CAPSTONE PROJECT REPORT

AMERICAN SIGN LANGUAGE GESTURE DETECTION SYSTEM

By

(101803552) Rudraksh Sharma

(101803335) Gemin Shrestha

(101803550) Sagaljyot Singh

(CPG: 4)

Under the Guidance of

Dr. Joohi Chauhan, Assistant Professor

Thapar Institute of Engineering and Technology, Patiala

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Submitted to the

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

In Partial Fulfilment of the Requirements for the Degree of

Bachelor of Engineering in Computer Engineering

at

Thapar Institute of Engineering & Technology, Patiala

May 2022

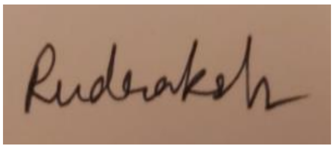
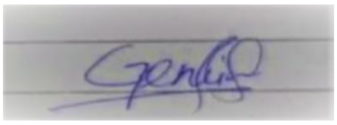
ABSTRACT

People suffering from communication disorders face various issues while communicating with others regularly. Most of the population doesn't have the foggiest idea about gesture-based communication. To help all these people we have created an American Sign Language detection system. Along these lines, this undertaking means to assemble a two-way communication channel for individuals suffering from communication disorders so that they can communicate effectively with others at an exceptionally negligible expense or practically free. This will accordingly dispose of the need to pay colossal sums to a translator. One point of interaction will change the ASL to text which can effectively be used by individuals with communication disorders and one more interface will be utilized by the ordinary individuals which will convert the sound to text and further convert that message into a sign which can effectively be deciphered, hence making the correspondence simpler for hard of hearing and unable to speak individuals on the everyday basis.

DECLARATION

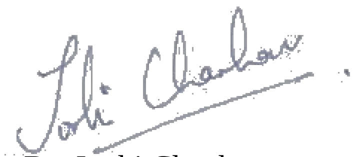
We hereby declare that the design principles and working prototype model of the project entitled American Sign Language Gesture Detection System is an authentic record four own work carried out in the Computer Science and Engineering Department, TIET, Patiala, under the guidance of Dr. Joohi Chauhan during 8th semester (2022).

Date: 10 May 2022

Roll Number	**Name**	**Signatures**
101803552	Rudraksh Sharma	
101803335	Gemin Shrestha	
101803550	Sagaljyot Singh	

Counter Signed By:

Faculty Mentor:



Dr. Joohi Chauhan

Assistant Professor

Computer Science & Engineering Department,

TIET, Patiala

ACKNOWLEDGEMENT

We would like to express our thanks to our mentor Dr. Joohi Chauhan. She has been of great help in our venture, and an indispensable resource of technical knowledge. She is truly an amazing mentor to have.

We are also thankful to Dr. Shalini Batra, Head, Computer Science and Engineering Department, entire faculty and staff of Computer Science and Engineering Department, and also our friends who devoted their valuable time and helped us in all possible ways towards successful completion of this project. We thank all those who have contributed either directly or indirectly towards this project. Lastly, we would also like to thank our families for their unyielding love and encouragement. They always wanted the best for us and we admire their determination and sacrifice.

Date: 10 May 2022

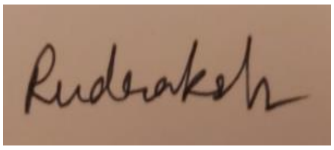
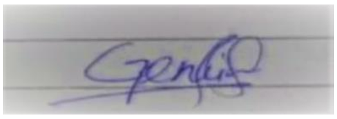
Roll Number	Name	Signatures
101803552	Rudraksh Sharma	
101803335	Gemin Shrestha	
101803550	Sagaljyot Singh	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1.1 Project Overview

Sign Language is the most natural and expressive way of communication for people suffering from a communication disorder. People who are not suffering from communication disorders never try to learn sign language for interacting with deaf people. This situation leads to the isolation of deaf people from the masses. But what if we can make a system that can translate sign language to text format? Surely it will minimize the communication gap between people without communication disorder and those who have a communication disorder. It is easy to find a large number of sign languages all over the world. Almost every spoken language has its respective sign language, so there are more than 200 languages available.

There are several sign languages available such as American, British, German, French, Italian, and Turkish Sign Language. American Sign Language (ASL) is well-known and the best-studied sign language in the world. Sign language is a way of verbal exchange between human beings suffering from some instance of communication disorder. Around 360 million human beings globally are affected by some kind of communication disorder. This number includes 328 million adults and 32 million children. Thus, with the growing range of people with deafness, there is a rise in demand for translators. Minimizing this communication gap can help us towards having effective communication among all human beings. Sign language translation is amongst the most growing line of research. It is a necessary tool of communication for humans with hearing impairments. A hand gesture recognition gadget can offer an opportunity for deaf people to talk with vocal humans without the need for an interpreter. This system is built for the automated conversion of ASL into textual content.

Gesture Detection System will act as an interpreter between the people suffering from communication disorder and those who are not. It is done by training a model using transfer learning. It captures the actions of deaf and dumb persons and converts that into corresponding text, which can be easily read by a normal person. The system will provide gestures of each alphabet of English along with some custom gestures like space, delete and nothing. This system can also take the speech of a person and convert that into sign language. Thus, it can help to interact with people suffering from a hearing disorder.

A model is trained using transfer learning to recognize alphabets from the signs captured by cameras and then combine these alphabets to form a word and then form a sentence from these words. The additional gesture is being used to mark the ending of the word.

A person who is deaf is not at all able to listen to what others say. So, there is a requirement to use sign language with them. But most people don't know sign language as it is very rarely taught as a part of the course curriculum. So, to make communication even better there is another option in which a person will just simply speak and that audio will be converted into text which further will be converted into sign language which will be shown on the screen and can be interpreted by the deaf/dumb person. Hence this project serves as a complete solution for speech impaired people to have a normal conversation with another person.

1.1.1 Technical terminology

- **Machine Learning (ML):** Machine learning is the logical study of calculations and algorithmic models that computer systems use to perform a particular assignment without utilizing explicit instructions, depending on inference and patterns instead.

- **Deep Learning:** Deep learning is a subset of ML in artificial intelligence (AI) that has networks equipped for unsupervised learning from data that is unstructured or unlabelled.

- **Data Analytics:** Data Analysis is a course of investigating, purifying, changing and displaying information fully intent on finding valuable data, illuminating conclusions and supporting decision-making.

- **Transfer Learning:** Transfer learning is a machine learning method where we reuse a pre-trained model as the starting point for a model on a new task. A model trained on one task is repurposed on a second, related task as an optimization that allows rapid progress when modelling the second task

1.1.2 Problem Statement

There are an enormous number of individuals that utilize gesture-based communication to express their feelings. They all deal with the issue of majority of the population not knowing sign language, and an individual translator is expensive and tedious. Even if someone masters lip reading, they can only get about 30% of the total communication. Accordingly, it is vital for new technologies and innovations to propose an answer and help these individuals.

1.1.3 Goal

Our fundamental objective is to make an application that will assist the people who are having hearing problems and are unable to speak to individuals in a two-way communication with others without wasting too much time and effort. We likewise need to make it simple for individuals to learn ASL so they can speak with individuals who experience issues with talking and hearing.

1.1.4 Solution

ASL GESTURE DETECTION SYSTEM is an application made by us which uses the ImageNet classification CNN model (published by Google) for transfer learning and redefining the last dense layer to output the probability of each of the 29 classes that we defined, this model is trained to recognize alphabets from the signs captured by cameras and then combine these alphabets to form a word and then form a sentence from these words. This system can also take the speech of a person and convert that into text/ sign language. Thus, creating a proper medium for two-way communication.

1.2 Need Analysis

Sign language is a necessary communication tool for people suffering from communication disorders.

The National Institute on Deafness and Other Communication Disorders (NIDCD) characterizes American Sign Language as a total, complex language that utilizes signs made by moving the hands and collaborating with expressions and postures of the body.

Since the affected people face difficulties in reading and writing, they usually use gesture-based communication. Therefore, they always need an interpreter which can translate their sign language to text and vice versa. These interpreters charge a heavy amount of money.

Gesture based communication isn't just about the hands. It is additionally about the development of an individual's arms, body and expressions.

Some children with autism spectrum disorder (ASD) struggle in developing verbal communication. Learning sign language can be a helpful communication tool for some children with ASD.

Some surgical procedures (such as oral surgery) impede a person’s ability to talk during recovery. With ASL, people can communicate without trying to talk.

Sign languages can be of used when speaking words is physically impossible, such as talking underwater, talking through the glass, from a distance, at a loud music concert, talking with your mouth full and so on.

With ASL, you can communicate easily across the room without disturbing patrons. Coaches of Sports teams need to communicate something to the players on the field and the players among themselves. With ASL you can communicate clearly and even include encoded messages for special plays.

Additionally, gesture-based communications can be an incredible method for gossiping without any other person knowing, and passing on secret data.

1.3 Research Gaps

TABLE 1: Research Gap

Research Gaps	**Deficiencies in research**	**References**
Difficulty in hand detection due to bad background.	If the colour of background matches with that of skin, then there is difficulty in detection of hand thus hampering gesture recognition.	“Hand Gesture Recognition Based on Computer Vision: A Review of Techniques”-by Munir Oudah, Ali Al-Naji, and JavaanChahl.
Combinations of many gestures are used in conversation.	Gestures could consist of a few developments, so we want to give some specific situation and perceive designs like moving fingers clockwise and showing a thumb could be utilized to stamp some predetermined number of documents or some limited area.	
Movement of hands and capturing it.	People that use sign language are usually performing it at a high speed thus making it hard to capture all the gestures.	
Speed and efficiency with less lag.	The gesture recognition framework should be intended to dispose of the slack between playing out a motion and its arrangement. Adoption of hand motions must be encouraged by demonstrating how consistent and quick it can be.	

Lighting issue	An appropriate lighting is expected to perceive the hand motion. Inadequate light can diminish the precision.	Study of Face Recognition Techniques: by Madan Lal, Kamlesh Kumar
Some gestures are based on movements.	Some of the gestures of the alphabets in the ASL are not static but rather a combination of hand movements, for example gesture of alphabets J and Z.	

1.4 Problem Definition and Scope

Since people suffering from communication disorders face difficulties in reading and writing. They usually use sign language for communication. Therefore, they always need an interpreter which can translate their sign language to text and vice versa. This makes them dependable on others and also generates a sense of insecurity, because they may get in a situation where it is hard to get a hold of someone who can understand them or interpret them. They should also have the means to communicate easily and instantaneously.

While there are new and accessible technologies emerging to help those with hearing disabilities, there is still plenty of work to be done. For example, advancements in machine learning algorithms could help the deaf and hard-of-hearing even further by offering ways to better communicate using computer vision applications.

Our project has a scope to not only solve the problems that these millions of people are facing but it can also be applied in a lot of other fields like using it in automobiles where gesture can detect what the driver wants or it can also be used in hospitals where surgeons can use it while keeping their hands sterilized.

1.5 Assumptions and Constraints

TABLE 2: Assumptions

S. No.	Assumptions
1	Capture the image from the front side of the person.
2	Capture the image in an environment with ample lighting. If you capture the image in night mode then hand gestures might not be recognized properly.
3	Capture the image of the single specific person at the spot.
4	Hand should be kept in the region specified by camera otherwise images will not be captured.

TABLE 3: Constraints

S. No.	Constraints
1	The person using this system must need to learn ASL and some signs are modified in order to avoid confusion between some symbols. The person must go through the documentation of the system to learn these signs.
2	The model had an accuracy of 90% which means it can identify 90 gestures correctly out of 100.

1.6 Standards

This section gives a brief introduction to the standards we are using.

IEEE

The institute of Electrical and Electronic Engineers Standards Association (IEEE-SA) is an association with IEEE that creates worldwide norms in an expansive scope of enterprises, including power and energy, biomedical and medical services, data innovation and advanced mechanics, media transmission and home computerization, transportation, nanotechnology, data confirmation and some more.

Web 2.0

Emphasize user generated content, simple to utilize, participatory culture and interoperability (i.e., viable with different items, frameworks, and gadgets) for end clients. It connects data sources together utilizing the model of the Web.

ISO

In the case of digital cameras, ISO sensitivity is a proportion of the camera's capacity to catch light. Advanced cameras convert the light that falls on the picture sensor into electrical signs for handling. ISO sensitivity is raised by enhancing the sign. Doubling ISO sensitivity pairs the electrical sign, halving how much light that needs to fall on the picture sensor to accomplish ideal exposure.

1.7 Approved Objectives

1) A system that can easily and accurately convert Sign Languages and other custom signs (e.g., Space, Delete, nothing) to text according to the user's input.

2) Audio to text/sign converter.

3) A web interface through which users can easily interact or launch software.

1.8 Methodology Used

1) Sign to text:

We have approached this problem by using transfer learning. The reuse of a pre-trained model on a new problem is known as transfer learning in machine learning. A machine uses the knowledge learned from a prior assignment to increase prediction about a new task in transfer learning. We have chosen to make our model using transfer learning as it is a good way to train a model with less data and also it can speed up the process a lot.

We have taken our dataset from Kaggle and using that we have classified it into a Training and Testing image dataset to perform data augmentation on our images, which will then be fed into our deep learning model and then the dimensions were manipulated accordingly. We are using the ImageNet classification CNN model (published by Google) for transfer learning and redefining the last dense layer to output the probability of each of the 29 classes that we have defined.

After this we exported this model in a special format (.tfjs format) so that it can be used on the browser with the help of TensorFlow.js.

2) Audio to Sign/Text:

To start with, the React Speech Recognition module is set up and installed. To utilize React Speech Recognition, we utilized the 'useSpeechRecognition' hook and the 'SpeechRecognition' object. Then the voice of the user is recorded using ‘SpeechRecognition.startListening()’.

We obtain the transcript of the user's voice using ‘const {transcript} = useSpeechRecognition()’.

In this module we have also tried to incorporate Audio to Sign where a person speaks a sentence and the sentence is converted into a relevant sign in the form of the GIF.

3) A Web interface through which we can easily do two-way communication:

Using basic HTML and React Js as our frontend interface, and also CSS inorder to beautify the web interface we have create a web interface in which user can use it for two way communications.We have incorporated the trained model in the form of .tfjs format which is used in the Website to predict the sign to text and also Speech to text/ Sign.We have deploy our model directly on the browser so that if the user have inbuilt microphone and camera can easily use it for communication between each other.

1.9 Project Outcomes & Deliverables

As it provides a both way mechanism for communication, ASL Gesture Detection System is going to be helpful for deaf and dumb people. It can help affected people to use the sign language which will be converted into text so that it can be easily read by the normal people. Also, normal people need not to learn the sign language and can simply use their audio which will be converted into text and will be easily interrupted by deaf and dumb people. Hence making the effective communication both ways.

1.10 Novelty of work

A Sign interpretation system is nothing new, even google is working on it but still people with hearing/speaking impairments has to face problems. This is mainly because the system is not readily available and most of the time does not help in two-way conversation. Our ultimate goal through this project is to tackle that problem by not only making a sign recognition system but also providing a speech to sign conversion system so that both the parties can interact with each other. With this project we not only hope to reduce the problems faced by these people and providing them with an easy and readily available solution, but also open the doors toward the discussion of accessibility features.

2.1 Literature Survey

2.1.1 Theory Associated with Problem Area

A Person who is either hard of hearing or dumb or both faces challenges in speaking with others in light of the fact that usually, individuals don't have any idea or comprehend gesture-based communication and it is difficult to constantly have an interpreter with them at all times. This makes the existence of speech impaired individuals reliant upon others.

Moreover, these sign language interpreters and translators charge a huge amount of money. The Bureau of Labour Statistics stated in May 2017 that interpreters and translators earn a median wage of $47,190 and also the demand for them is expected to increase by 18% between 2016 and 2026. This is where the ASL Gesture Detection System comes into the picture. This system completely vanishes the use of interpreters and translators.

Information from a survey conducted by a school in California suggests that there is a need for an artificial translator. A questionnaire was distributed to teachers in a residential school for the deaf in California. In addition to questions regarding the availability of sign language assessment at their schools, participants responded to items concerning their motivation to use a test for sign language measurement. Of the 100 distributed surveys, 85 were completed and returned. Results showed overwhelming agreement among respondents concerning the importance of sign language assessment.

ASL GESTURE DETECTION SYSTEM has two interfaces one for sign language to text conversion and the other for an audio to sign/text. These two interfaces can help a deaf person to communicate easily with a person who does not understand sign language. The first interface will take the signs from the deaf person and predict the corresponding sentences in the form of text. The second interface is used by other people who do not know sign language. That person will speak normally and that speech will be converted into text which will be further converted into GIFs of sign language.

2.1.2 Existing Systems and Solutions

There are many devices and software that are used for sign language to text conversion and vice versa.

One of the gadgets handles this issue by utilizing gloves. The structure incorporates a couple of gloves with small, stretchable sensors that run the length of all of the five fingers. These sensors, created utilizing electrically coordinating yarns, get hand developments and finger courses of action that address individual letters, numbers, words, and articulations. The contraption then changes the finger advancements into electrical signs, which are shipped off a dollar-coin-sized circuit board worn on the wrist. The board sends those transmissions remotely to a cell phone that makes an interpretation of them into verbally expressed words at the pace of around a single word each second. However, this system only solves the problem of only one side and cannot help in two-way communication.

Another device based on sign language is Hand Talk. It is a pocket app that automatically translates the oral language to both text and sign languages, such as ASL or Brazilian sign language. Users can use this app to learn sign language easily. However, it cannot help in two-way communication as it cannot convert the sign language to text.

Another system has been developed which can only recognize one sign at a time and shows the corresponding alphabet. However, this system cannot be directly used for the communication process.

2.1.3 Research Findings for Existing Literature

TABLE 4: Research Findings

S. No.	**Roll Number**	**Name**	**Paper Title**	**Tools/ Technology**	**Findings**	**Citation**
1	101803552	Rudraksh	A review of research on object detection based on deep learning	Google's object detection API.	Deep Learning, Target detection	Deng *et. al.* [5]
2	101803335	Gemin	Sign language recognition using a combination of new vision-based features.	Real time vision-based system	Image Thresholding	Zaki [8]
3	101803552	Rudraksh	Sign Language Recognition Using Convolutional Neural Networks.	Machine learning and Artificial Intelligence	Model training from the dataset.	Pigou L. [7]
4	101803550	Sagal	Data integration for the relational web	Data fusion tables	Data collection techniques	Cafarella *et. al.* [3]
5	101803335	Gemin	Comparisons of front-end frameworks for web applications.	HTML, CSS	Front-end Web development	Kaluza and Vukelic [6]
6	101803550	Sagal	A comparative Study of Various Techniques and Outcomes of Recognizing American Sign Language	Hand gesture techniques	Insight and optimisation of dataset	Shivashankar and Shrinath [4]

2.1.4 Problem Definition

So far, we have found that current systems can only be used for one-way communications, either it converts the sign to text or it converts audio to sign. These single systems cannot serve our purpose for two-way communication. Also, people need to be able to use these systems via webcam as we have seen during the pandemic most of the communication is done via video conferencing.

There is also a need for means of communication for the people who do not understand sign language with a deaf or dumb person. So, we added another functionality in which the normal person will only speak and that will automatically be converted into text or signs which will be displayed over the screen. The combination of both these interfaces will help to make communication better for deaf and dumb persons.

2.1.5 Survey of Tools and Technologies Used

Languages: - Python, HTML, CSS, JavaScript,React Js

Libraries used: TensorFlow (Tensorflow-gpu), TensorFlow.js, Numpy, and Matplotlib and Keras (Python Libraries for deep learning model implementation)

2.1.6 Summary

The current existing systems only solve half of the problem and they are really hard to afford or use. ASL GESTURE DETECTION SYSTEM application made by us is a one-step solution for all kind of communication that can be done by a deaf and dumb person and it also enables others to learn the basics of ASL quickly and easily. Apart from this our solution has no cost associated with it therefore people can use it freely without thinking of payments. Also, because we do not have any use of user data, we do not collect it thus making it a safer alternative to other applications.

2.2 Software Requirement Specification

2.2.1 Introduction

2.2.1.1 Purpose

The purpose of ASL GESTURE DETECTION SYSTEM is to reduce the communication gap for the people who use sign language to communicate with other persons. Since sign language is not known to everyone, it becomes difficult for these people to easily communicate with other people.

ASL GESTURE DETECTION SYSTEM will act as an interpreter between the people suffering from communication disorder and those who are not. It is based on the ImageNet classification CNN model (published by Google) for transfer learning. It captures the actions of deaf and dumb persons and converts that into corresponding text, which can be easily read by a normal person. The system will provide gestures of each alphabet of English along with some gestures of some commonly used words. It also provides the deaf/dumb person with a feature to add some custom gestures of some particular words according to the user wish. This system can also take the speech of a person and convert that into sign language. Thus, it can help to interact with people suffering from a hearing disorder.

2.2.1.2 Intended Audience and Reading Suggestions

Deaf and Dumb peoples who uses sign language to communicate with other people, will be using ASL GESTURE DETECTION SYSTEM in order to communicate better with other people. This document is intended for users and testers. The document represents the detailed idea being implemented through this project.

2.2.1.3 Project Scope

ASL GESTURE DETECTION SYSTEM is an important tool for Deaf and Dumb people who can’t use normal language to interact with people. Sign to Speech could be used at reception desks or during video conferences to allow signing people to speak with people who don't know the Sign Language. The main point of this project is to use cameras to recognize gestures from Sign Language to offer a new means of communication. The program will be able to transcribe gestures, done by a person using gestures, into written words, printed on the screen of the person who doesn't know sign language. Moreover, the person who doesn't know sign language can use audio feature to convert the audio to sign and then text to sign which can be easily be interpreted by the person who knows the sign language and doesn't know the normal language.

2.2.2 Overall Description

2.2.2.1 Product Perspective

Hand Gestures are a significant method for people suffering with communication disorders, this is the most well-known method of correspondence for almost totally deaf individuals with ordinary individuals. The fundamental point of our venture is to work on the correspondence between a signing and a non-signing individual. To do so, we developed a software that is capable of recognizing gestures, finding their meaning in a linked dictionary and sending the translated text to the person who doesn't know the sign language. This system also provides the interface for audio to sign language conversion to make both ways of communication. This system will help disabled people to be able to communicate with the normal people without any fear and let these people feel more like normal people and not feel isolated from other people.

2.2.2.2 Product Features

This project has four major functions:

1) Captures the real-time video feed using the webcam.
2) Performs image pre-processing and locates the hand position and captures image of the hand which will be fed to the model for prediction.

3) Prediction of Gesture corresponding to the hand gesture and convert it into corresponding text.
4) Converting the Audio capture through microphone into corresponding text or sign.

2.2.3 External Interface Requirements

2.2.3.1 User Interfaces

Visual interface:

This involves a user and a camera. Both a webcam or a separate camera can be used. Camera is used to collect the video feed that will be fed into the model. Users can easily adjust the camera according to its requirements.

Screen Interface:

The video feed taken by the camera will be displayed in a bounded box on the screen of the computer

Prediction Interface:

This interface will give the prediction of the hand gesture captured by the camera and will convert it into corresponding text message which can be displayed on the screen to be read by another person.

Audio to text/sign interface:

This interface will be used by the normal people which will capture the audio of the person through the in-built microphone of the device and convert this audio into text and then convert it to sign.

2.2.3.2 Hardware Interfaces

1) Webcam: Used for capturing the video feed of a specific person.

2) Microphone- To capture the audio feed through the system.

3) Computer: To process and analyse captured feed for required results.

4) **GPU** for the processing of the feed at a faster rate.

5) **Processor:** Intel CORE i5 processor with minimum 2.9 GHz speed.

6) **RAM:** Minimum 4 GB.

2.2.3.3 Software Interfaces

1) Using system having inbuilt Microsoft Windows 10 Basic will be use as operating system for development of this system Spyder will be use as development language

2) **TensorFlow v2.8.0:** TensorFlow is a start to finish open-source stage for AI. It has an exhaustive, adaptable ecosystem of instruments, libraries and local area assets that allows analysts to push the cutting edge in ML and designers effectively assemble and send ML controlled applications.

3) Keras: Keras is a deep learning API written in Python, running on top of the AI stage TensorFlow. Keras is the significant level API of TensorFlow 2.8: an agreeable, profoundly useful point of interaction for tackling AI issues, with an emphasis on present day deep learning. It gives fundamental deliberations and building blocks for creating and transporting AI arrangements with high cycle speed.

4) JavaScript (TensorFlow.js) will be used for Integrating different components of the system and will also be used for the deployment of this model to a website.

2.2.4 Other Non-functional Requirements

2.2.4.1 Performance Requirements

- **Usability:** The system can be used with any device having a good internet connection, camera, microphone in it.
- **Speed:** The system should have fast response so that information can be updated in real-time.
- **Minimum Delay:** The system should not take long with the identification/authentication procedures so that the user does not have to wait for a long time to enter.
- **Maximum security:** There is complete safety of users own personal customized expression, so that the user can be guaranteed maximum safety.
- **Cost-effective:** The solution should not be very expensive and not have high deployment costs, otherwise not everyone will be able to afford it.
- **Scalability:** The solution should be scalable to accommodate large numbers of users.

2.2.4.2 Safety Requirements

Camera that is used for this software should be plug-in properly. If the external camera is connected to the computer, then the connector and table must not have any kind of damage. If the camera is not needed for further usage, turn off the camera for privacy, if using an external camera then unplug it from the computer. If using an infrared camera then you should not use it for a long time. If you do not need to use software further then properly close the software for the safety of the environment which causes the increase of environment temperature (heating of the device).

2.2.4.3 Security Requirements

You can use the software without sending out any kind of data. All the data used is processed locally and is not sent to any kind of cloud storage. There might be times where this might be a necessity to follow local regulations, like GDPR for instance, or while handling any information that the client might need to keep on their machine and not shipped off to an outsider.

2.3 Cost Analysis

As our software doesn't use any kind of servers, the only cost expenditure is for the CDN for hosting the software on the platform. The expense of a CDN is a lot less expensive than keeping a server (possibly with a graphic card appended) running all day, every day. As most of the application works on the browser there is no need for extra server costs.

2.4 Risk Analysis

Sometimes the software predicts the letters quite quickly which results in wrong predictions to the user while changing the gesture. Although the user have the able to delete the wrong letters the output predicted by the software may result in a wrong sentence formation.

METHODOLOGY ADOPTED

3.1 Investigative Techniques

1) Fair Testing: It aims to establish relationships between various variables affecting the outcome. One by one variables are changed keeping others constant and therefore change due to each variable is calculated. It is best suited for technology investigations and in particular it is well suited to investigations involving measurements.

2) Identification and Classification: It is based on the approach of filtering events or objects into various groups and categories. Keys are often used as criteria to carry out the process. All the projects are based on contemporary techniques of Machine Learning, Artificial Intelligence and Deep Learning.

3) Modelling: Models are quite useful in understanding how a particular process works and also why it works in that way. Some models are already set or pre-defined while others need to be formulated. Segregation and classification-based projects. Majorly the projects involving use of databases and information retrieval.

4) Pattern Seeking: This approach involves observing and subsequently recording natural events, or carrying out experiments where variables cannot be easily controlled. Face recognition, Text recognition, Caption Generation and Automatic Image Tagging are just few of the projects using this technique.

5) Researching: It includes assembling and analysing others ' feelings or logical discoveries to respond to an inquiry or give the connected foundation data to help make sense of noticed occasions. Involves all the projects involving analysis and studying findings.

6) Descriptive: Different scientific and exploratory researches were found. After studying different techniques developed by many researchers it was established that there are multiple ways to implement the project and boost accuracy. Depending on the use case various techniques were found like using pre-trained embeddings of faces, feeding embeddings to further ML algorithms.

7) Comparative: Merging of special features of different techniques would lead to an increase in accuracy and fulfill use case requirements. Using only CNN to train the recognizers and using pre-trained models for the same purpose gave different results.

8) Experimental: For creating an appropriate model, different architectures can be used and tuned. The recognition module when made using deep learning model architecture has different transfer learning models to be used to get maximum accuracy

3.2 Proposed Solution

To solve all these problems, we have come up with the ASL Gesture Detection System.

ASL GESTURE DETECTION SYSTEM has two interfaces one for sign language to text conversion and the other for an audio to text and sign language. These two interfaces can help a deaf person to communicate easily with a person who does not understand sign language. The first interface will take the signs from the deaf person and predict the corresponding sentences in the form of text. The second interface is used by other people who do not know sign language. That person will speak normally and that speech will be converted into text which will be further converted into GIF of sign language.

Some of the features which ASL GESTURE DETECTION SYSTEM will be providing and some of their methodologies are as follows:

1) Accurate Conversion of Sign to Text

We are using the ImageNet classification CNN model (published by Google) for transfer learning and redefining the last dense layer to output the probability of each of the 29 classes that we defined. To compile our model, we decided that we will calculate its loss using categorical cross-entropy. The optimizer we used to update the weights was Adam, and we measured the success of the model using accuracy.

Transfer learning involves taking knowledge that has already been learned to help learn a different but similar thing. We humans do this all the time. You have a lifetime of experiences contained in your brain that you can use to help recognize new things you have never seen before. By only having to train the multi-layer perceptron at the very end of the network, it trains much faster than if you had to train the whole network from scratch.

2) Accurate Conversion of Audio to Sign/Text

First, the React Speech Recognition module is set up and installed. To use React Speech Recognition, we used the 'useSpeechRecognition' hook and the 'SpeechRecognition' object. Then the voice of the user is recorded using 'SpeechRecognition.startListening()'.

We obtain the transcript of the user's voice using 'const {transcript} = useSpeechRecognition()'.

In this module we have incorporated Audio to Sign where a person speaks a sentence and the sentence is converted into a relevant sign in the form of the GIF.

3) A web interface through which we can easily interact or launch our software.

Using basic HTML and React Js as our frontend interface, and also CSS inorder to beautify the web interface we have create a web interface in which user can use it for two way communications.We have incorporated the trained model in the form of .tfjs format which is used in the Website to predict the sign to text and also Speech to text/ Sign.We have deploy our model directly on the browser so that if the user have inbuilt microphone and camera can easily use it for communication between each other.

4) Deployment on Web

Our project deployment is browser based and does not use any kind of server-based deployment. All of our implementation will be client side and browser based.

TensorFlow.js likewise upholds numerous backends inside each of these environments (the genuine equipment-based conditions it can execute inside like the CPU or WebGL for instance. A "backend" in this setting doesn't mean a server-side environment - the backend for execution could be client side in WebGL for instance) to guarantee similarity and furthermore keep things running quickly.

3.3 Work Breakdown Structure

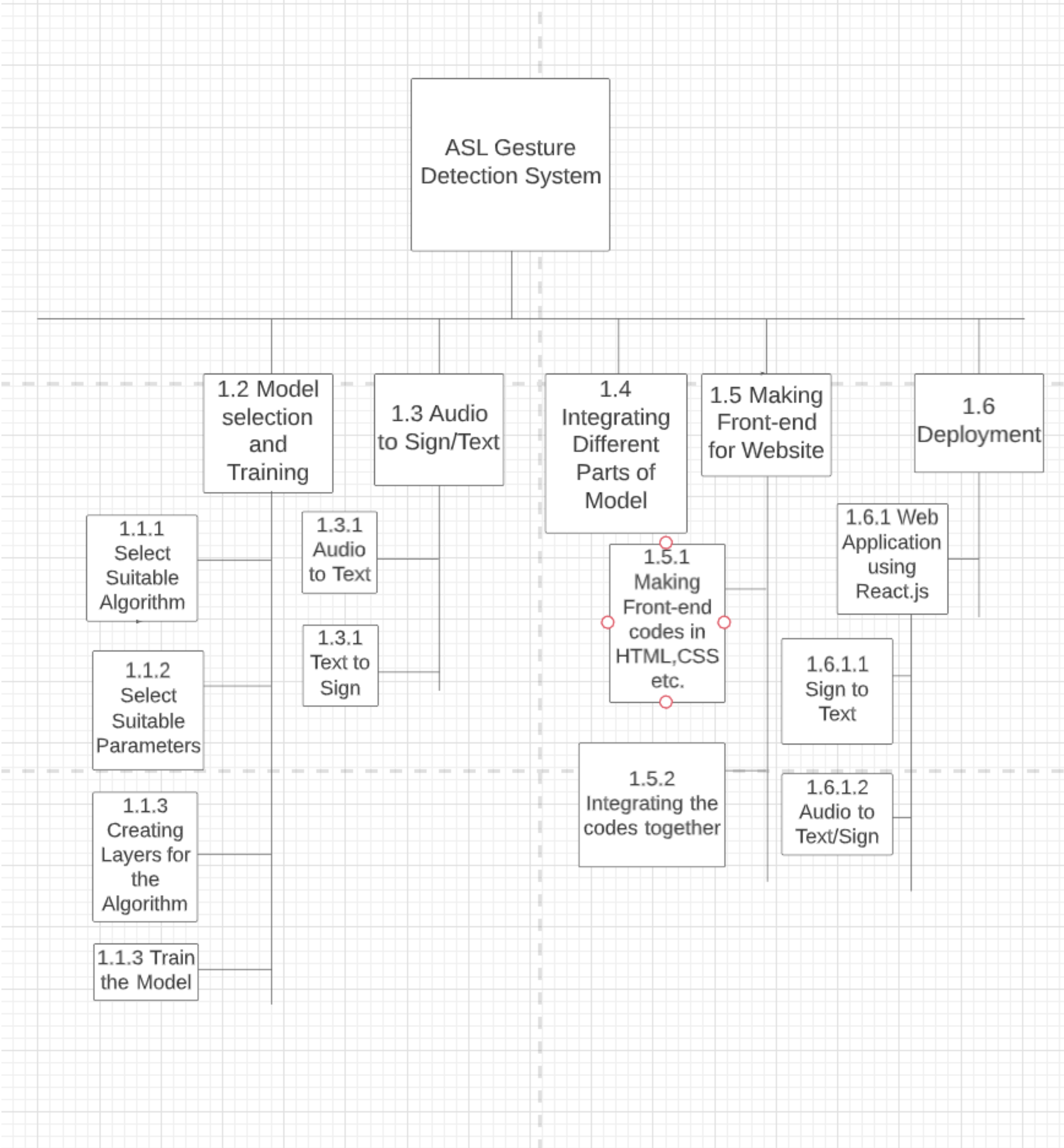


FIGURE 1: Work Breakdown Structure

3.4 Tools and Technology

Languages: - Python, HTML, CSS, JavaScript,React Js

Libraries used: TensorFlow (Tensorflow-gpu), TensorFlow.js, Numpy, and Matplotlib and Keras (Python Libraries for deep learning model implementation)

4.1 System Architecture

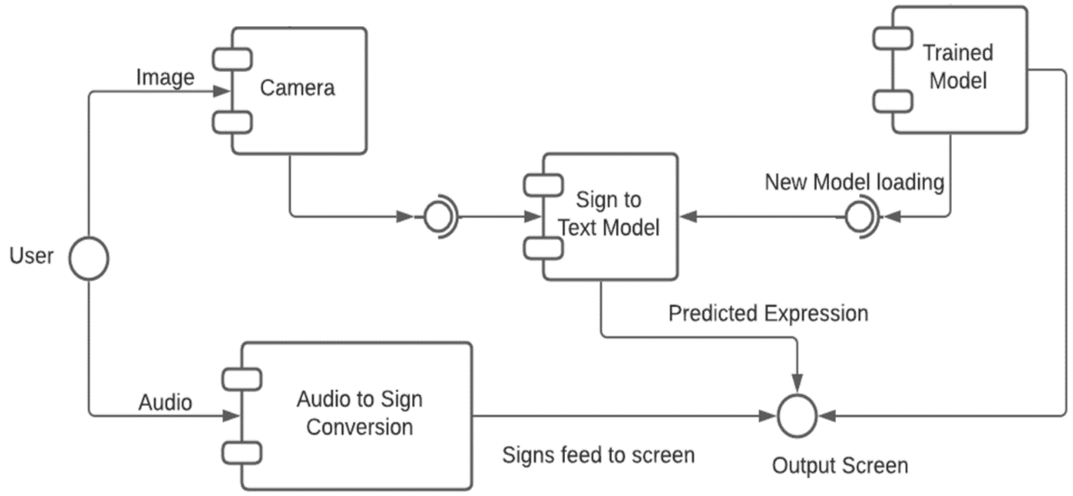


FIGURE 2: Component diagram/Architecture Design

An architecture diagram is a chart of a framework that is utilized to abstract the general blueprint of the product framework and the connections, limitations, and limits between parts. It is a significant instrument as it gives a general perspective on the actual sending of the product framework and its advancement guide. The different components of the framework and their cooperation with one another.

4.2 Design Level Diagrams

4.2.1 Data Flow Diagram

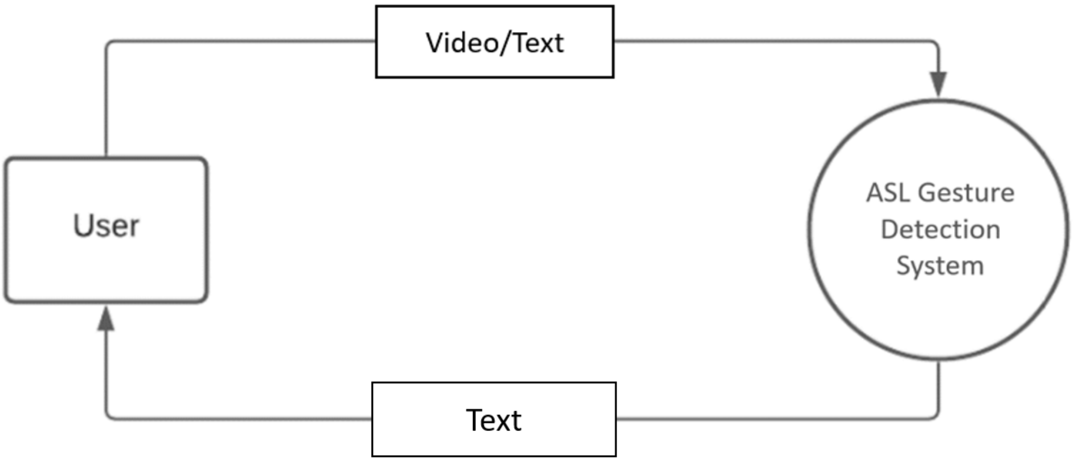


FIGURE 3: DFD Level 0

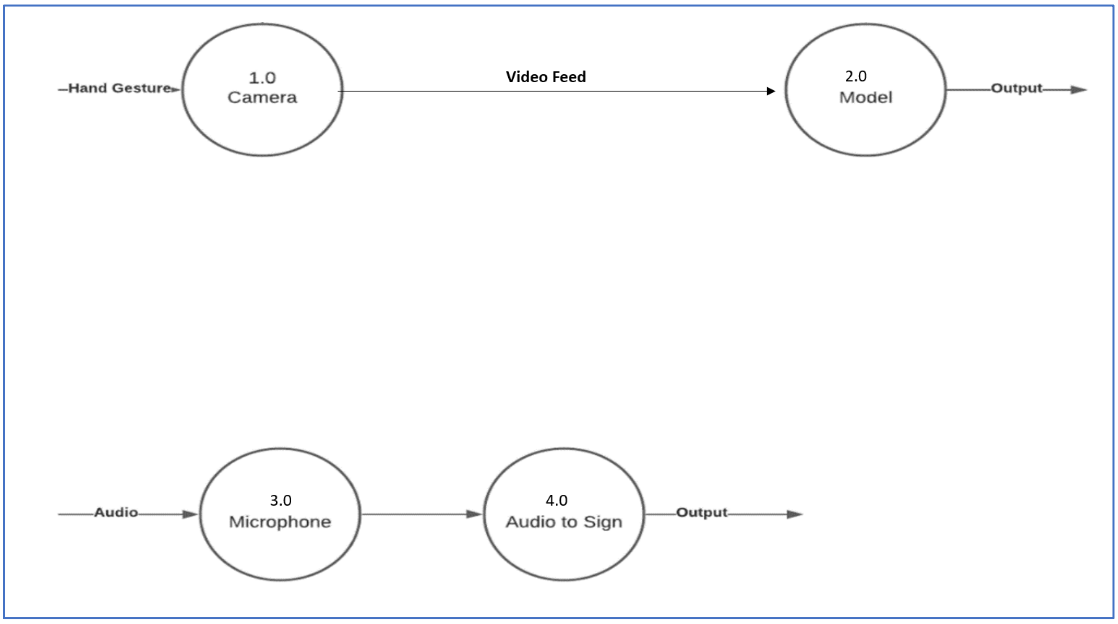


FIGURE 4: DFD Level 1

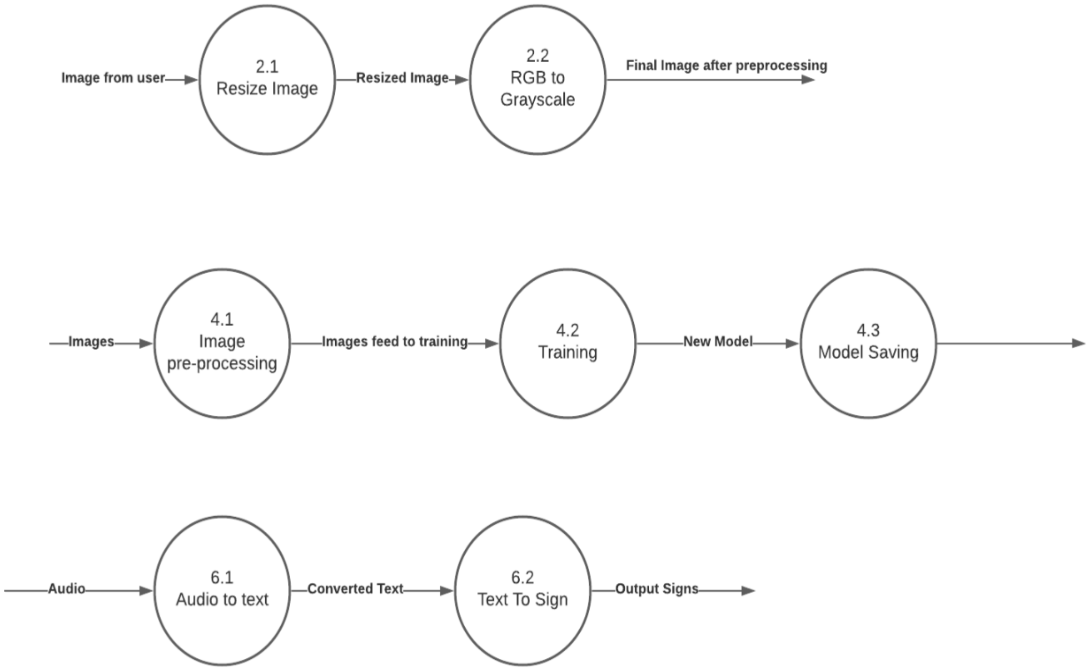


FIGURE 5: DFD Level 2

Level 0 is basic overview of the complete system how system interacts with the system

Level 1 provides a more detailed breakout of pieces of context of the level diagram. It shows the main functions which are an integral part of ASL Gesture Detection System. How different functions interact with each other and how data flows between them.

Level 2 further elaborates the functions which are in level 1

4.2.2 Activity Diagram

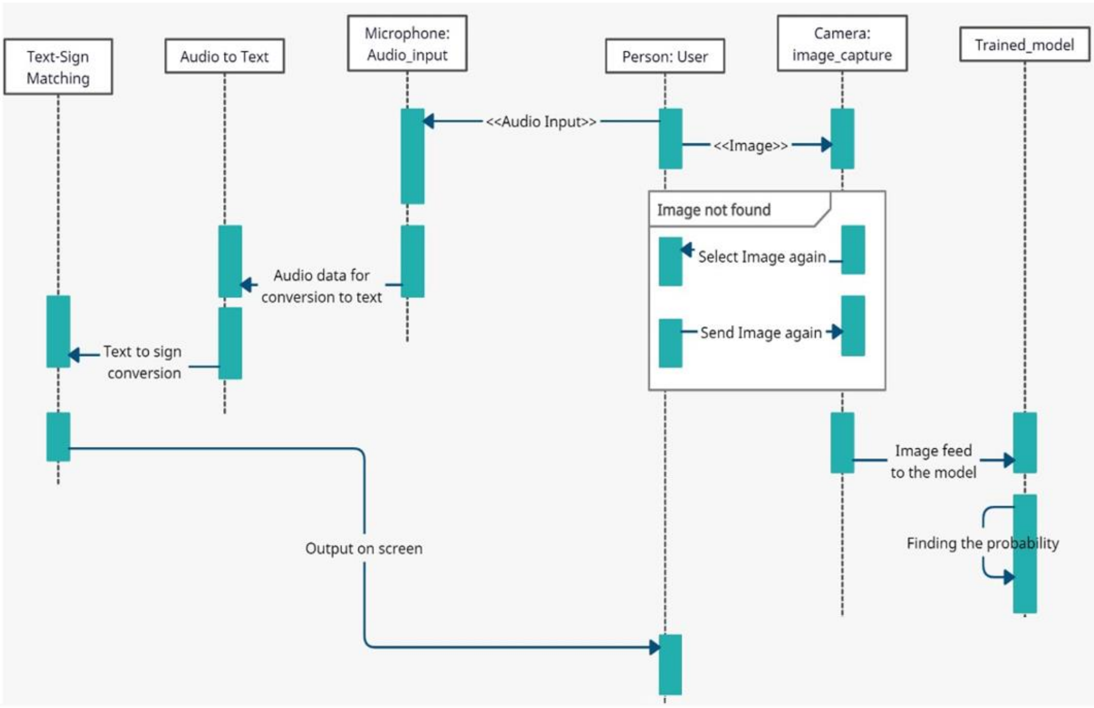


FIGURE 6: Sequence Diagram

4.2.3 State Chart Diagram

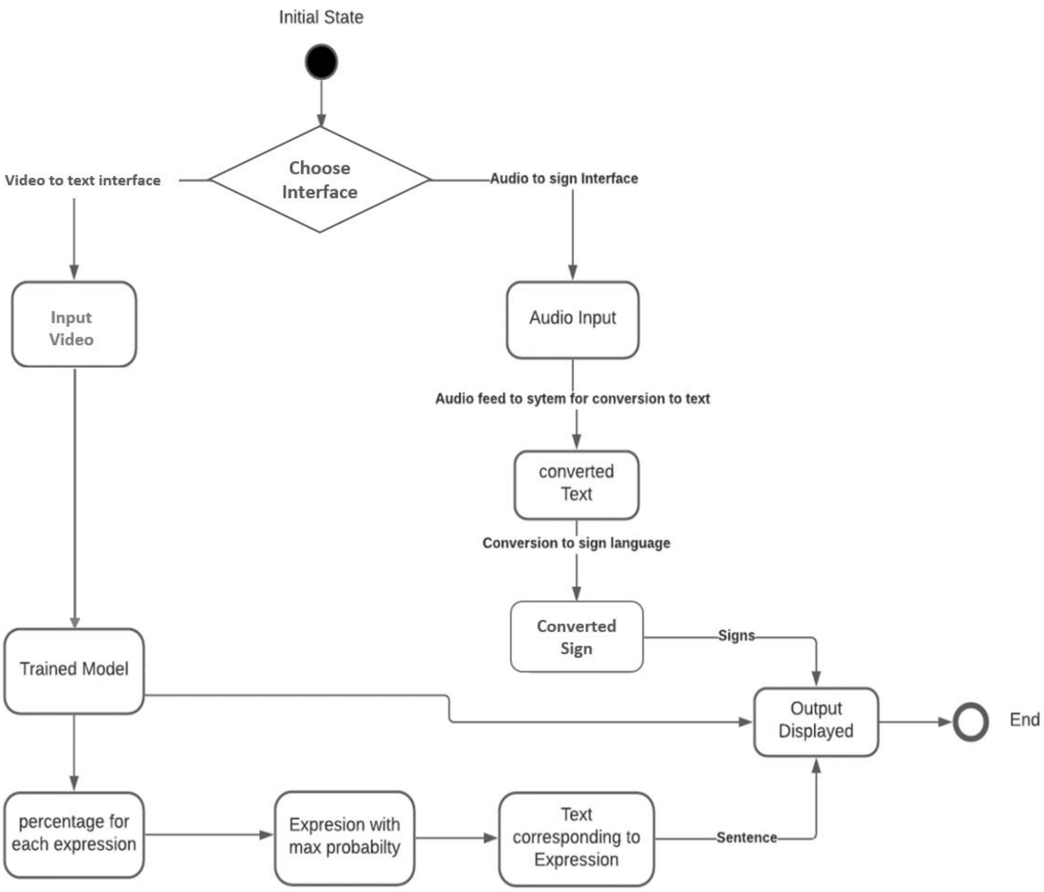


FIGURE 7: State Diagram

4.3 User Interface Diagrams

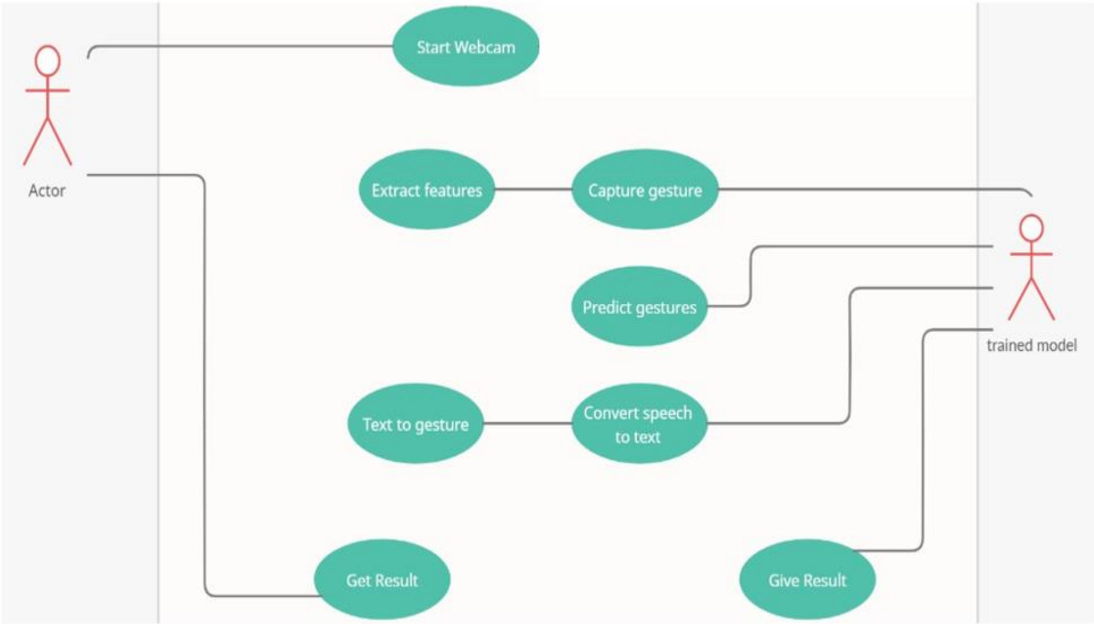


FIGURE 8: Use Case diagram

This use case diagram tells how the user interacts with system in different and gives a demonstration of system with respect to user. It gives a general overview to the user how the system interacts with the user in various scenarios.

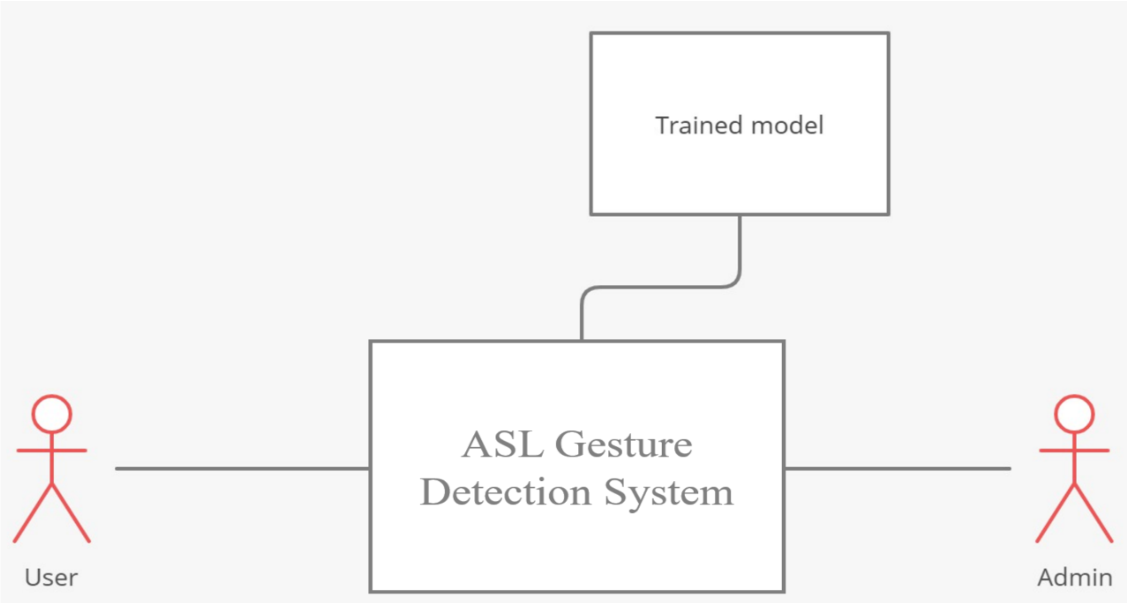


FIGURE 9: Product Perspective

The product perspective is the consumer perspective how it perceives the system and interacts with it.

IMPLEMENTATION AND EXPERIMENTAL RESULT

5.1 Experimental Setup

The project has been simulated on a laptop with 16GB internal RAM. The project is running on the browser on a local host. It uses the in-built camera of the laptop to capture hand gestures of the user and microphone of the laptop to record audio of the user. The steps for the simulation of the project are as follows:

1. The project is run on a browser and local host.

2. The camera captures the gesture and face of the user to perform recognition.

3. The microphone of the laptop is used to record audio of the user.

4. If an image of gesture is taken then it will display a sign corresponding to that gesture, if sound is recorded it will display text according to that speech.

5.2 Experimental Analysis

5.2.1 Dataset Description

For the initial phase of the testing, we focused on the module which needed hand gestures data, we collected the dataset from Kaggle. It had 29 labels from A-Z alphabets and custom gestures like space, delete, nothing.

Each expression had almost 3000 photos and was used as training data. So, there were a total of 87000 images for 29 gestures.

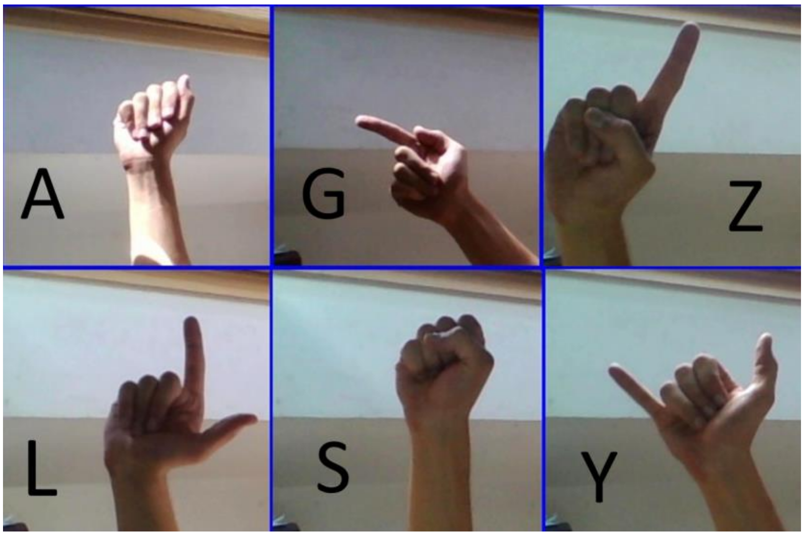


FIGURE 10: Dataset Images

For validating we have used two different datasets, first one contains one image for each gesture, and the second data contains 150 images per gesture. (All of the gesture data was collected from Kaggle).

5.2.2 Performance Parameters

The parameters we have considered for our project are accuracy of the model. For hand gestures it is important that the hand is detected accurately and after detection gesture is recognized accordingly. For this image of size 240*240 is captured and then trained Neural Network is used for recognition. Hand gestures are recognized via the images taken out of the video feed.

The React Speech Recognition module is used to convert the audio into text.

5.3 Working of the project

5.3.1 Procedural Workflow

Our project consists of two major workflows.

The first one is the Image to text interface module. It is responsible for detecting the gesture by hand. This gives us the name of the sign that the user is trying to draw.

The second flow is to convert audio to sign. Here React Speech Recognition module is used as an intermediary. All the sound is first recorded and then it is converted into text and then displayed.

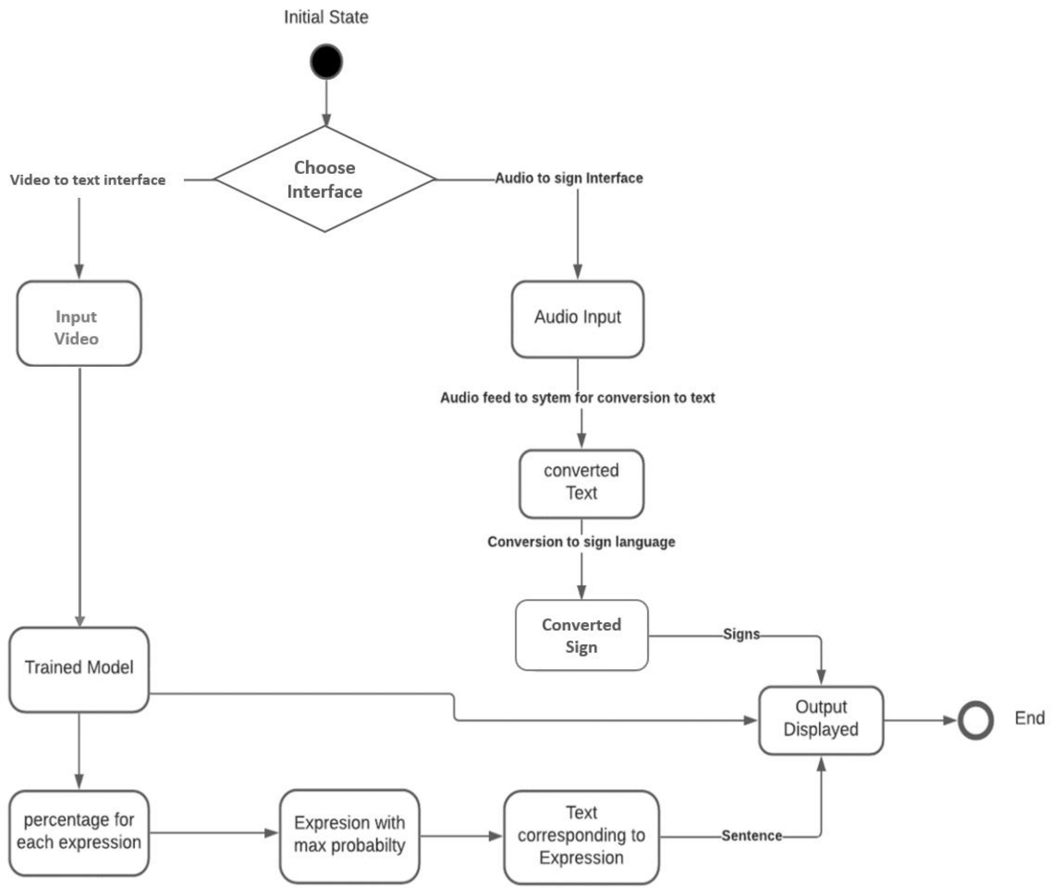


FIGURE 11: Workflow Diagram

5.3.2 Algorithmic Approaches Used

1) Hand Gesture Recognition

```
In [ ]: import tensorflow as tf

In [ ]: from tensorflow.keras.models import Sequential
        from tensorflow.keras.layers import Dense, Flatten, Conv2D, MaxPooling2D, Activation, Dropout
        from tensorflow.keras.losses import sparse_categorical_crossentropy
        from tensorflow.keras.optimizers import Adam
        from tensorflow.keras.preprocessing.image import ImageDataGenerator
        import tensorflow_hub as hub
        import numpy as np
        import matplotlib.pyplot as plt

In [ ]: import os

In [ ]: !pip install tensorflow_hub numpy matplotlib

In [ ]: !pip install scipy

In [ ]: img_size = 224
        batch_size = 32
        epochs = 15
        train_path = 'archive\\asl_alphabet_train\\asl_alphabet_train'
        test_path = 'archive\\asl_alphabet_test\\asl_alphabet_test'

In [ ]: augment_train_data = ImageDataGenerator(horizontal_flip=True,
                                               rotation_range=50,
                                               zoom_range=0.2,
                                               width_shift_range=0.2,
                                               height_shift_range=0.2,
                                               rescale=1./255)
        augment_test_data = ImageDataGenerator(rescale=1./255)
```

FIGURE 12: Code snippet for hand gesture recognition part 1

```
In [ ]: train_dataset = augment_train_data.flow_from_directory(train_path,
             shuffle=True,
             classes=['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K',
             'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y', 'Z', 'space', 'del', 'nothing'],
             target_size=(img_size, img_size),
             batch_size=batch_size)
        test_dataset = augment_train_data.flow_from_directory(test_path,
             classes=['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K',
             'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y', 'Z', 'space', 'del', 'nothing'],
             target_size=(img_size, img_size),
             batch_size=batch_size)

In [ ]: fig = plt.figure(figsize=(15, 10))
        for i in range(1,9):
            plt.subplot(4,2,i)
            plt.imshow(train_dataset[0][0][i-1])
        plt.show()

In [ ]: if os.name=='nt':
            !pip install wget
            import wget

In [ ]: url = "https://tfhub.dev/google/tf2-preview/mobilenet_v2/classification/4"
        download_model = hub.KerasLayer(url,input_shape=(img_size,img_size,3))
        model = Sequential([
             download_model,
             Dense(29),
             Activation("softmax")
        ])

In [ ]: # compiling model
        model.compile(optimizer=Adam(1e-3),
        loss="categorical_crossentropy",
        metrics=['accuracy'])
```

FIGURE 13: Code snippet for hand gesture recognition part 2

```python
#training model
print("\n Model summary: ")
print(model.summary())
print("\n Model Training: ")
model.fit(train_dataset,
batch_size=batch_size,
epochs=epochs)
```

```python
print("\n Model Evaluation: ")
model.evaluate(test_dataset)
```

```python
model.save("Models\\ASL_model\\h5_format\\asl_model.h5")
```

```python
load_model = tf.keras.models.load_model("Models\\ASL_model\\h5_format\\assets",custom_objects={"KerasLayer":hub.KerasLayer})
print(load_model.summary())
```

```python
print("\n Model Evaluation: ")
load_model.evaluate(test_dataset)
```

```python
!pip install tensorflowjs
```

```python
!tensorflowjs_converter --input_format keras \
                       Models\\ASL_model\\h5_format\\asl_model.h5 \
                       Models\\ASL_model\\tfjs_format
```

FIGURE 14: Code snippet for hand gesture recognition part 3

2) Audio to sign

```javascript
// Group 4: Rudraksh Sharma Gemin Shrestha Sagaljyot Singh

import { useEffect, useRef, useState } from 'react';
import service from '../services/service';
import 'regenerator-runtime/runtime'
//300px
const maxVideoSize = 300;
const LETTERS = ['A','B','C','D','E','F','G','H','I','J','K','L','M','N','O','P','Q','R','S','T','U','V','W','X','Y','Z','_NOTHING',

export default function Page() {
  const videoElement = useRef(null);
  const canvasEl = useRef(null);
  const outputCanvasEl = useRef(null);
  let [letter, setLetter] = useState(null);
  let [loading, setLoading] = useState(true);


  async function processImage() {
    if (
      videoElement !== null &&
      canvasEl !== null &&
      typeof videoElement.current !== 'undefined' &&
      videoElement.current !== null
    ) {
      let prev_Letter = '';
      let count = 0;
      let letter_words = '';

      while (true) {
        const ctx = canvasEl.current.getContext('2d');
        ctx.drawImage(videoElement.current, 0, 0, maxVideoSize, maxVideoSize);
        const image = ctx.getImageData(0, 0, maxVideoSize, maxVideoSize);

        const processImage = await service.imageProcessing(image);


        const ctxOutput = outputCanvasEl.current.getContext('2d');
        ctxOutput.putImageData(processedImage.data.payload, 0, 0);
```

FIGURE 15: Code snippet for audio to sign conversion part 1

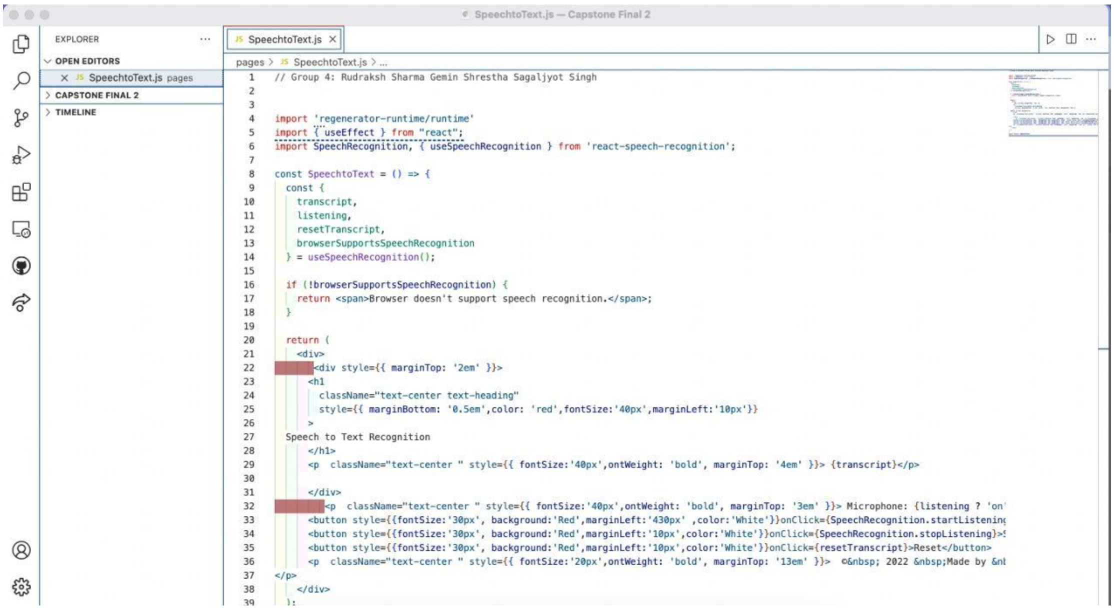


FIGURE 16: Code snippet for hand gesture recognition part 2

5.3.3 Project Deployment

All the components/ modules included in the architecture of our project are represented in the following diagram along with their relations

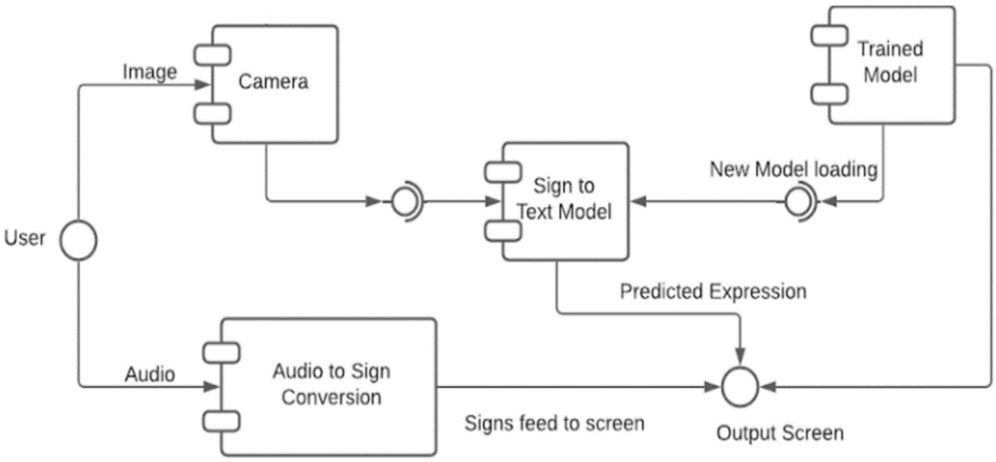


FIGURE 17: Deployment Diagram

5.3.4 System Screenshots

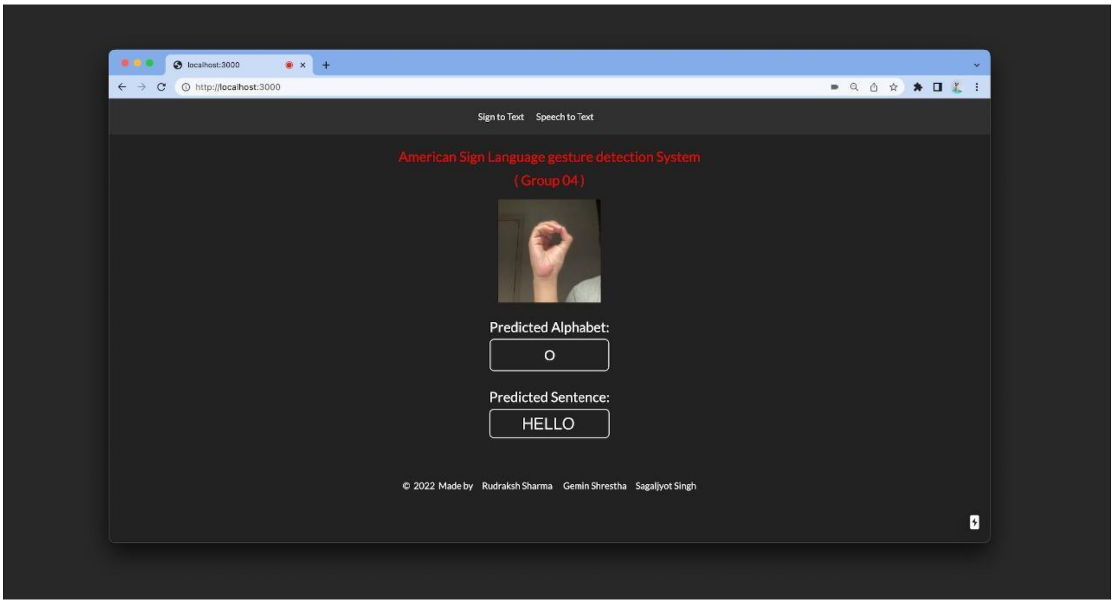


FIGURE 18: System screenshot Number 1

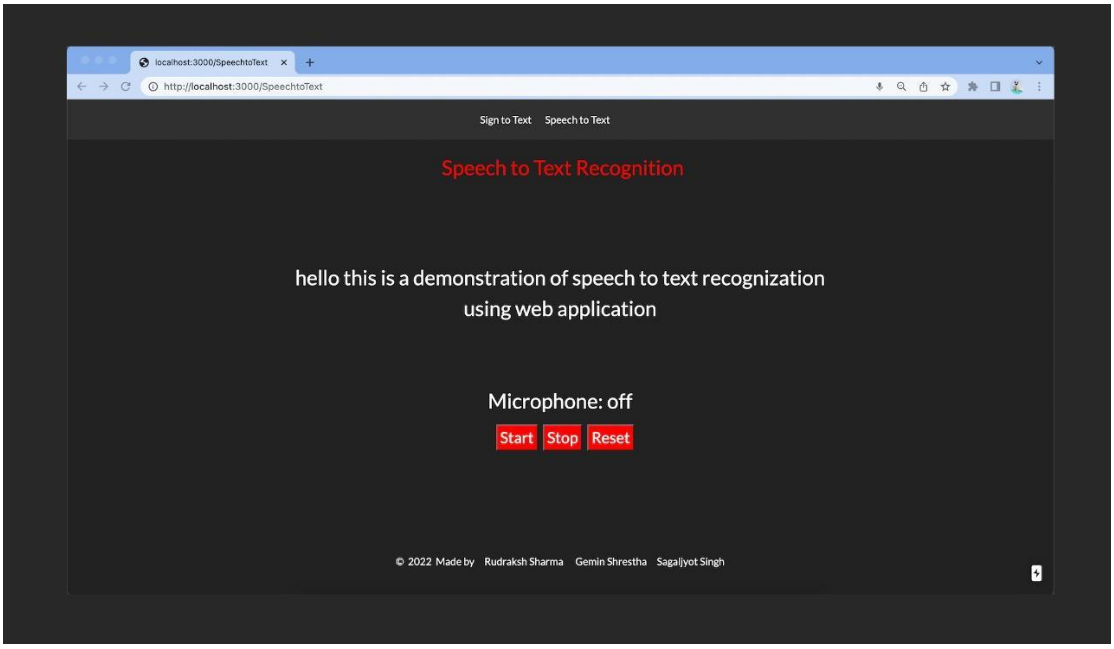


FIGURE 19: System screenshot Number 2

5.4 Testing Process

5.4.1 Test Plan

The application developed has been extensively tested. The various testing techniques and methods and their results have been documented in subsequent sections. All the modules were tested separately and cohesively.

5.4.2. Features to be tested

The tested modules include:

1) Hand Gesture Module (Detection, Recognition and Categorization)
2) Audio Capture
3) Audio to sign conversion

5.4.3 Test Strategy

The strategy for the test is fairly straightforward and simple. We test all the parts individually as well as collectively. This can be done using different types of testing techniques like Black Box testing (for testing each component separately), Integration Testing (for testing the integrated module to verify combined functionality). We have also used Unit testing to test individual parts.

5.4.4 Test Techniques

Unit Testing:

1) Module 1: Hand Gesture Module

Input: Input from web camera

Output: Name of Gesture displayed

2) Module 2: Audio Capture

Input: Sound

Output: Audio Successfully captured

3) Module 3: Audio to sign Conversion

Input: Captured Sound

Output: Text corresponding to the audio

Integration Testing:

1) Module 1: Main function

Input: Input from Web Camera

Output: Hand gesture recognized

2) Module 2: Audio capture and conversion

Input: Input from Web microphone

Output: text displayed on screen

5.4.5 Test Cases and Test Results:

TABLE 5: Test cases and results

S No.	Test case	Input	Expected Output	Actual Output	Result
T001	Hand Detection	Hand image from Web Camera	Successful detection	Hand detected and recognition started	Successful
T002	Gesture recognition	Hand Gesture Recognized	Gesture recognized	Sign label displayed	Successful
T003	Audio to sign conversion	Audio recorded via microphone on laptop	Text corresponding to that audio	Text corresponding to that audio	Successful

5.5 Results and Discussions

1) Most of the objectives that we started our project with, got fulfilled. All the components of the project are working.

2) The model detects the hand, recognizes the gesture and label is shown in real time which means the code is optimized enough.

3) The model used for hand gesture recognition is CNN made by using transfer learning. A pre-trained image classification model made by google was used and for capturing hand images webcam and TensorFlow.js was used.

4) For audio to sign conversion we are using the React Speech Recognition module which gives fast and accurate results.

5.6 Inferences Drawn

After running the start file on various types of people and also tested on unknown users on real time application:

1) The input data is a continuous and time variable data and even a small change in gesture while performing hand can result in recognizing a completely different gesture.

2) TensorFlow, Keras, and TensorFlow.js is a very impressive and extremely powerful library and modules which are capable of hand detection and recognition detection. It can be further trained using other Machine Learning models on top of them as used in the project.

3) React Speech Recognition module is an extremely versatile library which accurately records audio and converts them to text very fast.

4) Running TensorFlow.js in the internet browser on the client machine can prompt a few advantages like security, minimal expense, adaptability and usability.

5.7 Validation of Objectives

The validation of the proposed objectives is described in the table shown below.

TABLE 6: Validation of objectives

S No.	Objectives	Status
1	A system such that it can easily and accurately convert Sign Languages to text according to the user's expression.	Successful
2	Audio to sign converter	Successful
3	Web interface through which user can easily interact or launch software	Successful

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

The problem of Sign Language (SL) recognition using images is still a challenge. Similarity of gestures, user's accent, context and signs with multiple meanings lead to ambiguity. We concluded that a dataset to train and evaluate the system must have sufficient gesture variations to generalize each symbol. We extended the average method to determine if a dataset has enough gesture variations and we implemented a model trained by using transfer learning for recognizing hand gestures in images of American Sign Language letters. Currently, research works have focused mainly on the recognition of static signs of ASL from images that have been recorded under controlled conditions. We have noticed that by adding more symbol classes to recognize, the model precision decreases and thus the system will not be as good as before. The project is a simple demonstration of how a machine learning model can be used to solve computer vision problems with an extremely high degree of accuracy. A fingerspelling sign language translator is obtained which has an accuracy of 90%.. We could also use gloves to eliminate the problem of varying skin complexion and background of the signee. The lighting conditions of a room can also cause problems. The other issue that people might face is regarding their proficiency in knowing the ASL gestures. We have also learned how with the help of javascript and TensorFlow.js we can easily integrate machine learning models and create robust web applications.

6.2 Environmental, Economic and Societal Benefits

1. Further developed spelling Research is showing that youngsters who get familiar with the essential communication through signing abilities (like American Sign Language ASL) can have further developed spelling abilities. Signing assists with

giving children one more instrument for recalling spelling words and leaves a bigger engraving on the mind.

2. Improves classroom behaviours. A couple of teachers who have incorporated signal-based correspondence into their hearing classrooms are finding that they have better management of class. Involving signs for things, for example, questioning can diminish interferences and assist with keeping understudies on target during lectures.

3. Improves small motor skills. Gesture based communication assists with growing little coordinated abilities as a result of the adroitness expected for speaking with hand motions. The people who battle with little muscle strength and coordination can assemble these abilities as well as figure out how to impart all the more proficiently with another dialect.

4. Builds a better vocabulary. Individuals have a characteristic inclination to utilize their bodies to communicate, for example, pointing, shrugging shoulders, or wrinkling the nose. Communication via gestures takes that regular inclination and assists to reinforce vocabulary and the meanings of words.

6.3 Reflections

The team faced some moments of joy when things worked as desired and some moments of panic when nothing went according to agenda. In the course of this project a great level of learning has been achieved and several unexplored areas of engineering have been discovered.

6.4 Future Work

- The project includes only static image capturing, and can be enhanced to video capturing for dynamic gestures.

- The project can be extended to other sign languages by building the corresponding dataset and training the model.

- Sign languages are also spoken in a contextual basis where each gesture could represent an object, verb, so, identifying this kind of a contextual signing would require a higher degree of processing and natural language processing (NLP).

7.1 Challenges Faced

A lot of challenges were faced while accomplishing the objectives of our capstone project.

• Initially we explored a lot of possible solutions for the problem at hand, we had to dig deep enough to choose the strategies we ended up adopting to create the model. A lot of research papers were read to see how others had tried to come up with solutions to such setups.

• There isn’t a lot of data present in this field so initially monitoring was a pain in itself because we didn’t have a lot to compare with in the present conditions so we finally tried to come up with our own system.

• Recognition of hand and gestures required a favourable background, contrast and lighting conditions and these tampered with the experimental results, as the model would not be able to understand the position of the landmarks of the fingers.

• A lot of data was required to create a model that would have helped in converting audio to sign, so this was solved by making use of the React speech recognition module.

• During the making of the website, the major task was integration of components (sign to text and audio to sign) with our front end, and for that many backend languages and frameworks were tested and at last we decided to use TensorFlow,js.

7.2 Relevant subjects

The following course subject's knowledge will be used in the development of the proposed system:

TABLE 7: Relevant subjects and their description

Subject code	Subject name	Description
UCS507	**Software Engineering**	The methodology of agile will be used throughout the process of making the software for this project. As taught in this subject, all the well-suited diagrams and other different kinds of methods will be implemented like SDLC, scrum, UML diagrams, etc. Different kinds of testing strategies will be implemented which we learned in the practical implementation throughout this course.
UML501and UCS521	**Machine Learning and AI**	By virtue of these two concepts, different ML models will be used to perform hand gesture recognition as it is a subfield of AI which aims to teach computers the ability to do tasks with data, without explicit programming. The concept of training ML models to perform computational tasks is used here during the development of the device. All the conceptual and practical knowledge regarding the ML and AL gathered during this course will be implemented in this project.

UCS310	**Database Management System**	All the logical and physical database designs will be implemented as taught during this subject. Analysis of database design using E-R data model will be used as implemented during the project of this course. All the conceptual and practical knowledge regarding the database management gathered during this course will be implemented in this project.

7.3 Interdisciplinary Knowledge Sharing

The development of a Sign Language detection system requires knowledge from multiple domains. It required knowledge of machine learning models, deep learning, and web development for developing the software part. For web development knowledge of HTML, CSS and Javascript was required with some extra knowledge of Tensorflow.js. For deep learning knowledge of Python and Neural Networks was required. Also coordinating with teammates specializing in different domains was a challenge. Complete understanding and sharing of knowledge were necessary. Working together to achieve a common goal taught us a lot on how dividing the work is way too important. Overall, it was a great experience and we learned a lot from each other.

7.4 Peer Assessment Matrix

In the following peer assessment matrix, points were awarded out of 5(5 being the max). The legend followed is 5-Excellent, 4- Very Good, 3-Good 2-Satisfactory, 1-Unsatisfactory

TABLE 8: Peer Assessment Matrix

<table>
<tr><td></td><td colspan="4">Evaluation of</td></tr>
<tr><td rowspan="4">Evaluation By</td><td></td><td>Gemin Shrestha</td><td>Sagaljyot Singh</td><td>Rudraksh Sharma</td></tr>
<tr><td>Gemin Shrestha</td><td>**5**</td><td>**5**</td><td>**5**</td></tr>
<tr><td>Sagaljyot Singh</td><td>**5**</td><td>**5**</td><td>**5**</td></tr>
<tr><td>Rudraksh Sharma</td><td>**5**</td><td>**5**</td><td>**5**</td></tr>
</table>

7.5 Role playing and work schedule

The plan of attack towards this problem would be to break it down into smaller problems then pipeline them to form the bigger picture. Our problem could be divided into multiple major parts, being related to machine learning and the others containing the development part. In the same manner we have decided to divide our team taking into consideration their particular electives.

TABLE 9: Tasks and Individuals Involved

S. No	Task	Individuals involved
1.	Data Collection	Rudraksh, Gemin, Sagal
2.	Data Pre-Processing	Rudraksh
3.	Model Training	Rudraksh, Gemin
4.	Model Conversion	Rudraksh, Gemin
5.	Audio to sign Converter	Gemin
6.	Gesture detection using Object Detection API	Rudraksh
7.	Web Designing	Gemin
8.	Integration of ML models and Website	Gemin
9.	Documentation	Sagal

7.6 Students Outcome

TABLE 10: Student Outcomes

S. no.	Description	Outcome
A1	Applying numerical ideas to get insightful and mathematical arrangements.	Finding out landmarks in hand for recognition of hand gestures and

		calculating error and accuracy of model performance.
B1	Utilize proper strategies, instruments and methods for information assortment.	Collected hand gesture data from various users which helped us increase our model's accuracy a lot.
C1	Applying engineering techniques for solving computing problems.	The application of techniques that are going to solve the purpose of doing computations for the dataset.
C2	Can comprehend degree and requirements, for example, monetary, natural, social, political, moral, wellbeing and security, manufacturability, and supportability.	The team understands the scope of the project as one with the social, economic, health and safety benefits. Various imperatives have been assessed and arrangements are being explored for them.
D1	Can play different roles as a team player	Each member was easily able to switch from one responsibility to another responsibility as the need arose.
E1	Identify engineering problems.	Really took a look at plausibility and recorded distinguished issues in information assortment. Besides, concluding the degree was a significant piece of this.
F1	Showcase professional responsibility while interacting with peers and professional communities.	All of the members were punctual in attending the group meetings to discuss the upcoming responsibilities. The team was regular at arriving at the evaluation destination for panel and mentor evaluation as well.

G1	Produce a range of documents, for example, lab or task reports utilizing proper configurations.	Wrote SRS (Software Requirements Specification) along with an exhaustive project report which iterated throughout the duration of the project, following IEEE format. The report contains elaborate quantifications and explanations regarding the project.
G2	Deliver well-organized and effective oral presentations.	Delivered a presentation of the product on multiple occasions (project evaluation sessions and idea defending) in front of a panel.
H1	Aware of the environmental and societal impact of engineering solutions.	The project as a whole is a part to help people interact with the machines in their own language, so the team is successful in doing so.
I1	Able to explore and utilize resources to enhance self-learning.	Learnt different ways to tackle our problem statement by reading existing research work in the respective fields
K1	Write code in different programming languages.	Wrote code Python for ML modules. Many libraries and modules were also used.
K2	Apply different data structures and algorithmic techniques.	Used arrays, dictionaries, hash maps, strings etc. as per the requirements.
K3	Use software tools necessary for the computer engineering domain.	TensorFlow, Sklearn, Google Colab, Open CV, Speech Recognition and other python libraries are extensively used.

7.7 Brief Analytical Assessment

Q1. What sources of information did your team explore to arrive at the list of possible Project Problems?

Ans: The group was aware of the understanding of the Capstone requirement and some of the problems that need to be explored. Team explored the literature, mostly the technical journals and technical magazines from IEEE. The interfacing issues have been refreshed through textbooks or internet resources. However, the scope has been decided after consulting our supervisor.

Q2. What analytical, computational and/or experimental methods did your project team use to obtain solutions to the problems in the project?

Ans: Constructing a simple and effective model was the most challenging task. We collected hand gesture data of various users from Kaggle and converted them into embeddings of size 240 and built models to detect the hand and the landmarks of the fingers. Deep Learning and various modules of TensorFlow are extensively used for building models.

Q3. Did the project demand demonstration of knowledge of fundamentals, scientific and/or engineering principles? If yes, how did you apply?

Ans: In this technical project, we have used the principles of Deep Learning and Voice Analysis. Other skills used are software integration techniques. Design, architecture and documentation principles were taught in the subjects Software Engineering and Software Design and Construction.

Q4. How did your team share responsibility and communicate the information of schedule with others in the team to coordinate design and manufacturing dependencies?

Ans: The team of three members divided the project into subtasks, individuals taking one subtask at a time. Slack, a free online project management tool was used to assign tasks, deadlines and for other communication and coordination. Besides, we regularly connected via Google Meet to examine our progress.

Q5. What resources did you use to learn new materials not taught in class for the course of the project?

Ans: The team took various online tutorials and referred to the web to explore more about the topics on which we are working upon. Various courses on YouTube for learning how to make various models and their implementation were referenced.

Q6. Does the project make you appreciate the need to solve problems in real life using engineering and could the project development make you proficient with software development tools and environments?

Ans: The project addresses a real-life problem using engineering. Working on this project has made the group appreciate the need to solve real world problems and has motivated the team to take up new problems in diverse fields. The group was introduced to various new technologies like computer vision, deep learning and Web development.

APPENDIX A: REFERENCES

[1] B M Chethana Kumara, H S Nagendraswamy and R Lekha Chinmayi, "Spatial Relationship Based Features for Indian Sig Language Recognition", International Journal of Computing, Communications & Instrumentation Engineering, Vol. 3, Issue 2, ISSN 2349-1469, 2016.

[2] T. Yang, Y. Xu, and "A., Hidden Markov Model for Gesture Recognition", CMU-RI-TR-94 10, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 1994.

[3] M. J. Cafarella, A. Halevy, and N. Khoussainova, "Data integration for the relational web," PVLDB, vol. 2, no. 1, pp. 1090–1101, Aug. 2009.

[4] Shivashankar S, Srinath S, "A comparative Study of Various Techniques and Outcomes of Recognizing American Sign Language: A Review", International Journal of Scientific Research Engineering & Technology (IJSRET), Vol. 6, Issue 9, pp.1013-1023, 2017.

[5] Jun Deng, Xiaojing Xuan, Weifeng Wang, "A review of research on object detection based on deep learning", 7th IEEE International Conference on Communication Systems and Network Technologies (CSNT-2017), pp.293-299, 2017.

[6] Marin Kaluza and Bernad Vukelic "Comparisions of front-end frameworks for web applications", vol 6 (2018), pp-261-286

[7] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham

[8] Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision-based features. Pattern Recognition Letters 32(4), 572–577 (2011)

[9] N. Mukai, N. Harada and Y. Chang, "Japanese Fingerspelling Recognition Based on Classification Tree and Machine Learning," 2017 Nicograph International (NicoInt), Kyoto, Japan, 2017, pp. 19-24. doi:10.1109/NICOInt.2017.9

[10] Byeongkeun Kang, Subarna Tripathi, Truong Q. Nguyen" Real-time sign language fingerspelling recognition using convolutional neural networks from depth map" 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)

[11] TensorFlow, https://en.wikipedia.org/wiki/TensorFlow, 29 Nov 2021, [22 April 2022].

[12] CNN, https://en.wikipedia.org/wiki/Convolutional_neural_network, 8 Nov 2021, [20 April 2022]

[13] OpenCV Documentation https://docs.opencv.org/2.4/doc/tutorials/imgproc/, 31 Dec 2021, [11 March 2022]

[14] Research - Sign Language Recognition, http://www-i6.informatik.rwthaachen.de/~dreuw/database.php, 7 Aug 2021, [15 March 2022]

[15] *TensorFlow* [Online]. Available: URL https://www.tensorflow.org/, 14 September 2021, [16 March 2022]

[16] Build and deploy a custom object detection model with TensorFlow Lite (Android) [Online]. Available: URL https://developers.google.com/codelabs/tflite-object-detection-android, 31 December 2021, [13 March 2022]

[17] Nicholar Renotte. *Learn Machine Learning* [Online]. Available: URL https://www.youtube.com/c/NicholasRenotte/featured, 23 February 2022, [25 April 2022]

[18] Real-Time Object Detection Using TensorFlow [Online]. Available: URL https://www.mygreatlearning.com/, 20 February 2022, [11 March 2022]

APPENDIX B: PLAGIARISM REPORT

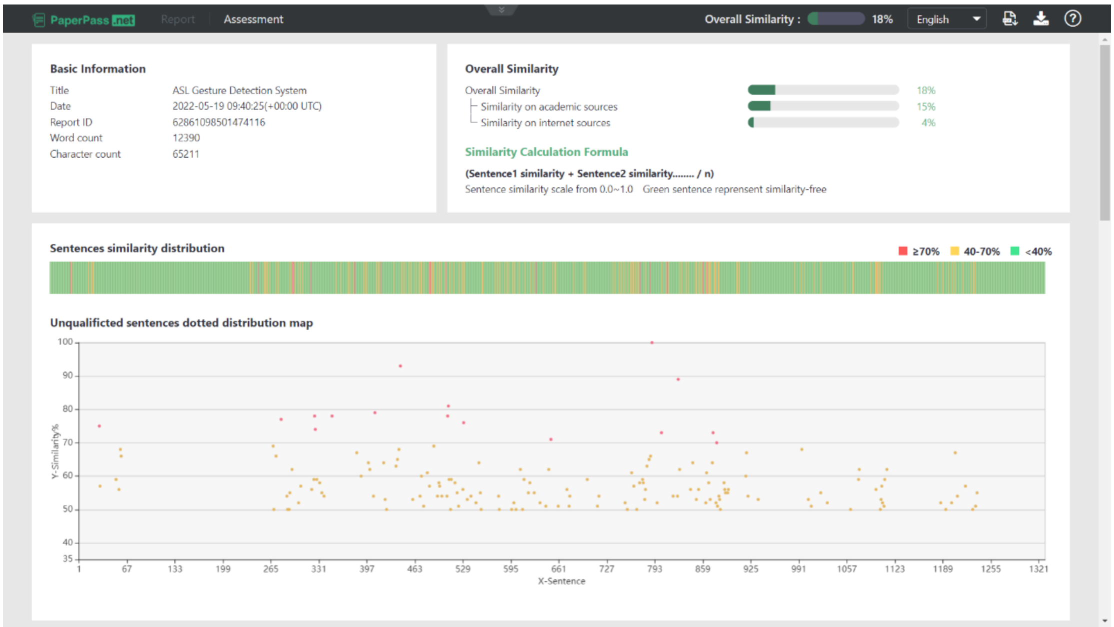


FIGURE 20: PaperPass Originality Report