

**A REAL TIME COMMUNICATION SYSTEM POWERED BY AI  
FOR SPECIALLY ABLED**

**TEAM ID**

**PNT2022TMID12678**

**TEAM MEMBERS**

**R KAVYA SREE**

**D JANANI**

**P LOGA SHENEHA**

**A PRIYADHARSHINI**

## **Table Of Contents**

### **1. INTRODUCTION**

- a. Project Overview
- b. Purpose

### **2. LITERATURE SURVEY**

- c. Existing problem
- d. References
- e. Problem Statement Definition

### **3. IDEATION & PROPOSED SOLUTION**

- f. Empathy Map Canvas
- g. Ideation & Brainstorming
- h. Proposed Solution
- i. Problem Solution fit

### **4. REQUIREMENT ANALYSIS**

- j. Functional requirement
- k. Non-Functional requirements

### **5. PROJECT DESIGN**

- l. Data Flow Diagrams
- m. Solution & Technical Architecture
- n. User Stories

## 6. PROJECT PLANNING & SCHEDULING

- o. Sprint Planning& Estimation
- p. Sprint Delivery Schedule
- q. Reports from JIRA

## 7. CODING & SOLUTIONING

- r. Libraries to be installed
- s. Real time sign to speech
- t. Facial EmotionDetection
- u. Language Customization
- v. Real time speechto text

## 8. TESTING

- w. Test Cases
- x. User Acceptance Testing

## 9. RESULTS

- y. Performance Metrics

## 10. ADVANTAGES & DISADVANTAGES

## 11. CONCLUSION

## 12. FUTURE SCOPE

## 13. APPENDIX

- z. Source Code
- aa. GitHub &Project Demo Link

# **Real time communication system powered by AI for specially abled**

## **INTRODUCTION**

### **1.1 Project Overview**

The aim of this project is to create a software that does not only convert sign language into text and speech but also translates speech into sign language in real time and as quick as the person speaks. We will be using a deep learning model like CNN for this project. CNN is used for image classification and classifies the object into the respective classes and does the object detection accordingly. An app is built which uses this model. This app enables deaf and mute people to convey their information using signs which gets converted to human-understandable language and speech is given as output.

### **1.2 Purpose**

According to the times now survey, the Indian population consists of about 30 percent disabled people, and of that 20 percent are deaf and mute. The only chance of communication is the sign language but it's practically not feasible that everyone studies the sign language. Technology has risen to unprecedented rates which also comes with a leeway for the disabled people. With the help of technology, 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.

## **2. Literature Survey**

### **2.1.Existing Problem**

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 complicated 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 handy to have a proper conversation between a normal person and an impaired person in any language.

### **2.2 References**

Design of Communication Interpreter for Deaf and Dumb Person was published by Pallavi Verma (Electrical and Electronics Department, Amity University, Greater Noida, Uttar Pradesh, India), Shimi S. L (Assistant Professor, NITTTR, Chandigarh, India ), Richa Priyadarshani ( Electrical and Electronics Department, Amity University, Greater Noida, Uttar Pradesh, India).

International Journal of Science and Research (IJSR) · Jan 2013

Development of full duplex intelligent communication system for deaf and dumb people was published in the year January 2017

DOI:10.1109/CONFLUENCE.2017.7943247 At 7th International Conference on

Cloud Computing, Data Science & Engineering - Confluence (Confluence) by Surbhi Rathi Department of Information Technology, Yeshwantrao Chavan College of Engineering Nagpur, India and UjwallaGawande, Department of Information Technology Yeshwantrao Chavan College of Engineering Nagpur,India.

A Review Paper on Sign Language Recognition for The Deaf and Dumb published by R Rumana(B.E Graduate(IV year), Department of Computer Scienceand Engineering, SCSVMV,Kanchipuram) , ReddygariSandhya Rani(B.E Graduate(IV year), Department of Computer Science and Engineering,SCSVMV, Kanchipuram) , Mrs. R. Prema(Assistant Professor, Department of Computer Scienceand Engineering, SCSVMV,Kanchipuram).

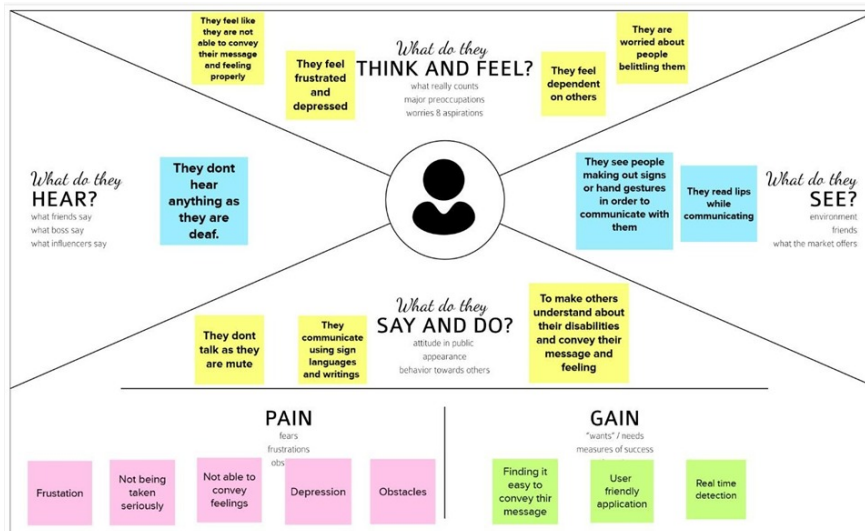
Published (First Online):01-11-2021

## **2.3 Problem Statement Definition**

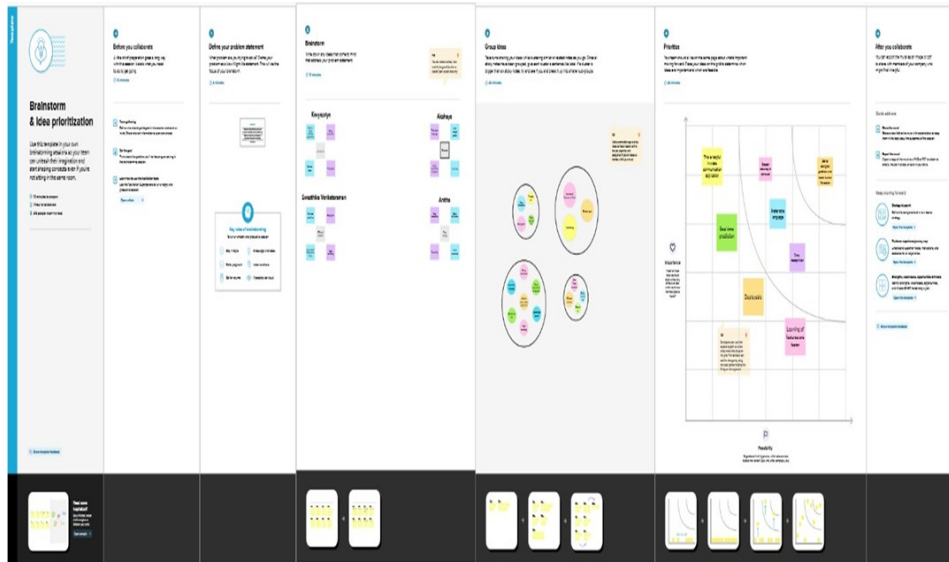
People with disabilities are a part of our society. Even though technology is constantly evolving, little is being done to improve the lives of these people. Communication with a deaf-mute person has always been difficult. Because hand sign language is not taught to the general public, it can be difficult for silent people to communicate with non-mute people. In times of crisis, they may find it difficult to communicate. When other modes of communication, such as speech,are unavailable, the human hand has remaineda popular method of information transmission. A voice conversion system with hand gesture recognition and translation will be very helpful in establishing proper communication between a normal person and a handicapped person in any language.

## 3. Ideation and Proposed Solution

### 3.1 Empathy Map Canvas



### 3.2 Ideation and Brainstorming



## 3.3 Proposed Solution

S No.	Parameter	Description
----------	-----------	-------------



1.	<p>Problem Statement (Problem to be solved)</p>	<p>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.</p>
2.	<p>Idea / Solution description</p>	<p>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.</p>

		given as output.
3.	Novelty / Uniqueness	<ul style="list-style-type: none"> <li>• Facial Emotion Detection</li> <li>• Language customization</li> <li>• User-friendly interface.</li> <li>• Greater accuracy.</li> </ul>
4.	Social Impact / Customer Satisfaction	<p>The proposed solution is keen on providing a friendly user interface and user experience. User Interface (UI) is aimed to be developed in such a that way that it can be very handy and easy to learn. The system is also aimed to be light weight which would make the system provide faster and accurate results and hence it provides a better User Experience (UX).</p>



## 4. Requirement Analysis

### 4.1 Functional Requirement

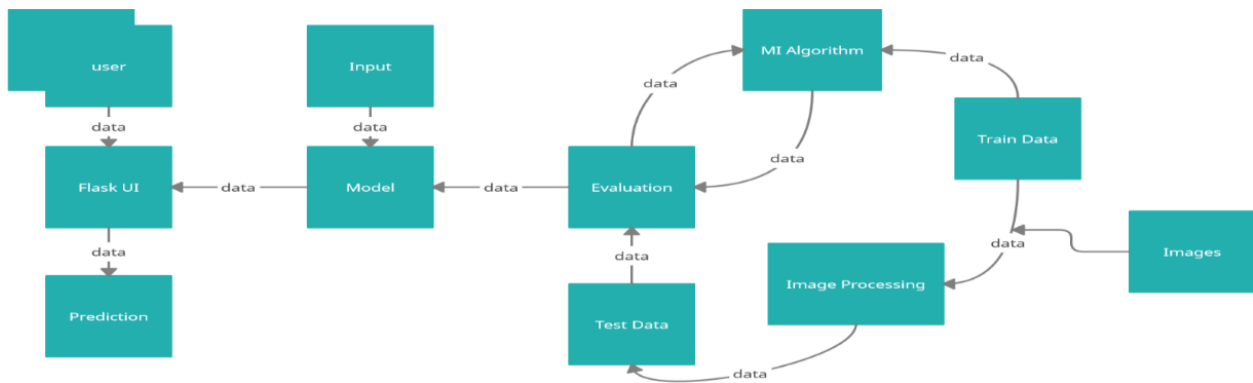
FR No.	Functional Requirement(Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Language customization	The user performs language customization.
FR-2	User Options	The user either chooses to convert speech to sign language and sign language to speech.
FR-3	Test Inputs	The real time video and audio data is collected and fed into the machine learning model.
FR-4 x	Result	The conversion will take place simultaneously and will be displayed on the screen.

## 4.2 Non-Functional Requirements

<b>FR No.</b>	<b>Non-Functional Requirement</b>	<b>Description</b>
NFR-1	<b>Usability</b>	The user will have access to all the resources present in that website.
NFR-2	<b>Security</b>	User information is protected.
NFR-3	<b>Reliability</b>	It offers accurate results.
NFR-4	<b>Performance</b>	The web application makes use of light weight model hence the result will be accurate and fast.
NFR-5	<b>Availability</b>	The web application can be accessed 24/7 from anywhere when connected to the internet.
NFR-6	<b>Scalability</b>	The trained ML model can provide accurate results whenever the size of the dataset and the number of users is extended.

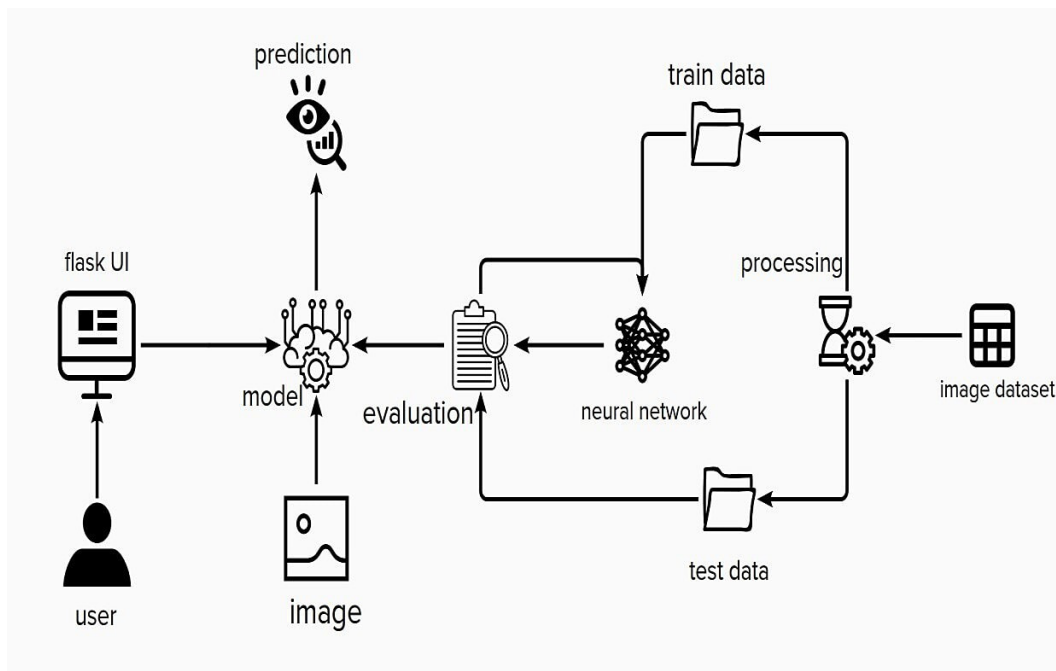
## 5. Project Design

### 5.1 Dataflow Diagram

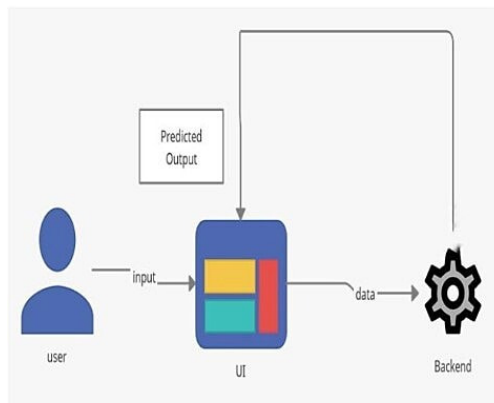


### 5.2 Solution and Technical Architecture

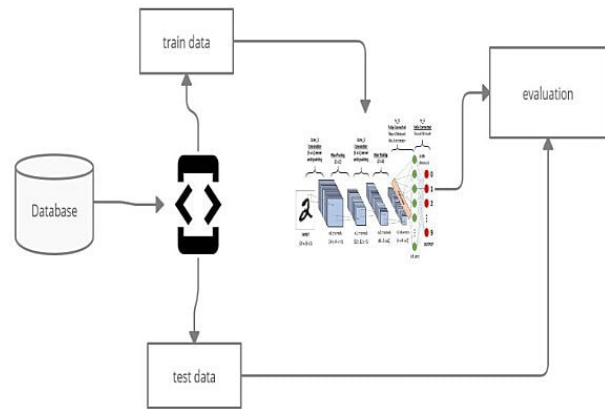
#### Solution Architecture



## Technical Architecture



DEPLOYMENT



TRAINING AND EVALUATION

### 5.3 User Story

User	Functional Requirement (Epic)	User Story Number	User Story/ Task	Acceptance criteria	Priority	Release
Customer	Uploading the real time data.	USN-1	The user will be presented with two options. 1. Speech to sign language conversion. 2. Sign language to speech conversion.	They can access the portal	High	Sprint-1
		USN-2	Language selection	They can access the portal	Low	Sprint-1

		USN-3	The deaf-mute person will choose the speech to sign language conversion which would take them into a portal that collects the realtime data (sign language recognition) and converts it into speech simultaneously.	Video processing	High	Sprint-2
		USN-4	Emotion detection	Video processing	Medium	Sprint-1
		USN-5	Normal person would choose speech to sign language which would take them into a portal where their speech is converted into sign language simultaneously.	Video and audio processing	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	Choose Available Options	USN-1	The user will be presented with two options. 1. Speech to sign language conversion. 2. Sign language to speech conversion	1	Low	Janani. D
Sprint-1	Dashboard	USN-2	Go to dashboard and see the available features	1	Low	
Sprint-2	Language selection	USN-3	The user can select any one of the	2	Medium	Loga sheneha.P

			available options according to their requiremen t.			
Sprint-3	Convert from one language toanother	USN-4	The deaf-mute person will choose the speech to sign language conversion which would take them into a portal that collects the real time data (sign language recognition) andconverts it into speech simultaneous ly	2	High	Kavya sree R

Sprint	Functional Requirement(Epic)	User Story Number	User Story /Task	Story Points	Priority	Team Members
Sprint-4	Emotion detecton	USN-5	By processing the video it detects the emotion ofthe user.	2	High	Priyadharshni A
Sprint-1	Exit	USN-6	Click exit button to exit from theapplication	1	Low	Janani D

## 6.2 Sprint Delivery Schedule

### 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 iterationunit (story pointsper day)

### Burndown Chart:

A burndown chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time

## **7 . Coding and Solutioning**

### **7.1 Libraries to be installed (feature 1)**

```
pip install fer  
pip install  
flaskpip  
installcv2  
pip install  
numpypip  
installkeras  
pip install  
tensorflowpip  
installcvzone pip  
installpytsx3  
pip install scikit-image
```

### **7.2 Real time sign to speech(feature 2)**

Sign language is generally used by the people who are unable to speak, for communication. Most people will not be able to understand the Universal Sign Language (unless they have learnt it) and due to this lack of knowledge about the language, it is very difficult for them to communicate with mute people. A device that helps to bridge a gap between mute persons and other people forms the crux of this project. Our system makes use of a model build using CNN that is capable of detection sign languages real time.

### **7.3 Facial Emotion Detection**

Our system makes use of the FER model. Facial Emotion Recognition (commonly known as FER) is one of the most researched fields of computer vision till date and is still in continuous evaluation and improvement. The model is a convolutional neural network with weights saved to HDF5 file in

the `data` folder relative to the module's path. It can be overridden by injecting it into the `FER()` constructor during instantiation with the `emotion_model` parameter.

## Language Customization

Google Translate is a free multilingual machine translation service. It can translate the Website's text content from one language to another. It offers a huge list of languages to translate and has an efficient, reliable and easy way to translate the webpage in whatever language the user wants. It supports over 100 languages. Use this website translator to convert webpages into your choice of language.

## 8. Testing

### 8.1 Test Cases

- Verify if user can see the options when user clicks the URL
- Verify if the UI elements are getting displayed properly
- Verify if the user can choose any languages
- Verify if the user is getting redirected to the sign to speech page
- Verify if the application can convert the sign to speech
- Verify if the user can exit the sign to speech page
- Verify if the user is getting redirected to the speech to sign page
- Verify if the UI elements are being displayed
- Verify if the application can convert speech to text on clicking voice to text button.
- Verify if the user can exit the speech to sign page.

## 8.2 UAT Testing

### Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	11	7	4	2	24
Duplicate	1	0	2	0	3
External	2	3	2	1	8
Fixed	10	5	3	14	32
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	1	0	0	0	1
Totals	25	15	13	18	71

## 10 .Advantages and Disadvantages

### Advantages:

- i. Real time sign to speech detection.
- ii. Model provides good accuracy.
- iii. Real time facial emotion detection.
- iv. Language Customization.
- v. Real time speech to text conversion.
- vi. Friendly UI
- vii. Data privacy

### Disadvantages:

- viii. At times the website may lag.
- ix. Model is not tested on a wide set of data set, having all the signs.
- x. Sign language customization feature is not available.
- xi. User cannot take notes while using the app.
- xii. User cannot make calls using the app.

## **11. Conclusion**

Communication is crucial for self-expression. Additionally, it meets one's necessities. Effective communication is necessary for career advancement. Effective communication skills can make your personal life easier and improve your interactions with others by facilitating mutual understanding. A system that translates speech into acceptable sign language for the deaf and dumb has been developed as part of our project. It also translates sign language into a human hearing voice to communicate with average people. A convolution neural network has been used to build a model that is trained on various hand motions. Utilizing this concept, an app is created. Through the use of signs that are translated into speech and human-understandable English, this software aids deaf and dumb individuals to communicate easily.

## **12. Future Scope**

The following are the features that can be added in our application:

- i) A communication app can be built with the same set of features. The user can choose the appropriate mode (speech to sign or sign to speech) and accordingly the real time detection would take place on both the endusers' application.
- ii) The accuracy of the model shall be increased.
- iii) Customization of languages shall be added.
- iv) Users shall be allowed to write notes while on call.
- v) Customization of signs can also be added as a feature.

## **13 APPENDIX**

### **GITHUB LINK**

<https://github.com/IBM-EPBL/IBM-Project-26692-1660033657>





