REAL-TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY ABLED

(SIGN LANGUAGE DETECTION USING MEDIA PIPE HOLISTICS)

TEAM ID: PNT2022TMID04376

Kirthiknimalan S K (737819CSR089)

Kavin M (737819CSR081)

Kavin Bharathi K C(737819CSR082)

Kavin T (737819CSR083)

TABLE OF CONTENTS

CHAPTER.NO	TITLE	PAGE.NO
1.	1. INTRODUCTION	
	1.1 Project Overview	
	1.2 Purpose	
2.	2. LITERATURE SURVEY	
	2.1 Existing problem	
	2.2 References	
	2.3 Problem statement definition	
3.	3. IDEATION & PROPOSED SOLUTION	
	3.1 Empathy map canvas	
	3.2 Ideation & brainstorming	
	3.3 Proposed solution	
	3.4 Problem solution fit	
4.	4. REQUIREMENT ANALYSIS	
	4.1 Functional requirement	
	4.2 Non-functional requirements	
5.	5. PROJECT DESIGN	
	5.1 Data flow diagrams	
	5.2 Solution & technical architecture	
6.	6. PROJECT PLANNING & SCHEDULING	
	6.1 Sprint planning & estimation	
	6.2 Sprint delivery schedule	
7.	7. CODING & SOLUTIONING	
	7.1 Feature 1	
	7.2 Feature 2	
8.	8. RESULTS	
	8.1 Performance metrics	
9.	9. ADVANTAGES & DISADVANTAGES	
10.	10. FUTURE SCOPE	
11.	11. APPENDIX	

CHAPTER 1

INTRODUCTION

1.1 PROJECT OVERVIEW

Sign language is a mode of communication that uses visual ways like expressions, hand gestures, and body movements to convey meaning. This is extremely helpful for people who face difficulty with hearing or speaking.

Sign language recognition refers to the conversion of these gestures into words or alphabets for deaf people of existing formally spoken languages or gestures into audio for blind people. Thus, the conversion of sign language into words or audio by an algorithm or a model can help bridge the gap between people with hearing or speaking impairment and the rest of the world.

We start by collecting key points from media pipe holistic and collect a bunch of data from key points. We then build an LSTM model and train with our stored data which helps us to detect action with a number of frames. Once training is done, we can use this model for real-time hand gesture detection and simultaneously convert the gesture to speech using OpenCV.

1.2 PURPOSE

Sign language is manual communication commonly used by people who are deaf. Sign language is not universal. People who are deaf from different countries speak different sign languages. The gestures or symbols in sign language are organized in a linguistic way.

The purpose of sign language recognition (SLR) systems is to provide an efficient and accurate way to convert sign language into text or voice aids for the hearing impaired for example or enable very young children to interact with computers (recognizing sign language), among others.

Signs give extra visual information about the words used in the message making it easier to understand. The extra visual cues given in signs supports the learning of new words and helps to model how and when to use them.

CHAPTER 2

LITERATURE SURVEY

2.1 EXISTING PROBLEM

Communication between specially-abled and ordinary people has always been a challenging task. Ordinary persons cannot learn the way to communicate with specially-abled persons easily.

As communication needs to be faster in order to obtain the exact meaning, the system which is to be developed needs to be faster and more accurate. This system should also convey the message at the time of emergency. The message should be transferred from one person to another without any change in the meaning.

2.2 REFERENCES

Liang and Ouhyoung proposed a sign language recognition system using Hidden Markov Model and an integrated statistical approach used in computational linguistics. Real-time continuous gesture recognition system for sign language by Rung-Huei Liang, Ming Ouhyoung intended to recognize a large set of vocabularies in sign language by recognizing constructive postures and context information. The system uses position, orientation, and motion models, in addition to the posture model, which is implemented to enhance the performance of the system.

Starner and Pentland's American sign language system could recognize short sentences of American Sign Language (ASL) with 40 vocabularies, each being attached to its part of speech, which greatly reduced the computational complexity. The feature vector was fed to a Hidden Markov Model (HMM) for recognition of the signed words. This system gracefully integrated a useful concept in computational linguistics into gesture recognition. Furthermore, Nam's system tried to recognize hand movement patterns. An HMM-based method for recognizing the space-time hand movement pattern was proposed, and 10 kinds of movement primes could be recognized successfully.

Fel’s Glove Talk focused on a gesture-to-speech interface. Moreover, a multilayer perceptron model was used in Beale and Edward’s posture recognizer to classify sensed data into five postures in ASL. To help people with disabilities, Newby worked on the recognition of the letters and numbers of the ASL manual alphabet based on statistical similarity. A simplified method, using an appropriate spline, was proposed by Watson. Gestures are represented by a sequence of critical points (Local minima and maxima) of the motion of the hand and wrist. This approach is more flexible in matching a gesture both spatially and temporally and thus reduces the computational requirement.

2.3 PROBLEM STATEMENT DEFINITION

Communication between specially abled and an ordinary people has always been a challenging task. Ordinary persons cannot learn the way of communication between specially abled persons easily. As communication needs to be faster in order to obtain the exact meaning, the system which is to be developed needs to be faster and accurate. This system should also convey the message at the time of emergency. The message should be transferred from one person to another without any change in the meaning.

CHAPTER 3

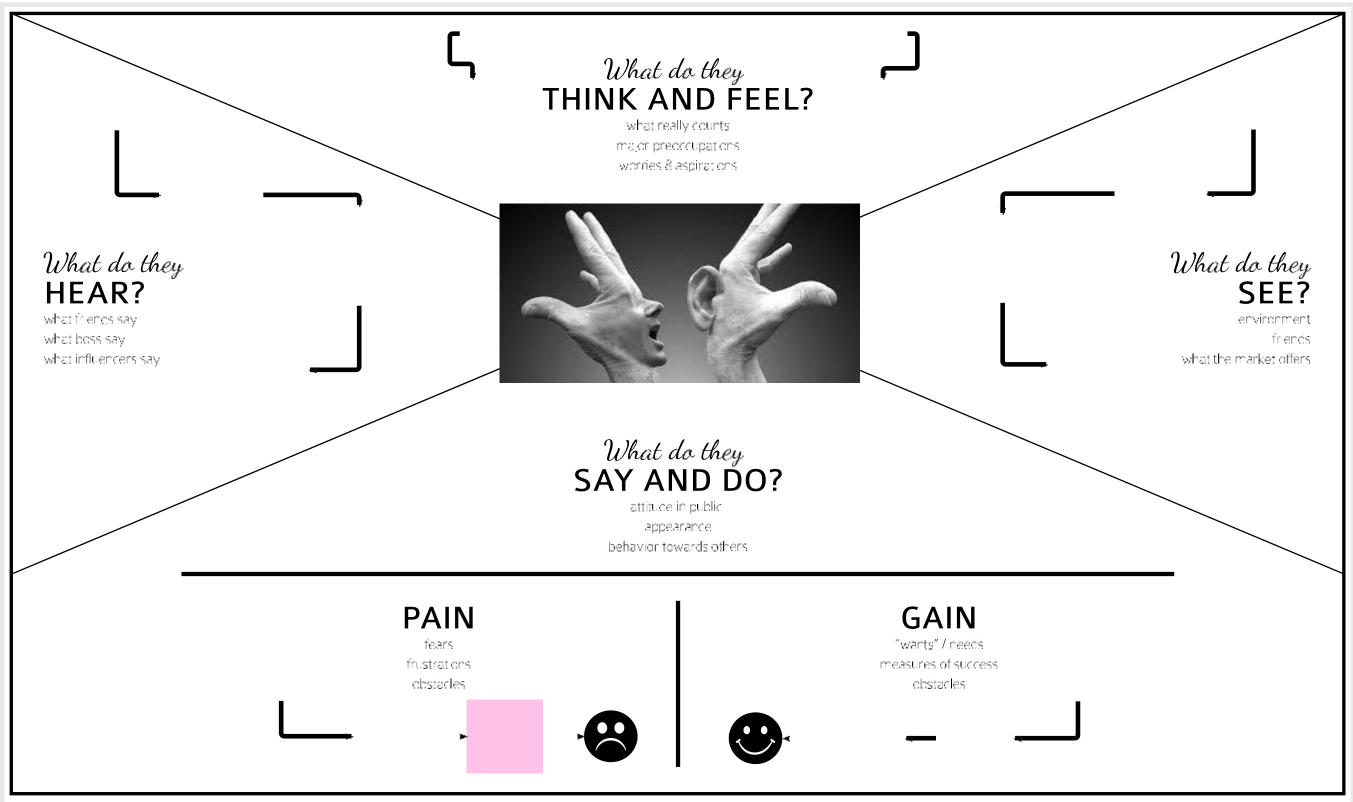
IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

An empathy map is a collaborative tool team can use to gain a deeper insight into their customers. This tool helps to understand the reason behind some actions a user takes deeply. This tool helps build empathy towards users and helps design teams shift focus from the product to the users who are going to use the product.

Deaf and mute people think how do I communicate with others? and is speaking the only way to communicate? They feel Am I a normal person. Deaf and mute people communicate using pen and paper and using proper aids (devices in the market).

Pains are Abnormal looks, Improper communication, and Socializing issues. The gains are Hassle-free communication, independent, and a Cheaper solution.



3.2 IDEATION & BRAINSTORMING

There are various ideas to implement sign language recognition.

- A study on-manual sign involves the facial region, including the movement of the lead, eye blinking, eyebrow movement, and mouth shape. This can be traced and interpreted to show communication.
- The recognition of signs with facial expressions, hand gestures, and body movement simultaneously with better recognition accuracy in real-time with improved performance helps in better communication.
- Blind people can use smart sticks to enable visually impaired people to find difficulties in detecting obstacles and dangers in front of them during walking and to identify the world around them and it acts like an artificial vision and alarm unit.
- The keyboard for the deaf feature can support the sign language images and symbols in the keyboard as a different feature to convert between the normal person language and the deaf language.
- The deaf person faces a very difficult problem to understand or identify the medicine's instructions. Idea is to prepare a sign language video have all the instructions on the medicine and what is the quantity of the medicine that should be taken by the deaf person.
- Object detection models can be used in order to specify the objects in front of the people with the positions of the objects which can be said in text/audio as per the need.

After brainstorming, selecting the best idea to propose the sign language recognition. Choosing recognition of signs with facial expression, hand gestures, and body movement simultaneously with better accuracy in real-time with improved performance.

3.3 PROPOSED SOLUTION

Problem Statement (Problem to be solved)

Communication between deaf-mute and a normal person. It is often difficult for mute person to convey their information to normal people in emergency as well as in normal times since normal people are not trained in sign language. Hence there is a need for a system which enables them to communicate with normal people.

Idea / Solution description

To develop a system that converts sign language into human hearing voice so that it can be conveyed to normal people.

Novelty / Uniqueness

Convolution neural network is used to create the model and it is trained on different hand gestures and an app is built.

Social Impact / Customer Satisfaction

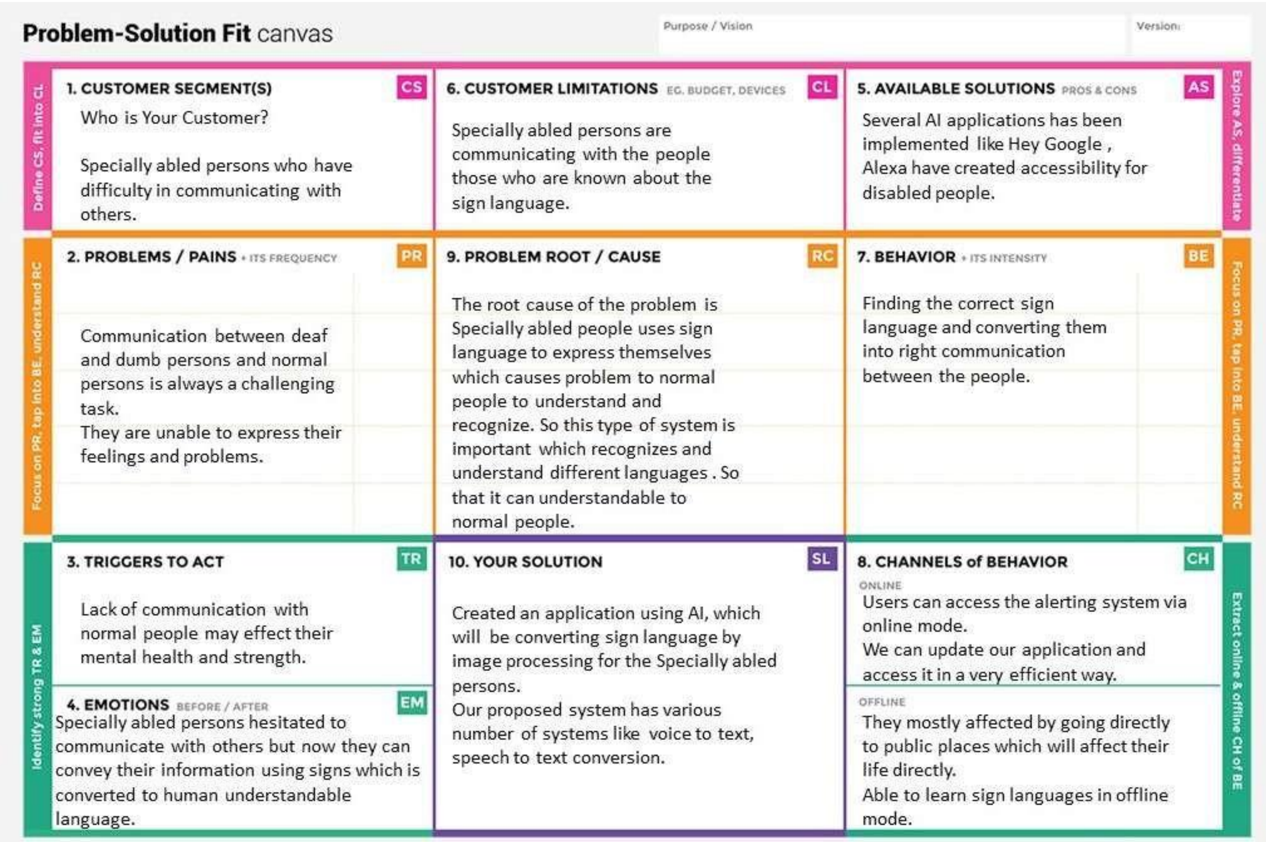
Communication process is carried without the help of additional human intervention and there is no additional hardware support needed.

Business Model (Revenue Model) And Scalability of the Solution
The app can be made available to more groups which will increase its growth.
Easy to handle and it can be accessed from any device and by everyone as it is hosted in IBM cloud.

3.4 PROBLEM-SOLUTION FIT

Communication between specially-abled and ordinary people has always been a challenging task. We take this problem and give a solution by recognizing words or sentences using sign language. This solution is extremely helpful for people who face difficulty with hearing or speaking. Hearing disabilities and speaking problems are becoming common among kids.

The recognition of signs with facial expressions, hand gestures, and body movement simultaneously with better recognition accuracy in real-time with improved performance helps in better communication. Deaf and mute people face difficulties in communicating with normal people, not being understood, and being left out of important discussions. Sign language recognition is the task of recognizing sign language glosses from video streams and the glosses are converted into audio. It can bridge the communication gap between deaf and mute people, facilitating the social inclusion of hearing-impaired people.



CHAPTER 4

REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

A functional requirement defines a system or its component. It may involve calculations, technical details, data manipulation and processing, and other specific functionality that define what a system is supposed to accomplish.

User communication – The user must know the sign language because if the user may be deaf and mute means they have to share their message into sign language. Different countries have different sign languages, so know their native sign language for collecting information from that.

User communication – The user has to communicate in front of the camera. The camera covers all the body parts because it takes input from facial expressions, hand gestures, eyebrows, and body movements.

4.2 NON-FUNCTIONAL REQUIREMENT

Non-functional requirements are a set of specifications that describe the system's operational capabilities and constraints and attempt to improve its usability, performance, scalability, and reliability.

Usability – The camera captures all expressions including facial expressions and hand gestures which can be easily used by all age groups.

Reliability – The system is very liable; it can last for long amounts of time if well maintained.

Performance – The cost-effective nature of the system makes it extremely liable and thus, efficient.

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.

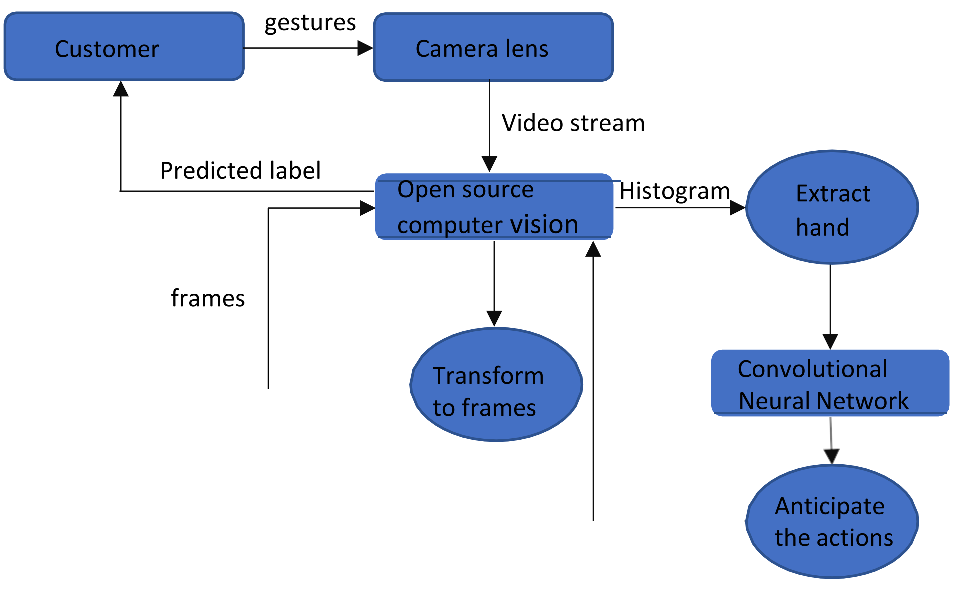
Scalability – The system gives output rapidly. It also predicts quickly when it gets so many inputs at a time. It predicts different types of sign language at a time.

CHAPTER 5

PROJECT DESIGN

5.1 DATA FLOW DIAGRAM

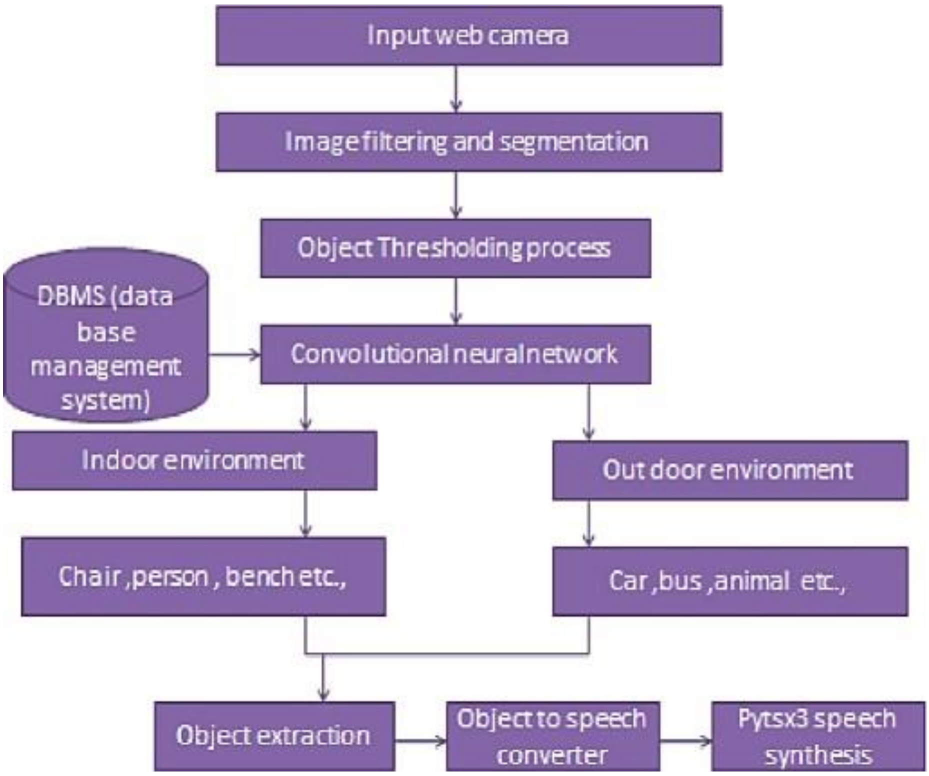
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. The data-flow diagram is a tool that is part of structured analysis and data modelling. When using UML, the activity diagram typically takes over the role of the data- flow diagram. A special form of data-flow plan is a site-oriented data-flow plan. Data flow (flow, dataflow) shows the transfer of information (sometimes also material) from one part of the system to another. The symbol of the flow is the arrow. The flow should have a name that determines what information (or what material) is being moved.



5.2 SOLUTION & TECHNICAL ARCHITECTURE

Technical Architecture (TA) is **a form of IT architecture that is used to design computer systems**. It involves the development of a technical blueprint with regard to the arrangement, interaction, and interdependence of all elements so that system-relevant
requirements are met.

Technical Architecture :



CHAPTER 6

PROJECT PLANNING & SCHEDULING

6.1 SPRINT PLANNING & ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password	2	High	4
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	4
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	4
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	4
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	4
	Dashboard					4

Sprint-3	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	4
Sprint	**Functional Requirement (Epic)**	**User Story Number**	**User Story / Task**	**Story Points**	**Priority**	**Team Members**
Sprint-3	Upload Image	USN-3	As a User, I can upload the sign language image for translating into audio format	2	High	4
Sprint-3	Solution	USN-4	As a User , If user get any query ,Then they get suggestion through help desk	2	Medium	4
Sprint-4	Manage	USN-5	Do-it-Yourself Service for delivery everything	1	High	4

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	10	6 Days	24 Oct 2022	29 Oct 2022	10	29 Oct 2022
Sprint-2	10	6 Days	31 Oct 2022	05 Nov 2022	5	05 Nov 2022
Sprint-3	10	6 Days	07 Nov 2022	12 Nov 2022	6	12 Nov 2022
Sprint-4	10	6 Days	14 Nov 2022	19 Nov 2022	6	19 Nov 2022

CHAPTER 7

CODING & SOLUTIONING

7.1 FEATURE 1

We used flask for Web UI as a user interface. We use python language for processing all the data. We start by collecting key points from mediapipe holistic and collect a bunch of data from key points. We then build an CNN model and train with our stored data which helps us to detect action. Using this model sign language is converted into text. By using google speech API is used to convert text into speech. A machine learning model is used to predict and recognize sign language.

7.2 FEATURE 2

Python and Jupyter are the technology used for creating an application. These technologies are open-source frameworks. Using HTML, CSS, and Python are used to build scalable architecture. The client layer is used for the web interface. The data layer is used for dataset storage and processing. The application layer is used to model building and training. IBM cloud used for high availability. High performance for accurate prediction of signs and less prediction time.

CHAPTER 8

RESULTS

8.1 PERFORMANCE METRICS

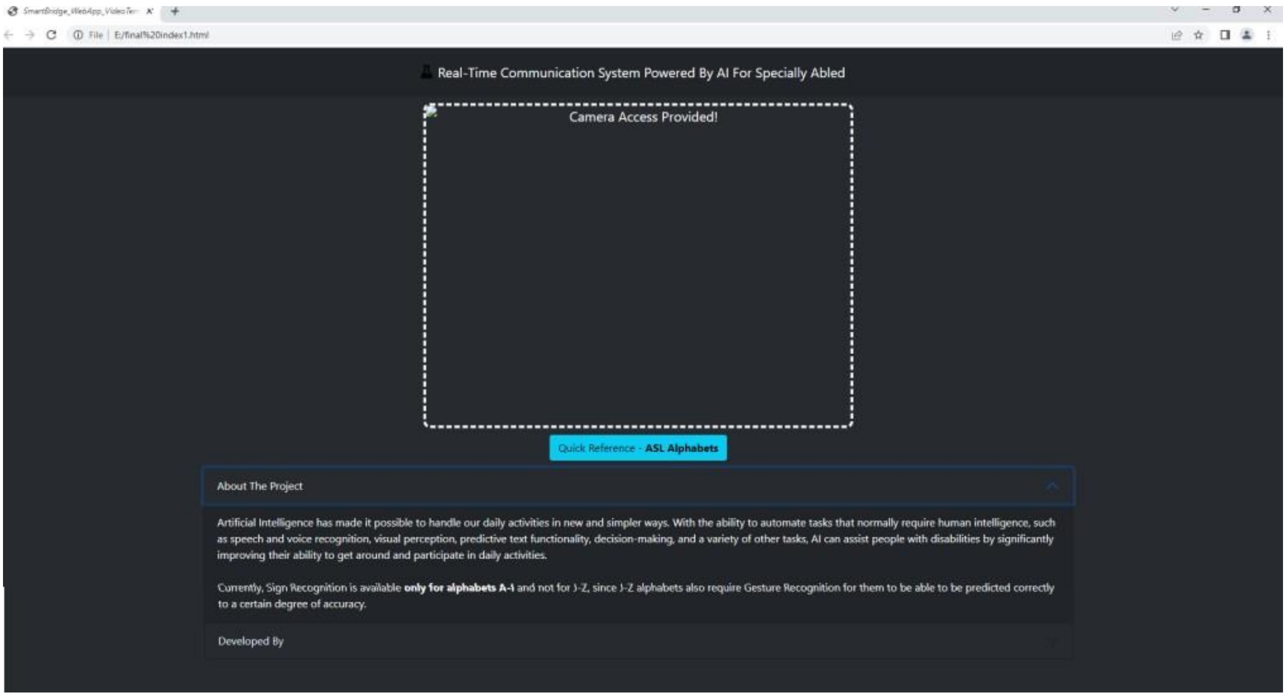
Project metrics are key indicators that help to track a project's performance. To be a successful project manager, one must monitor the team's progress and lead the efforts to the project's goals. Metrics also help to implement corrective measures in case the numbers don't align with the expectations. Accuracy describes the closeness of values to a true value – in other words, how correct they are compared to your target or goal.

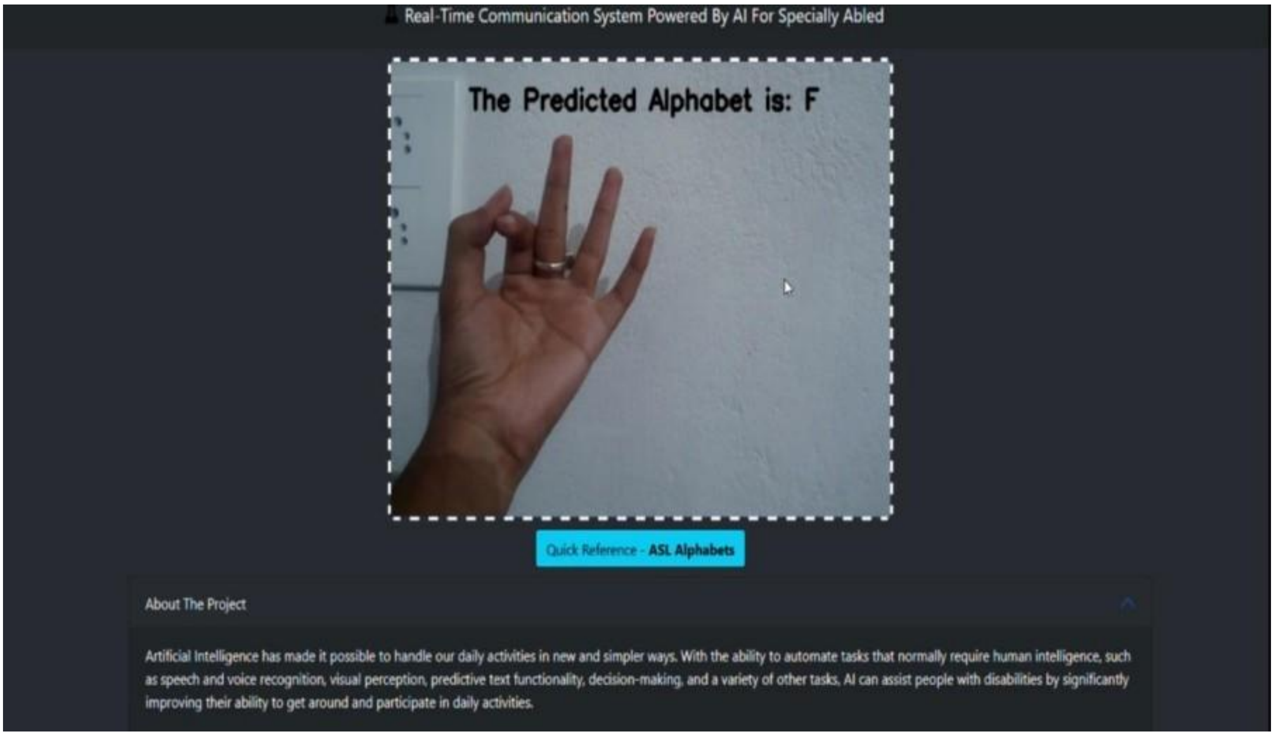
When you measure your results and find them very close to your target value, they are accurate. Accurate project estimates help identify cost and schedule requirements with relative precision, and reduce the risk of running out of time, resources, and budget during a project.

Training Accuracy -

0.9956 % Validation

Accuracy – 0.9756 %


E:/final%20index1.html
Real-Time Communication System Powered By AI For Specially Abled
Camera Access Provided!
Quick Reference - ASL Alphabets
About The Project
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.
Currently, Sign Recognition is available only for alphabets A-I 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.
Developed By


Real-Time Communication System Powered By AI For Specially Abled
The Predicted Alphabet is: F
Quick Reference - ASL Alphabets
About The Project
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.

CHAPTER 9

ADVANTAGES & DISADVANTAGES

9.1 ADVANTAGES

9.1.1 STRONGER BOND BETWEEN PARENTS AND INFANTS

Young children's inability to communicate about what they are feeling can often be a source of anger and frustration, both for the parents and the children. When a child is distraught, the ability to sign can prevent tantrums and frustration by allowing them to communicate their basic needs, whether they're hungry or in pain.

Sign language can also increase opportunities for parents and children to bond, because of the eye-to-eye and tactile contact it requires.

9.1.2 ENHANCED ABILITY TO INTERPRET BODY LANGUAGE

It involves facial expressions and body language as well as hand gestures, learning sign language could also enhance your ability to recognize and interpret body language. Body language includes a range of nonverbal signals that people use to communicate their feelings and intentions, such as posture and facial expressions.

9.1.3 EASY UNDERSTANDABLE

Nowadays normal person doesn't know the meaning of sign language, but deaf-mute people are needed to share their opinions and emotions with others or parents.

9.2 DISADVANTAGES

9.2.1 HANDS-ON SPEECH

Sign language requires the use of hands to make gestures. This can be a problem for people who do not have full use of their hands. Even seemingly manageable disabilities such as Parkinson's or arthritis can be a major problem for people who must communicate using sign language. Having a broken arm or carrying a bag of groceries can, for a deaf person, limit communication. The amount of light in a room also affects the ability to communicate using sign language.

9.2.2 DEVICE MANAGEMENT

Deaf-mute people need one device with a camera option all the time to use a sign language recognition application to share their information with others. Because that application only translating the sign language into text or audio to others who doesn’t understand the sign language. The devices like mobile phones, system, tablet etc.

CHAPTER 10

FUTURE SCOPE

Applying augmentations to the dataset can make the model training more accurate but also stabilize it at higher accuracies. Thereby depicting its caliber to make highly accurate predictions with an accuracy rate of 99%. we examined and assessed the deep learning techniques used to classify a sign language. The project aims to develop a system that converts the sign language into a human hearing voice or text in the desired language to convey a message to normal people, as well as convert speech or text into understandable sign language for the deaf and dumb. The Deaf/Dump people needs a way to communicate easily and quickly with the normal people, so that the Deaf/Dump people feel confident enough to express there thought, ideas, and can make conversation with the normal people.

Designing and implementing a system using artificial intelligence, Deep Learning algorithms and image processing concepts to take input as hand gestures (or) sign language and It generates recognizable outputs in the form of speech. The system uses neural networks and Computer vision to recognizes the video or image of sign language then smart deep learning algorithms translate it into text. As the specially abled people feel very difficult to convey their message to normal people in emergency times as well as in normal times. The main purpose of this application is to make deaf-mute people feel independent and more confident. They can participate in daily activities rather than being inactive and can get good job opportunities.

Adaptive learning platforms also provide personalized learning experiences tailored to the specific needs of students with disabilities. This application aims to help deaf and dumb by providing them with an attractive communication.

CHAPTER-11

APPENDIX

11.1 Source Code

Real-Time Communication System Powered by AI for Specially Abled Project
Image Preprocessing

Image Preprocessing Import

ImageDataGenerator Library And

Configure It from

tensorflow.keras.preprocessing.im

age import ImageDataGenerator

train_datagen=ImageDataGenerat

or(rescale=1./255,horizontal_flip=

True,vertic al_flip=True,zoom_ra

nge=0.2)

test_datagen=ImageDataGenerato

r(rescale=1./255) Apply

ImageDataGenerator

Functionality To Train And Test

Set

x_train=train_datagen.flow_from_

directory(r"C:\Users\Acer\Downl

oads\convers ation engine for deaf

and

dumb\Dataset\training_set",target

_size=(64,64),

class_mode="categorical",batch_size=30) Found 15750 images belonging to 9 classes. I x_test=test_datagen.flow_from_directory(r"C:\Users\Acer\Downloads\conversat ion engine for deaf and dumb\Dataset\test_set",target_size=(64,64), class_mode="categorical",batch_size=30) Found 2250 images belonging to 9 classes. Model Building Import The Required Model Building Libraries 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 Initialize The Model model=Sequential() Add The Convolution Layer model.add(Convolution2D(32,(3,3),activation="relu",input_shape=(64,64,3))) #No of feature detectors, size of feature detector, image size, activation function Add The Pooling Layer model.add(MaxPooling2D(pool_size=(2,2))) Add The Flatten Layer model.add(Flatten()) Adding The Dense Layers model.add(Dense(200,activation='relu')) model.add(Dense(200,activation='relu')) model.add(Dense(9,activation="softmax")) Compile The Model model.compile(loss="categorical_crossentropy",metrics=["accuracy"],optimiz er='adam') len(x_train) 525 len(x_test) 75 Fit And Save The Model Fit the neural network model with the train and test set, number of epochs, and validation steps. The weights are to be saved for future use. The weights are saved in signlanguage.h5 file using save(). model.fit(x_train,epochs=9,validation_data=x_test,steps_per_epoch=len(x_tra in),validation_steps=le n(x_test)) Epoch 1/9 525/525

[=======================

======] - 329s 616ms/step -

loss: 0.3160 - accuracy: 0.8886 -

val_loss: 0.1389 - val_accuracy:

0.9644 Epoch 2/9 525/525

[=======================

======] - 251s 478ms/step -

loss: 0.0592 - accuracy: 0.9810 -

val_loss: 0.2418 - val_accuracy:

0.9662 Epoch 3/9 525/525

[=======================

======] - 271s 515ms/step -

loss: 0.0345 - accuracy: 0.9886 -

val_loss: 0.2308 - val_accuracy:

0.9680 Epoch 4/9 525/525

[=======================

======] - 240s 457ms/step -

loss: 0.0244 - accuracy: 0.9923 -

val_loss: 0.1640 - val_accuracy:

0.9711 Epoch 5/9 525/525

[========================

=====] - 217s 412ms/step - loss:

0.0258 - accuracy: 0.9914 -

val_loss: 0.0888 - val_accuracy:

0.9769 Epoch 6/9 525/525

[=========================

=====] - 267s 509ms/step - loss:

0.0171 - accuracy: 0.9942 -

val_loss: 0.2250 - val_accuracy:

0.9782 Epoch 7/9 525/525

[========================

======] - 344s 655ms/step -

loss: 0.0139 - accuracy: 0.9955 -

val_loss: 0.1629 - val_accuracy:

0.9773 Epoch 8/9 525/525

[========================

======] - 356s 678ms/step -

loss: 0.0107 - accuracy: 0.9964 -

val_loss: 0.1430 - val_accuracy:

0.9631 Epoch 9/9 525/525

[========================

======] - 363s 692ms/step -

loss: 0.0136 - accuracy: 0.9956 -

val_loss: 0.2175 - val_accuracy:

0.9756 model.save("signlanguage-

new.h5") Test The Model Import

The Packages And Load The

Saved Model from keras.models

import load_model import numpy

as np import cv2 from

tensorflow.keras.models import

load_model from

tensorflow.keras.preprocessing

import image import numpy as np

model=load_model("signlanguage

.h5") Load The Test Image, Pre-

Process It And Predict

img=image.load_img("16.png",tar

get_size=(64,64)) img type(img)

PIL.Image.Image x =

image.img_to_array(img) x

array([[[0., 0., 0.], [0., 0., 0.], [0.,

0., 0.], ..., [0., 0., 0.], [0., 0., 0.],

[0., 0., 0.]], [[0., 0., 0.], [0., 0., 0.],

[0., 0., 0.],[0., 0., 0.], [0., 0., 0.],

[0., 0., 0.]], [[0., 0., 0.], [0., 0., 0.],

[0., 0., 0.], ..., [0., 0., 0.], [0., 0.,

0.], [0., 0., 0.]], ..., [[0., 0., 0.], [0.,

0., 0.], 61 [0., 0., 0.], ..., [0., 0.,

0.], [0., 0., 0.], [0., 0., 0.]], [[0., 0.,

0.], [0., 0., 0.], [0., 0., 0.], ..., [0.,

0., 0.],[0., 0., 0.], [0., 0., 0.]], [[0.,

0., 0.], [0., 0., 0.], [0., 0., 0.], ...,

[0., 0., 0.], [0., 0., 0.], [0., 0., 0.]]],

dtype=float32) x.shape (64, 64, 3)

x = np.expand_dims(x,axis=0)

x.shape (1, 64, 64, 3) pred_prob =

model.predict(x) 1/1

[========================

=====] - 1s 1s/step pred_prob

array([[1., 0., 0., 0., 0., 0., 0., 0.,

0.]], dtype=float32)

class_name=["A","B","C","D","E

","F","G","H","I"] pred_id =

pred_prob.argmax(axis=1)[0]

pred_id 0 print("the alphabet is

",str(class_name[pred_id])) the

alphabet is A