SIGN TO TEXT AND SPEECH TRANSLATOR USING MACHINE LEARNING

GUIDED BY
Anoop P V

GROUP MEMBERS:
Daniya Mariya Jiju
Jishna Babu
Diniya P
Divya A

TABLE OF CONTENTS

- INTRODUCTION

- PROBLEM DEFINITION

- OBJECTIVES

- LITERATURE SURVEY

- FEASIBILITY STUDY

- MODULES

- DFD

- BLOCK DIAGRAM

- SRS

- USECASE DIAGRAM
- SYSTEM ARCHITECTURE
- WORK ALLOCATION
- GANT CHART
- STATUS
- DATASETS
- METHODOLOGY
- IMPLEMENTATION
- SNAPSHOTS
- CONCLUSION

INTRODUCTION

❑ Sign language recognition (SLR) is a computer vision technique that uses machine learning to recognize and identify signs.

❑ SLR systems typically use a combination of hand shape, hand position, and movement to recognize signs.

❑ An accumulative video motion (AVM) technique to encode the sign motions in the video stream into a single image

PROBLEM DEFINITION

❑ The challenge lies in bridging these gaps to create a more inclusive and accessible environment for everyone.

❑ One significant obstacle is the lack of seamless communication tools that can effortlessly translate sign language into text and speech.

❑ Enabling meaningful and real-time conversations between individuals with diverse communication needs.

OBJECTIVES

❑ **Develop a Reliable Sign Language Translation System:**

To bridge the communication gap between deaf and mute individuals and those who do not understand sign language.

❑ **Establish a User-Friendly Interface :**

To make the computer understand human sign language. Develop a mobile and desktop application

LITERATURE SURVEY

SLNO	PAPER	YEAR	AUTHOR	HIGHLIGHTS	LIMITATIONS
1	An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion	2022	Hamzah Luqman	Use CNN algo for higher accuracy, Robustness to Variability	Computational expensive to train and run, Bias
2	Sign Language to Text-Speech Translator Using Machine Learning	2021	Akshatha Rani K 1 , Dr. N Manjanaik	Learn discriminative features from a limited dataset	Less accessible to people with limited resource, accurate
3	Hand gesture recognition Using PCA	2015	Dr. Amardeep Singh	Improve the recognition accuracy in low brightness	Limited ability to complex gestures, cost is high

4	Myo-Sign Enabling end to end Sign language recognition with wearable	2019	Dong Wang	63 sign evaluated 93% accurate, Mobility, Customization	Signers have small hand it will fail.
5	Use deep convolution network for gesture recognition in ASL	2017	Vivek Bhedha, N. Dianna Radpour	Reduce the risk of overfitting, Understand complex hand shapes	Less accurate
6	Hand gesture recognition based on Karhumen_Lore transform	2013	Karhumen-Lore	Improve the recognition accuracy	Tested for 10 different hand gestures

FEASIBILITY STUDY

1. Technical feasibility :

- Resource availability

- Technology

2. Economic feasibility :

- Provide economic merit to firms

- Improve QoS

HARDWARE REQUIREMENTS

- Processor : Pentium Dual Core 2.46 or above
- Hard disk : 320 GB and above
- RAM : 4 GB or above
- Android phones with 4gb RAM (above)

SOFTWARE REQUIREMENTS

- OS : Windows 7 or above
- Coding Language : Python, Java, HTML
- Framework : Django
- Backend : MySQL
- IDE : Pycharm

CONT…..

- Android Studio : Sdk with API 4.0 OR higher

- Packages : Numpy, OpenCv, tensorflow, Mediapipe, tkinter,

  hunspell, pyttsx, pillow, Django.

MODULES

> ADMIN
>> Can Manage signs (add, view, delete), View users, feedback, complaint,
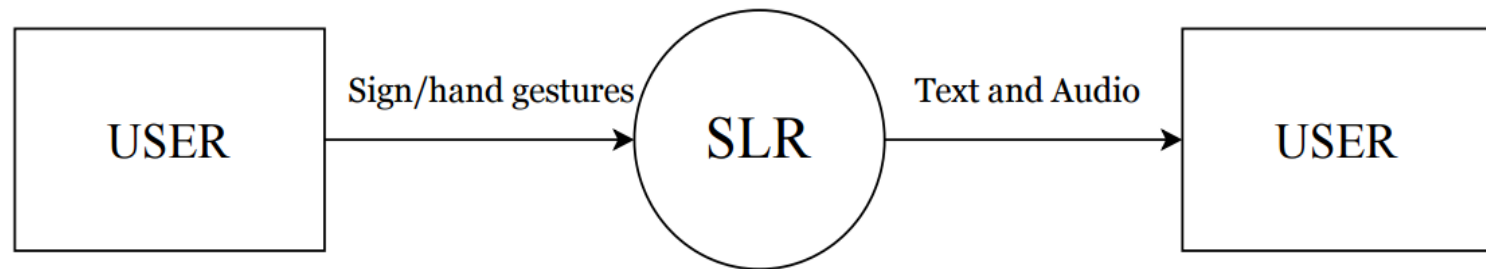>> Send reply etc.

> USERS
>> Can Register, login, Sign recognition, Send feedback and Complaints,
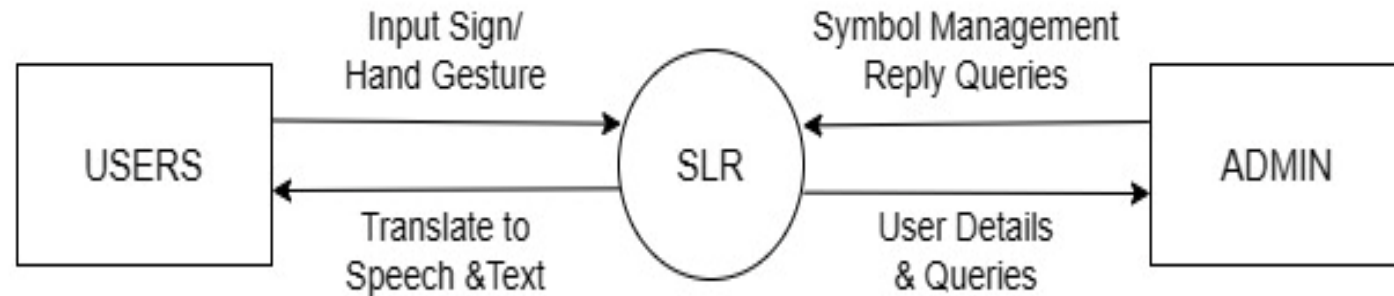>> View replies and sign

> SYSTEM
>> Can Capture (video accumulation and upload via files) image given by user,
>> Preprocess, Segment, Extract features, Recognize gesture, Translation,
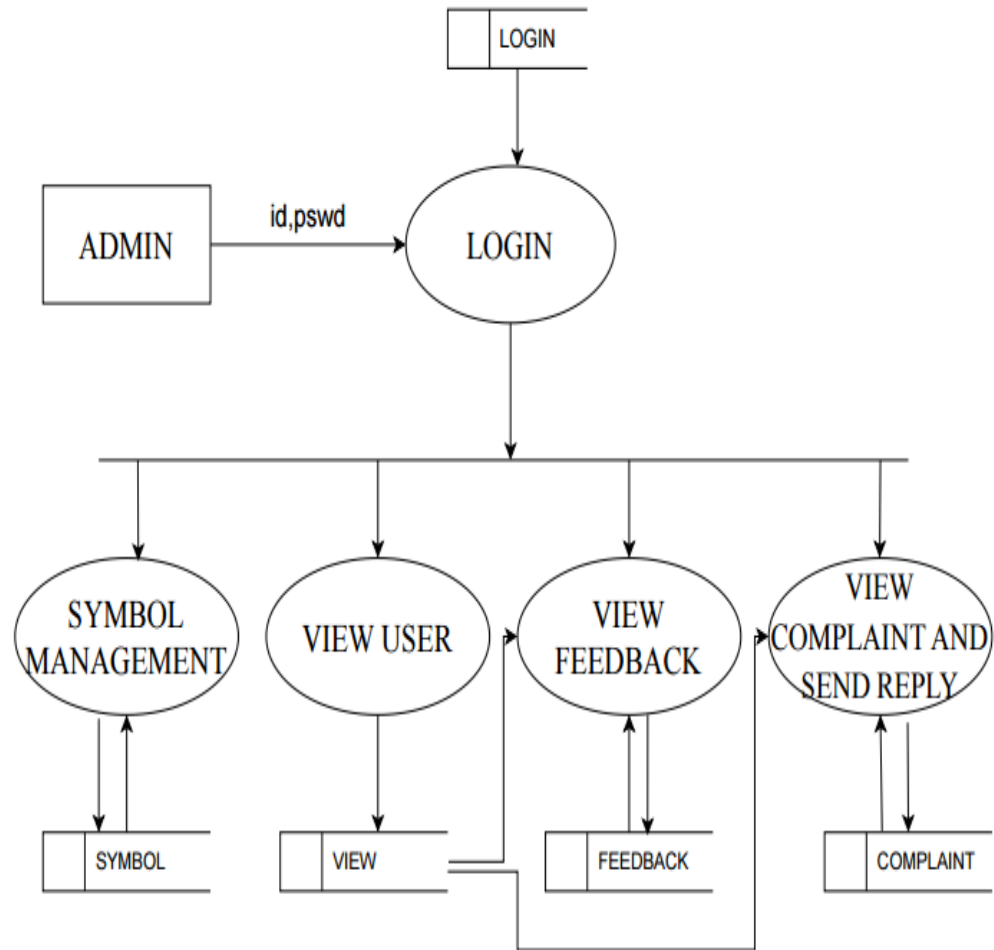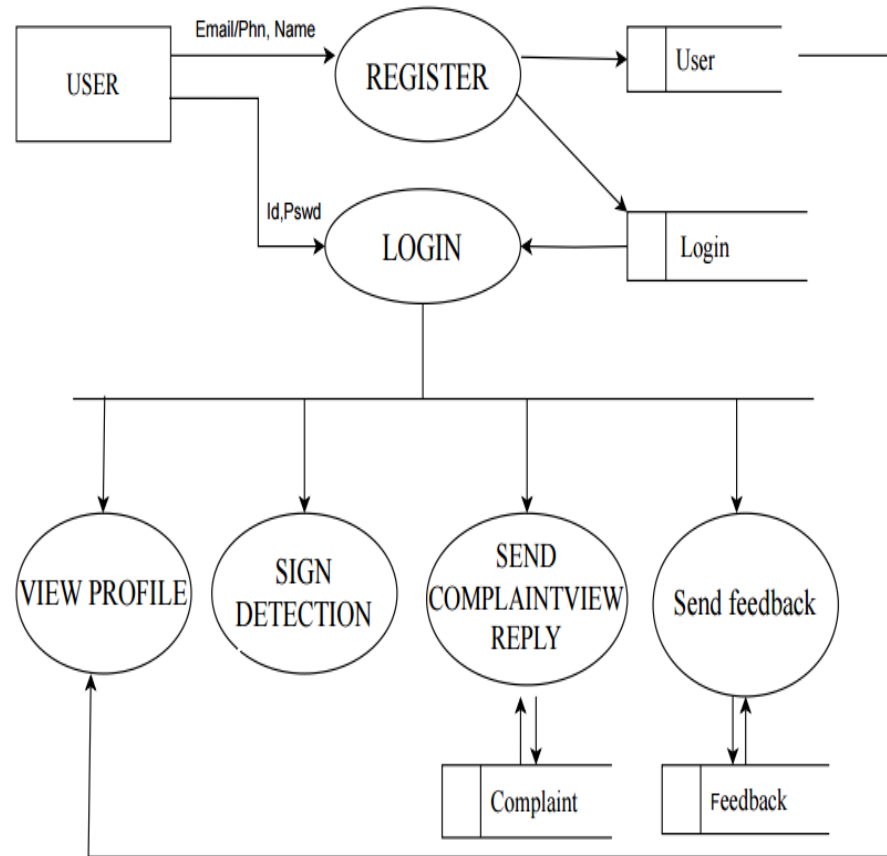>> Result displayed.

DATA FLOW DIAGRAM
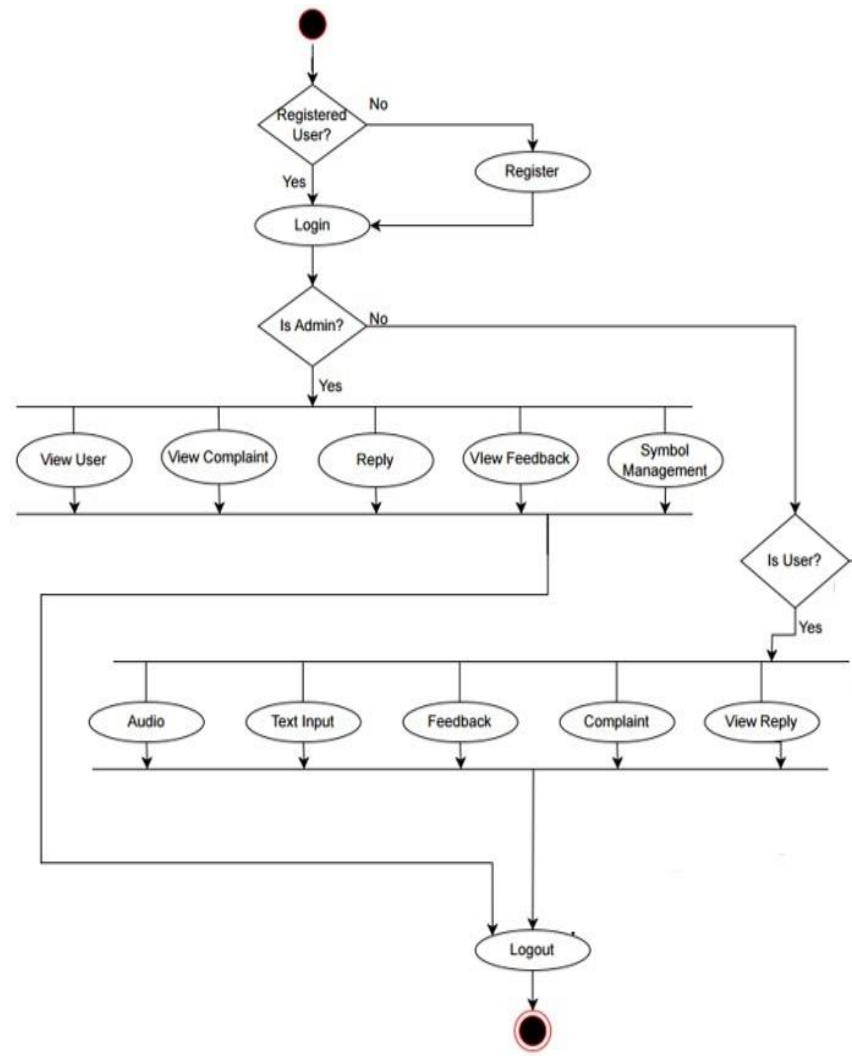
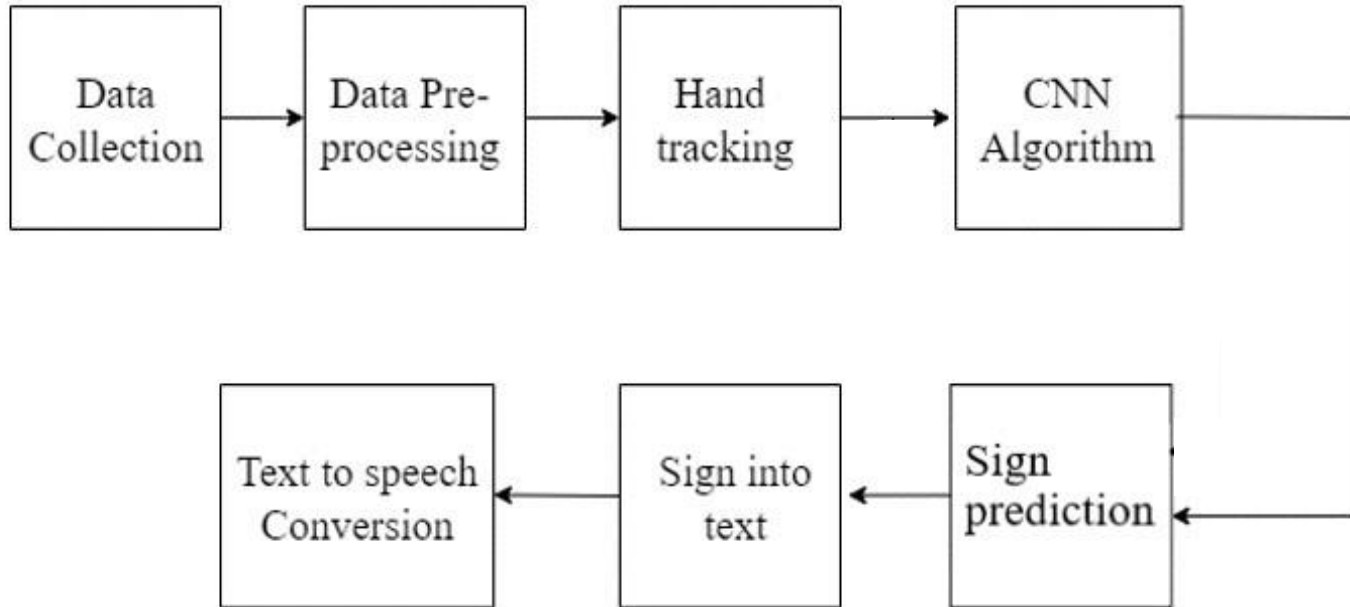# LEVEL0

LEVEL 1

LEVEL 1.1 for ADMIN

LEVEL 1.2 for USER

ACTIVITY DIAGRAM

BLOCK DIAGRAM

CONTROL FLOW DIAGRAM
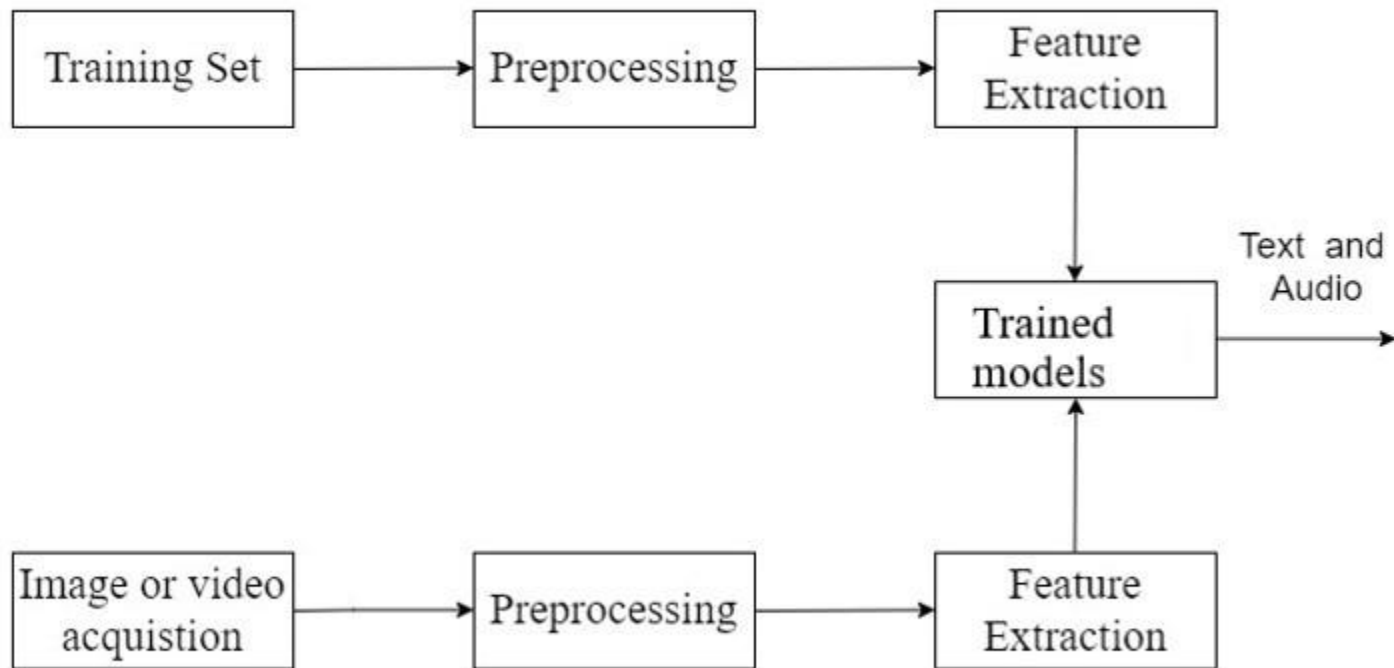
USECASE
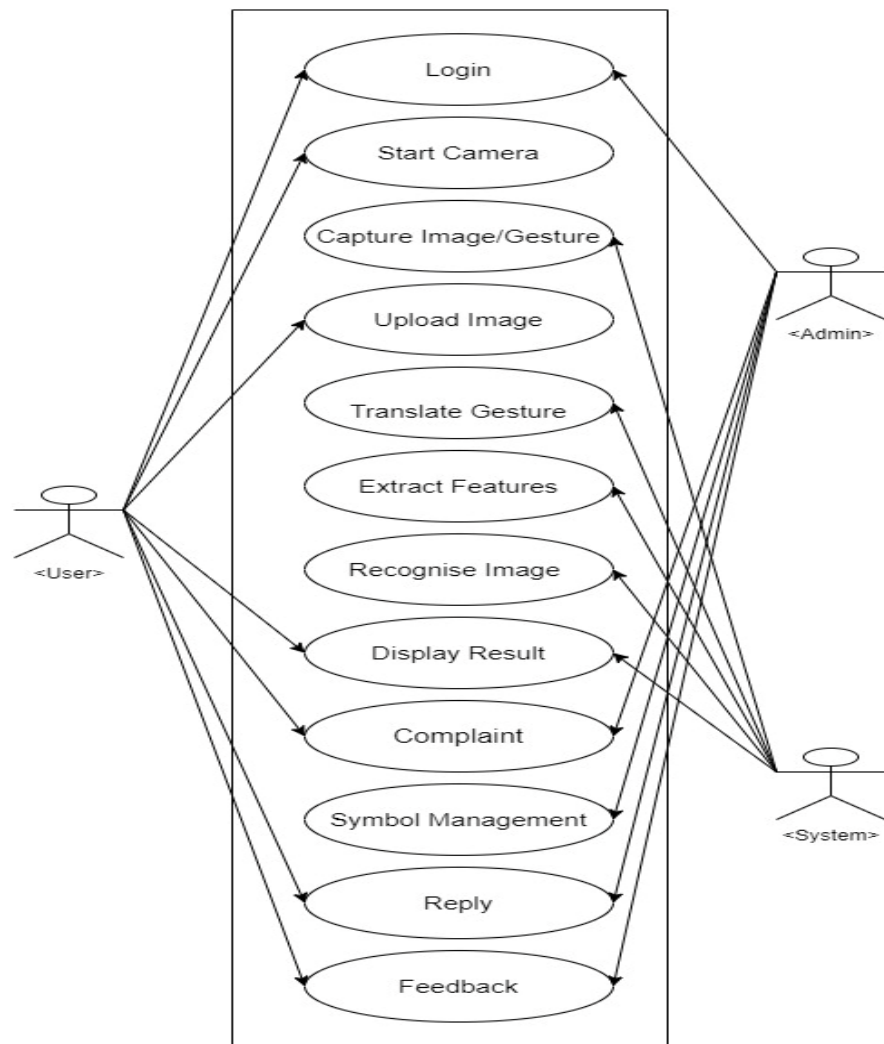
SYSTEM ARCHITECTURE
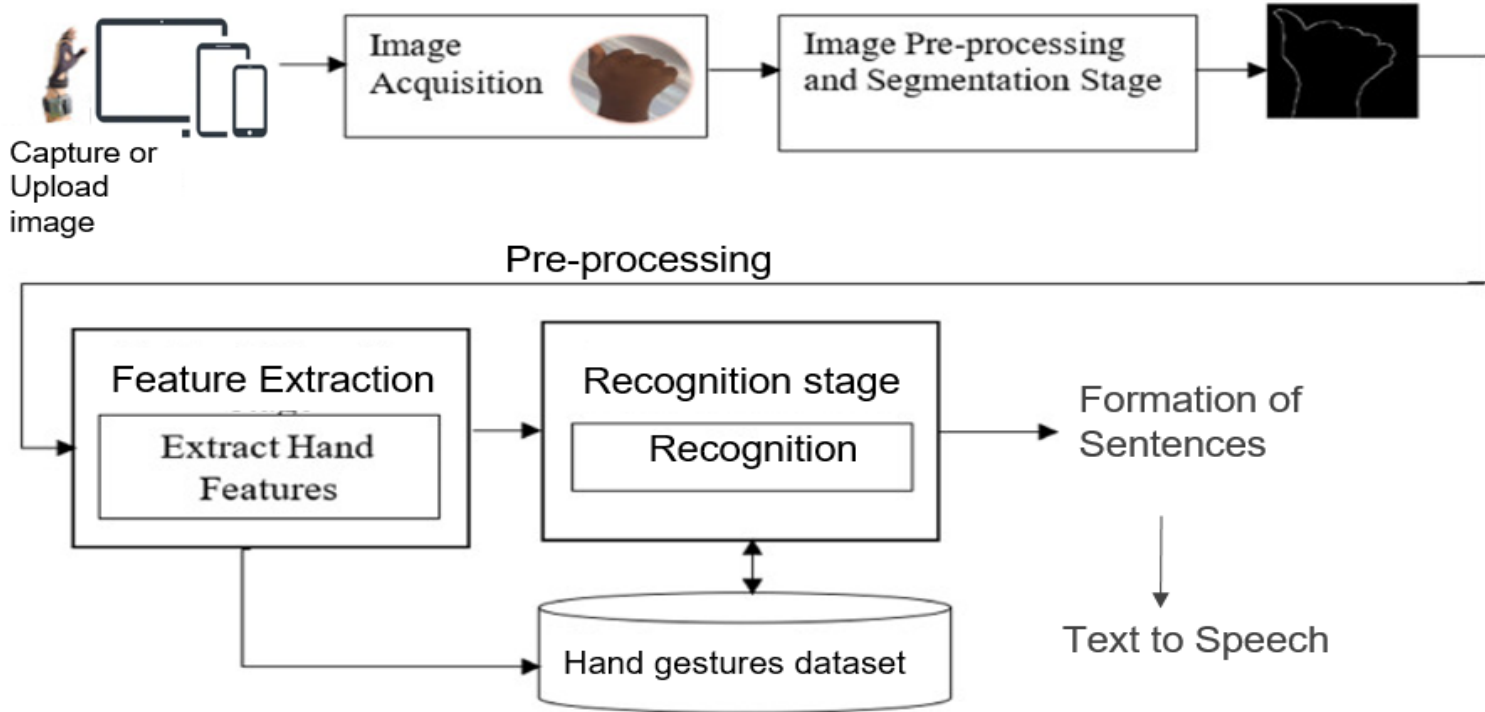
WORK ALLOCATION

SL NO	NAME	WORK ALLOCATED
1	Daniya Mariya Jiju	Dataset collection(ASL), Literature survey, Project Report, Admin Module, Main Portion
2	Diniya P	Literature survey, Dataset collection(ASL), Project Report, Main Portion, User module
3	Divya A	Literature survey, Documentation(ppt), Dataset collection, Main Portion Admin Module
4	Jishna Babu	Literature survey, Documentation(ppt), Dataset collection, Main Portion User module

GANTT CHART
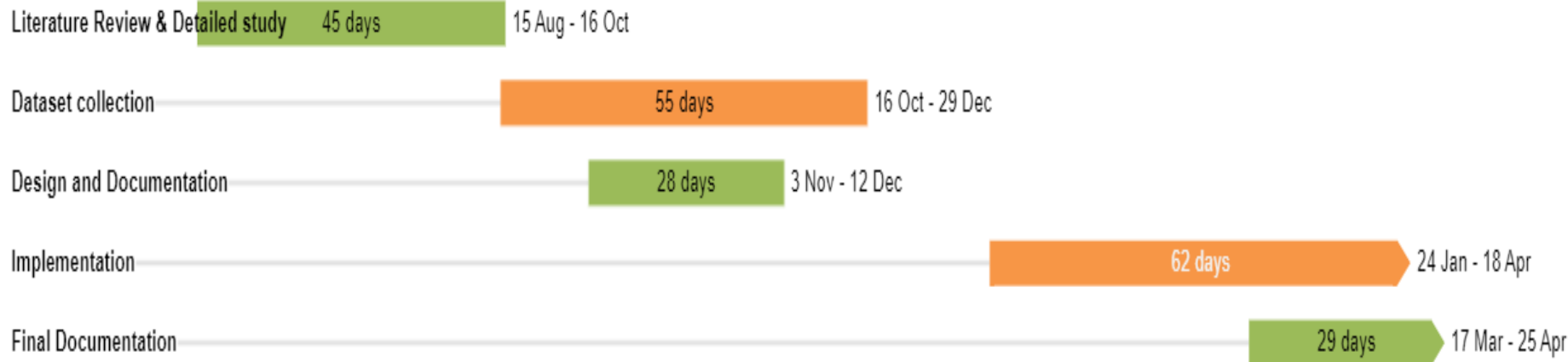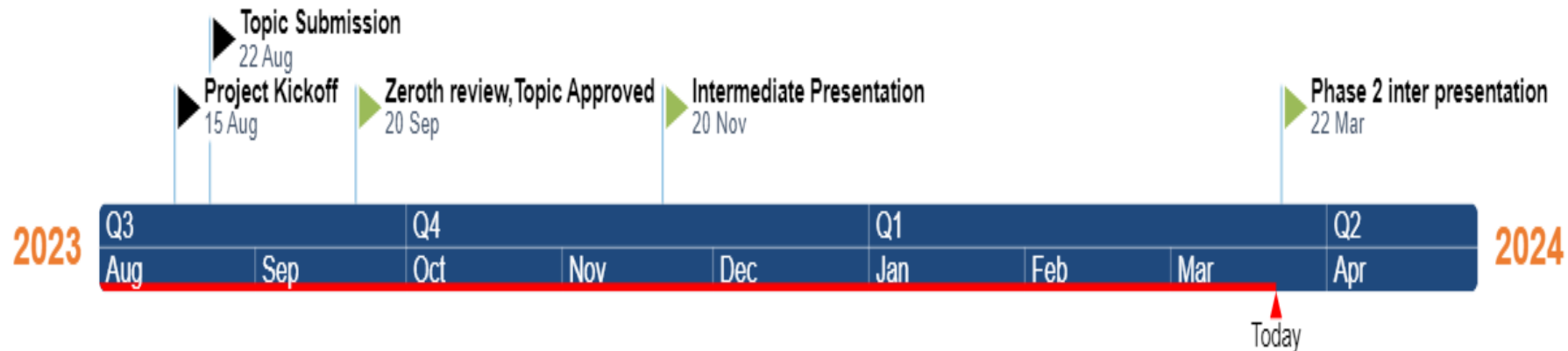
DATASETS
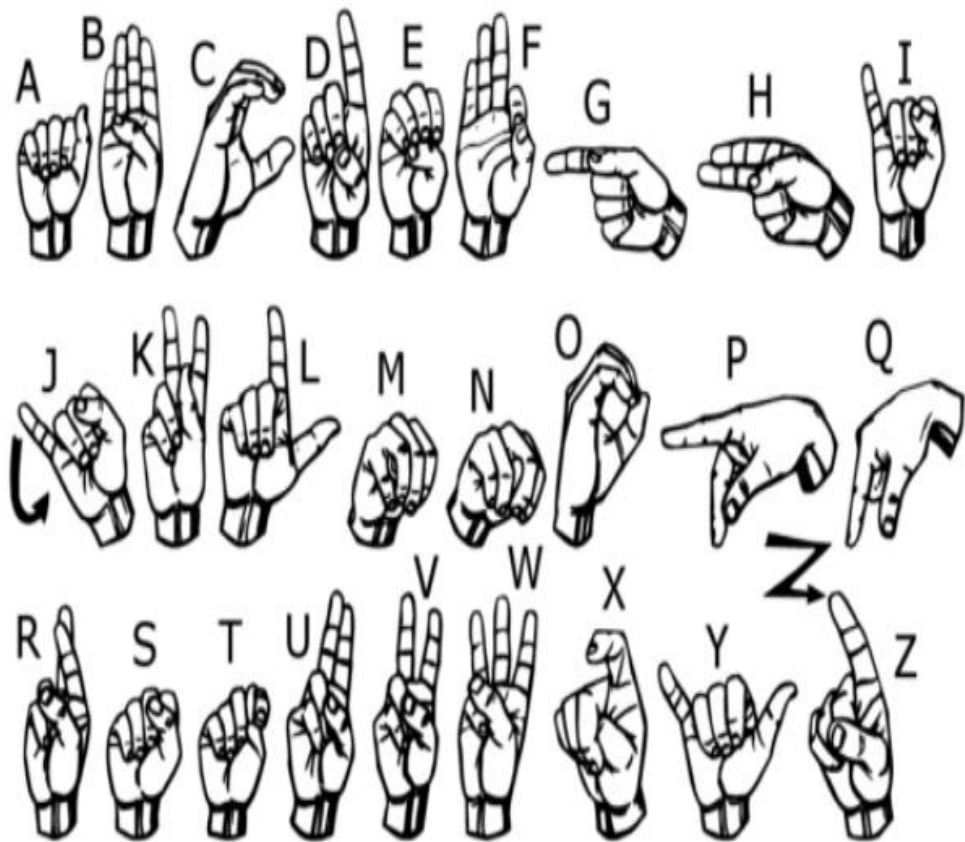
METHODOLOGY

- In Desktop application users can input images via video accumulation, then system capture gestures and extract features, result will displayed if given image is matched with trained data. Also users can put alphabet as input system will generate corresponding output as video(combining given text)

- In mobile application users can upload images via files. Pre-processing and training as like Desktop application. Here users can view all sign as per words and ASL.

GESTURE CLASSIFICATION
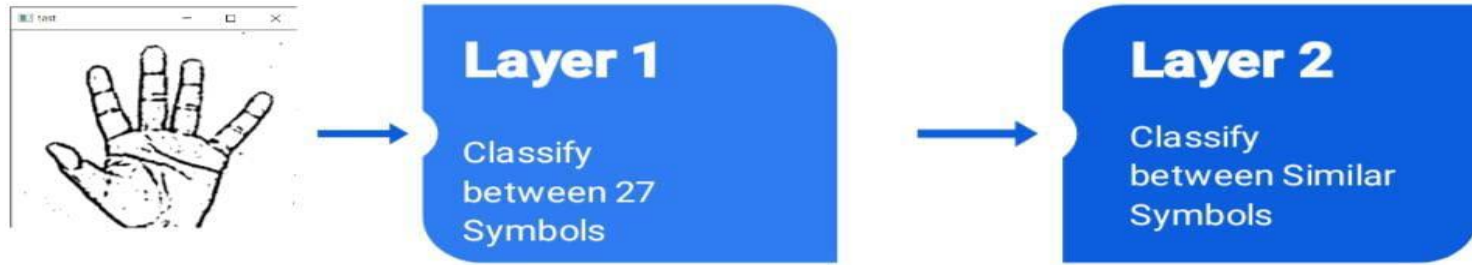
❑ALGORITHM LAYER 1:

i. Apply gaussian blur filter and threshold to the frame taken with opencv to get the processed image after feature extraction.

ii. This processed image is passed to the CNN model for prediction and if a letter is detected for more than 50 frames then the letter is printed and taken into consideration for forming the word.

iii. Space between the words are considered using the blank symbol.

❑ ALGORITHM LAYER 2:

i.   We detect various sets of symbols which show similar results on getting detected.

ii.   We then classify between those sets using classifiers made for those sets only.

iii.   In our testing we found that following symbols were not showing properly and were giving other symbols also :

   a.   For D   : R and U

   b.   For U   : D and R

   c.   For I   : T, D, K and I

   d.   For S   :   M and N

IMPLEMENATION

❑ System use Django framework for web application.

❑ TensorFlow framework used for bulid and train the models.

❑ Sklearn is a python package used to find accuracy and train and test dataset

❑ Matplotlib, Seaborn, TensorBoard(visualize graph) used for evaluation and optimization.

❑ Graph plotted using CNN algorithm.

❑ Pillow for image processing

❑ Pyttsx used to convert text to speech

Cont…

❑ MediaPipe framework for performing computer vision (video accumulation)

❑ tkinter used to create GUI

❑ OpenCv for image processing, video capture, analysing features.

❑ Softmax and relu as activation functions.

❑ Hunspell used to suggest correct alternatives for each input word.(spell checker).

❑ Adam optimizer is used as optimization techniques.

Case id	Description	Expected Result	Actual result	Status
1	Upload trained image as input	Proper Speech and text translation displayed	Proper Speech and text translation displayed	PASS
2	Upload a non-trained image as input	Output will displayed as per matching trained dataset	Output will displayed as per matching trained dataset or Bad request	PASS
3	Capturing images without background substration	Accuracy of recognition varies	Accuracy of recognition varies	PASS
4	Capturing images with background substration	Display correct output	Display correct output	PASS
5	Sending a empty complaint and feedback by user	Error message will be displayed with suggestions	Error message will be displayed with suggestions	PASS
6	Enter valid key for getting their sign	Result will displayed as video	Result will displayed as video	PASS
7	Login and sign_up with invalid credentials	Error message will displayed with suggestions	Error message will displayed with suggestions	PASS
8	Enter invalid key for getting their sign	Attribute Error	Attribute Error	PASS

SNAPSHOTS OF WORKING & RESULTS

Welcome

EMAIL

PASSWORD

LOGIN

**Don't have an account?
Please Sign up!**

SIGN UP

If you already has an
account, just sign in.

SIGN IN

Create your Account

NAME

Diniya

EMAIL

din

⚠ Please include an '@' in the email address. 'din' is missing an '@'.

8592011725

PASSWORD

••••••••

REGISTER

Add Symbol

View Symbol

SL NO	USER_NAME	FEEDBACKS	RATING
1	user	good	★★★★☆
2	diniya	nice	★★★☆☆
3	divya	good	★★★☆☆
4	Jishna	gd	★★★★☆
5	daniya	user friendly	★★★★☆
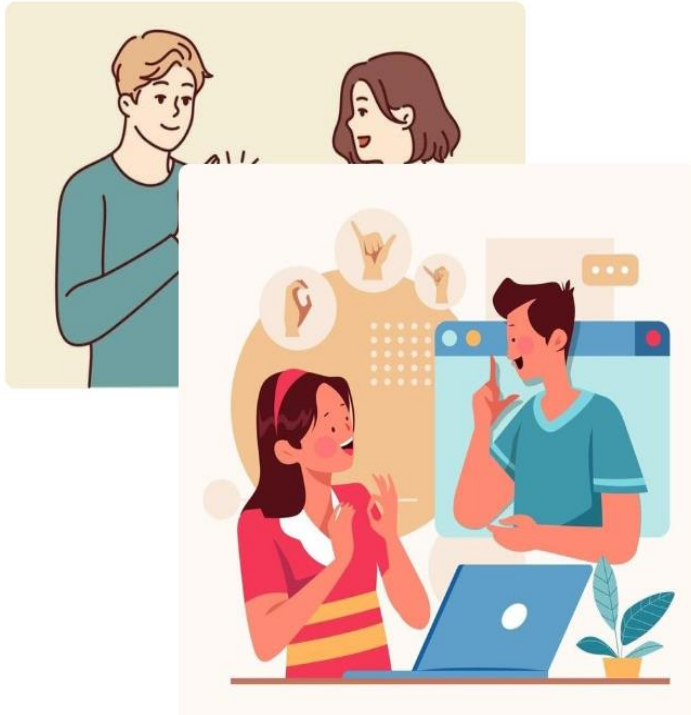
About Us

About Sign Translator And It's Innovative IT Solutions

This website is dedicated to providing sign language and speech synthesis services. Our mission is to make communication more accessible for everyone, regardless of their abilities or limitations.

Our Features:

- Text-to-Speech Conversion:Convert written text into spoken words.
- Sign Language Translation: Translate text into sign language gestures.
- User-Friendly Interface: Our platform is designed to be intuitive and easy to use for all users.
- Accessibility: We strive to ensure our website is accessible to all users, including those with disabilities.

Signs   Sign Translator   Feedback   Complaint ⌄   Logout

Enter the text....

Submit

Sign Language To Text Conversion
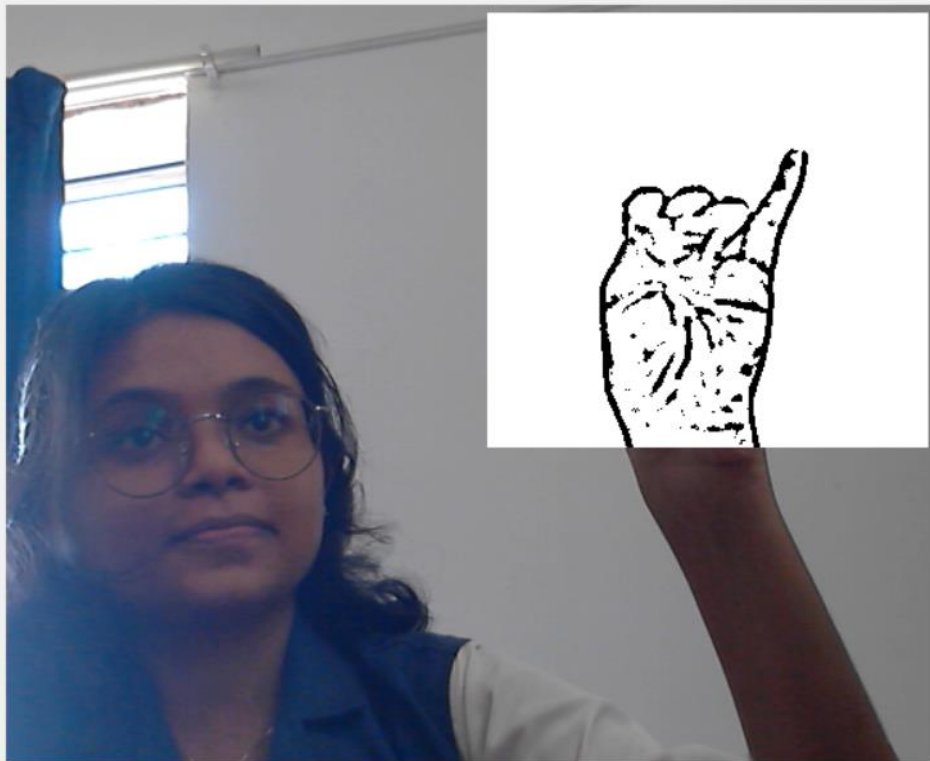
Character :

I

Clear

Speak

Word :   HAI
Sentence :

SLR

Please

UPLOAD

SLR

Login

maya@gmail.com

••••••••

LOGIN

SLR

Enter your complaint

Required

SUBMIT
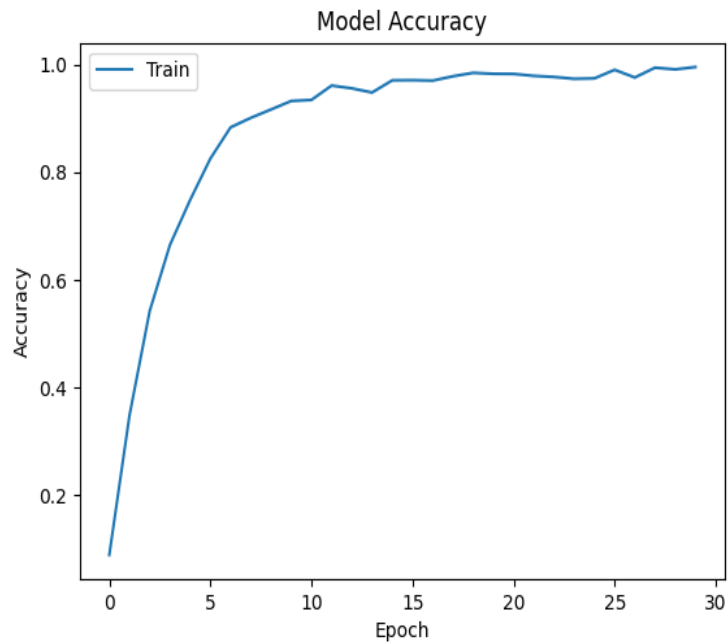
Accuracy curve for ASL
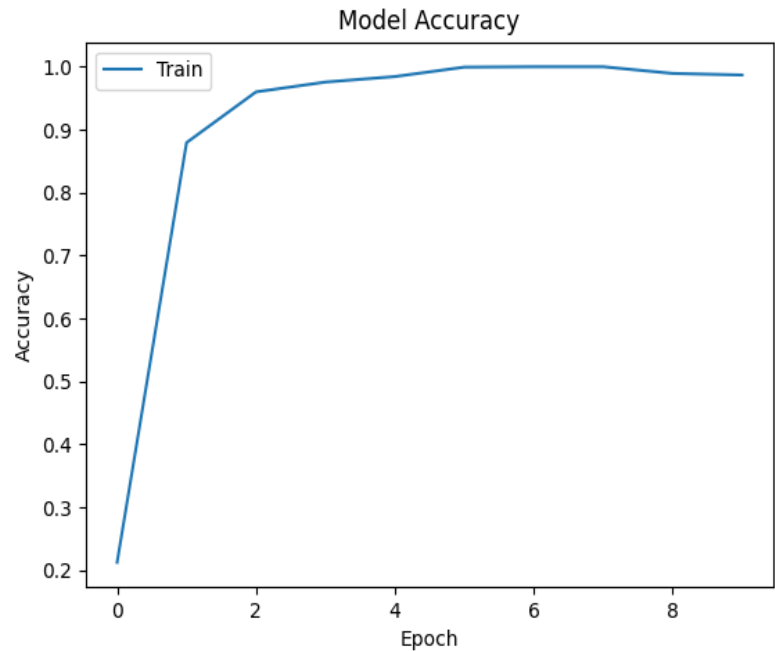
Accuracy curve for Words

```
[[234   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0 256   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0 215   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0 241   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0 224   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0 263   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0 243   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0 240   0   0   0   0   0   5   0   0   0]
 [  0   0   0   0   0   0   0   0 213   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0 227   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0 248   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0 220   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0 260   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0 239   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0 267   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 235   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0 240]]
```

Confusion Matrix for Hand poses(words)

147	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	2	0	0
0	139	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0
0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	135	0	0	0	0	0	4	0	0	0	0	0	0	0	0	3	10	0	0	0
0	0	0	0	0	0	150	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	7	143	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	150	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	2	0	152	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	152	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	154	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	147	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	150	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	10	0	0	0	133	0	0	0	0	8	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	151	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	150	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	149	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	148	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Confusion Matrix for ASL

FUTURE SCOPE

❑ More accurate prediction than the current system could be obtained by training it with a large dataset.

❑ Background elimination.

❑ Expansion to Multiple Sign Languages

❑ Improvements in computational efficiency and model optimization.

❑ Personalized Learning and Adaptation.

❑ Fast access and translation

CONCLUSION

❑ Exhibited robust performance in accurately translating sign language gestures into text and speech representations.

❑ Focuses on recognizing ASL alphabet gestures, pocket sign(hand pose) and achieves a commendable accuracy rate of 99.0% on our dataset.

❑ Provides a valuable tool for facilitating communication between individuals with hearing and speech impairments and those who do not understand sign language.

REFERENCES

1) Sign language detection and conversion to text using CNN and OpenCV (https://ieeexplore.ieee.org/document/10091051)

2) Sign Language to Text and Speech Translation in Real Time Using Convolutional Neural Network" (https://www.ijert.org/research/sign-language-to-text-and-speech-translation-in-real-time-using-convolutional-neural-network-IJERTCONV8IS15042

3) www.slideshare.com

THANK YOU…!!!!