ISL to Speech & Vice Versa

Minor Project I (15B19CI591)

Group No.: 63

Submitted By:
Rhythm Agrawal (9918103080)
Eshan Dhawan (9918103066)
Kavya Vishnoi (9918103188)

Under the Supervision of Dr.Raju Pal

Department of CSE/IT
Jaypee Institute of Information Technology University, Noida

October 2020

TABLE OF CONTENTS

ABSTRACT

Communication plays an important role for human beings. Communication is treated as a life skill.Human beings interact with each other to convey their ideas, thoughts, and experiences to the people around them. But this is not the case for deaf-mute people. Sign language paves the way for deaf-mute people to communicate. Through sign language, communication is possible for a deaf-mute person without the means of acoustic sounds.

The aim behind this work is to develop a system for recognizing the sign language, which provides communication between people with speech impairment and normal people, thereby reducing the communication gap between them. Compared to other gestures (arm, face, head and body), hand gesture plays an important role, as it expresses the user's views in less time. Here we will be translating Sign language to text and speech. We also aim to convert text and speech to sign language.This might help to ease the miseries of the specially abled.

Keywords:

{handicapped aids;sign language recognition;speech synthesis;deaf-mute people;sign language recognition;speech impairment;communication gap;hand gesture;English alphabets recognition;English words recognition;text-to-speech synthesizer;Gesture recognition;Speech;ASL;Gesture recognition module;Text-to-speech synthesis module}

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

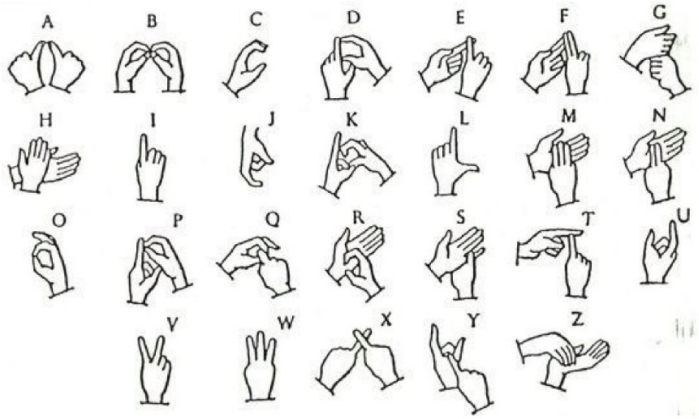
ISL	Indian Sign Language
ASL	American Sign Language
SVM	Support Vector Machine
CNN	Convolutional neural network
GMM	Gaussian Mixture Model
MOG	Mixtures if Gaussian
CGI	Common Gateway Interface
HOG	Histogram of Oriented Gradient

1. INTRODUCTION

A Sign Language is a language in which communication between people is made by visually transmitting the sign patterns to express the meaning. It is a replacement of speech for hearing and speech impaired people. Thus, because of which it has attracted many researchers in this field for a long time. Many researchers have been working in different sign languages like American Sign Language, British Sign Language, Taiwanese Sign Language, etc. but few works have made progress on Indian Sign Language. The hearing impaired people become neglected from society because the normal people never try to learn ISL nor try to interact with the hearing impaired people. This becomes a curse for them and so they mostly remain uneducated and isolated. Thus recognition of sign language was introduced which has not only been important from an engineering point of view but also for the impact on society. Our paper aims to bridge the gap between us and the hearing impaired people by introducing an inexpensive Sign Language Recognition technique which will allow the user to understand the meaning of the sign without the help of any expert translator.Instead of using high-end technology like gloves or kinect, we aim to solve this problem using state of the art computer vision and machine learning algorithms. Computers are used in communication paths which help in capturing of the signs, processing it and finally recognizing the sign.Several techniques have been used by different researchers for recognizing sign languages or different hand gestures.We will be working on a real time video capturing.

2. CHALLENGES

The ISL lags behind its ASL Counterpart as the research in this field is hampered by the lack of standard datasets. Unlike American Sign Language, it uses both hands for making gestures which leads to occlusion of features. ISL is also subject to variance in locality and the existence of multiple signs for the same character. Also some characters share the same alphabet(E.g V and 2 have the same sign, similarly W and 3 have the same sign) and the resolution of the sign is context dependent.



3. BACKGROUND STUDY

3.1. devmesh.intel.com/projects (*ROHIT GHUMARE*)

Image acquisition (vision based static hand gesture recognition using webcam)

Data set (https://github.com/Iron-Stark/Indian-Sign-Language/tree/master/data/images)

Animated Letters (https://github.com/Shubh-Yadav/Automatic-Indian-Sign-Language-Translator/tree/master/letters)

Now images are preprocessed to suppress unwanted distorsions using GMM. A GMM is known as a parametric probability density function which is signified as a weighted sum of a Mixture of Gaussians (MOG) parameters, and *M* component.

For Segmentation k-means clustering algorithm is used which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. In cluster analysis, the k-means algorithm can be used to partition the input data set into k partitions. Label every pixel in the image using results from K means. Then a blank cell array is created to store the results of clustering. Followed by creating an RGB label using pixel labels. The selection of appropriate clusters is another important aspect. The cluster which displays the maximum disease affected part is to be selected. In the next step of feature extraction, the features of the selected cluster are extracted. The features of the selected cluster are extracted. The selected image is converted to grayscale since the image is in RGB format. At the next step the Gray Level Co-occurrence. The following 13 features that are extracted and evaluated: Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS. Variance, Smoothness, Kurtosis, Skewness. The thirteen features are stored in an array. In ML, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observations. We will use the **SVM classification algorithm** as it gives the best output and has the highest accuracy among all other algorithms.

Conversion of Text to Speech few words are trained and stored in a database. Google Text to Speech API will be used here.

TechStack used:

Python 3.6 (Keras, Tensorflow, Numpy, Pandas, Matplotlib, Scikit learn)

Google Text to Speech API

Common Gateway interface(CGI) for everything. CGI is one of the essential parts of HTTP.

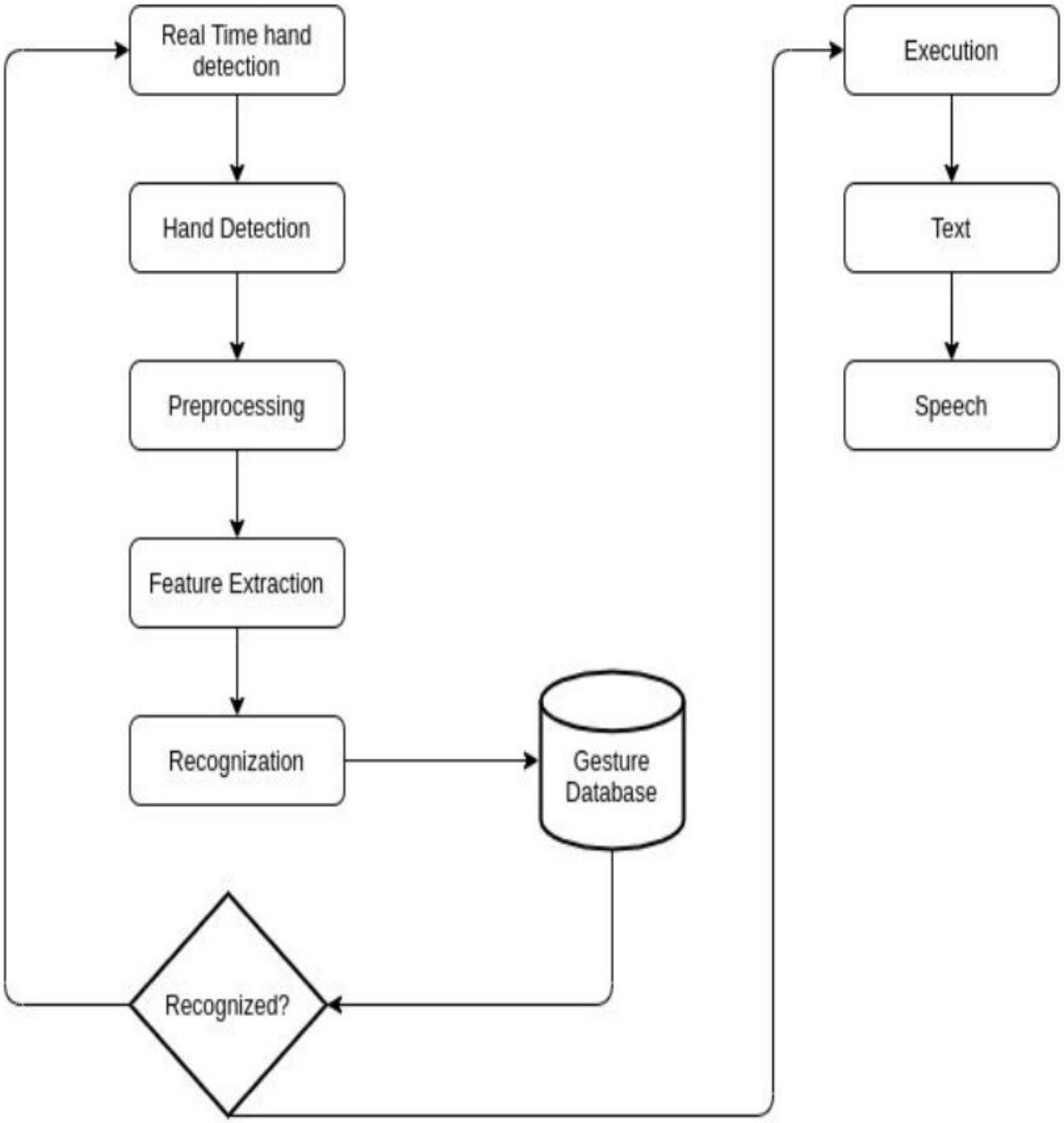
IDE: Jupyter Notebook

Google Cloud Platform (GCP)

Google Colab

SQL & Firebase

Overview of the system

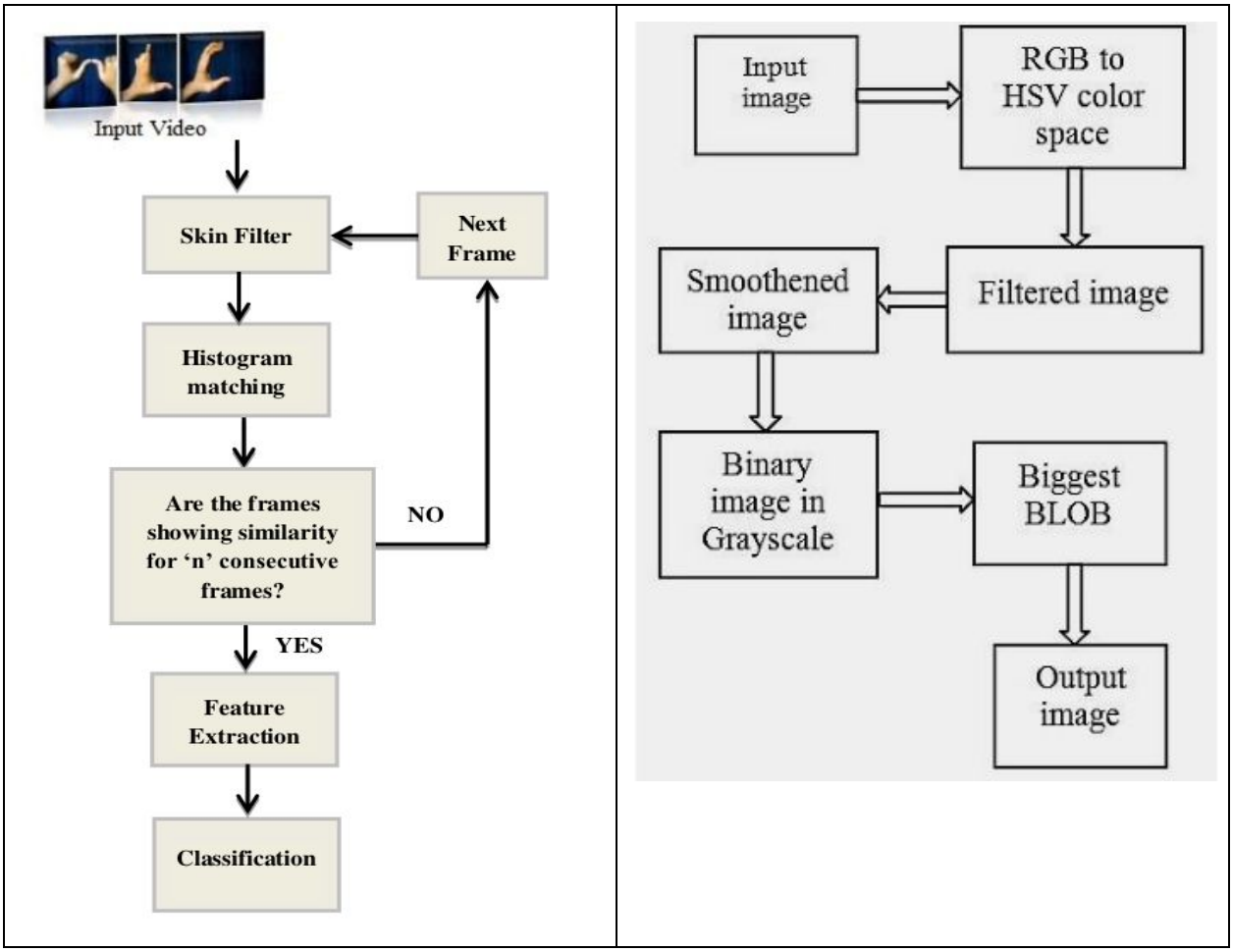


3.2. Recognition of Indian Sign Language in Live Video

Joyeeta Singha (Dept. of ECE DBCET Assam Don Bosco University Guwahati, Assam)

Karen Das (Dept. of ECE DBCET, Assam Don Bosco University Guwahati, Assam)

This research paper proposed the system which comprises of 3 major stages-preprocessing stage which includes the skin filtering and histogram matching to find out the similarity between frames, Feature Extraction stage in which the Eigenvalues and Eigenvector are being considered as features and finally Eigen value weighted Euclidean distance based classification technique as used here.The first step for the proposed system is the capturing of the video using webcam where different alphabets were taken into consideration. After extracting out the skin colored regions from the background, histogram matching is done in the next step where the similarities of the consecutive frames are checked by finding out the difference of their histogram. If the difference is found to be above a certain threshold, they are considered as similar. This difference is found for 'n' number of frames. They have chosen the threshold 'n' to be 17. If all the 'n' frames show similarities, then it is considered to be an unidentified sign and further steps of feature extraction and classification is carried on. A fast, novel and robust system was proposed for recognition of different alphabets of Indian Sign Language for video sequences.Eigen vectors and Eigenvalues were considered as the features and finally effective classification was achieved using Eigen value weighted Euclidean Distance based classifier. The overall recognition rate was calculated and found to be 96.25%.



3.3. Indian Sign Language Character Recognition - IIT Kharagpur

Methodology

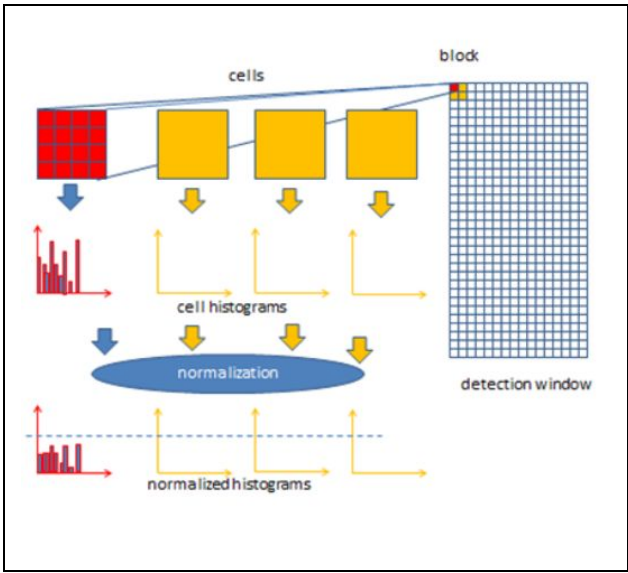
1 Image Segmentation

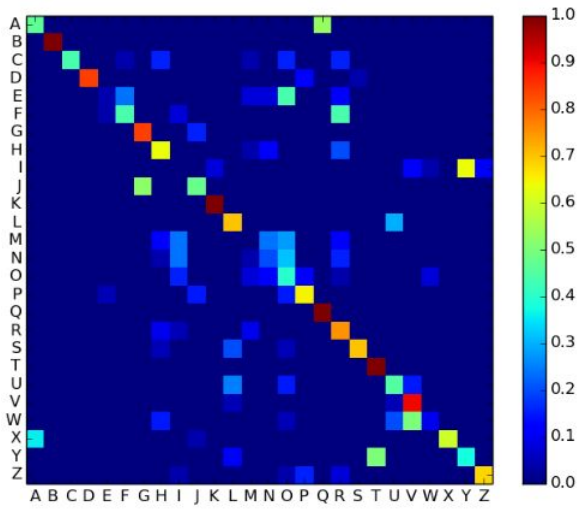
- 1.1 Training on skin segmentation dataset
- 1.2 HSV model
- 1.3 YIQ and YUV model(Final Approach)

2 Feature Extraction

- 2.1 Bag of Visual words
- 2.2 Histogram of Oriented Gradient(HOG) Features with dimension reduction
- 2.3 Histogram of Oriented Gradient Features(without dimensionality reduction)

3 Machine Learning on Feature Vectors





4. REQUIREMENT ANALYSIS

4.1 Libraries used:

- Numpy
- Pandas
- MatplotLib
- Sklearn
- OpenCV
- pyttsx3

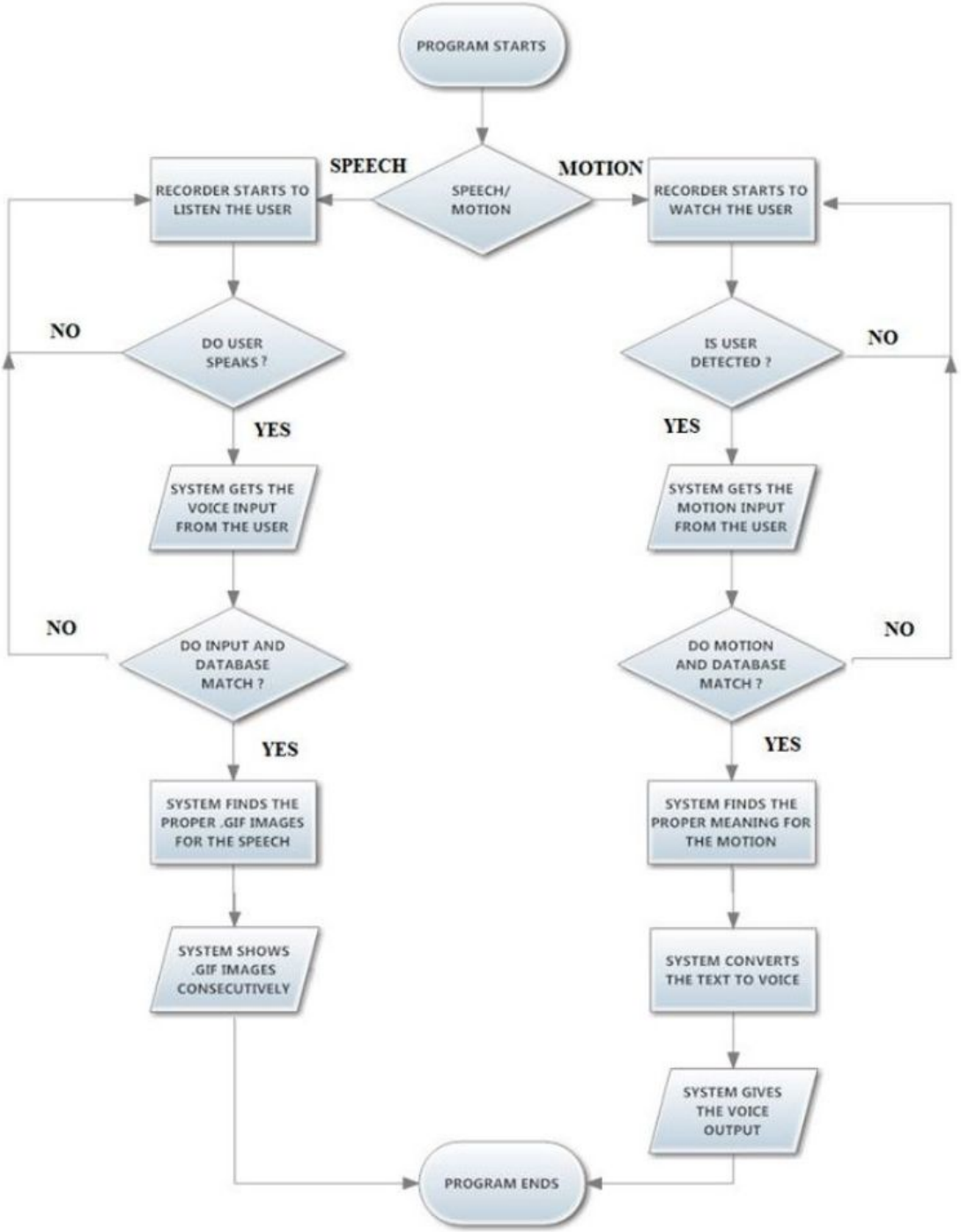
4.2 Software Requirements:

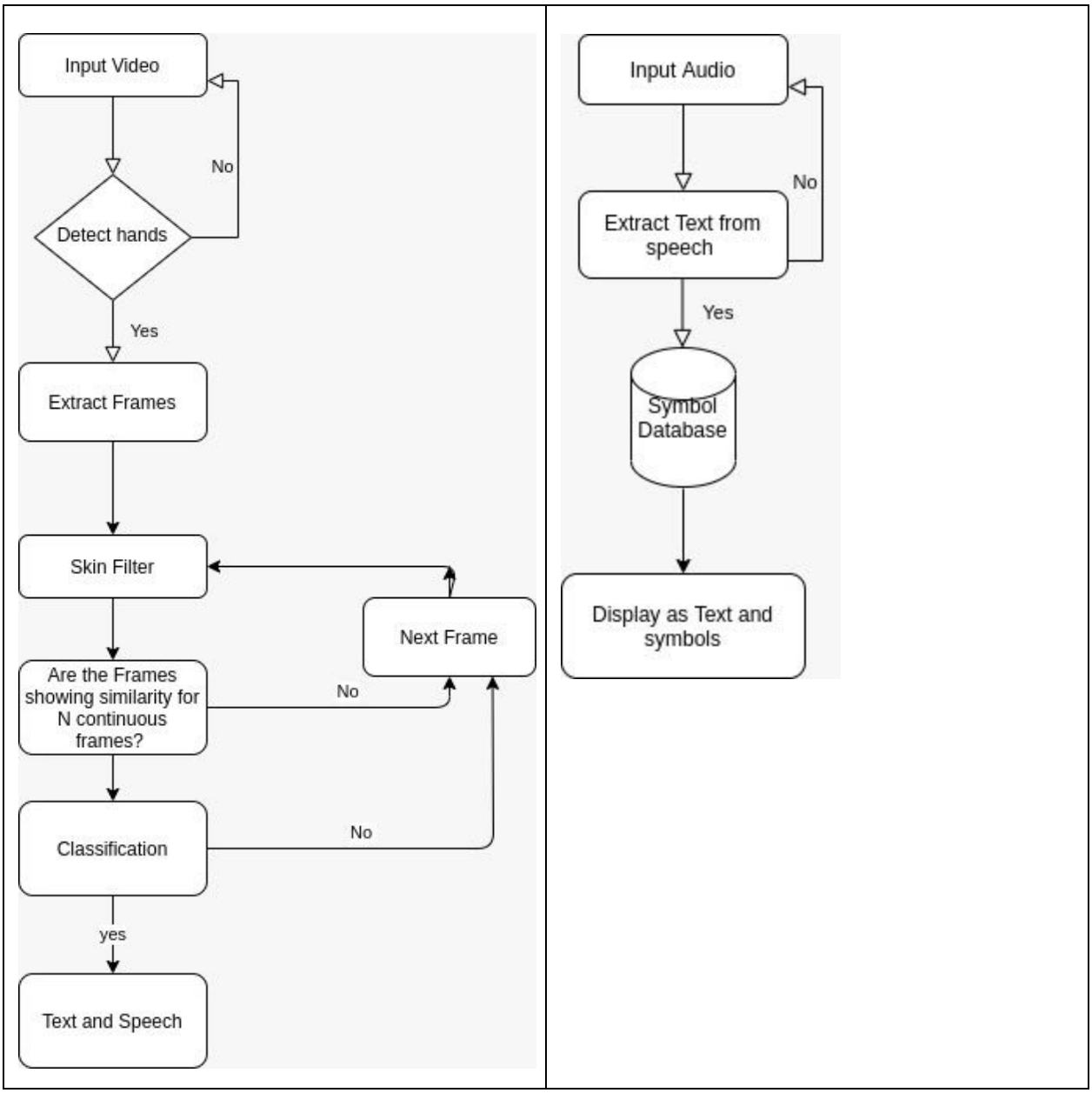
- Anaconda (Jupyter Notebook, Spyder)
- TensorFlow
- Python 3(or higher)

4.3 Hardware Requirements:

- Processor- Ryzen 5 (4th gen)
- RAM- 8 GB
- Disk space- 1 TB + 256 GB(SSD)

5. DETAILED DESIGN




Input Video
No
Detect hands
Yes
Extract Frames
Skin Filter
Next Frame
Are the Frames showing similarity for N continuous frames?
No
Classification
No
yes
Text and Speech
Input Audio
No
Extract Text from speech
Yes
Symbol Database
Display as Text and symbols

6. IMPLEMENTATION

Github Repository(Private) *Code Snippet- How to run project*

git clone https://github.com/eshandhawan51/isl_convertor.git

pip3 install pyaudio speech_recognition pyttsx3 numpy opencv-python

cd isl_convertor-master

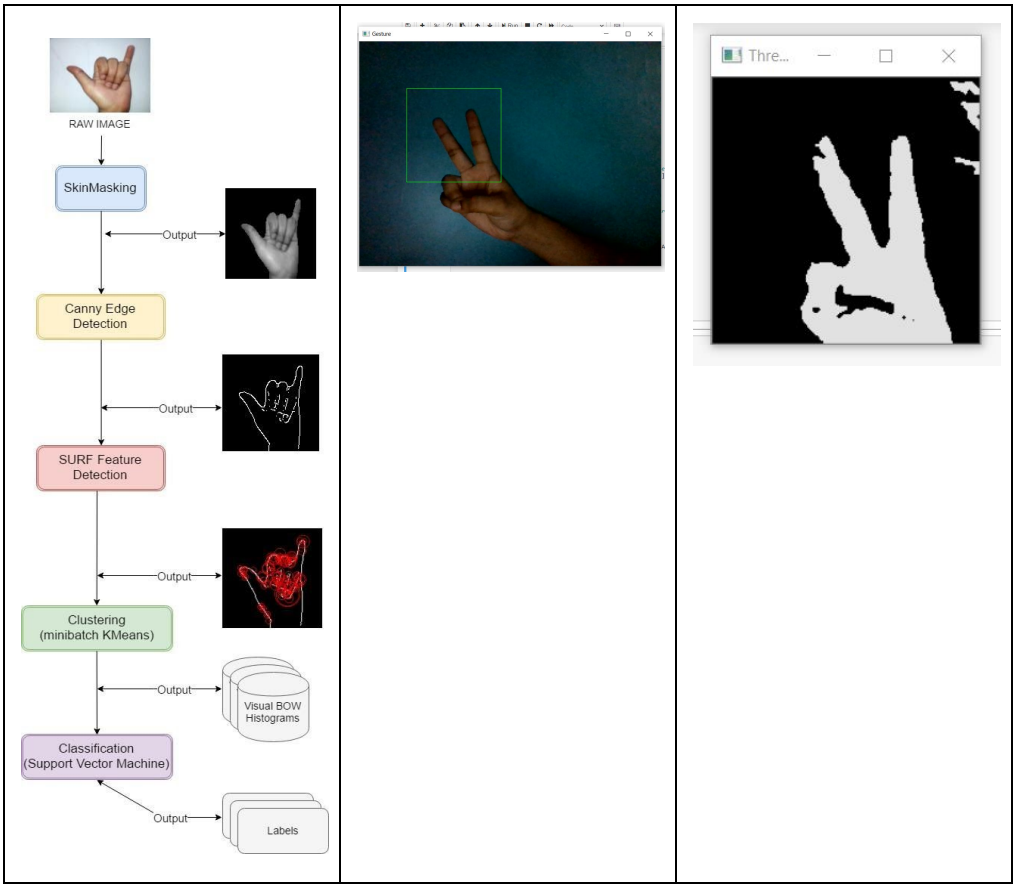
python3

nano main.py

from isl_converter import *

main.py to call all modules and integrate them

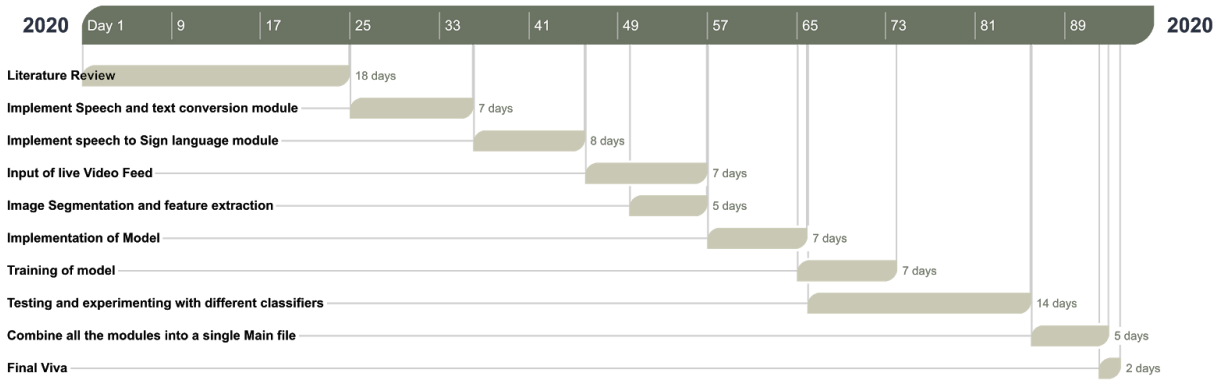
6.1 TESTING REPORTS



7. CONCLUSION

The project gives us the many advantages of the usage area of sign language. After this system, it is an opportunity to use this type of system in many places such as schools, doctor offices, colleges, universities, airports, social services agencies, community service agencies and courts, briefly almost everywhere. One of the most important demonstrations of the ability for communication to help sign language users communicate with each other occurred. Sign languages can be used everywhere when it is needed and it would reach various local areas. The future works are about developing mobile applications of such systems that enable everyone to be able to speak with deaf people.

8. WORK PLAN



Minor Project Timeline

Tabular Representation Of Work Plan

Title	Start date	End date	Duration (in days)
Literature Review	09/01/2020	09/24/2020	18
Implement Speech and text conversion module	09/25/2020	10/05/2020	7
Implement speech to Sign language module	10/06/2020	10/15/2020	8
Input of live Video Feed	10/16/2020	10/26/2020	7
Image Segmentation and feature extraction	10/20/2020	10/26/2020	5
Implementation of Model	10/27/2020	11/04/2020	7
Training of model	11/04/2020	11/12/2020	7
Testing and experimenting with different classifiers	11/05/2020	11/24/2020	14
Combine all the modules into a single Main file	11/25/2020	12/01/2020	5
Final Viva	12/01/2020	12/02/2020	2

REFERENCES

[1] Journal Article

https://ieeexplore.ieee.org/document/7569545 [1]

[2] Published Article

https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf[2]

[3] Research Paper

https://www.researchgate.net/publication/335433017_CONVERSION_OF_SIGN_LANGUAGE_TO_TEXT_AND_SPEECH_USING_MACHINE_LEARNING_TECHNIQUES[3]

[4] Github Repo

https://github.com/imRishabhGupta/Indian-Sign-Language-Recognition[4]

[5] Research Paper

https://research.ijcaonline.org/volume70/number19/pxc3887306.pdf[5]

[6] Technical Report

https://towardsdatascience.com/american-sign-language-hand-gesture-recognition-f1c4468fb177 [6]

[7] Technical Report

https://cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf[7]

[8] Intel Developer Article

https://devmesh.intel.com/projects(Article Removed)
https://github.com/eshandhawan51/isl_convertor/blob/master/Intel%20Blog.md

PPT

https://docs.google.com/presentation/d/1o7NqgoGOwPakD7tsvj0njIwe5cmJmBoUy1DpakZay2Y/edit?usp=sharing

---

[1] "Sign language to speech conversion - IEEE Xplore." https://ieeexplore.ieee.org/document/7569545. Accessed 17 Oct. 2020.
[2] "Conversion of Sign Language into Text - Research India ...." https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf. Accessed 17 Oct. 2020.
[3] "(PDF) CONVERSION OF SIGN LANGUAGE TO TEXT AND ...." 27 Aug. 2019, https://www.researchgate.net/publication/335433017_CONVERSION_OF_SIGN_LANGUAGE_TO_TEXT_AND_SPEECH_USING_MACHINE_LEARNING_TECHNIQUES. Accessed 17 Oct. 2020.
[4] "imRishabhGupta/Indian-Sign-Language-Recognition ... - GitHub." https://github.com/imRishabhGupta/Indian-Sign-Language-Recognition. Accessed 17 Oct. 2020.
[5] "Recognition of Indian Sign Language in Live Video - arXiv." https://arxiv.org/pdf/1306.1301. Accessed 17 Oct. 2020.
[6] "American Sign Language Hand Gesture Recognition | by ...." 13 Dec. 2019, https://towardsdatascience.com/american-sign-language-hand-gesture-recognition-f1c4468fb177. Accessed 17 Oct. 2020.
[7] "Indian Sign Language Character Recognition - CSE-IITK." https://cse.iitk.ac.in/users/cs365/2015/_submissions/vinsam/report.pdf. Accessed 17 Oct. 2020.