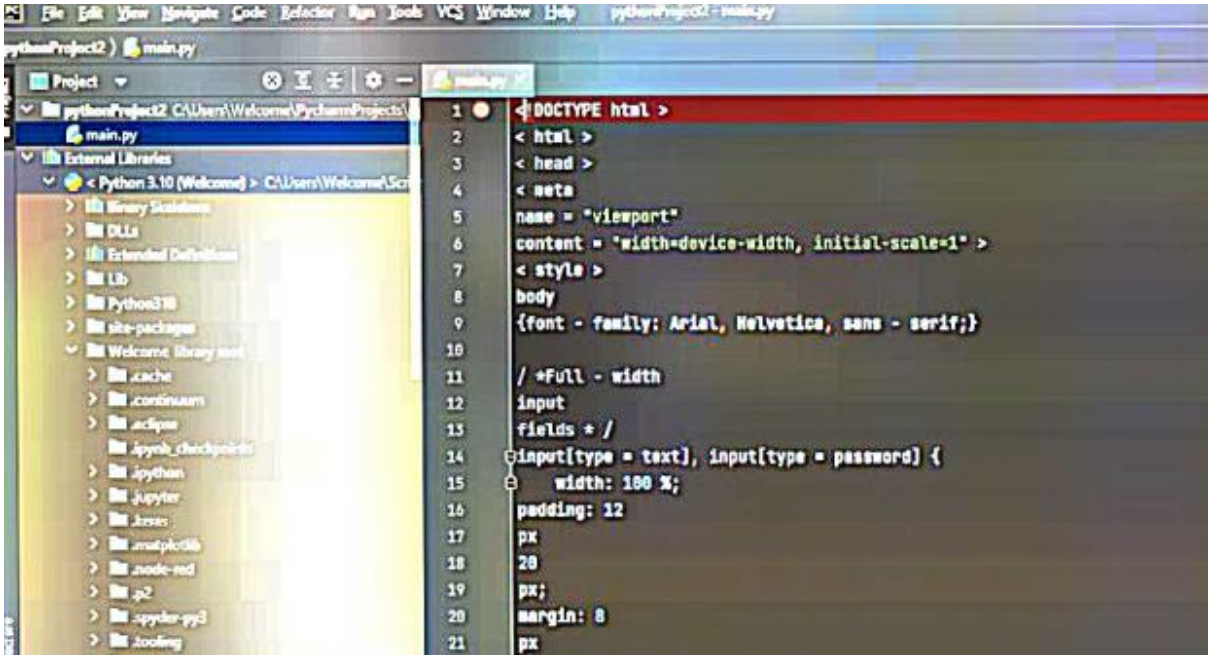


Fig 7.1 Coding & Solutioning

7.3 Database Schema

There are two database

1.Train.db

- It contains all images for real time communication of different abled Person .

2.Test.db

- It contains the testing of all image processing.

8. TESTING

8.1 Test Cases

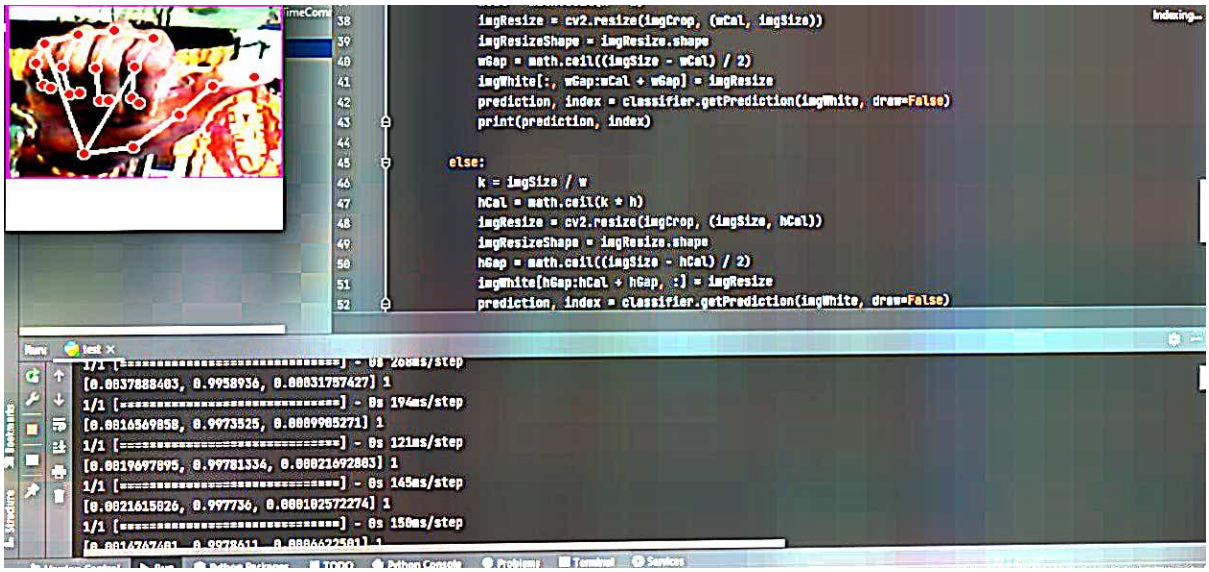


Fig 8.1 A.) Testing cases

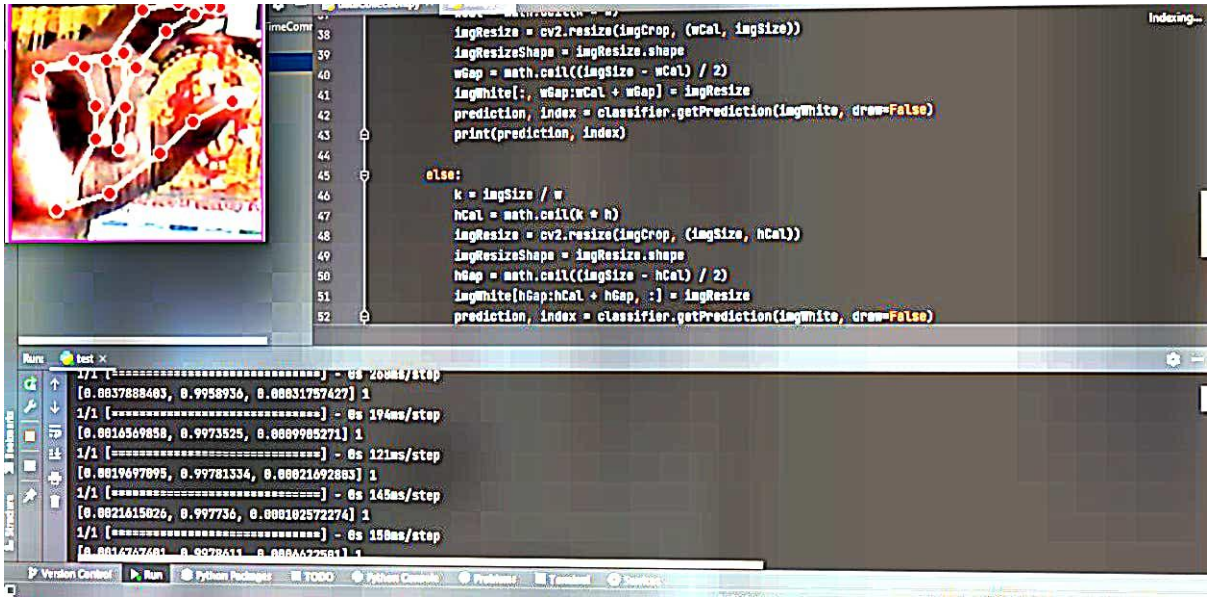


Fig 8.1 B.) Testing Cases

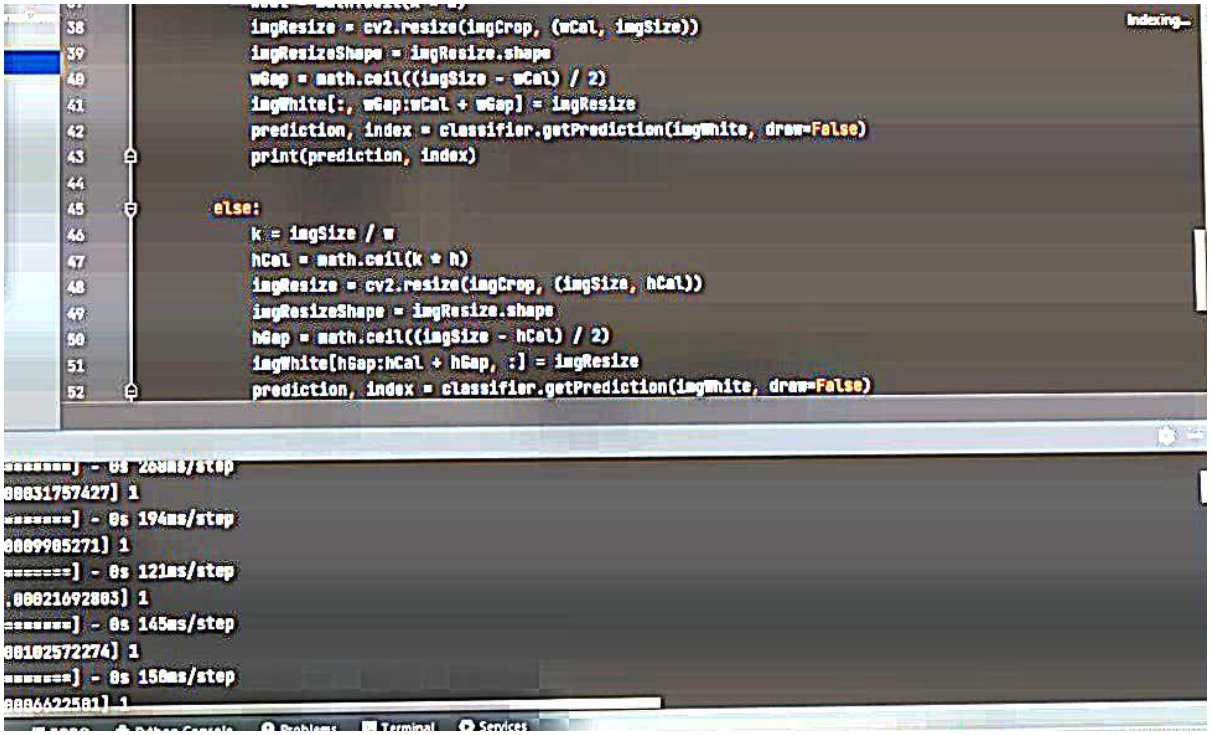


Fig 8.3 C.) Testing Cases

9. RESULTS

9.1 Performance Metrics

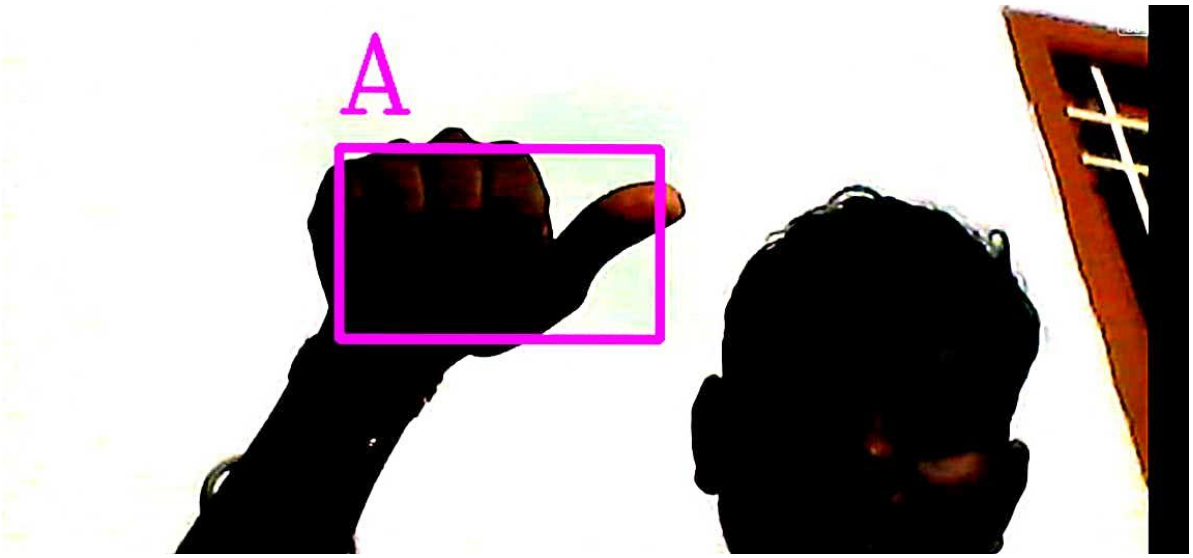


Fig 9.1 Screenshot Of The Result A

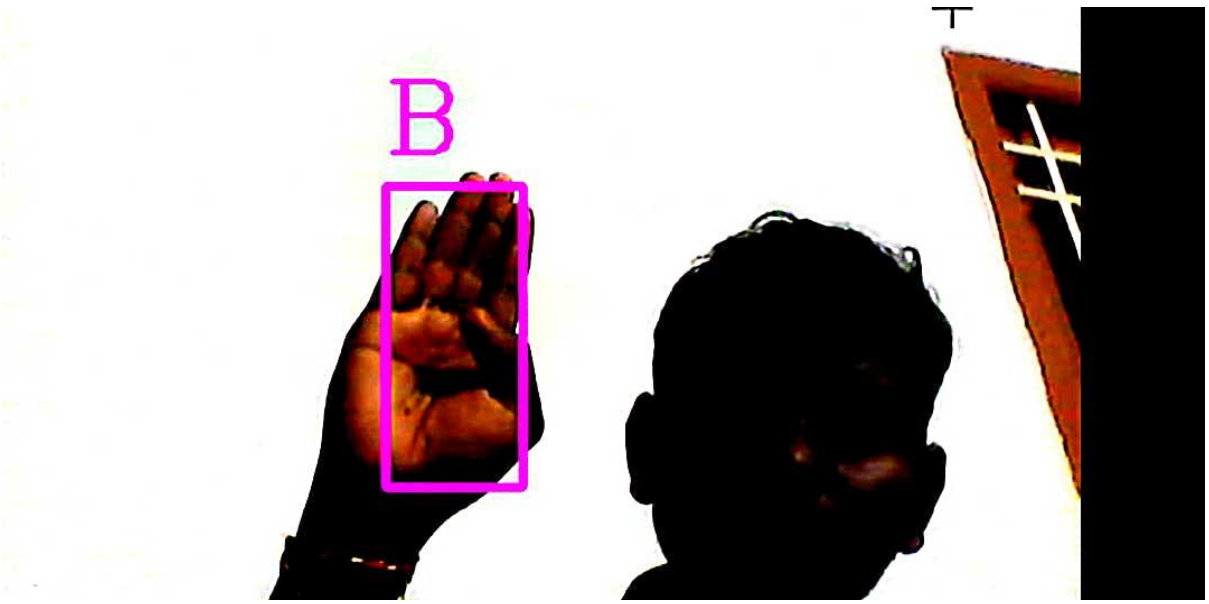


Fig 9.2 Screenshot of the Result B

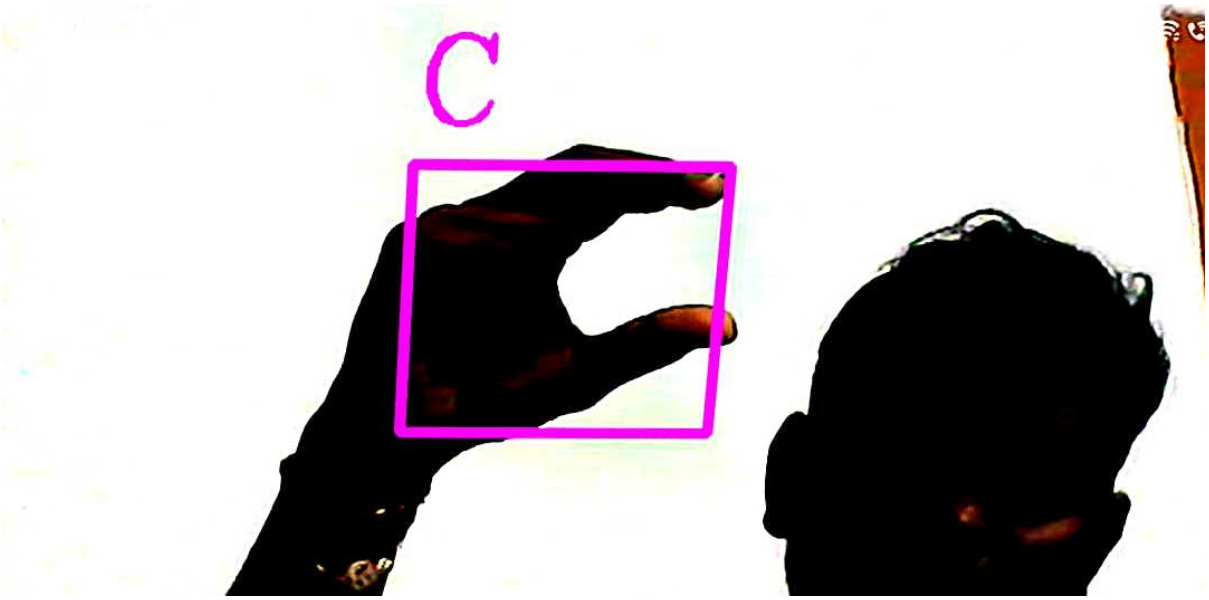


Fig 9.3 Screenshot of the result C

10. ADVANTAGES & DISADVANTAGES

Advantage

- Real Time Communication Saves TimeAnother benefit offered by real-time communication is that it can help you save time. The ability to meet with individuals remotely cuts down the amount of time you'll need to spend in larger, full-team meetings

- enables employees from across the world to communicate with each other 24×7 and share ideas or solve problems quickly. It is a cost-effective way of getting several people from different locations to attend meetings and conferences – without having to spend time or money on travel, and accommodation

Disadvantage

- The biggest disadvantage of communication is that it takes a lot of time to listen, speak, read, or write to someone. While trying to do one thing you can accidentally hurt another person's feelings by not listening or paying attention. This could result in damaging your relationship with them.

- harder to use for people of older generations.cyber-bullying and cyber-skyping made possible.not always reliable.makes communciation less personal.not always sure who you are communicating

11.CONCLUSION

Nowadays, applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted so as to perform various applications. When an image is transformed from one form to another such as digitizing, scanning, and communicating, storing, etc. degradation occurs. Therefore ,the output image has to undertake a process called image enhancement, which contains of a group of methods that seek to develop the visual presence of an image. Image enhancement is fundamentally enlightening the interpretability or awareness of information in images for human listeners and providing better input for other automatic image processing systems. Image then undergoes feature extraction using various methods to make the image more readable by the computer.Sign language recognition system is a powerful tool to preparae an expert knowledge, edge detect and the combination of inaccurate information from different sources. \the intend of convolution neural network is to get the appropriate classification.

12.FUTURE SCOPE

For any object there are many features, interesting points on the object, that can be extracted to provide a "feature" description of the object. SIFT image features provide a set of features of an object that are not affected by many of the complications experienced in other methods, such as object scaling and rotation. The SIFT approach, for image feature generation, takes an image and transforms it into a "large collection of local feature vectors". Each of these feature vectors is invariant to any scaling, rotation or translation of the image.

Scale-Space Extrema Detection. This stage of the filtering attempts to identify those locations and scales that are identifiable from different views of the same object. This can be efficiently achieved using a "scale space" function.

Keypoint Descriptor. The local gradient data, used above, is also used to create keypoint descriptors. The gradient information is rotated to line up with the orientation of the keypoint and then weighted by a Gaussian with variance of 1.5 * keypoint scale.

13.APPENDIX

Source Code

```
import cv2
from cvzone.HandTrackingModule import HandDetector
import numpy as np
import math
import time
cap = cv2.VideoCapture(0)
detector = HandDetector(maxHands=1)
offset = 20
imgSize = 300
folder = "Data/C"
counter = 0
while True:
success, img = cap.read()
hands, img = detector.findHands(img)
if hands:
hand = hands[0]
x, y, w, h = hand['bbox']
imgWhite = np.ones((imgSize, imgSize, 3), np.uint8) * 255
imgCrop = img[y - offset:y + h + offset, x - offset:x + w + offset]20
imgCropShape = imgCrop.shape
aspectRatio = h / w
if aspectRatio > 1:
k = imgSize / h
wCal = math.ceil(k * w)
imgResize = cv2.resize(imgCrop, (wCal, imgSize))
imgResizeShape = imgResize.shape
wGap = math.ceil((imgSize - wCal) / 2)
imgWhite[:, wGap:wCal + wGap] = imgResize
else:
k = imgSize / w
```

```
hCal = math.ceil(k * h)
imgResize = cv2.resize(imgCrop, (imgSize, hCal))
imgResizeShape = imgResize.shape
hGap = math.ceil((imgSize - hCal) / 2)
imgWhite[hGap:hCal + hGap, :] = imgResize
cv2.imshow("ImageCrop", imgCrop)
cv2.imshow("ImageWhite", imgWhite)
cv2.imshow("Image", img)
key = cv2.waitKey(1)
if key == ord("s"):
counter += 1
cv2.imwrite(f'{folder}/Image_{time.time()}.jpg',imgWhite)
print(counter)21
```

Flask Application:

```
from flask import Flask, render_template,Response
from flask import Flask, Response, render_template
import cv2
from cvzone.HandTrackingModule import HandDetector
from cvzone.ClassificationModule import Classifier
import numpy as np
import math
cap = cv2.VideoCapture(0)
detector = HandDetector(maxHands=1)
classifier = Classifier("Model/keras_model.h5", "Model/labels.txt")
offset = 20
imgSize = 300
folder = "Data/C"
counter = 0
labels = ["A", "B", "C"]
while True:
success, img = cap.read()
imgOutput = img.copy()
hands, img = detector.findHands(img)
if hands:
```

```
hand = hands[0]
x, y, w, h = hand['bbox']
imgWhite = np.ones((imgSize, imgSize, 3), np.uint8) * 255
imgCrop = img[y - offset:y + h + offset, x - offset:x + w + offset]
imgCropShape = imgCrop.shape
aspectRatio = h / w
if aspectRatio > 1:
k = imgSize / h
wCal = math.ceil(k * w)
imgResize = cv2.resize(imgCrop, (wCal, imgSize))
imgResizeShape = imgResize.shape
wGap = math.ceil((imgSize - wCal) / 2)
imgWhite[:, wGap:wCal + wGap] = imgResize
prediction, index = classifier.getPrediction(imgWhite, draw=False)
print(prediction, index)
else:
k = imgSize / w
hCal = math.ceil(k * h)
imgResize = cv2.resize(imgCrop, (imgSize, hCal))
imgResizeShape = imgResize.shape
hGap = math.ceil((imgSize - hCal) / 2)
imgWhite[hGap:hCal + hGap, :] = imgResize
prediction, index = classifier.getPrediction(imgWhite, draw=False)
cv2.rectangle(imgOutput, (x - offset, y - offset-50),
(x - offset+90, y - offset-50+50), (255, 0, 255), cv2.FILLED)
cv2.putText(imgOutput, labels[index], (x, y -26), cv2.FONT_HERSHEY_COMPLEX,
1.7, (255, 255, 255), 2)23
cv2.rectangle(imgOutput, (x-offset, y-offset),
(x + w+offset, y + h+offset), (255, 0, 255), 4)
cv2.imshow("ImageCrop", imgCrop)
cv2.imshow("ImageWhite", imgWhite)
cv2.imshow("Image", imgOutput)
cv2.waitKey(1)
```

GitHub:

https://github.com/IBM-EPBL/IBM-Project-47750-1660801981

Demo Link:

https://drive.google.com/file/d/1jFSYIRF1mbLAcRXxU4QR6EyXsuJrMNvd/view?usp=drivesdk