MINOR PROJECT 1 (15B19CI591)

ISL ⇌ Speech

An initiative for specially abled people.

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

November 2020

ACKNOWLEDGEMENT

I would like to place on record my deep sense of gratitude to Dr. Raju Pal, ASSISTANT PROFESSOR, Jaypee Institute of Information Technology, India for his generous guidance, help and useful suggestions.

I express my sincere gratitude to Mrs. Payal Khurana Batra Ma'am and Mrs. Varsha Garg Ma'am, Dept. of CSE/IT, for his/her stimulating guidance, continuous encouragement and supervision throughout the course of present work.

I also wish to extend my thanks to my friends and other classmates for their insightful comments and constructive suggestions to improve the quality of this project work.

Signature(s) of Students

Rhythm Agrawal (9918103080)

Eshan Dhawan (9918103066)

Kavya Vishnoi (9918103188)

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and beliefs, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma from a university or other institute of higher learning, except where due acknowledgment has been made in the text.

Place: Jaypee Institute of Information Technology University, Noida

Date: 29th November, 2020

Name: Rhythm Agrawal

Enrolment No.: 9918103080

Name: Kavya Vishnoi

Enrolment No.: 9918103188

Name: Eshan Dhawan

Enrolment No.: 9918103066

CERTIFICATE

This is to certify that the work titled "**ISL ⇌ Speech**" submitted by Rhythm, Kavya, Eshan of B.Tech of Jaypee Institute of Information Technology, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of any other degree or diploma.

Raju Pal

Digital Signature of Supervisor

Name of Supervisor DR. Raju Pal

Designation Assistant Professor

Date 01-12-2020

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}

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ISL Indian Sign Language

ASL American Sign Language

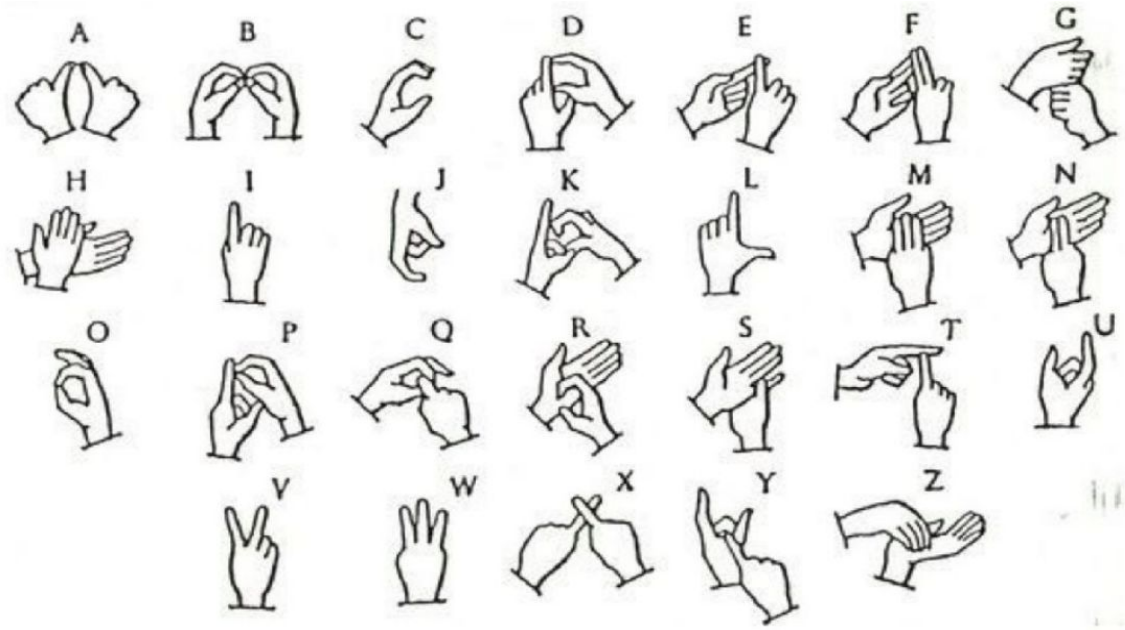
SVM Support Vector Machine

CNN Convolutional neural network

GMM Gaussian Mixture Model

1.1. 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.



1.1 ISL letters

2.1 BACKGROUND STUDY

2 -A . 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)

AnimatedLetters
(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 **K Nearest Neighbour algorithm** as it gives the **best output** and has the **highest accuracy** among all other algorithms Tested .

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)

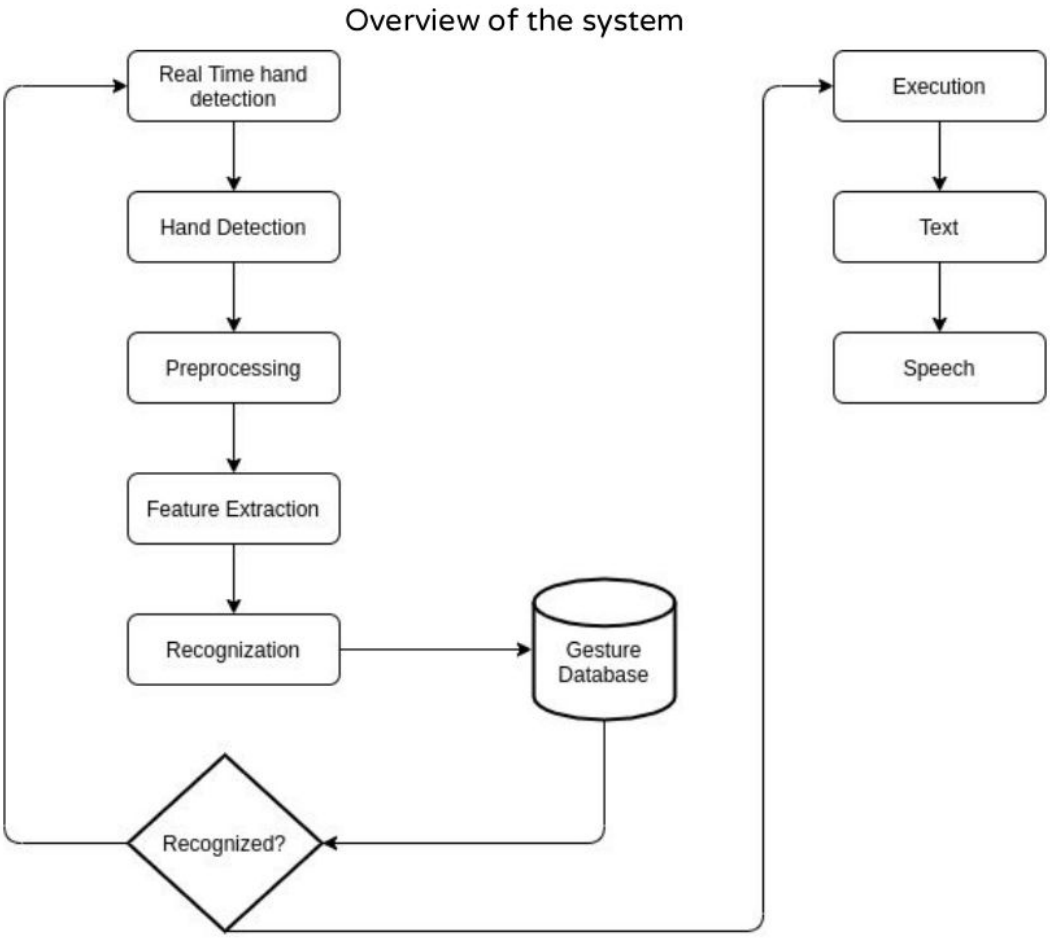
Google Text to Speech API

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

IDE: Jupyter Notebook & Visual Studio Code

Google Cloud Platform (GCP)

Google Colab



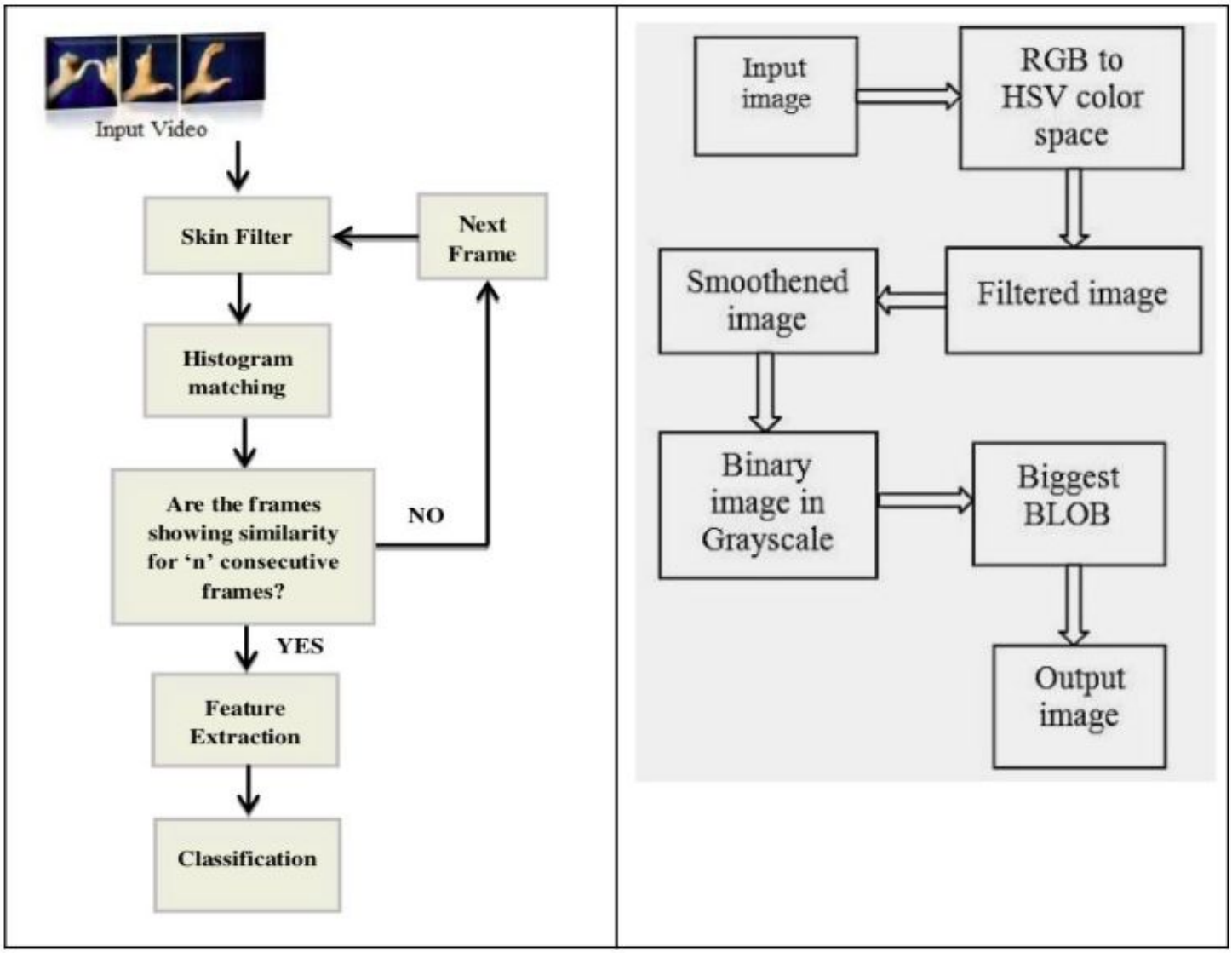
2.1 Overview Of System

2 -B. 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%.



2.1 Flowchart of image processing

3.1 REQUIREMENT ANALYSIS

1 Libraries used:

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

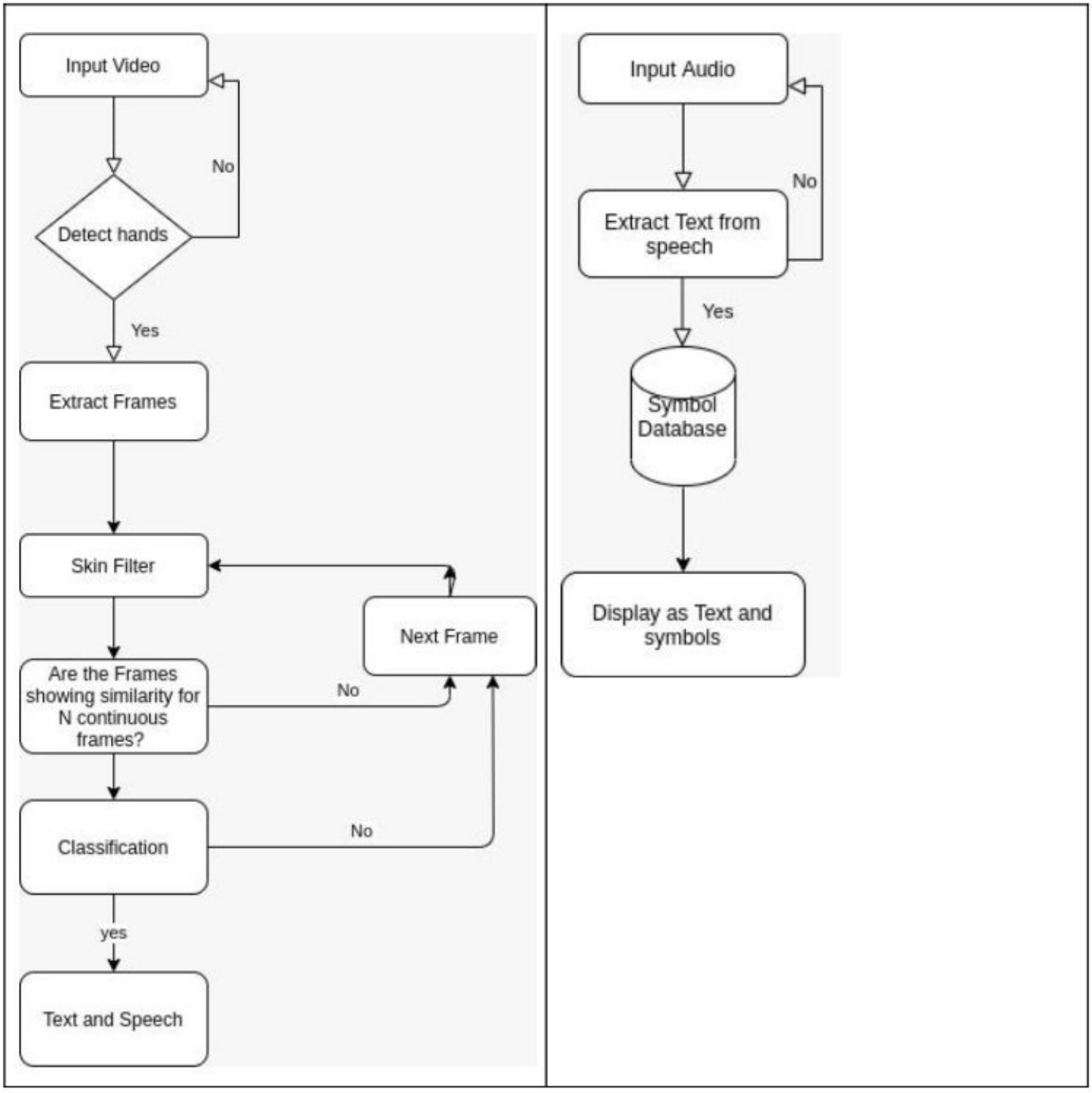
2 Software Requirements:

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

3 Hardware Requirements:

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

4.1 DETAILED DESIGN



Design Flowchart for video to text and speech

Design flowchart for speech to text and symbol

Detailed design flowchart

4.1-A) Image Processing :

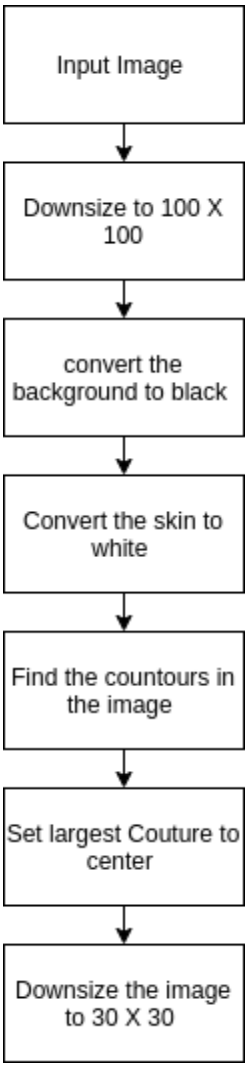


Image processing flowchart

For Image processing all images are downsized to 100 X 100 for easier processing and then HSV filter is applied to the image to change the skin to white and the rest of the background to black. Then contours are identified and the largest contour is taken on the frame and rest of then are removed then the image is again downsized to 30 X 30

4.2.1 Stages of image processing

4.1-B) Data Preprocessing :

For data preprocessing Image transformation is applied to all the images in the dataset and then flattened to one dimension and stored in either test.csv or train.csv for training and testing of the models

4.1-C) Training of Models

We used Sklearn library for models

Tested Models :

i. Logistic Classifier

ii. Support Vector Machine (SVC)

iii. K Nearest Neighbour

4.1-D) Video Input and processing

Open-CV library is used for video Input and then Image Transformation is applied on the frames then the frames are sent to the model to predict the letter.

5.1 IMPLEMENTATION

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

pip3 install pyaudio speech_recognition pyttsx3 numpy opencv-python pickle

cd isl_convertor

python3 main.py

6.1 TESTING REPORTS

The testing report contains details and model score from various models used for testing.

- Model used **knn classifier** model details :

  KNeighborsClassifier(n_neighbors=10)

KNN Model metrics

knn model score from test train split data

0.8896103896103896

Classification report :

	precision	recall	f1-score	support
a	0.71	1.00	0.83	5
b	1.00	1.00	1.00	5
d	1.00	1.00	1.00	12
e	0.75	0.43	0.55	7
f	0.71	1.00	0.83	10
g	1.00	1.00	1.00	7
h	0.88	0.88	0.88	8
j	0.80	1.00	0.89	4
k	0.86	0.86	0.86	7
m	0.86	0.67	0.75	9
n	1.00	0.91	0.95	11
o	1.00	1.00	1.00	8
p	0.78	1.00	0.88	7
q	0.73	0.89	0.80	9
r	1.00	0.71	0.83	7
s	0.91	1.00	0.95	10
t	1.00	1.00	1.00	9
x	1.00	0.60	0.75	5
y	1.00	1.00	1.00	7
z	1.00	0.71	0.83	7
accuracy			0.89	154
macro avg	0.90	0.88	0.88	154
weighted avg	0.90	0.89	0.89	154

knn model score from External data

0.18485523385300667

Classification report :

	precision	recall	f1-score	support
a	0.09	0.50	0.15	22
b	0.67	0.14	0.23	57
d	0.45	0.45	0.45	22
e	0.00	0.00	0.00	11
f	0.36	0.36	0.36	22
g	0.20	0.50	0.29	22
h	0.00	0.00	0.00	22
j	0.00	0.00	0.00	0
k	0.55	0.27	0.36	22
m	0.00	0.00	0.00	22
n	0.00	0.00	0.00	21
o	0.60	0.95	0.74	22
p	0.00	0.00	0.00	11
q	0.80	0.20	0.32	20
r	0.00	0.00	0.00	11
s	0.00	0.00	0.00	11
t	0.67	0.09	0.15	47
x	0.00	0.00	0.00	52
y	0.00	0.00	0.00	0
z	0.00	0.00	0.00	32
accuracy			0.18	449
macro avg	0.22	0.17	0.15	449
weighted avg	0.30	0.18	0.18	449

5.1.1 KNN Model Stats

- Model used **logistic classifier** model details :

  LogisticRegression(max_iter=10000)

Logistic classifier Model metrics

logistic model score from test train split data

0.9285714285714286

Classification report :

	precision	recall	f1-score	support
a	1.00	0.75	0.86	8
b	1.00	1.00	1.00	6
d	1.00	1.00	1.00	5
e	1.00	0.93	0.97	15
f	1.00	1.00	1.00	7
g	1.00	1.00	1.00	8
h	0.83	1.00	0.91	5
j	1.00	1.00	1.00	5
k	1.00	1.00	1.00	6
m	1.00	0.88	0.93	8
n	0.91	1.00	0.95	10
o	1.00	1.00	1.00	4
p	0.92	1.00	0.96	11
q	1.00	0.70	0.82	10
r	0.67	0.75	0.71	8
s	0.90	1.00	0.95	9
t	0.91	1.00	0.95	10
x	1.00	1.00	1.00	4
y	0.89	0.89	0.89	9
z	0.71	0.83	0.77	6
accuracy			0.93	154
macro avg	0.94	0.94	0.93	154
weighted avg	0.94	0.93	0.93	154

logistic model score from External data

0.1603563474387528

Classification report :

	precision	recall	f1-score	support
a	0.02	0.05	0.02	22
b	0.76	0.23	0.35	57
d	1.00	0.18	0.31	22
e	0.00	0.00	0.00	11
f	0.62	0.59	0.60	22
g	0.24	0.50	0.33	22
h	0.58	0.50	0.54	22
k	0.00	0.00	0.00	22
m	0.00	0.00	0.00	22
n	0.00	0.00	0.00	21
o	0.50	0.55	0.52	22
p	0.06	0.09	0.07	11
q	0.03	0.05	0.04	20
r	0.00	0.00	0.00	11
s	0.05	0.09	0.07	11
t	0.29	0.09	0.13	47
x	0.00	0.00	0.00	52
y	0.00	0.00	0.00	0
z	0.00	0.00	0.00	32
accuracy			0.16	449
macro avg	0.22	0.15	0.16	449
weighted avg	0.28	0.16	0.18	449

- Model used **SVM classifier** model details :

  SVC(degree=1, kernel='poly')

SVC Model metrics

svm model score from test train split data

0.7857142857142857

Classification report :

	precision	recall	f1-score	support
a	1.00	0.80	0.89	10
b	0.88	1.00	0.93	7
d	1.00	1.00	1.00	3
e	1.00	0.75	0.86	4
f	1.00	1.00	1.00	7
g	1.00	1.00	1.00	6
h	0.75	1.00	0.86	9
j	0.67	1.00	0.80	2
k	0.33	0.17	0.22	6
m	0.19	0.71	0.30	7
n	1.00	0.60	0.75	15
o	1.00	1.00	1.00	7
p	1.00	0.54	0.70	13
q	1.00	0.55	0.71	11
r	0.64	0.82	0.72	11
s	1.00	0.78	0.88	9
t	1.00	1.00	1.00	5
x	1.00	1.00	1.00	8
y	1.00	0.90	0.95	10
z	1.00	0.75	0.86	4
accuracy			0.79	154
macro avg	0.87	0.82	0.82	154
weighted avg	0.89	0.79	0.81	154

svm model score from External data

0.17371937639198218

Classification report :

	precision	recall	f1-score	support
a	0.03	0.18	0.05	22
b	0.71	0.09	0.16	57
d	0.50	0.41	0.45	22
e	0.00	0.00	0.00	11
f	0.54	0.64	0.58	22
g	0.29	0.50	0.37	22
h	0.52	0.50	0.51	22
j	0.00	0.00	0.00	0
k	0.00	0.00	0.00	22
m	0.00	0.00	0.00	22
n	0.00	0.00	0.00	21
o	0.59	0.91	0.71	22
p	0.09	0.09	0.09	11
q	0.06	0.10	0.07	20
r	0.00	0.00	0.00	11
s	0.00	0.00	0.00	11
t	0.06	0.02	0.03	47
x	0.00	0.00	0.00	52
y	0.00	0.00	0.00	0
z	0.00	0.00	0.00	32
accuracy			0.17	449
macro avg	0.17	0.17	0.15	449
weighted avg	0.22	0.17	0.16	449

7.1 EXPERIMENTAL RESULTS & ANALYSIS

After testing various models and techniques We came to the conclusion that K nearest Neighbour Classifier yields the best results with the highest model scores and accuracy

6.1.1 Model Scores Over Train test Split

Models	**Test results over Train test split**
Logistic Regression	0.9285714285714286
K-Nearest Neighbors (K-NN)	0.8896103896103896
Support Vector Machine (SVM)	0.7857142857142857

6.1.2 Model Score Over External test data

Models	**Test Results Over External Data**
Logistic Regression	0.1603563474387528
K-Nearest Neighbors (K-NN)	0.18485523385300667
Support Vector Machine (SVM)	0.17371937639198218

Since **K Nearest Neighbour** has the highest Accuracy in the external Test data so we selected that model.

8.1 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. 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.

8.2 FUTURE SCOPE

Although we tried our best but we know there is a lot of scope to improve our model as well as its accuracy. First challenge which we faced was that of the limited dataset we have, we all know in the Data Science fiel only data matters, i mean without sufficient data we can’t predict or speculate anything. The root cause behind is due to covid we were not able to visit divyang centre where we can increase our existing dataset. We may also use deep learning to increase our model efficiency using various kind of neural network algorithms. We may also use some hardware gloves which have sensors hence it will give us the most accurate features having features clearly extracted. Also technology may evolve further where we don;t even have to wear those gloves which are not convenient to carry instead we may develop some kind of stickers which contain sensors.

REFERENCES

[1] P Vijayalakshmi & M Aarthi. (2016). Sign language to speech conversion. 2016 International Conference on Recent Trends in Information Technology (ICRTIT)

https://ieeexplore.ieee.org/document/7569545

[2] Adewale, Victoria & Olamiti, Adejoke. (2018). CONVERSION OF SIGN LANGUAGE TO TEXT AND SPEECH USING MACHINE LEARNING TECHNIQUES. 5. 58-65.
Link

[3] imRishabhGupta.(October, 2020) Indian-Sign-Language-recognition. Github repository

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

[4] ROHIT GHUMARE. Indian-Sign-language-to-speech-conversion

https://devmesh.intel.com/projects(Article Removed)

https://github.com/eshandhawan51/isl_convertor/blob/master/Intel%20Blog.md

[5] Iron-stark. (November,2020). Indian-Sign-Language

https://github.com/Iron-Stark/Indian-Sign-Language