REVA UNIVERSITY
Bengaluru, India

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

A PROJECT REPORT
ON

"FULL DUPLEX MODE OF COMMUNICATION BETWEEN IMPAIRED PEOPLE AND OTHERS"

Submitted in fulfillment of the requirements for the award of the Degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING

Submitted by

Mahendra Nath Reddy E	(R18CS483)
Dinesh B	(R18CS119)
Chethan Kumar S P	(R18CS481)
Vishakh V Badami	(R18CS539)

Under the guidance of

Prof. Kanaiya V Kanzaria
Asst. Prof., School of CSE
REVA University

May 2021 – 2022

Rukmini Knowledge Park, Kattigenahalli, Yelahanka, Bengaluru-560064

www.reva.edu.in

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

A PROJECT REPORT

ON

"FULL DUPLEX MODE OF COMMUNICATION BETWEEN IMPAIRED PEOPLE AND OTHERS"

Submitted in fulfillment of the requirements for the award of the Degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

Mahendra Nath Reddy E	(R18CS483)
Dinesh B	(R18CS119)
Chethan Kumar S P	(R18CS481)
Vishakh V Badami	(R18CS539)

Under the guidance of

Prof. Kanaiya V Kanzaria

Asst. Prof., School of CSE

REVA University

May 2021 – 2022

Rukmini Knowledge Park, Kattigenahalli, Yelahanka, Bengaluru-560064

www.reva.edu.in

DECLARATION

We, **Mr. Mahendra Nath Reddy E**(R18CS483), **Mr. Dinesh B (**R18CS119), **Mr. Chethan Kumar S P** (R18CS481), **Mr. Vishakh V Badami** (R18CS539), students of B. Tech, belongs to School of Computer Science and Engineering, REVA University, declare that this Project Report entitled "**Full Duplex Mode of Communication Between Impaired People and Others**" is the result the of project work done by us under the supervision of **Prof. Kanaiya V Kanzaria**, School of CSE, REVA University.

We are submitting this Project Report in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** by the REVA University, Bengaluru during the academic year 2021-22.

We declare that this project report has been tested for plagiarism and has passed the plagiarism test with the similarity score less than 25% and it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

We further declare that this project report or any part of it has not been submitted for award of any other Degree / Diploma of this University or any other University/ Institution.

(Signature of the Students)
Signed on

Mahendra Nath Reddy E	(R18CS483)
Dinesh B	(R18CS119)
Chethan Kumar S P	(R18CS481)
Vishakh V Badami	(R18CS539)

*Certified that this project work submitted **Mahendra Nath Reddy E** (R18CS483), **Dinesh B (**R18CS119), **Chethan Kumar S P** (R18CS481), **Vishakh V Badami** (R18CS539), has been carried out under my guidance and the declaration made by the candidate is true to the best of my knowledge.*

Signature of Guide
Date:

Signature of Director of School
Date:
Official Seal of the School

REVA UNIVERSITY
Bengaluru, India

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the project work entitled "**Full Duplex Mode of Communication Between Impaired People and Others**" carried out under my guidance **Mahendra Nath Reddy E** (R18CS483), **Dinesh B (**R18CS119), **Chethan Kumar S P** (R18CS481), **Vishakh V Badami** (R18CS539), bonafide students of school of CSE, REVA University during the academic year 2021-22, is submitting the project report in partial fulfillment for the award of **Bachelor of Technology In Computer Science Engineering** during the academic year **2021–22**. The project report has been tested for plagiarism and has passed the plagiarism test with the similarity score less than 25%. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said Degree.

Signature with date

Prof. Kanaiya V Kanzaria
Guide

Signature with date

Dr. Ashwinkumar U M
Director

External Examiner:

SL. No	Name of Examiner & Affiliation	Signature & Date
1		
2		

ACKNOWLEDGEMENTS

The euphoria and satisfaction of this project will be incomplete without thanking the personalities responsible for this venture, which otherwise would not have become a reality.

We are grateful to **Dr P. Shyama Raju**, The Founder and honorable Chancellor, REVA University, Bangalore, for his encouragement.

We are grateful to **Dr M. Dhanamjaya**, Vice-Chancellor REVA University, Bangalore, for his extended support.

We are grateful to **Dr Ashwinkumar U M**, Director, School of Computer Science and Engineering, REVA University, Bangalore, for his valuable support and encouragement.

We express our profound gratitude to our project guide, **Prof. Kanaiya V Kanzaria,** Asst. Professor, School of Computer Science and Engineering, REVA University, Bangalore, for his inspirational guidance, encouragement, support, and ongoing mentoring during the course of work.

We would also like to thank all my faculty members without whose support this project work would have been incomplete. We would like to thank one and all who directly or indirectly helped me in completing the project successfully.

Students Name &SRN:

Mahendra Nath Reddy E (R18CS483)
Dinesh B (R18CS119)
Chethan Kumar S P (R18CS481)
Vishakh V Badami (R18CS539)

Table of Contents

LIST OF FIGURES

ABSTRACT

We hope to solve and demonstrate a real-world application that allows mute people to communicate more readily with this system. The most common expressions used in communication between citizens are speech and gestures. Deaf-Dumb persons utilize gestures to communicate, but they find it difficult and unpleasant to do so with those who are unfamiliar with signing symbols. As a result, the communication channel gap becomes a task. And to bridge this gap a reliable system is needed. In this study, we propose an application for converting Indian signing symbols into text and then to speech, we used Convolutional Neural Network in this system. This system can be used for reverse conversion also i.e., Spoken English phrases can be translated into Indian Sign gestures.

Keywords: CNN, Gestures, Communication, Indian Sign Language.

CHAPTER 1

INTRODUCTION

Sign Language is a language incorporating non-verbal correspondence that takes its origin in visual cues and hand gestures. Sign Language is composed of various gestures formed by shapes of the hand, its movements and orientations. Presently, sign is the primary language of those people who have hearing aid or who are hard of hearing and used by those who can hear but cannot physically speak.

It is a complex but complete language which involves and is not limited to the movement of hands, facial expressions, and postures of the body. Sign language is not universal. Every country has an unique sign language. Each sign language has its own rules of grammar and word orders. The problem arises when the differently abled try to communicate using sign language with the people who are unaware of this language and its grammar.

Sign Language recognition relies on gestures. The strategy for communication between incommunicative and general folks is signing, which is an abstract entity that will be mapped to language communication whose origin depends on the "Sign" or "Gestures." There are many one-of-a-kind signal languages used across the world, every with its personal vocabulary. This is due to the fact signal languages have been evolved certainly with the aid of using human beings belonging to one-of-a-kind ethnic groups.

Examples of diverse sign languages include Indian Sign Language (ISL), American Sign Language in North America (ASL), British Sign Language in the United Kingdom (BSL), and South Africa. Sign language (SASL).

The Indian Sign Language has a low literacy rate, which is thought to be due to the following factors:
1) Insufficient ISL interpreters

2) The ISL device is not available.

3) There aren't enough studies on Indian Sign Language.

Hard-of-hearing people use Indian Sign Language (ISL), a visible language focused mostly on gestures and motion, as their major mode of communication. Indian Sign Language is the most widely used sign language among the Indian subcontinent's deaf community.

In addition to fingerspelling, ISL includes gestures at the sentence level. For sentences, which does not have direct sign language conversion, for which the signer is unaware of the movements, and for interpretation of a particular word, fingerspelling is utilized.

So, it becomes necessity to develop an automatic Sign Language translator to minimize the gap in communication. Sometimes, the people with hearing and speech impediment seek the help of an intermediary sign language interpreter, a person proficient in both sign and regular speech, to translate their thoughts to common people and vice versa.

However, this way turns out to be very costly and does not work out all the time. Thus, the need to introduce a system which can automatically recognize the sign language gestures becomes a priority. Introducing such a system would significantly bridge the gap between differently abled and people who use speech as a form of communication in the society.

Hand gestures in Indian Sign Language are classified into two groups:
They are:

- Manual (hand shape, hand location, orientation and movements)
- Non-manual (facial expression, eye gaze and head/body posture)

In ISL, there are one handed and two-handed signs which can be static and dynamic. Fig. 1 shows ISL manual alphabet set.

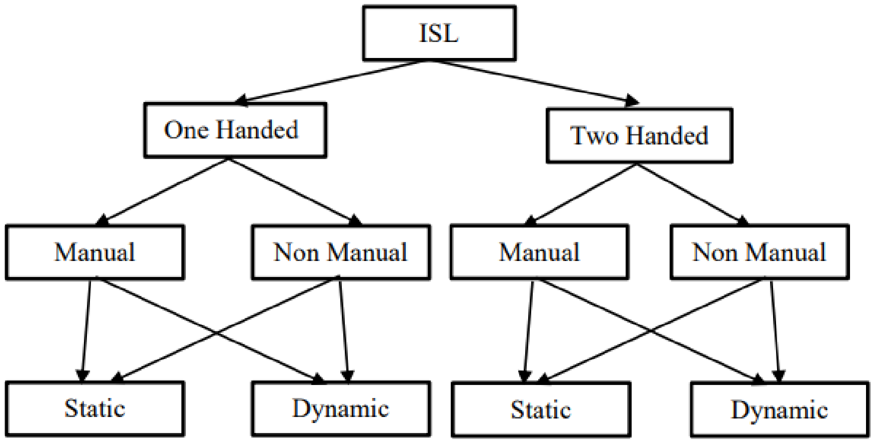


Fig 1: A Classification of ISL

The English alphabets of static Indian Sign Language hand gestures (A-Z) are showed in Fig. 2.

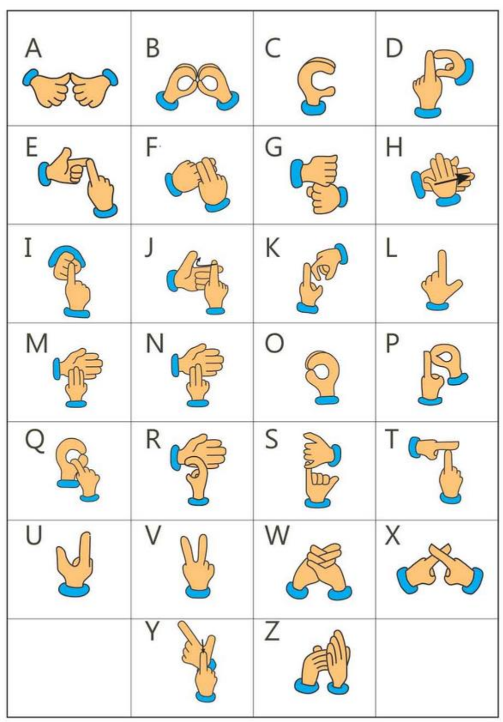


Fig 2: Poster of the Manual Alphabet in Indian Sign Language (ISL).[14]

People with speech and hearing disabilities tend to isolate themselves and feel lonely because they have difficulty talking to others. This has a significant impact on their lives. Because of the above-mentioned difficult scenarios that specially challenged people confront.

An effective solution to this problem is one based on computer vision-oriented gesture recognition, which involves image processing techniques. Consequently, this category faces more complexity. Sign languages are fully functional yet, are not universally used nor understood, and the easiest way to perceive them is by creating a computer recognition model that through video Sequence captures spatial movements and records them, predicting the translation system for sign language using Convolutional Neural Networks (CNN). This is divided into three main segments: the system design, the data set, and the deep learning model training and evaluation.

This translator enables people with disabilities to communicate with others successfully. Indian Sign Gestures to Text, Speech and Speech to Indian Sign Gesture translation is performed using the described methodology. We used a CNN network i.e., model which is trained on preprocessed sign image data set without any complex pre-processing wherein we can directly input live video stream with camera by a signer and fed to the network. The whole idea behind this dissertation is to detect and predict different signs from a set of predefined signs of Indian sign language.

And Google Speech Recognition API to translate spoken words in English to ISL gestures. This study focuses on the most accurate translation of Indian Sign Gestures to Text, Speech, and Spoken English words to Indian Sign Language movements as accurately as feasible.

1.1. PROBLEM DEFINITION

A Sign Language is a language which is used to communicate between people by using sign patterns to express the meaning. It is used by hearing impaired people. Thus, it has attracted many researchers to this field.

Many researchers have been working on different sign languages like American Sign Language, British Sign Language, Taiwanese Sign Language, etc. But a very few progresses have been made on Indian Sign Language.

The hearing-impaired people are being neglected from the society because the normal people neither try to learn Sign Language nor try to interact with the hearing-impaired people. This has become the greatest hurdle for them which has made them remain uneducated and isolated.

Our project aim is to develop a Two-way sign language translation system/application that transforms Two-Handed Indian sign language Gestured to text, then to speech, and vice versa that minimize the gap of communication between the hearing-impaired people and the others.

The approach is to build a cost-effective model that is readily accessible. This would not only solve the collective problem of lack of availability of Indian Sign Language speakers and provide a vast scope of research and development in this particular field but also for the better impact on society

1.2. OBJECTIVES

- To minimize the verbal exchange gap among listening to impaired and regular humans, to make conversation effective among all. Because India lacks many institutions for developing Indian sign language [other than ISLRTC], there is a lack of comprehension among many people, and some institutions recommend ASL over ISL.

- To build an efficient and accurate system of bi-directional communication for Mute people and others to make achieve efficient communication.

- This application/system could be used by especially abled person to be able to convey their hand sign or gesture language into speech and aid an ordinary person to translate speech to gesture or hand sign language to make the communication more fluent between a person with hearing impairment or speaking disability but has knowledge about Indian sign language.

- Most of the programs present usually have only American sign language as their region of processing, but this application is specially designed for Bi-Directional Indian Sign Language (ISL) Translation.

CHAPTER 2

LITERATURE SURVEY

The intention of the proposed work lies in the mere framework of developing an system that translates Indian Sign Language gestures to text, audio and vice versa for having the efficient communication between impaired people and others. Some similar thoughts are perceived in the papers referred below.

MS Anand et al. **[1]** proposed a double phase ISL conversion system. In speech-to-sign module the input voice was fed to the noise removal sub-module. After that, the output was sent into a speech recognizer, which decoded the input voice and converts to phrases sequence. A rule-based technique was used to turn the sequence of phrases into a sequence of sign symbols in a natural language. Finally, using the sign animation module the animated signs were displayed with text annotation respectively.

Kusurnika et al, **[2]** proposed a system, that can be used to get speech as output. They made a system that captures double handed Indian sign language as a series of images and then processed with the help of MATLAB for the translation of double handed Indian sign language to speech and text.

Paras Vij et al. **[3]** designed a two-phase sign language generation technique The initial phase involved pre-processing 2 Hindi Sentences and translating them into ISL grammar. WordNet and Dependency Parser have been combined to complete this task. In the Dependency Parser, dependency graphs represented words and their relationships between head words and words that affect those heads. HamNoSys was utilised in the second phase to translate this syntax into respective Sign Language. The symbols generated were translated by SIGML to form XML tags. A 3D rendering software can then read the XML tags form.

Pooja Balu et al, **[4]** goal in this study is to build an autonomous communication system for deaf and hard of hearing people. The given input text is translated into sign symbols by this system. It accepts text (alphabets/numbers) as input and transforms it to the appropriate sign code, which it then displays on the screen. The proposed method correctly transcribed 26 alphabets and 9 numbers.

Kshitij et al, **[5]** builds a CNN model named Inception was used to extract features from a video stream, and the output was then evaluated using soft max and CNN's Pool Layer. The authors used a camera to capture motions, which were then converted into frame sequences and put into a CNN pool layer, which was then evaluated by RNN (Recurrent Neural Network). On the supplied data set, Gesture Detection with Inception Model reported 99 percent accuracy.

Supriya Pawar et al. **[6]** proposed a camera-based hand gesture recording system. Image processing of the captured gesture is performed. Then, sign recognition and conversion of signs to text and voice are done. The final result is amplified voice equivalent to each processed gesture. This project aims to develop a useful tool that uses gesture recognition to reduce communication barriers between the deaf and dumb community. This project is a prototype to verify the feasibility of gesture recognition by image processing.

Dhruv et al, **[7]** because of raises in the hardware cost, the authors believe that utilizing a web app to recognize gestures is more efficient than using a hand wearable device. The idea behind the web app is to use the built-in webcam to collect motions. It initially gathers photographs in RGB format, then scales and changes them as needed. This system employs a convolutional neural network (CNN).

Ankita Harkude et al. **[8]** built a system that takes input as voice, transforms that to text, and a set of pre-labelled Indian Sign Language Pictures or GIFs are displayed. Communication between normal and deaf community is made easier using this system. After then, all the words in the sentence are compared to words in a lexicon that includes images and GIFs that represent those words. synonym’s will be used If the terms aren't recognized. The system has a set of gestures that are pre-labelled.

Dhivyasri et al. **[9]** proposed Support Vector Machine, for gesture to text conversation and for speech to gesture conversion, Google Speech Recognition API were used. The proposed application can read Indian Sign Language. As a result, a more valid sign language interpreting application was created.

Pankaj et al, **[10]** aims to develop an end-to-end human interface framework which is capable of recognising and understanding spoken language and then exhibit the relevant ISL gestures to enable a very convenient, real-time form of dialogue between the disabled and the general public. They used the depth sensing and motion capturing capabilities of the Microsoft Xbox Kinect 360 to gather motion data for all of the different ISL motions, then Unity3D is used for creating all the animations of relevant sign symbols. And at the final stage the whole system is ultimately packaged into an Android Application.

2.1 Technologies Used

Python

Fig 3: Python Logo

Python is an interpretive, high-level and general-purpose programming language. Created by Guido van Rossum and first published in 1991, Python's design Philosophy emphasizes the readability of code with its impressive use of large white space. Python is interpreted, there’s no need to compile the software until you run it. This is equivalent to both PERL and PHP. It can prompt and communicate directly with the interpreter to write your programs. Also, it is Object-oriented it supports Object-Oriented style or programming technique that encapsulates code inside items.

OpenCV

OpenCV (Open-Source Computer Vision Library) is a library of programming functions specifically designed for real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage and Itseez (later acquired by Intel). The library is cross-platform and free to use under the open-source Apache 2 License.

Fig 4: OpenCv

OpenCV features GPU acceleration for real-time operations beginning in 2011. [15] Officially launched in 1999, the OpenCV project was originally an Intel Research effort to advance CPU-intensive applications, part of a series of initiatives including real-time ray tracing and 3D display walls. The key contributors to the project included a range of Intel Russia optimization experts and the Intel Performance Library Team. If the library detects Intel's Integrated Performance Primitives on the device, these proprietary tailored routines will be used to speed up the operation.

Deep Learning

Fig 5: Deep Learning

In deep learning, a Convolutional Neural Network (CNN, or ConvNet) is a class of deep neural networks most widely used for visual imaging analysis. They are also known as shift invariant or space invariant artificial neural networks (SIANN) based on their shared-weight architecture and translation invariance characteristics. CNNs are regularized variants of multi-layer perceptron's. Multilayer perceptron's typically mean completely linked networks, that is, each neuron in a single layer is connected to all neurons in the next layer. The "full-connectivity" of these networks makes them vulnerable to data overfitting. Typical methods of regularization include introducing some type of weight calculation to the loss function. CNNs take a different approach to regularization i.e., they take advantage of the hierarchical structure in the data and assemble more complicated patterns using smaller and simpler patterns. As a consequence, on the scale of connectivity and complexity, CNNs are at the lower end. Convolutional networks are influenced by biological processes in that the pattern of connectivity between neurons resembles the organization of the animal visual cortex. Individual cortical neurons reply to stimuli only in a restricted area of the visual field known as the receptive field. The receptive fields of various neurons overlap partially in such a way that they occupy the entire visual field.

TensorFlow

Fig 6: TensorFlow Logo

TensorFlow is a machine learning and artificial intelligence software library that is free and open source. It can be used for a variety of applications, but it focuses on deep neural network training and inference. [11]. The Google Brain team created TensorFlow for internal Google use in research and production. [12] [13]. In 2015, the first version was released under the Apache License 2.0. In September 2019, Google launched TensorFlow 2.0, an improved version of TensorFlow. TensorFlow is compatible with a wide range of programming languages, including Python, JavaScript, C++, and Java. This adaptability lends itself to a wide range of applications in a variety of industries.

CHAPTER 3

Proposed Work

3.1. METHODOLOGY

This project work proposes a schema technique that mainly consists of 2 Phases:

Phase 1: Indian Sign Language to Text, Speech.

Phase 2: Spoken English to Indian Sign Gestures

3.2 ARCHITECTURE

Phase 1: Indian Sign Language to Text, Speech

3.2.1 Phase 1: DATA GATHERING & PRE-PROCESSING

The Data pre-processing takes place steps as shown below:

- In the first step the vast amount of data collected from various sources for the prediction of Indian Sign Symbols is first to be stored in in a particular folder, so that this data could be used to train a model to get meaningful output.

- In order to reduce noise, differentiate background form object, getting contours and highlight edges so that the model could predict correct target labels with great efficiency. For all this image processing algorithms were performed like BGR to Gray conversion, BGR to HSV conversion, skin masking.

  After the Image processing the skin is shown in white color and rest of the background is shown as black. The below are the images before and after image processing.

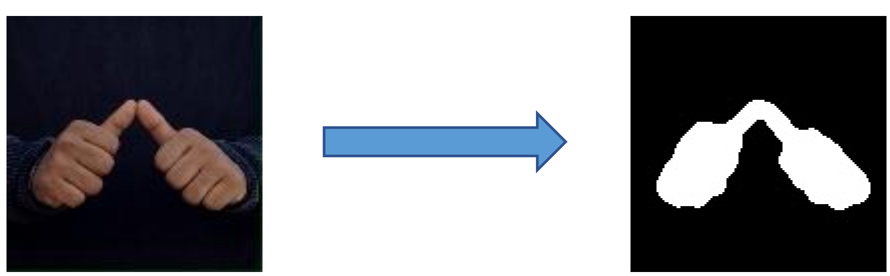


Fig 7: Data Pre-Processing (Before and After)

The dataset we examined comprises 1200 images for each class of Alphabets from A to Z, we sought to create a model for prediction which identifies the user's input sign gesture and predicts the appropriate label.

3.2.2 Phase 1: SPLITTING THE DATA INTO TRAINING AND TESTING SET

The collected dataset after data pre-processing has been split into 2 parts i.e., "test" and "train" with a ratio of 75:25 from the whole data set. If a model is trained properly, the output of testing and training data will be the same.

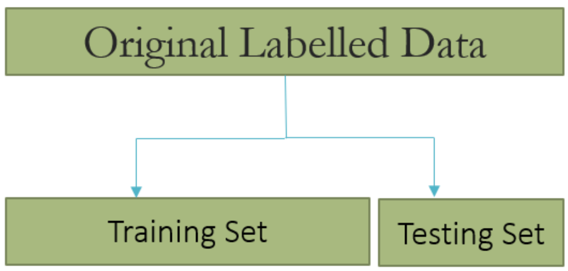


Fig 8: Data Splitting

3.2.3 Phase 1: Convolutional Neural Network

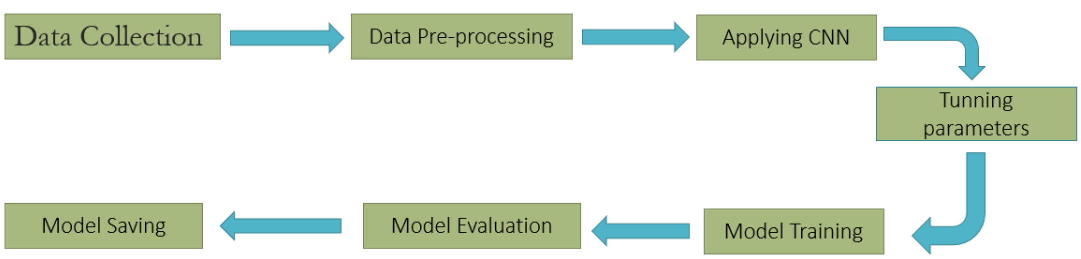


Fig 9: Flow Diagram of CNN

After splitting the data into train and test sets, we will apply the CNN model layer by layer using keras. After adding certain activation functions such as relu and softmax, our model is ready for training. We will fit the model and train it for 5 epochs to get the maximum accuracy. Then the model will be evaluated on the test data after achieving the desired accuracy we will save our model for the further use in real time application.

3.2.4 Phase 1: Real-Time Implementation

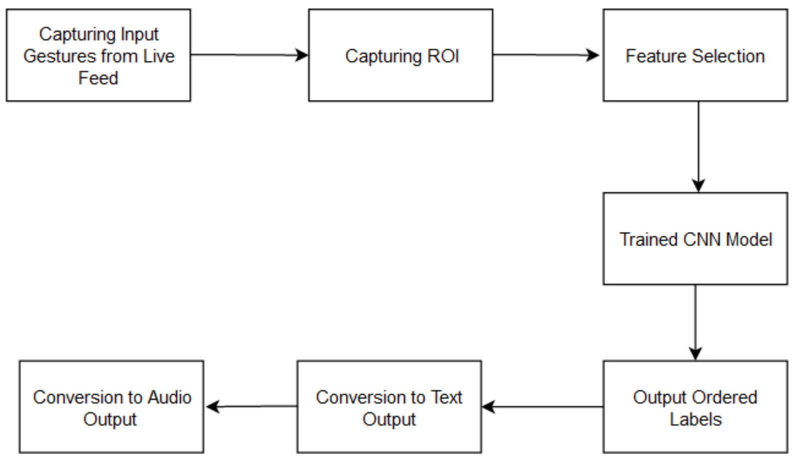


Fig 10: Use Case Diagram of Sign to text

This stage collects images from the user's device (Camera) as input. Those frames are sent to the data preparation for skin Segmentation, which cleans up the noise and tracks the important points in Region of Interest (ROI).

Then important points that are extracted are fed into Trained CNN Model (final_model.h5), such that if frame is a match and classified then it returns the corresponding alphabet for that signed gesture. After the text has been recognized it is then converted into audio using gtts.

Algorithm:

Step 1: Reading image from camera.

Step 2: Extract the data from Region of Interest (ROI).

Step 3: Finding contours of hand for getting shape of hand.

Step 4: if image (sign) ==database image:

Fetch the equivalent text!

Send text to the voice manager.

else

Display the necessary error message.

Exit

END

Phase 2: Spoken English to Indian Sign Language Gestures

3.2.5 Phase 2: Real-Time Implementation

The user's speech is received through the device's microphone and passes through the voice-to-text model. The textual content generated from the speech is then analysed to extract letters from the word. After speech recognition, the spoken characters generated are mapped with predefined labelled sign gesture images and displays the relevant sign symbols respective to the characters sequence of the speech received with a certain delay. Since the speech-to-text domain has already been explored extensively, we have used a pre-existing API developed by Google.

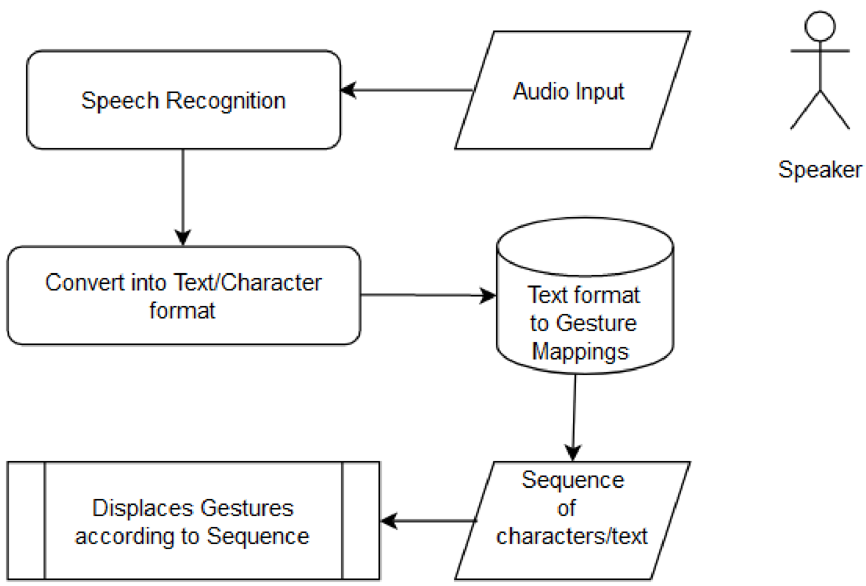


Fig 11: Use case Diagram of Speech to Sign

Algorithm:

Step1: Start

Step2: Getting the Input

1. Listen the Speech using Microphone.

Step3: Recognize the Speech.

Step4: Convert Speech to Text.

1. Make the Text to lowercase for further manipulation.

Step5: Detected Text

1. Count the Letters of the Word/Phrase.
    a. Display the text as "Text: Recognized Words"
    b. Display the Visual of the phrase recognized with some delay of Actions.
2. Continue all the steps from Step 3 and continue till the Speech Ends. 3.

If Error in Step 2, That is if no Speech Detected then display error message "Sorry, I did not get that, please try again"

3.3 Implementation Code Screen Shots

3.3.1 Phase 1: Indian Sign Language to Text

CNN Model:

```
In [14]: #initializing the CNN

#sequential function provided by keras lets us build a model layer by layer
model = Sequential()

#convolution layer and pooling
model.add(Convolution2D(filters=32,kernel_size=3,padding='same',activation='relu',input_shape=(128,128,1)))
#pooling layer helps with downsampling i.e dimensionality reduction
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Convolution2D(filters=32,kernel_size=3,padding='same',activation='relu'))
model.add(Convolution2D(filters=32,kernel_size=3,padding='same',activation='relu'))
#converting the data into a 1D array before inputting to the fully connected layer
model.add(Flatten())

#connected layer

model.add(Dense(units=128, activation='relu'))
model.add(Dropout(0.40))
model.add(Dense(units=96, activation='relu'))
model.add(Dropout(0.40))
model.add(Dense(units=64, activation='relu'))

model.add(Dense(units=35, activation='softmax'))

#compiling
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy',AUC(),Precision(),Recall(),FalsePositives(),
```

Fig 12: Implemented CNN Model

Model Fitting:

```
In [16]: #fitting the model

         history=model.fit(training, steps_per_epoch=len(training),epochs=5,validation_data=testing,validation_steps=len(testing))

Epoch 1/5
985/985 [==============================] - 447s 452ms/step - loss: 0.3899 - accuracy: 0.8854 - auc_1: 0.9950 - precision_1: 0.9
506 - recall_1: 0.8581 - false_positives_1: 1406.0000 - true_negatives_1: 1069594.0000 - true_positives_1: 27030.0000 - false_n
egatives_1: 4470.0000 - val_loss: 0.0249 - val_accuracy: 0.9921 - val_auc_1: 1.0000 - val_precision_1: 0.9954 - val_recall_1:
0.9897 - val_false_positives_1: 48.0000 - val_true_negatives_1: 356952.0000 - val_true_positives_1: 10392.0000 - val_false_nega
tives_1: 108.0000
Epoch 2/5
985/985 [==============================] - 425s 432ms/step - loss: 0.0645 - accuracy: 0.9807 - auc_1: 0.9992 - precision_1: 0.9
847 - recall_1: 0.9775 - false_positives_1: 477.0000 - true_negatives_1: 1070523.0000 - true_positives_1: 30792.0000 - false_ne
gatives_1: 708.0000 - val_loss: 0.0150 - val_accuracy: 0.9950 - val_auc_1: 1.0000 - val_precision_1: 0.9952 - val_recall_1: 0.9
950 - val_false_positives_1: 50.0000 - val_true_negatives_1: 356950.0000 - val_true_positives_1: 10447.0000 - val_false_negativ
es_1: 53.0000
Epoch 3/5
985/985 [==============================] - 455s 462ms/step - loss: 0.0449 - accuracy: 0.9863 - auc_1: 0.9995 - precision_1: 0.9
883 - recall_1: 0.9841 - false_positives_1: 366.0000 - true_negatives_1: 1070634.0000 - true_positives_1: 30999.0000 - false_ne
gatives_1: 501.0000 - val_loss: 0.0232 - val_accuracy: 0.9930 - val_auc_1: 0.9999 - val_precision_1: 0.9947 - val_recall_1: 0.9
910 - val_false_positives_1: 55.0000 - val_true_negatives_1: 356945.0000 - val_true_positives_1: 10406.0000 - val_false_negativ
es_1: 94.0000
Epoch 4/5
985/985 [==============================] - 480s 487ms/step - loss: 0.0342 - accuracy: 0.9889 - auc_1: 0.9996 - precision_1: 0.9
908 - recall_1: 0.9874 - false_positives_1: 290.0000 - true_negatives_1: 1070710.0000 - true_positives_1: 31103.0000 - false_ne
gatives_1: 397.0000 - val_loss: 0.0078 - val_accuracy: 0.9985 - val_auc_1: 1.0000 - val_precision_1: 0.9990 - val_recall_1: 0.9
973 - val_false_positives_1: 11.0000 - val_true_negatives_1: 356989.0000 - val_true_positives_1: 10472.0000 - val_false_negativ
es_1: 28.0000
Epoch 5/5
985/985 [==============================] - 484s 491ms/step - loss: 0.0304 - accuracy: 0.9910 - auc_1: 0.9996 - precision_1: 0.9
926 - recall_1: 0.9899 - false_positives_1: 231.0000 - true_negatives_1: 1070769.0000 - true_positives_1: 31181.0000 - false_ne
gatives_1: 319.0000 - val_loss: 0.0191 - val_accuracy: 0.9960 - val_auc_1: 0.9991 - val_precision_1: 0.9969 - val_recall_1: 0.9
948 - val_false_positives_1: 32.0000 - val_true_negatives_1: 356968.0000 - val_true_positives_1: 10445.0000 - val_false_negativ
es_1: 55.0000
```

Fig 13: CNN Model Fitting

3.3.2 Phase 2: Spoken English to Indian Sign Language gestures

```
Spyder (Python 3.8)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\91703\Desktop\New folder\MajorProject\Code\Spyder\Gesture-Speech-main\speech to gesture.py
temp.py    speech to gesture.py

#import library
import cv2
import speech_recognition as sr
import os

# Initialize recognizer class (for recognizing the speech)

r = sr.Recognizer()

# Reading Microphone as source
# listening the speech and store in audio_text variable

with sr.Microphone() as source:
    print('Turn on data traffic for accessing speech recogniser :)')
    print("Say Something")

    audio_text = r.listen(source)
    print("Time over, thanks")
# recoginize_() method will throw a request error if the API is unreachable, hence using exception ha

    try:
        # using google speech recognition
        text = r.recognize_google(audio_text)
        print("Text: "+text)
    except:
         print("Sorry, I did not get that, please try again")

for elem in text:
    path = 'images'
    if elem ==' ':
        full_path = os.path.join(path,'blank')+'.jpg'
        print(full_path)
        img = cv2.imread(full_path ,cv2.IMREAD_GRAYSCALE)
        cv2.imshow('blank',img)
        cv2.moveWindow('blank',40,30)
        cv2.waitKey(1000)
        cv2.destroyWindow('blank')
    else:
        full_path = os.path.join(path,elem)+'.jpg'
        print(full_path)
        img = cv2.imread(full_path ,cv2.IMREAD_GRAYSCALE)
```

Fig 14: Implementation of Speech to Sign

3.4 REQUIREMENTS

Software Requirements:

Operating System: Window 10 Home/Mac OS

Platform: Spyder or Jupyter IDE

Language: Python 3.x

Libraries: Matplotlib, NumPy, Pillow, OpenCV-python, TensorFlow, gTTS, tKinter, speech_recognition, pyAudio

Hardware Requirements:

Intel Core i5 processor, 8^{th} generation.

1 TB HDD/ 256 SSD

Min 4Gb RAM

For better quality and results an external web cam and microphone is recommended:

Logitech C310 HD Webcam

Noise Cancellation Microphone

CHAPTER 4

4.1 RESULT ANALYSIS

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_6 (Conv2D)           (None, 128, 128, 32)      320

 max_pooling2d_2 (MaxPooling  (None, 64, 64, 32)       0
 2D)

 conv2d_7 (Conv2D)           (None, 64, 64, 32)        9248

 conv2d_8 (Conv2D)           (None, 64, 64, 32)        9248

 flatten_2 (Flatten)         (None, 131072)            0

 dense_8 (Dense)             (None, 128)               16777344

 dropout_4 (Dropout)         (None, 128)               0

 dense_9 (Dense)             (None, 96)                12384

 dropout_5 (Dropout)         (None, 96)                0

 dense_10 (Dense)            (None, 64)                6208

 dense_11 (Dense)            (None, 35)                2275

=================================================================
Total params: 16,817,027
Trainable params: 16,817,027
Non-trainable params: 0
```

Fig 15: Model Summary of Trained CNN

Phase 1: Indian Sign Language to Text

Indian Signs have been recorded and recognized in Real-Time:

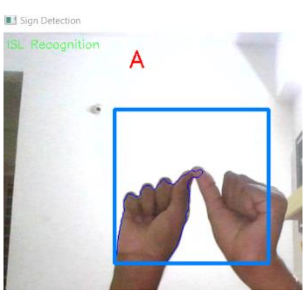


Fig 16(i): A

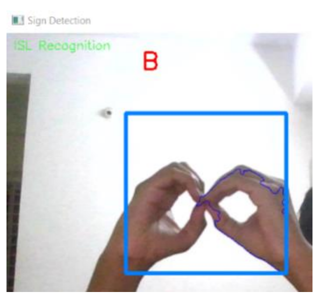


Fig 16(ii): B

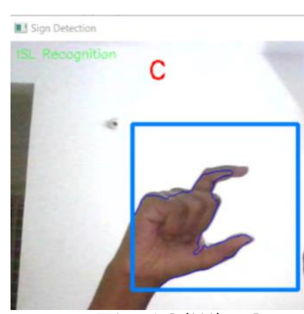


Fig 16(iii): C

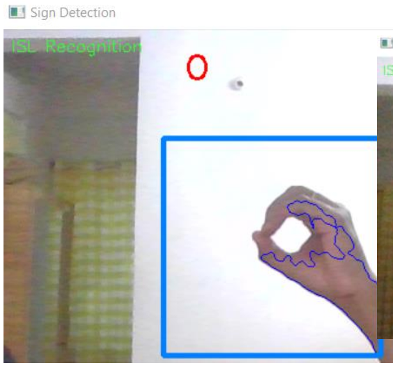


Fig 16(iv): O

Fig 16(v): U

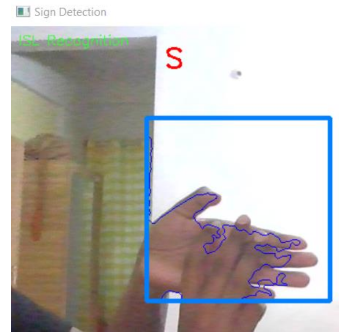


Fig 16(vi): S

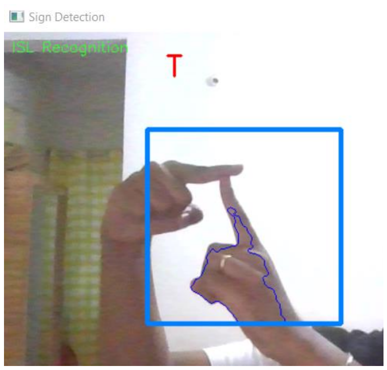


Fig 16(vii): T

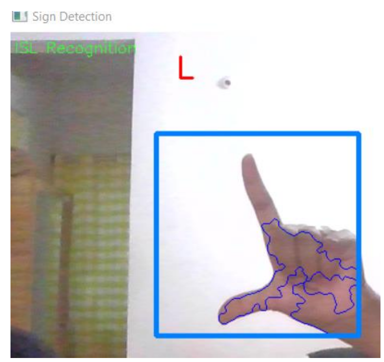


Fig 16(viii): L

Phase 2: Spoken English to Indian Sign Language gestures

When nothing is spoken, it will display the above error message.

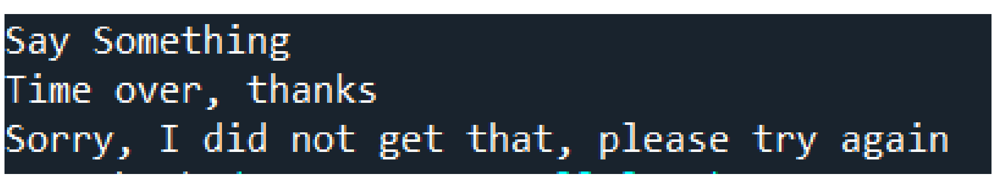


Fig 17(i): Error Message

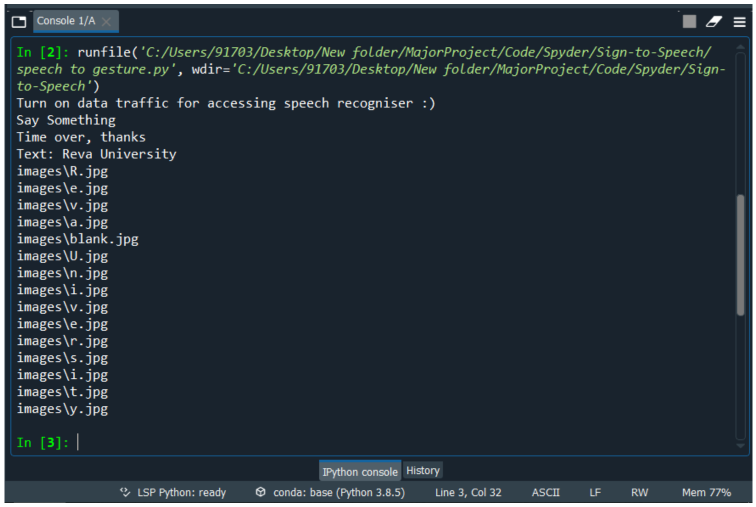
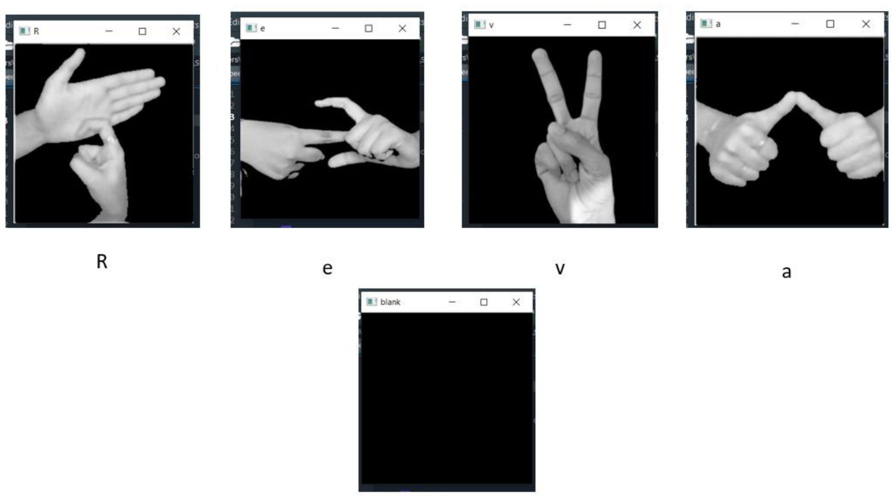


Fig 17(ii): Displays Text

When microphone identifies speech, it will display the text of that speech and relevant names of sign symbols that will be displayed.

Speech Input is given as (Reva University), the below are the gestures to the characters format of the string respectively.



R e v a

Blank

Fig 17(iii): Displaying Reva in Signs

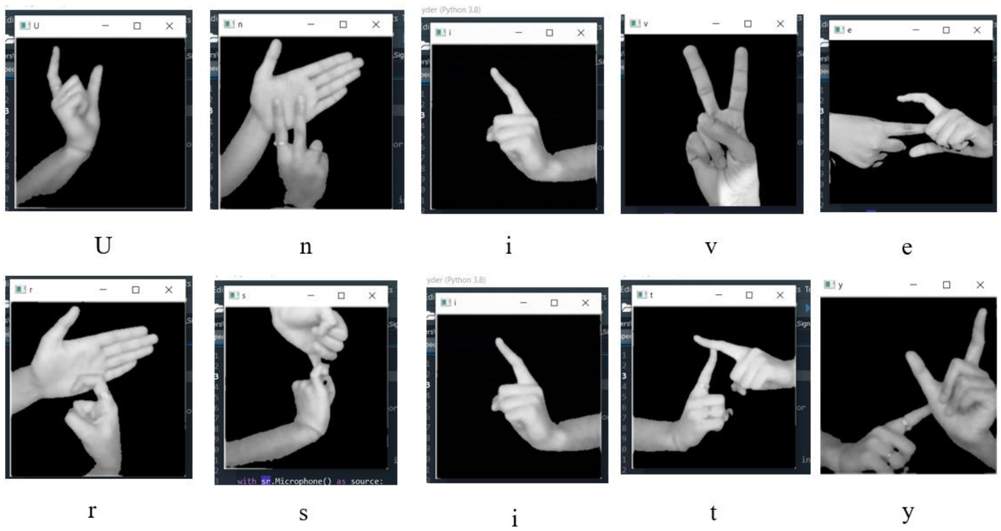


Fig 17(iii): Displaying University in Signs

4.2 Accuracy:

Our Model has achieved the accuracy around 99%.

```
Accuracy on test set: 99.6%
ROC(Receiver Operation Characteristic) AUC(Area Under Curve): 0.9991403818130493
Precision: 99.69%
Recall: 99.48%
F1-score: 0.9958525844961417
Specificity: 0.9999103641456583
```

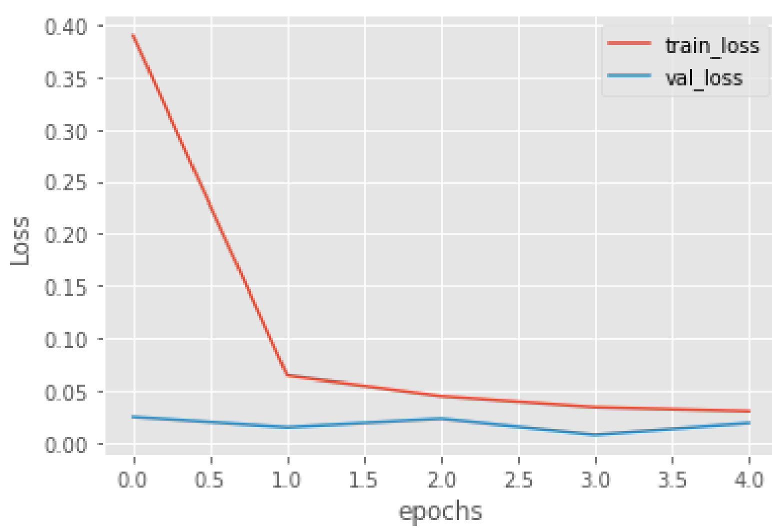


Fig 18: Train Loss Vs Val Loss

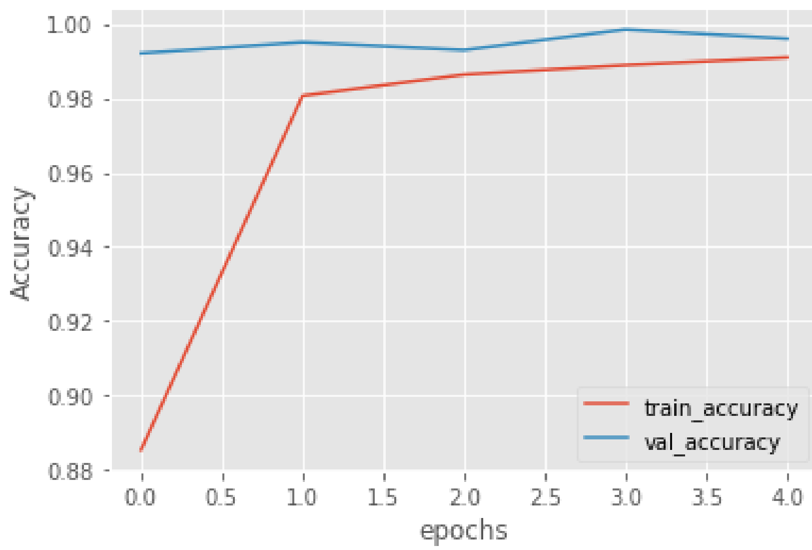


Fig 19: Train Accuracy Vs Val Accuracy

4.3 Limitations

- Translation is not possible for dynamic signs.
- Results are dependent on the lighting condition.
- Translation of signs has a minor delay.
- Because of the less diversity in the dataset sometimes the model predicts wrong output for the signs.

CHAPTER 5

5.1 Cost Analysis

Logitech C310 HD Webcam with inbuilt microphone: ₹ 2,069.00

Noise Cancellation Microphone: ₹ 1,000.00

Paper Publication: ₹ 7,000.00

Report Binding: ₹ 2,450.00

Miscellaneous: ₹ 1,609.00

Total Cost: Approx. ₹ 14,500.00

5.2 CONCLUSION

In this project, we attempted to address a few the most pressing communication challenges faced by impaired people. We prefer to focus on the root cause of impaired people's inability to communicate with others. The end consequence is that the other side of the audience does not appear to be prepared to assess what these people are seeking to say or what message they need to send.

Our project, "Full Duplex Mode of Communication Between Impaired people and others", aims to extend help for hearing impaired people to communicate. In this project we developed a translation system that can translate the shown Indian Sign Language gestures to text in Real Time with an accuracy of 99%. And it also can be used for reverse conversion i.e., Spoken English Phrases are converted and displayed with Indian Sign Language gestures. This project is the prototype that is designed to be a standalone device translating ISL to audio and text and vice versa. This project easily enables communication for the hearing-impaired people. However, the limitation with this project is to maintain constant lighting condition.

5.3 Further Work

The goal for future is to build an end-to-end system that converts Indian Sign Language Gestures to Speech with at most accuracy and then integrating both Sign language Gestures to Speech and Speech to Sign Gestures as one application. We need an active internet connection to use google api for spoken English phrases to Indian Sign Gestures, which can be developed further to work in offline mode as well. We also aim to expanding the project by incorporating facial expressions as well. This application/system is beneficial for two-handed Sign gestures in both directions for efficient communication between disabled and non-impaired people.

REFERENCES

[1] M.Suresh Anand, A.Kumaresan, Dr.N.Mohan Kumar, An Integrated Two Way ISL (Indian Sign Language) Translation System A New Approach, Volume 4, No. 2, Jan-Feb 2013 International Journal of Advanced Research in Computer Science.

[2] Kusurnika Krori Dutta 1, Satheesh Kumar Raju K2, Ani! Kumar G S3, Sunny Arokia Swarny B4, "Double Handed Indian Sign Language to Speech and Text" 2015 Third International Conference on Image Information Processing.

[3] Paras Vij, Parteek Kumar, ―Mapping Hindi Text To Indian sign language with Extension Using Wordnet‖, AICTC '16: Proceedings of the International Conference on Advances in Information Communication Technology & Computing August 2016 Article No.: 38 Pages 1–5.

[4] Pooja Balu Sonawane, Anita Nikalje, "Text to Sign Language Conversion by Using Python and Database of Images and Videos", International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE) Vol 5, Issue 2, February 2018.

[5] Kshitij Bantupalli Department of Computer Science Kennesaw State University Kennesaw, USA, Ying Xie Department of Computer Science Kennesaw State University Kennesaw, USA, "American Sign Language Recognition using Deep Learning and Computer Vision", 2018 IEEE International Conference on Big Data (Big Data).

[6] Prof. Supriya Pawar1, Sainath Muluk2, Sourabh Koli3 Assistant Professor, Dept. of Computer Engg, D.Y., Patil College of Engg, Ambi, Pune, India1 Dept. of Computer, Engg, D.Y., Patil College of Engg, Ambi, Pune, India 2, 3, "Real Time Sign Language Recognition using Python", International Journal of Innovative Research in Computer and Communication Engineering - Vol. 6, Issue 3, March 2018.

[7] Dhruv Vaghela1, Ankit Vaity2, Rajvi Makwana3, Afreen Shaikh4 , Chanda Chouhan5, "Hand Gesture Recognition To Speech Conversion", IOSR Journal of

Engineering (IOSR JEN), ISSN (e): 2250-3021, ISSN (p): 2278-8719, PP 21-25, 2019.

[8] Ankita Harkude , Sarika Namade, Shefali Patil, Anita Morey #4 1,2,3,4#Department of Information Technology, Usha Mittal Institute of Technology, SNDT Women's University, “Audio to Sign Language Translation for Deaf People”, ISO 9001:2008 Certified International Journal of Engineering and Innovative Technology (IJEIT) Volume 9, Issue 10, April 2020.

[9] Dhivyasri S, Krishnaa Hari K B, Akash M, Sona M, Divyapriya S, Dr. Krishnaveni V, “An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning”, 2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 13 – 14 May 2021 | Coimbatore.

[10] Pankaj Sonawane, Karan Shah, Parth Patel, Shikhar Shah, Jay Shah, “Speech To Indian Sign Language (ISL) Translation System", 2021International Conference on Computing, Communication, and Intelligent Systems (ICCCIS).

[11] TensorFlow: Open-source machine learning. *Google. 2015.* Archived *from the original on November 11, 2021*. "It is machine learning software being used for various kinds of perceptual and language understanding tasks" – Jeffrey Dean, minute 0:47 / 2:17 from YouTube clip.

[12] Video clip by Google about TensorFlow 2015 at minute 0:15/2:17

[13] Video clip by Google about TensorFlow 2015 at minute 0:26/2:17

[14] http://islrtc.nic.in/poster-manual-alphabet-isl - 27-04-2022, 10:50AM

[15] https://www.quora.com/What-is-the-logic-behind-OpenCV-logo at 24-05-2022, 1:30 Pm

CONFERENCEPAPERS PUBLISHED

This Project Paper is Selected in the Conference IACIT for the Publication.

Screen Shot:

I2 IACIT 2022
To: E. MAHENDRANATH REDDY
Cc: Kanaiya V Kanzaria

Thu 5/12/2022 9:25 PM

Dear Author/Authors,

Greetings from IACIT-2022 | School of Computer Science and Engineering, REVA University, Bengaluru.

Congratulations!!

We would like to inform you that your paper is accepted for presentation in "**4th International Virtual Conference on Advances in Computing and Information Technology (IACIT-2022)**" and publication in the Scopus Indexed Journal.

Paper Id: IACIT-2022-PID-592

Paper Title: Spoken English to Indian Sign Language Translator

Fig 20: IACIT Acceptance

Full duplex mode of communication between impaired people and others

ORIGINALITY REPORT

14%	8%	11%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Pankaj Sonawane, Karan Shah, Parth Patel, Shikhar Shah, Jay Shah. "Speech To Indian Sign Language (ISL) Translation System", 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2021 Publication	2%
2	Joyeeta Singha, Karen Das. "Recognition of Indian Sign Language in Live Video", 'Foundation of Computer Science', 2013 Internet Source	2%
3	Submitted to University of Iowa Student Paper	1%
4	awesomeopensource.com Internet Source	1%
5	Anjali Kanvinde, Abhishek Revadekar, Mahesh Tamse, Dhananjay R. Kalbande, Nida Bakereywala. "Bidirectional Sign Language Translation", 2021 International Conference	1%

on Communication information and Computing Technology (ICCICT), 2021
Publication

6	doi.org Internet Source	1%
7	ijirset.com Internet Source	1%
8	Kusumika Krori Dutta, Satheesh Kumar Raju K., Anil Kumar G.S., Sunny Arokia Swamy B.. "Double handed Indian Sign Language to speech and text", 2015 Third International Conference on Image Information Processing (ICIIP), 2015 Publication	1%
9	Muthukumar. K, Poorani S, Gobhinath S. "Extraction of Hand Gesture Features for Indian Sign languages using Combined DWT-DCT and Local Binary Pattern", International Journal of Engineering & Technology, 2018 Publication	1%
10	Dhivyasri S, Krishnaa Hari K B, Akash M, Sona M, Divyapriya S, Krishnaveni V. "An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning", 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021 Publication	<1%

11	Submitted to University of the Western Cape Student Paper	<1%
12	www.ijsrd.com Internet Source	<1%
13	Gunasagari G. S, Poornima N, Abhijna Yaji, Achuth M, Anisha Maria Dsilva, Chethana S R. "Review on Text and Speech Conversion Techniques based on Hand Gesture", 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021 Publication	<1%
14	Kshitij Bantupalli, Ying Xie. "American Sign Language Recognition using Deep Learning and Computer Vision", 2018 IEEE International Conference on Big Data (Big Data), 2018 Publication	<1%
15	caeaccess.org Internet Source	<1%
16	Elsayed E. Hemayed, Allam S. Hassanien. "Edge-based recognizer for Arabic sign language alphabet (ArS2V-Arabic sign to voice)", 2010 International Computer Engineering Conference (ICENCO), 2010 Publication	<1%
17	M. Sumana, Sahana S. Hegde, Shreehari N. Wadawadagi, D. V. Sujana, Vaibhav Gubbi	<1%

Narasimhan. "Chapter 7 Smart Tutoring System for the Specially Challenged Children", Springer Science and Business Media LLC, 2022
Publication

18	www.tcetmumbai.in Internet Source	<1 %
19	Amandeep Singh Dhanjal, Williamjeet Singh. "Comparative Analysis of Sign Language Notation Systems for Indian Sign Language", 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), 2019 Publication	<1 %
20	BiswaRanjan Samal, Mrutyunjaya Panda. "Integrative Review on Vision-Based Dynamic Indian Sign Language Recognition Systems", 2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology(ODICON), 2021 Publication	<1 %
21	"English text to Indian Sign Language Machine Translation: A Rule Based Method", International Journal of Innovative Technology and Exploring Engineering, 2019 Publication	<1 %
22	"Proceedings of Data Analytics and Management", Springer Science and Business	<1 %

Media LLC, 2022

Publication

23	Archana S. Ghotkar, Dr. Gajanan K. Kharate. "STUDY OF VISION BASED HAND GESTURE RECOGNITION USING INDIAN SIGN LANGUAGE", International Journal on Smart Sensing and Intelligent Systems, 2017 Publication	<1 %
24	V Aiswarya, N Naren Raju, Singh S Johanan Joy, T Nagarajan, P Vijayalakshmi. "Hidden Markov Model-Based Sign Language to Speech Conversion System in TAMIL", 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), 2018 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		