SIGN LANGUAGE IDENTIFIER

Capstone Project Report

END SEMESTER EVALUATION

Submitted by

101703476 Samiksha Kapoor

101703482 Sanjana Vashisth

101703488 Sargunpreet Kaur

101703545 Simran

BE Fourth Year- COE

CPG No. 76

Under the Mentorship of

Dr. Seema Bawa

Professor

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

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

December, 2020

ABSTRACT

"Sign Language Identifier" is a web-based application which is an interpreter for deaf and dumb people that converts sign language to text; and speech to ISL.

Deaf people do not have many options for communicating with a hearing person, and all of the alternatives do have major flaws. Interpreters aren't usually available, and also could be expensive. The pen and paper approach is also a bad idea as it is not practical. So a framework is required that provides a helping-hand for deaf and dumb people to communicate using sign language.

Deaf and Dumb people simply just have to perform the actions of Indian Sign Language and in front of the webcam which will be converted to text; and speech to ISL. Hence, the communication gap between the normal people and Deaf and Dumb people will be reduced. They will be also able to express themselves through this, hence reducing the feeling of isolation felt by them.

DECLARATION

We hereby declare that the design principles and working prototype model of the project entitled Sign Language Identifier is an authentic record of our own work carried out in the Computer Science and Engineering Department, TIET, Patiala, under the guidance of Dr. Seema Bawa, Professor, T.I.E.T. during 6th semester (2020).

Date: 30 November, 2020

Roll No.	**Name**	**Signature**
101703476	Samiksha Kapoor	
101703482	Sanjana Vashisth	
101703488	Sargunpreet Kaur	
101703545	Simran	

Counter Signed By:
Faculty Mentor:
Dr. Seema Bawa
Professor
CSED,
TIET, Patiala

ACKNOWLEDGEMENT

We would like to express our thanks to our mentor “Dr. Seema Bawa Mam”. 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 would also like to express our thanks to “Ms. Sawinder Kaur”, PHD student T.I.E.T. Her guidance was of immense help to us.

We are also thankful to, “Dr. Maninder Singh”, 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: 20 December, 2020

Roll No.	**Name**	**Signature**
101703476	Samiksha Kapoor	
101703482	Sanjana Vashisth	
101703488	Sargunpreet Kaur	
101703545	Simran	

TABLE OF CONTENTS

7. Project Metrics

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SLI	Sign Language Identifier
ISL	Indian Sign Language
SRS	Software Requirement Specification
DL	Deep Learning
CNN	Convolutional Neural Network

INTRODUCTION

The language barrier or communication barrier has been overcome by making learning a common language or using language to language digital translator. But for the people who are not able to hear and speak, such technologies and ideas make no difference. Therefore "Sign Language Identifier" is an interface that will get rid of this communication divide.

1.1 Project Overview

A web-based application called "Sign Language Identifier" is an interpreter for deaf and dumb people that converts sign language to text, and speech to ISL. Inorder to know it thoroughly we need to know its background and current practices.

Sign language is any means of communication that uses body movements, especially hands and arms when spoken communication is impossible or not desirable. Therefore sign language is the bridge gap wherever vocal communication is not possible. Only 5% of the population is hearing impaired, so most of the people don't learn sign language and the medium of its education is also less. Therefore the communication gap between the two communities still persists. Due to this many deaf and dumb people were left uneducated and jobless. The pen and paper approach is a bad idea for communication as it is very time consuming and also fails while communicating with uneducated people. The schools are made for the deaf and dumb people where they are tought sign language and they learn many skills and do activities using sign language. But this defines the boundary for them and they are left segregated from the rest of the world. This not only affects them but also their country, as it loses many great thinkers and intellectuals which could have helped in uplifting the nation just because of the communication barrier.

Medical field has worked on helping the deaf and dumb people. Cochlear implant is the cure fort the deaf and dumb children under the age of six but it also has strict selection process for the surgery and is also expensive.

Then there came the person interpreters who translate the communication for hearing impaired people and became their speakers. But interpreters aren't usually available, and also could be expensive. So Deaf people do not have many options for communicating with a hearing person, and all of the alternatives do have major flaws.

So a framework needs to be built that provides a helping-hand for deaf and dumb people to communicate using sign language. This would be a user-friendly environment for the user by providing text output for a sign gesture input and speech output for ISL input. There are some frameworks that are built for the same purpose, but they have many flaws which make them imperfect to use.

Basic sign language system is based on the 5 parameters and they are hand and head recognition, hand and head orientation, hand movement, shape of hand and location of hand and head. Capture of hand movements is via webcam. Complex video backgrounds in another gap in sign language recognition where it poses a major challenge to extract signers hand shapes accurately. This SLR system focuses on putting a simple constant background with the signer's shirt matching the background. In cluttered video backgrounds tracking hands has become simpler but tracking each finger movements remains quite a challenging task. So by using different technologies of machine and deep learning and computer vision in “Sign Language Identifier” will eradicate the above complexities.

This system will be user friendly and all the complexities in recording will be solved using different technologies. Therefore, users will not have to be careful or take certain measures to use the system. The system is based on Indian Sign Language, it is basically made for the indian users or for those who use indian sign language. Previously no system was designed for ISL, in order to use it at work or while communication people usually have to learn other sign language. It has been observed that people who learn sign language during adolescence or teenage are not able to grap it better than those who learnt it when they were born. The World Bank strongly stated that children whose first language is not used at school experience a low level of learning and are much less likely to contribute to the country's economy and intellectual development. The main purpose of education is to contribute skilled manpower and it's hard to get that kind of man if they learn in their second language. Above all language is the instrument used by people to think, analyse and relate to other people and environment. Language is therefore the basis of self awareness, for sensing of meaning and for intellectual development. Therefore switching from one sign language to another will degrade their performance at work, their creativity and they will find it difficult to express themselves even if they are using a sign language translator.

Therefore, The Sign Language Interpreter will provide the people using indian sign language, a user friendly interface that will help them to have communication with people who have no knowledge about the sign language in a much more natural way and this will boost their confidence and productivity. Following are benefits that users will get.

1. Since deaf people are usually deprived of normal communication with other people, they have to rely on an interpreter or some visual communication. Now the interpreter can not be available always, so this project can help eliminate the dependency on the interpreter.

2. The system can be extended to incorporate the knowledge of facial expressions and body language too so that there is a complete understanding of the context and tone of the input speech.

3. A progressive web based application will increase the reach to more people.

4. Integrating hand gesture recognition system using computer vision for establishing 2-way communication system.

5. It will be a scalable project which can be extended to capture the whole vocabulary of ISL through manual and non manual signs.

1.1.1 Technical terminology

I. Machine Learning: To get better insight about how "sign language identification" problems can be looked at and hence, solved.

II. Deep Learning: To apply the most suitable deep learning algorithm (CNN) for sign language identification, and also knowledge of optimizers, loss functions.

III. Natural Language Processing: For matching gif or image corresponding to speech from a bag of words.

IV. Image Processing: To capture the image from the camera and further removal of noise using noise reduction filters and further enhancing the input for better output and ease of prediction for the trained model.

V. Software Engineering: Software Development Lifecycle, Preparation of SRS, Working as per Scrum model, Shaping functional and non-functional requirements, Understanding and communication of ideas through UML Diagrams.

VI. CNN: Convolution Neural Network is a class of deep neural networks, applied to analyzing visual imagery.

VII. OpenCV: OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. It provides a common infrastructure for computer vision applications and to accelerate the use of machine perception.

VIII. Tkinter: Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit.

IX. Flask: Flask is a web application framework written in Python.

1.1.2 Problem statement

To design “Sign Language Identifier”, a web app for deaf and dumb people is to be deigned keeping in view that sign language is used by the hearing impaired people for their communication. These kinds of systems create a gap between normal people and hearing impaired people. Sign Language Identifier is a helping aid for speech impaired to communicate with the rest of the world using sign language. This application will convert the sign language gestures to text/speech and the speech/text back to sign language.

1.1.3 Goal

1. To survey and analyze the current status of usage and effectiveness of Indian Sign Language (ISL) solutions.
2. To propose and implement an interactive solution that will convert sign language to text and speech and vice versa.
    A. To collect ISL dataset and create ISL dictionary for commonly used words and to preprocess them.
    B. To apply Deep Learning Model for training on collected dataset.
3. To test and demonstrate the usability and effectiveness of the proposed solution for targeted user groups (Deaf and dumb persons).

1.1.4 Solution

Solution to the problem of communication of the deaf and dumb people is to build a computer that can be programmed in such a way that it can translate sign language to text format; and speech to ISL format, the difference between the normal people and the deaf community can be minimized. Indian Sign Language (ISL) Interpretation system is a good way to help the Indian hearing impaired people to interact with normal people with the help of a computer. Compared to other sign languages, ISL interpretation is an area which gained less attention by the researchers since American Sign Language (ASL) has most of the signs, single handed and thus, complexity is less. Also, ASL already has a standard database that is available for use. When compared with ASL, Indian Sign Language relies on both hands and thus, an ISL recognition system is more complex.

1.2 Need Analysis

Sign Language is the most natural and expressive way for hearing impaired people. People who are not deaf, never try to learn sign language for interacting with the deaf people. India does not have a single well-reputed junior or senior college for them. The deaf and mute people have to drop out of education or take admission for long distance courses or they have to take admission in colleges meant for those without any physical challenges. In regular colleges, they face a series of hurdles. To overcome low esteem born out of their disability while studying in colleges, where most of the students have no physical challenges, is pretty difficult. And most of them do not seek admission or prefer to drop out after joining. Besides, none of these colleges avail interpreters or lecturers who can teach physically challenged students in sign language. The students, therefore, have to get help from students or close friends/relatives to study. They aren't able to become self dependent in most of the cases and rely on mercy of others.

Every day thousands of local businesses around the globe face problems with providing their services to them. People who are talented and capable of bringing a good change in society are hindered by their disability, which is not just their loss but even nation's loss.

Not even here, everywhere they feel the disability they are born with. This leads to isolation of the deaf people. But if the computer can be programmed in such a way that it can translate sign language to text format; and speech to ISL format, the difference between the normal people and the deaf community can be minimized. Indian Sign Language (ISL) Interpretation system is a good way to help the Indian hearing impaired people to interact with normal people with the help of a computer. Compared to other sign languages, ISL interpretation is an area which gained less attention by the researchers since American Sign Language (ASL) has most of the signs, single handed and thus, complexity is less. Also, ASL already has a standard database that is available for use. When compared with ASL, Indian Sign Language relies on both hands and thus, an ISL recognition system is more complex.

1.3 Research Gaps

Sign language plays a vital role in the life of any auditory impaired individual but it does not get the same status as any other language. Hand gestures are most widely used as a medium of sign language based communication framework among various forms of gestures. Recognizing gestures under dynamic parameters such as lighting conditions, multiple hands as background, left handed person, right handed person, size of finger etc. are thought provoking fields in automate Indian sign language recognition . A review is based on datasets of hand gestures, alphabets, digits, words and sentences with various real time conditions used by different researchers at national and international level is summarized in Table 1.

Table 1: Research Gap, Deficiencies and References

Research Gap	Deficiencies	References
Incomplete dataset	• There is no complete dataset or dictionaries made or used for complete real time conversation • Many dataset contain only alphabets or specific words, which could only help in identifying the words but will not be able to frame a proper grammatical sentence	Naoum, Reyadh, Hussein H. Owaied, and Shaimaa Joudeh. Development of a new Arabic sign language recognition using k-nearest neighbor algorithm. (2012) Christopher Lee and Yangsheng Xu, “Online, interactive learning of gestures for human robot interfaces” Carnegie Mellon University, the Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996
Less research in sign languages using both hands and hand and	• There is no research made to identify the word or the alphabet made using both hands	Richard Watson, “Gesture recognition techniques”, Technical report, Trinity College, Department of

head coordination	● Indian Sign Language contains gestures involving both hands. Therefore no system was made to identify it. ● Sign language involve the words or sentences made using both hands, hand and head coordination, these words and sentences could not be identified	Computer Science, Dublin, July, Technical Report No. TCD-CS-93-11, 1993 Mandeep Kaur Ahuja, Amardeep Singh, “Hand Gesture Recognition Using PCA”, International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015
Purposed researches and model are background specific	● In order to identify the gesture, the background should be of specific color or the person should wear a specific colored jacket, This makes the system extremely rigid, practically impossible and decreases its scope of use.	P.V.V. Kishore, Pullakura Naoum, Reyadh, Hussein H. Owaied, and Shaimaa Joudeh. Development of a new Arabic sign language recognition using k-nearest neighbor algorithm.“ (2012)Rajesh Kumar Published 2012 Computer Science International Journal of Advanced Computer Science and Applications

Rigid system of maintaining the specific distance from the webcam while recording	● In order to record the video, specific distance should be maintain from the camera which give rise to the complication in recording the video ● It is also not possible for the user to calculate the distance very time before recording the video ● This will give rise to inappropriate results and also efficiency of the system will decrease	Alina Kuznetsova, Laura Leal-Taix´ e, Bodo Rosenhahn Institute fuer Informationsverarbeitung, Leibniz University Hannover Appelstr. 9A, Hannover, 30167, Germany P.V.V.Kishore P.Rajesh Kumar Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text
Less Efficiency on video	● The system only worked on the images. It was not able to identify the signs in the video ● Therefore, no use in the practical world	Naoum, Reyadh, Hussein H. Owaied, and Shaimaa Joudeh. Development of a new Arabic sign language recognition using k-nearest neighbor algorithm.“ (2012)

1.4 Problem Definition and Scope

To bridge the gap between deaf and dumb and the rest , the main purpose of this project is to design a web app that can convert sign language to english text; and speech to ISL. Sign Language is an incredible advancement that has grown over the years. It is primarily used by deaf and dumb to communicate with rest of the world. Unfortunately, not everyone can understand and interpret sign language which leads to a communication gap between deaf and dumb and the rest. The main purpose of the project is to eliminate this communication gap so that sign language can be understood by common people without the help of any interpreter. This project aims at building a web based application that can convert sign language to text;and speech to ISL. Anyone willing to communicate through sign language will be able to login through the app. Hand gestures of the signer will be captured by the camera of the device which will be further converted to text and speech recorded by the audio will be converted to ISL. Signer will be at a fixed distance from the camera. The dataset for training the model will contain gestures from Indian Sign Language only.

1.5 Assumptions and Constraints

The following are the general assumptions and constraints applied while building the Sign Language Identifier web app (as shown in Table 2).

Table 2: Assumptions and Constraints

S. No.	Assumptions and Constraints
1.	Indian Sign Language is used.
2.	The user must have a user id to login to the web app.
3.	The device must be connected to a webcam and speaker.
4.	The signer must be at a fixed distance from the webcam.
5.	User has a steady pace while making signs so that the camera could detect the signs properly.

6.	The signs are made in a controlled environment keeping a fixed background.

1.6 Standards

1. ISL(Indian Sign Language) standard: Used to create Indian Sign Language dictionary.
2. Web 2.0 : Internet technology to create online applications that behave dynamically.
3. Camera measurement standards :

   3.1 Resolution Measurements - ISO 12233

   3.2 Noise Measurements – ISO 15739

   3.3 Sensitivity Measurements – ISO 12232

   3.4 OECF (tone reproduction) Measurements – ISO 14524

   3.5 Chromatic Displacement Measurements – ISO 19084

   3.6 Color Characterization Test Procedures – ISO 17321

   3.7 Image Stabilization Measurements – ISO 20954

   3.8 Low Light Measurements – ISO 19093

   3.9 Shooting Time Lag Measurements – ISO 15781

   3.10 Shading Measurements – ISO 17957

1.7 Approved Objectives

1. To create a web app to enable Indian Sign Language to be converted into English text; and speech to Indian Sign Language.
2. To perform tasks for conversion of Indian Sign Language to English

   2.1 Capture Video

   2.2 Generate text

   2.3 Display text
3. To perform the task for conversation of Speech to ISL

   3.1 Record the voice

   3.2 Generate the ISL gif

   3.3 Display the ISL gif
4. To design and implement the web application.
5. To test and validate the web app as per functionality and user experience.

1.8 Methodology Used

ISL to text

For the sign language recognition system, a vision based approach will be used. The system will contain four modules/derived algorithms: preprocessing and segmentation, feature extraction, sign recognition and sign to text. Indian Sign Language is the area of concern. The signs will be captured using a webcam where the signer must be directly in front of the webcam. Then the output will be in the form of text. Following is methodology for sign language to text.

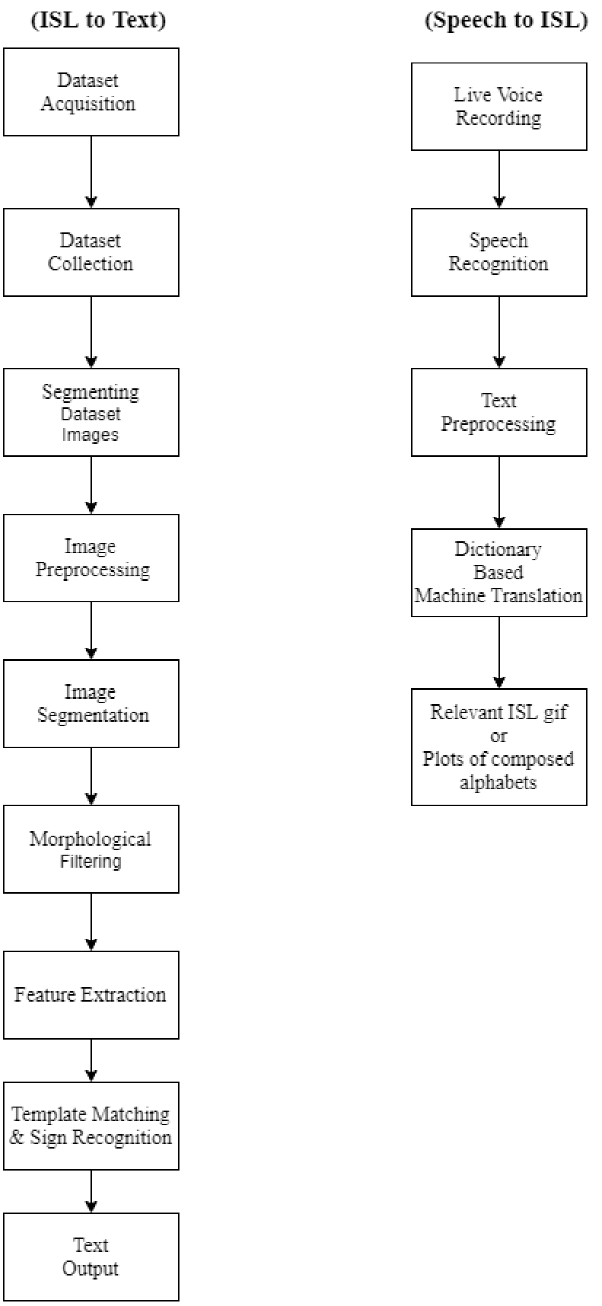


Figure 1: Sign Language Identifier Flowchart

Figure 1 shows the flowchart for the methodology so adopted.

1.8.1 Preparing the dataset

A dataset of 26 English language alphabets in Indian Sign Language will be created. The images in the dataset are to be segmented. The dataset will then be divided for training and testing. A dictionary for ISL will be created so that idea extends to sentences as well.

1.8.2 Image Acquisition

The gestures of sign language will be captured through the webcam for testing. The gesture should be made in the rectangular region defined in the web app. This will be in the form of a video, from which suitable frames will be extracted.

1.8.3 Image Preprocessing

The BGR image is now acquired, followed by segmentation and morphological filtering methods. These operations are done on the acquired image.

a. **Segmentation:** Here the image is converted into small segments in order to get more accurate image attributes. Here the segmentation of hands is carried out to separate objects and the background. To make segmentation more robust, so that the feature extraction and sign recognition of the image is accurate, the BGR image is converted to HSV, then fed into inRange() function which performs segmentation.

b. **Morphological filtering:** The image attributes extracted from the segmentations consist of noise, therefore morphological filtering removes the noise and gives a smooth contour. The output of these filterings are image attributes which are useful for feature extraction and sign recognition. Dilation and Erosion are used for morphological filtering.

1.8.4 Feature Extraction

The feature is nothing but the thresholded and morphological filtered image which came as output of thresholding and morphological filtering operation.

1.8.5 Template Matching and Sign Recognition

The image so generated is matched with labels with help of predict_class() method of CNN model. Say the sign for 'A' is performed, then the class is 'A' which is the folder's name in the dataset.

1.8.6 Text Generation

The predicted text as predicted by predict_class() method, is displayed on screen.

Speech to ISL

Firstly a dataset of alphabets and gifs is created. Post that, the coding is done for the conversion of speech to ISL. Firstly, the microphone is activated and input voice is thresholded. Then, the voice recognition takes place using Google API. The spoken word is processed to text and then text is searched with labelled folders of the dataset. If word exists, then corresponding GIF is displayed, else the plot of alphabets of which letter is composed of is displayed on the screen.

1.9 Project Outcomes and Deliverables

Sign language converter will help deaf and dumb people to communicate with other people around them. The deaf and dumb people do not have many options to communicate with normal people. Interpreters are also not usually available. Also, this app will be helping them a lot as it will save time and will not be complex to use.

The deaf and dumb person will simply perform the actions according to sign language and that will be converted to text which could be easily understandable by the other person. Hence this app will assist them and translate the sign language to text; and speech to ISL as fast as the person speaks and not make them feel isolated from society.

1.10 Novelty of Work

The novelty of the work lies in the fact that the sign language used is Indian Sign Language (ISL). Though a lot of work is done in American Sign Language (ASL), ISL has a bit less research when compared with ASL and the work done in ISL is restricted to 26 alphabets of English or maximum to maximum, Classification and Feature Extraction takes place. Only a few out of which employ Image Processing tasks to make it efficient. Hardly, any work is done by combining all techniques of Image Processing, Machine Learning, Neural Network and Natural Language Processing (NLP). Generally, only one or combination of two techniques is used/ solution is hardware based/ glove based, which makes it less easily accessible, lesser ease of avail, so it's not able to reach more people, who are in need. Also, lesser systems are practical and if they are practical, they have less accuracy, so the "Sign Lnaguage Interpreter", works on making it more feasible and practical keeping in view the needs of deaf and dumb.

In SLI various Image Processing techniques are used like applying different filters to extract frames out of videos, followed by image pre processing operations like Image Segmentation, Morphological Processing (Dilation and Erosion), then some of Machine Learning and Deep Learning tasks like Template Matching, Feature Extraction, then NLP tasks and at last SLI gets a text, which brings in novelty of the project.

REQUIREMENT ANALYSIS

2.1 Literature Survey

Sign language has been one of the major areas of research from global sign languages (like American Sign Language) to the local languages using various technologies. The work has been done by the scientists in their native languages but not much work is done on the Indian Sign Language (ISL) as it is complex and the complete dictionary of ISL has not been devised. So in this project, the main focus is on the ISL (Indian Sign Language) using different technologies like Deep Learning and Image Processing etc.

2.1.1 Theory Associated With Problem Area

The problem of communication gap between the hearing and hearing impaired people exists for a very long time. It led to segregation of the two communities as they were not able to understand each other. Hearing impaired children were not able to get education, thus slowing their and their country's growth. The sign language was developed and after some time it became official in school as well. But due to less population of the hearing impaired people not much attention was given to develop the medium or school where they could learn using sign language. But still this could not bridge the gap between hearing and hearing impaired people. As the technology developed, the first attempt to make sign language identifiers was a glove based approach that identifies the sign made by the hand wearing a glove. But this approach was expensive and not efficient. Therefore, it could not be used by the common people.

Later when the AI era came, many apps, system for the same were made but none could succeed to provide real time communication. It either failed in having a proper dataset or had a lot of constraint while using the system. Nowadays a lot of efforts are made to make such devices that overcome all the constraints and make a user friendly interface for the hearing impired people.

2.1.2 Existing Systems and Solutions

Many researchers have tried to develop the sign language identifier system, but very few were able to make it with high efficiency and success rate. Some have come up with great solutions but it has not been implemented yet. The following are the existing systems and solutions developed for sign language identifiers.

MotionSavvy: It is a tablet app that understands sign language. It is the software that converts the american sign language to text or voice. It also has a voice recognition system which helps hearing people to convert their voice to American Sign Language(ASL). It currently identifies 100 words. There are many signs in american language itself. The MotionSavvy case embeds the Leap, and the MotionSavvy software leverages the Leap's 3D motion recognition, which detects when a person is using ASL.

Re-voice Glove: It is the glove based approach and is a data glove. It is an electronic device that converts the gestures made by the gloves into text or voice. It has a screen and speaker attached to the back of the glove, which makes it portable. It has many sensors attached to it that detects its motion. It could be attached to the smart devices and then make the gesture and type the word it represents. This way the glove learns to identify the signs. Therefore, it is not specific to a particular sign language and it also bridges the communication gap between the people that uses different sign language.

AI interpreter: It translates the user's sign in real time. It is based on the chinese sign language. Its dataset contains 900 chinese phrases and can detect up to 1000 words. It has one drawback that it requires white background. The Tecent YouTu Lab has decided to install it in every railway station, bus stands and airport to help the hearing impaired people get their tickets or get queries without any problem.

AR App: Augmented reality app is the mobile app used to convert American Sign Language to text and voice. It also has a voice recognition system that converts the voice to sign language. It uses a mobile camera for recording the gestures and mobile speaker for voice recognition. It is basically a mobile solution as an instant sign language interpretation app that leverage computer vision models over the cloud, and present users the visual translations in augmented reality to empower both the sign language users and nonusers for more collaborations.

Google Translator: It is an AI app that converts the sign language to text/speech. It works by placing a smartphone in front of the user while the app translates gestures or sign language into text and speech. The app, called GnoSys, uses neural networks and computer vision to recognise the video of sign language speakers, and then smart algorithms translate it into speech. The translation software in the market are either slow or expensive, or rely on old technology which does not allow scaling to other markets outside the country of origin. But this app is a compellingly fast, easy, comfortable and economical solution. It can translate as quickly as the person speaks, translate any sign language and can be plugged into many products, such as video chat applications, AI assistants, etc. The pocket interpreter for the deaf relies on superior new technology: AI and neural networks. All the translation happens in the cloud. It just requires a camera on the device facing the signing person, and a connection to the internet.

2.1.3 Research Findings for Existing Literature

Many researches have been done on sign language Interpretation. Some of them are as follows

The work in this field began using the VPL data glove. It is the most commercially available glove. It was developed by Zimmerman[11] during the 1970s. It is based upon patented optical fiber sensors along the back of the fingers. Star-ner and Pentland developed a glove-environment system capable of recognizing 40 signs from the American Sign Language (ASL) with a rate of 5Hz. Then after many years in 1996, Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power Glove[1]. Then in 1999, another research was done by Hyeon-Kyu Lee and Jin H. Kim. Hyeon-Kyu Lee et al presented work on real-time hand-gesture recognition using HMM (Hidden Markov Model). Kjeldsen and Kendersi devised a technique for doing skintone segmentation in HSV space, based on the premise that skin tone in images occupies a connected volume in HSV space. They further developed a system that used a back-propagation neural network to recognize gestures from the segmented hand images[1]. Next in

2001,[7]Mohamed proposed a vision based recognizer to automatically classify Arabic sign language. A set of statistical moments for feature extraction and support vector machines was used for classification which provided an average recognition rate of 87%. Then in 2012, Reyadh [8] worked on an alphabet sign recognition system with a recognition rate with naked hand of 50%, red hand of 75%, black Hand of 65% and white hand of 80%. The sign alphabet images were mapped to histograms to uniquely represent the sign images. KNN algorithm which measures the distance between these sign histograms was used for classification. The process was simple and effective only on images and produced a very low recognition rate on video data. A lot of research was done in 2015. The method proposed [10] involves extracting the hand gestures from original color images. The segmented hand positions shape modulated using Chan-Vese (CV) active contour model and obtained 92.1% recognition rate. Kishore PVV, proposed 4-Camera model. The segmented hand gestures with extracted shapes created a feature matrix described by elliptical Fourier descriptors which were classified with back propagation algorithms trained using artificial neural networks. The normal recognition rate in the proposed 4 Camera model for sign language recognition is about 92.23%. Then, Etsuko Ueda and Yoshio Matsumoto presented a novel technique-a hand-pose estimation that can be used for vision-based human interfaces. In this method, the hand regions are extracted from multiple images obtained by a multi viewpoint camera system, and constructing the “voxel Model”[2]. Also then, Chan Wah Ng, Surendra Ranganath presented a hand gesture recognition system. They used image furrier descriptors as their prime feature and classified them with the help of RBF network. Their system’s overall performance was 90.9%. Claudia Nolker and Helge Ritter presented a hand gesture recognition model based on recognition of finger tips.In their approach they found full identification of all finger joint angles and based on that a 3D model of hand was prepared using neural networks. Basic sign language system is based on the 5 parameters and they are hand and head recognition, hand and head orientation, hand movement, shape of hand and location of hand and head. Among the 5 parameters above, the above system only met one parameter that also partially as they were not able to recognize the gestures formed using both hands. Hence they didn't get much success. Next research was also done in Hand Gesture Recognition Using PCA in [6]: In this paper author presented a scheme using a database driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching which can be effectively used for human robotics applications and similar other applications.. Initially, the hand region was segmented by applying a skin color model in YCbCr color space. In the next stage thresholding is applied to separate foreground and background. Finally, template based matching technique was developed using Principal Component Analysis (PCA) for recognition. Next, after 1 year in 2016, the dynamic time wrapping based level building (LB-DTW) algorithm was proposed [9] to solve sign sequence segmentation and sign recognition. This LB-DTW introduces two problems in recognition. One was under the bad relationship the recognition rate was very low and HMM was incorporated to improve recognition rate by calculating the similarity between sign model and testing sequence. On the other hand, the grammar constraint and sign length constraint were employed to improve recognition rate whereas in experiments with a KINECT data set of chinese sign language containing 100 sentences composed of 5 signs each, the proposed method showed superior recognition performance and lower computation compared to other existing techniques.

2.1.4 Problem Definition

There is a need for “Sign Language Identifier”, a web application for deaf and dumb people as sign language is used by the hearing impaired people for their communication. The existing systems don't have the proper dataset and dictionary for the normal conversation. Also there is a lot of difficulty in using these systems as it requires some specific coloured background or the electronic glove, which is expensive, difficult to carry and require a lot of care and maintenance. The efficiency and accuracy of the existing system is not sufficient that it can be used for the real time conversation. Therefore a system is required that will bridge the gap between normal people and hearing impaired people. It should be the helping aid for speech impaired to communicate with the rest of the world using sign language. The application should convert the sign language gestures to text and the speech to sign language and have sufficient dataset and a dictionary for the normal conversation. It should be user friendly.

2.1.5 Survey of Tools and Technologies Used

Table 3 shows Tools and Technologies used in various works carried out by researchers (as discussed in literature survey)

Table 3: Tools and their working

Tools	Working
VPL	Virtual Programming Language glove is a dataglove that is used to give input to the user. It has bundles of optical fibre with capture the orientation of the glove and give that a signal to the computer
Glove based Gesture Recognition	It has an electronic glove worn in one hand and all the gestures or movement made by the hand are captured based on the technology used to make that glove. It can be connected to a computer using a wire or wireless module.
Vision based Gesture Recognition	In Vision based Gesture Recognition, a camera is required to capture the movements of the person performing sign language. Then using different technology the captured video or photos are decoded to the meaning of the word performed by the sign language.

Table 4 provides an insight to technologies and working.

Table 4: Technologies and their working

Technology	Working
Optical Fibre Sensors	Fiber optic sensors work based on the principle that light from a laser or any superluminescent source is transmitted via an optical fiber,

	experiences changes in its parameters either in the optical fiber or fiber Bragg gratings and reaches a detector which measures these changes
Hidden Markov Model	It is a machine learning model that requires any dataset to train its model. It is used to predict future events. It has initial and transition probability and bunch of observation on which it works
Support Vector Machine	It is a supervised machine learning algorithm used for classification, regression and to find the outliers
Feature Extraction	Feature extraction is an image processing technique which is used to get the important features of the image, either this is extracted feature is used to recognize the image or it is used to modify that feature in the image.
KNN	K Nearest Neighbour is a supervised machine learning algorithm. It can be used for both classification and regression. In case of classification the k-nearest neighbour having maximum common class is the output and in case of regression the average of all the k-neighbours.
ANN and Back Propagation Algorithm	Artificial Neural Network is also a machine learning algorithm that uses nodes to represent the data, which helps the algorithm to get a deep understanding of the data. Output and input nodes are connected through hidden nodes which are used to simplify the data using weights and an evaluation function.

	Back Propagation Algorithm is used to adjunct the weights in order to improve the efficiency of the model.
Voxel Model	A voxel represents a single sample, or data point, on a regularly spaced, three-dimensional grid. This data point can consist of a single piece of data, such as an opacity, or multiple pieces of data, such as a color in addition to opacity.
RBF network	RBF network is an artificial neural network with an input layer, a hidden layer, and an output layer. The Hidden layer of RBF consists of hidden neurons, and the activation function of these neurons is a Gaussian function. Hidden layer generates a signal corresponding to an input vector in the input layer, and corresponding to this signal, the network generates a response.
PCA	Principal component analysis is a statistical approach used for reducing the number of variables in face recognition. In PCA, every image in the training set is represented as a linear combination of weighted eigenvectors called eigenfaces.
LB-DTW	Dynamic Time Warping (DTW) is an algorithm that aligns and compares two time series by calculating a matching error for them. Generally, the error measures obtained with DTW are more intuitive than other measures, e.g. Euclidean distance

Table 5: Tools and Technologies used in various works carried out by researchers

Author	Tool	Accuracy	Technology
Thomson Zimmerman (1970)	VPL(virtual programming language) glove Gesture recognition system	83.1%	Optical fiber sensors
Cristopher Lee and Yangsheng Xu (1996)	Glove based Gesture recognition system	90.9%	Hidden Markov Model
Hyeon-Kyu Lee and Jin H. Kim (1999)	Glove Based Real time hand gesture recognition system	90.9%	Hidden Markov Model
Mohandes, Mohamed (2001)	Vision Based Gesture recognition system	87%	Support Vector Machine, Feature extraction
Reyadh (2012)	Vision Based Alphabet Sign Recognition System	naked hand-50% red hand-75% black hand-65% white hand-85%	KNN
Kishore PVV, P. Rajesh Kumar(2015)	Vision Based 4-Camera Model	92.23%	Back propagation algorithm, ANN
Etsuko Ueda and Yoshio Matsumoto (2015)	Vision Based Hand Pose recognition using voxel model	90.1%	Voxel Model
Chan WahNG,	Vision Based Hand Gesture Recognition	90.9%	Image Furrier descriptor, RBF

Surinder Ranganath (2015)			network
Claudia Nolker and Helge Ritter (2015)	Vision Based Hand Gesture Recognition	80%	Neural Network
Mandeep Kaur Ahuja (2015)	Vision Based Hand Gesture Recognition	70%	PCA(Principal Component Analysis),
Yang, Wenwen, Tao, Jinxu (2016)	Vision based Sign Sequence Segmentation and Sign Recognition	83.7%	LB-DTW, Fast Hidden Markov Model

Table 6 displays the research papers, findings and citation.

Table 6: Research Paper and its findings and technology

S. No	Name and Roll No.	Paper Title	Tools/ Technology	Findings	Citation
1	Samiksha Kapoor 101703476	A Survey of Gesture Recognition Techniques Technical Report TCDCS	VLP data glove	Optical fibre when ran along the back of the hand, is cable of capturing the gesture made by the hand and its orientation	https://www.scss.tcd.ie/publications/tech-reports/reports.93/TCD-CS-93-11.pdf
2		Online, interactive	Sayre Glove, Dexterous	In online mode it updates the	https://ieeexplore.ieee.or

		learning of gestures for human/robot interfaces	Hand Master and Power Glove	model of each gesture	g/document/509165
3	Sanjana Vashisht 101703482	Automatic translation of arabic sign to arabic text (ATASAT) system	Vision based recognizer	SVM model can give 87% accuracy in vision based detection system	https://www.airccj.org/CSCP/vol6/csit65211.pdf
4		Development of a new arabic sign language recognition using k-nearest neighbor algorithm	KNN algorithm	Gestures made by Colored hands can be recognized better than naked hands. Signed language images can be mapped to histogram and then KNN classifier can help to recognize the sign	http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.460.1595
5	Sargun preet Kaur 101703488	Segment, Track, Extract, Recognize	Chan-Vese (CV) active contour model	Hand could be segmented to get the correct gesture	http://citeseerx.ist.psu.edu/viewdoc/download?do

		and Convert Sign Language Videos to Voice/Text			=10.1.1.685.1887&rep=rep1&type=pdf#page=45
6		Hand Pose Estimation for Vision Based Human Interface	Voxel Model	Hand regions can be captured by the multiview point images clicked by the camera	https://issuu.com/ijrat/docs/paper_id-35201536
7	Simran 101703545	Hand Gesture Recognition Using PCA	PCA and template matching based on skin color	It is a database driven hand recognition system which recognize the gesture based in the skin color and then separate foreground and background bases on PCA	https://pdfs.semanticscholar.org/43b7/e88adbcdfcf5b25d337821c94a0f0ee525b3.pdf
8		Continuous sign language recognition using level building based on fast Hidden	LB-DTW algorithm, HMM	LB-DTW algorithm can help us to solve the sign sequence segmentation and sign recognition. If	https://www.sciencedirect.com/science/article/abs/pii/S0167865516300344

		Markov Model		grammar constraints and sign length constraints are applied, the efficiency of the recognition system increases.	

2.1.6 Summary

Concluding the work of all the researchers, the methods of making sign language identifiers can be broadly divided into a glove based approach and a vision based approach irrespective of the sign language being used. SLI is built based on the later method. Although accuracy achieved is quite high using the glove based approach but the inconvenience, cost, more prone to damage, uncertainty, tricky and risky to use dominate its pros. Further dissecting the vision based work, a technology is needed to identify and reliably distinguish the signs. Mainly researchers have used ANN, KNN, HMM, SVM etc which all are the machine learning methods to classify the objects. SLI also has a machine learning method for recognition of the signs but a more sophisticated and rigid in terms of identifying and classifying the image, that is CNN (convolution neural network). All the methods as mentioned above used by the researchers are good enough to identify the image but require more preprocessing, additional techniques and more micro details in order to make it work with sufficient accuracy. Whereas CNN is the evolved version of the ANN specifically for the image related classification that reduces the work, boasts accuracy and makes it flexible for any addition of the property without distorting the entire neural network. Now talking in terms of the dataset required for the testing and training. Researchers have used colored images, dictionaries according to the technique they have used. In SLI for testing and training, the dataset of morphologically processed images is used. The benefit of doing so is the speed, precision and increased accuracy which gives the result close to glove based approach.

In addition to the above similarities in one way or another, SLI has additional features of converting the english text to sign language for the two way communication. Adding this feature has made it a complete project for the deaf and dumb people to have communication reliably and without any difficulty, which is the main motive behind making this Sign Language Identifier. For more convenience we have it in the form of a website to make it portable as nowadays everyone has a mobile phone and its interface is user friendly.

2.2 Standards

1. ISL (Indian Sign Language) standard: Used to create Indian Sign Language Dictionary.
2. Web 2.0 : Internet technology to create online applications that behave dynamically.
3. Camera measurement standards :
   3.1 Resolution Measurements - ISO 12233
   3.2 Noise Measurements – ISO 15739
   3.3 Sensitivity Measurements – ISO 12232
   3.4 OECF (tone reproduction) Measurements – ISO 14524
   3.5 Chromatic Displacement Measurements – ISO 19084
   3.6 Color Characterization Test Procedures – ISO 17321
   3.7 Image Stabilization Measurements – ISO 20954
   3.8 Low Light Measurements – ISO 19093
   3.9 Shooting Time Lag Measurements – ISO 15781
   3.10 Shading Measurements – ISO 17957
4. Video quality: atleast 360 pixels.

2.3 Software Requirement Specification

Clear requirements help to get a better vision for the project and helps developers to make code in a certain way. SRS for SLI are given below.

2.3.1 Introduction

This document lays out a project plan for the development of a “Sign Language Interpreter” system.

The intended readers of this document are current and future developers working on “Sign Language Interpreter” and the sponsors of the project. The plan will include, but is not restricted to, a summary of the system functionality, the scope of the project from the perspective of the “Sign Language Identifier” team, scheduling and delivery estimates, project risks and how those risks will be mitigated, the process by which the project, and metrics and measurements will be developed will be recorded throughout the project.

2.3.1.1 Purpose

Sign Language is the most natural and expressive way for hearing impaired people. Unfortunately, most of the common people are not familiar with sign language. This leads to a large communication gap between deaf and dumb and the rest. The SLI project aims at eliminating this communication gap. The main purpose of this project is to build a web based app that can convert ISL into text; and speech to ISL.

2.3.1.2 Intended Audience and Reading Suggestions

Intended Audience for the “SLI” project includes

1. Deaf and dumb (signers) who will perform gestures (using Indian Sign Language) that will be detected by the camera.
2. People who are not much familiar with the Indian Sign Language. These people will be able to understand the sign language from the text generated by the web app and will be able to communicate back using speech to the ISL interface of the web app.

Reading Suggestions

About SLI

SLI (Sign Language Identifier) is a web app designed to convert the Indian Sign Language to text and speech to ISL. The main objective of the app is to reduce the communication gap between deaf and dumb and the rest of the world. You simply need to perform signs in front of the camera, and SLI will provide you output w.r.t. corresponding sign.

How to use SLI

SLI will record the gestures made by the signer and will convert it to text. SLI also has the feature to convert live voice (speech) to Indian sign language (ISL).

1. In order to use the app, it is necessary for every user to login. If any user is using the app for the first time, then he/she must sign up first and then login.
2. The user has to press the start button before recording and stop button when done recording.
3. The text corresponding to the signs made by the signer will be displayed on the screen.
4. The web app can also convert speech to ISL. The user interface for the same is very easy to use. The user has to record the speech, he/she wants to convert to ISL.

For more accurate results following measures should be taken

1. The signer must be at a fixed distance from the webcam.

2. The signer must have a steady pace while making signs so that the camera could detect the signs properly.

3. The signs should be made in a controlled environment keeping a fixed background.

4. The user must keep the camera at a fixed angle so that it can record the signs properly and accurately.

2.3.1.3 Project Scope

To bridge the gap between deaf and dumb and the rest , the main purpose of this project is to design a web app that can convert sign language into text; and speech to ISL. Anyone willing to communicate through sign language will be able to login through the app. Hand gestures of the signer will be captured by the camera of the device which will be further converted to text . The quality of the webcam will be more than 360 pixels. Signer will be at a fixed distance from the camera. The dataset for training the model will contain gestures from Indian Sign Language only.

2.3.2 Overall Description

Overall description includes product perspective, product functions, user characteristics, general constraints, assumptions, release, features and user interface, which are discussed in detail in sections as under.

2.3.2.1 Product Perspective

The project is a web application implementing a client server model. The web app requires the webcam and speaker for its working. Figure 2 shows the functionality area and supported systems of the web app (after login).

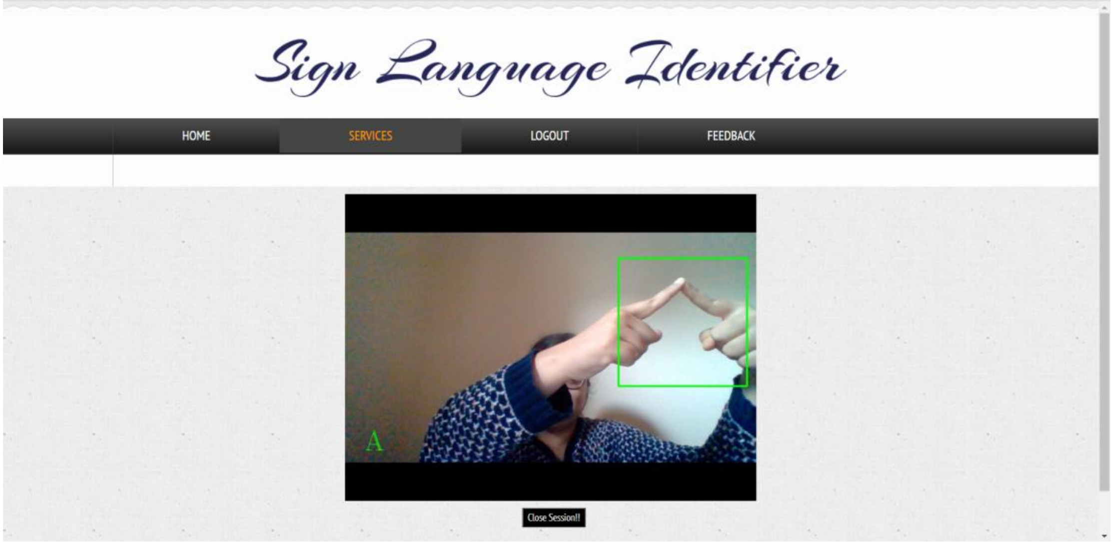


Figure 2: User interface of the web app

First the user will enter their username and password, then sql server will check for its validity from the "user" database. Authenticated users will be allowed to access the web app otherwise the user has to make a user id before using the product. The web app will ask for the permission of the webcam. Figure 3 shows the Layered architecture of the product. The user will press the start button of the webcam for the video input and stop button once they're done. This recorded video is then converted into an image sequence, which is preprocessed and through template matching from “the sign to text” database, corresponding text is saved. On pressing the “convert sign language to “text” button, the saved text will appear in the text box. In order to convert it to speech to ISL, a person has to press on the live voice button and start saying the text what he/she wants to convert to ISL.

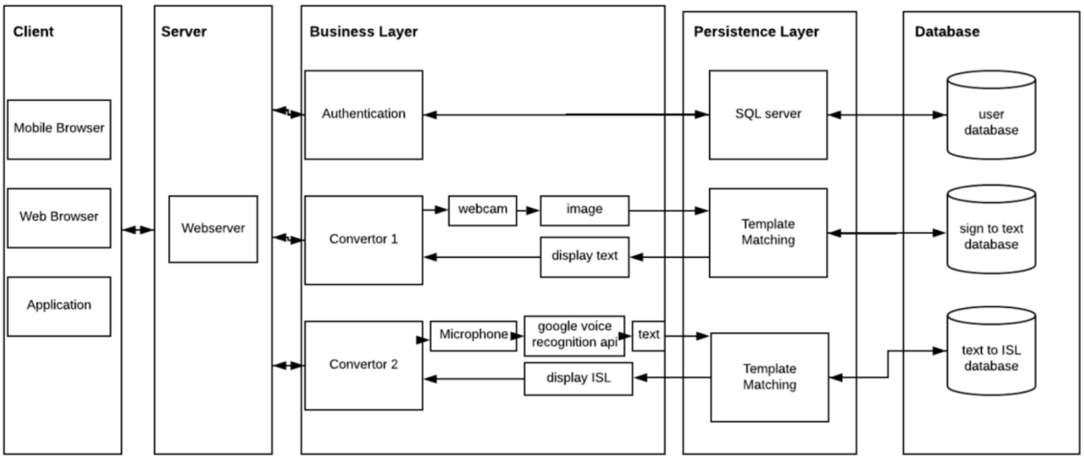


Figure 3: Layered Architecture of the product

2.3.2.1.1 Product Functions

1. The web app should be able to authenticate the user by matching username and password by the values kept in the database.
2. The web app should be able to make video recordings of the person doing sign language on pressing start button and save the recording on pressing stop button.
3. The web app should be able to convert the recorded video into a sequence of images which will be further preprocessed before checking it for the related text in the database.
4. The web app should be able to display the corresponding text on the screen on pressing the conversion button,
5. The web app should be able to convert speech to ISL on pressing the live voicing.

2.3.2.1.2 User characteristics

The web app is a software, designed for the deaf and dumb people to help them in having conversion with the world without any barrier. It converts the sign language to text; and speech to ISL. No prior expertise is required to use this software except for knowledge of indian sign language.

2.3.2.1.3 General Constraints, Assumptions and Dependencies

The following are the general constraints, assumption and dependencies applied in building the sign language identifier web app:

1. The user must have a user id to login to the web app.
2. The device must be connected to a webcam and speaker.
3. The quality of the webcam should be more than 360 pixels.
4. The conversion between the languages should be less than 2 seconds.

2.3.2.2 Product Features

Product features provide an insight to how a product must work and what objectives it must satisfy. How the releases are handled and various other aspects are stated below.

2.3.2.2.1 Objectives

The web app, sign language identifiers to fill the gap in communication between the hearing impaired and other people. Table 7 shows visions, goals, initiatives and customers for SLI.

Table 7: Objectives for SLI

Visions	“Sign Language Identifier” to be used all over India and by the people who use indian sign language
Goals	To make the sign language identifier reachable to all in need
Initiatives	All the initiative with respect to Indian sign language
Customers	Deaf and Dumb people

2.3.2.2.2 Release

Table 8 provides an insight to release details.

Table 8: Release Details

Release	Sign Language Identifier-Release 1
Date	1 December, 2020
Features	• Recording signs of the language • Converting it to text • Speech to Sign Language
Milestones	1. A-Z ISL alphabet to text 2. Speech to ISL
Dependencies	• Keras • Tensorflow • CV2 • Numpy • OS • Sklearn • Speech_recognition • Matplotlib.pyplot • Flask • Pillow • HTML and CSS • Itertools • String • sklearn

2.3.2.2.3 Features

Table 9 provides an insight to the features of SLI, it's description, purpose, user value and assumption.

Table 9: Features

Features	• Recording signs of the language • Converting it to text • Speech to Sign Language
Description	• The user will perform a sign in front of the webcam. • Corresponding text will appear. • For speech to ISL, the user simply needs to click "Live Voice" and they say something. • Corresponding gif or plots of alphabets of which word is composed of will be displayed.
Purpose	To make the interface user friendly and help the user get output with convenience so that it can have the real time conversation with hearing person
User problem	Users were not able to have real time conversation due to difficulty in recording the video with accuracy, maintaining certain distance the webcam and having particular color background
User Value	Now user can record the video without worrying about the distance and color of the background
Assumption	User must have webcam and the recorded frame should only contain the user

2.3.2.2.4 User Interface

1. The user uses various buttons of the application in the system to operate it.

   a. START BUTTON: This would start the application and hence the user (speech-impaired person) would give input to the application through gestures.

   b. STOP BUTTON: This would make the application stop and hence the user can close the application.

   c. Live Voicing: This would start recording the voice of the speaker to convert it to ISL.

   d.Done All: This would exit the speech to ISL interface.

2. Input for the application would be provided by any camera and speaker attached to the system to recognize gestures.

Output

1. Output would be provided for various gestures in a separate window in the form of text in case of ISL to text.

2. Output will be the gif of sign language for speech to ISL .

2.3.3 External Interface Requirements

External interface requirements specify hardware, software, or database elements with which a system or component must interface.

2.3.3.1 User Interfaces

User Interface (UI) is an important part of any software system. A software is widely accepted if it is easy to operate and quick in response and provides a consistent user interface. The user interface must be completely secure. Security requirements include the need to protect authorization information from unauthorized access, the maintenance of customer confidentiality. Along with that, The user interface should be efficient in both speed and use. User interface requirements are briefly mentioned below

- Responsive
- Easy interface
- Consistent UI elements

- Strategic use of color and texture
- Feedback mechanism
- Purposeful Layout
- Efficient

2.3.3.2 Hardware Interfaces

This project is a software based project, only working camera is required at user end, hence no other hardware interfaces are required for SLI.

2.3.3.3 Software Interfaces

The software requirements are description of features and functionalities of the target system. Requirements convey the expectations of users from the software product. Software requirements are categorized into functional and non-functional requirements.

Functional Requirements

- Authentication
    a. New users will have to create an id using attributes: username, email address and password.
    b. The user will use username and password to sign in to the web application.
- Capturing Video
    a. Web app will start capturing video of a person doing sign language on pressing the webcam button.
    b. Web app will stop the webcam and store the captured video on the repressing of the same button.
    c. Web app will alert the user if distance between user and webcam is either too much or too less.
- Text generation
    a. Web app will convert the sign language video to text.
    b. Web app will give an alert message if text related to sign language is not found.

- Display Generated Text
    a. Web app will display the text generated corresponding to the video on the screen.
    b. Web app will not display anything but give an alert message if text is not generated.
    c. The generated text will be displayed in natural english language.
- Record Speech
    a. Web app will record the voice of the person pressing the Live voicing button.
    b. Web app will close the recording interface on speaking “good bye”.
- Display GIF
    A. Web app opens a separate window displaying the output.
    B. Web app will show the gif of the person doing ISL matching the recorded speech.

2.3.4 Other Non-Functional Requirements

To judge the operation of the system (SLI), rather than specific behavior, there are certain non functional requirements given below.

2.3.4.1 Performance Requirements

The web app basically has a sign language to text; and speech to ISL conversion feature. The performance measuring variable will be the time taken and quality of the video captured by the webcam on pressing the webcam button and displaying the desirable/ corresponding output. The requirement is to have good quality video with minimum video quality being 360 pixels with plain background and to generate output in real time, with time delay less than 2 sec.

Another variable to define performance is the correct conversion of sign language to text; and speech to ISL with proper context, which depends on the dataset's efficiency, Machine Learning Model so employed, Image Processing and NLP techniques, so employed.

2.3.4.2 Safety Requirements

Safety of the web application is highly needed, since the sign language may have some sensitive information. The web app must take care of safety: risk encryption, handling vulnerabilities.

The user login and sign up just requires email id, username and password, hence no personal details are asked for, ensuring no pose to security threat. Passwords could be stored securely, by firstly modifying password with a unique way and then applying hashing method to make it more secure. Hashing and salting passwords could be used. For more sensitive information, asymmetric encryption could be employed. But since, it is planned to keep it simple in the initial stage, so only, limited attempts are permitted for wrong username or password. More features as proposed could be added later on.

2.3.4.3 Security Requirements

As the web app requires internet or network connectivity so the risk to users' privacy and threat to their personal information are high. So in order to overcome such problems

SLI will use a firewall authentication system for user login and WAF (web application firewall) to protect the database.

2.4 Cost Analysis

Since there is no hardware involved, only the user must have a device with a working camera and microphone, so major cost goes into database, HA Storage, Network and Web Hosting. Table 10 shows insight of pricing in detail.

Table 10: Cost Analysis as per Budget Requirement

Service	**Budget**
Database	• $0.204 per GB/month for SSD storage capacity • $0.096 per GB/month for backups (used)
HA Storage (High Availability Storage)	• $0.408 per GB/month for SSD storage capacity • $0.096 per GB/month for backups (used)
Network	• Ingress to Cloud SQL: Free • Note: Egress charges may apply on the source. For example, egress from Compute Engine is charged at the external IP addresses rate. • IPv4 addresses: $0.012 per hour while idle.
Web Hosting	• $17 per month (yearly)

Note: These prices are for Google Cloud. The above cost analysis has been done with reference to https://cloud.google.com/sql/pricing#sql-server.

Egress from Cloud SQL: See Network Egress Pricing in the above link.

2.5 Risk Analysis

There is already interest among researchers in various fields in applying different methods to sign language applications. In particular, some applications / methods use gesture recognition touch on sign language recognition as a domain. These methods work on algorithms to detect fingers, hands, and human gestures. However, by framing sign language recognition as an application area, they risk misrepresenting sign language recognition as a gesture recognition problem, ignoring the complexity of sign language as well as the broader context within which such systems must function. In this work, the sign language will be translated into text; and speech to ISL without changing the context with the help of natural language processing techniques.

The only drawback with this approach is the risk of losing information. The gestures of different letters are very similar and differ only in the location of the thumb and contours of hand. If the silhouette of the hand was taken, it would not be possible to determine which letter was being presented. Contours are similar to silhouettes but focus on extracting edges on an image. Some methods derive the edges from the silhouette, making them equivalent; However, edge detection techniques can also be applied to images.

METHODOLOGY ADOPTED

3.1 Investigative Techniques

Table 11 displays investigative techniques for SLI.

Table 11: Investigative Techniques

S. No.	Investigative Projects Techniques	Investigative Techniques Description	Investigative Projects Examples
1.	Descriptive	There are projects based on new system modules, concepts and algorithms which help deaf and mute people live life normally.	The easy-to-use innovative digital interpreter dubbed as "Google translator for the deaf and mute" works by placing a smartphone in front of the user while the app translates gestures or sign language into text and speech. The app, called GnoSys, uses neural networks and computer vision to recognise the video of sign language speakers, and then smart algorithms translate it into speech.
2.	Comparative	Many different solutions have been developed for converting sign language to text. Some solutions used a hardware based approach and some used	Re-voice glove (An electronic device that converts gestures made by glove into text and speech)

		a software based approach. Hardware based solutions are expensive and not affordable by common people whereas software based approaches are cheap and affordable.	MotionSavvy (A tablet app that understands sign language)
3.	Experimental	The main focus of the project is to design a web app that converts sign language to text and speech. Currently there are many solutions for the same but most of these solutions are developed for American Sign Language. The goal is to develop a solution for Indian SIgn Language. After some experimental investigation, it is found out that the most reliable solutions are AI based solutions which apply deep learning techniques to achieve the desired results.	Google Translator (An AI app that converts sign language to text and speech) AI interpreter (It translates user’s sign in real time)

3.2 Proposed Solution

ISL to text

For the sign language recognition system, a vision based approach will be used. The system will contain four modules/derived algorithms: preprocessing and segmentation, feature extraction, sign recognition and sign to text. Indian Sign Language is the area of concern. The signs will be captured using a webcam where the signer must be directly in front of the webcam. Then the output will be in the form of text. Following is methodology for sign language to text.

The main steps for converting sign language to text are given below.

3.2.1 Preparing the dataset

A dataset of 26 English language alphabets in Indian Sign Language will be created. The images in the dataset are to be segmented. The dataset will then be divided for training and testing. A dictionary for ISL will be created so that idea extends to sentences as well.

3.2.2 Image Acquisition

The gestures of sign language will be captured through the webcam for testing. The gesture should be made in the rectangular region defined in the web app. This will be in the form of a video, from which suitable frames will be extracted.

3.2.3 Image Preprocessing

The BGR image is now acquired, followed by segmentation and morphological filtering methods. These operations are done on the acquired image.

c. **Segmentation:** Here the image is converted into small segments in order to get more accurate image attributes. Here the segmentation of hands is carried out to separate objects and the background. To make segmentation more robust, so that the feature extraction and sign recognition of the image is accurate, the BGR image is converted to HSV, then fed into inRange() function which performs segmentation.

d. **Morphological filtering:** The image attributes extracted from the segmentations consist of noise, therefore morphological filtering removes the noise and gives a smooth contour. The output of these filterings are image attributes which are useful for feature extraction and sign recognition. Dilation and Erosion are used for morphological filtering.

3.2.4 Feature Extraction

The feature is nothing but the thresholded and morphological filtered image which came as output of thresholding and morphological filtering operation.

3.2.5 Template Matching and Sign Recognition

The image so generated is matched with labels with help of predict_class() method of CNN model. Say the sign for ‘A’ is performed, then the class is ‘A’ which is the folder's name in the dataset.

3.2.6 Text Generation

The predicted text as predicted by predict_class() method, is displayed on screen.

Speech to ISL

Firstly a dataset of alphabets and gifs is created. Post that, the coding is done for the conversion of speech to ISL. Firstly, the microphone is activated and input voice is thresholded. Then, the voice recognition takes place using Google API. The spoken word is processed to text and then text is searched with labelled folders of the dataset. If word exists, then corresponding GIF is displayed, else the plot of alphabets of which letter is composed of is displayed on the screen.

Figure 4 shows a block diagram for ISL to text conversion as well as speech to ISL.

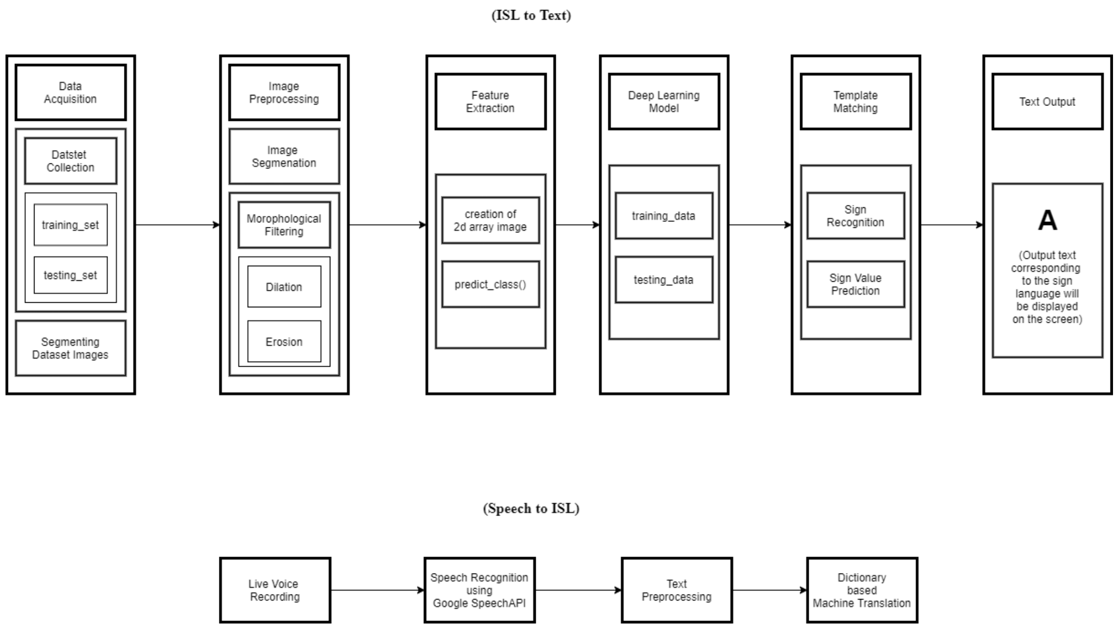


Figure 4: Block Diagram for ISL to text; speech to ISL

3.3 Work Breakdown Structure

“Sign Language Identifier” has various modules with major modules being Data Acquisition, Image preprocessing, Feature Extraction, Deep Learning Model, Template Matching and Text Output. They further have submodules as shown in Figure 5 given below. Along with that, for speech to ISL, the modules are shown in the same figure i.e. Figure 5.

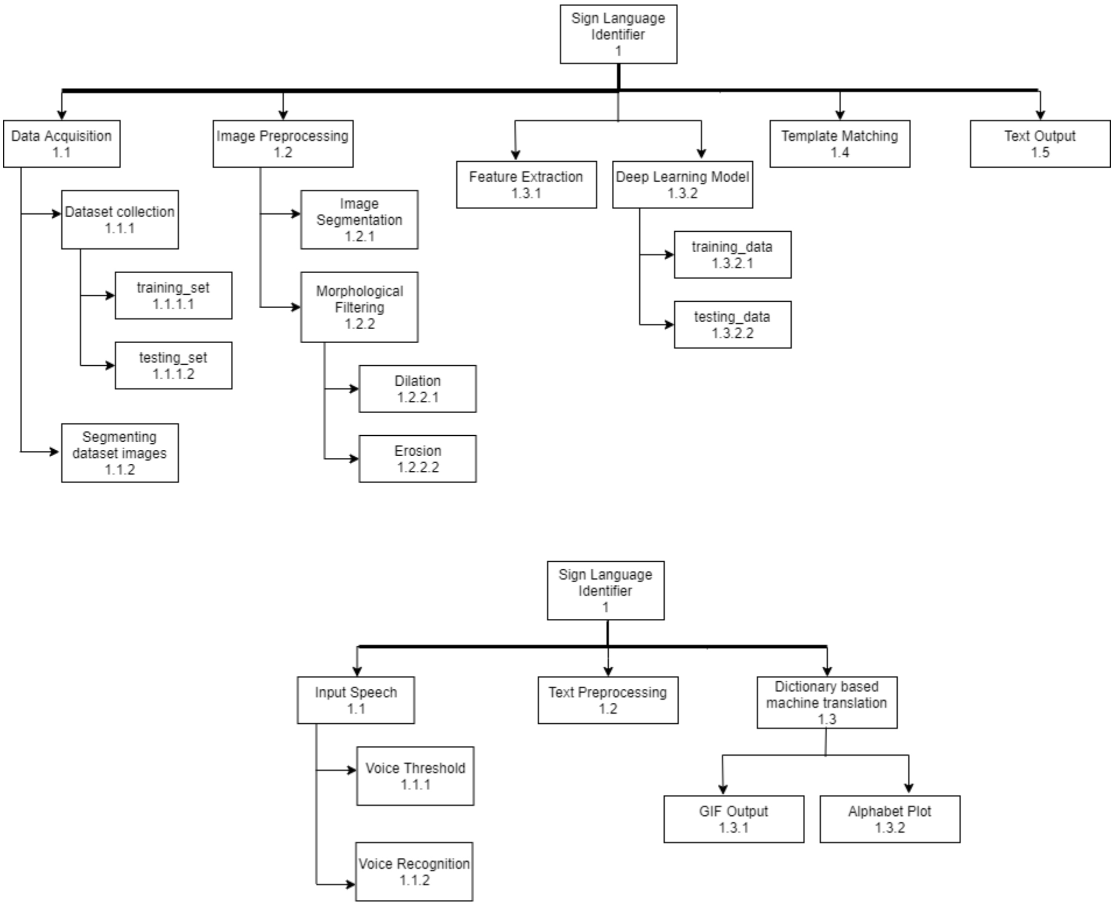


Figure 5: Work Breakdown Structure

Following Gantt Chart (Figure 6) shows tentative project timeline and %age of tasks completed along with start date and end date which is shown in detail in Gantt Chart (Table 12) along with Task Description and Status (in terms of progress).

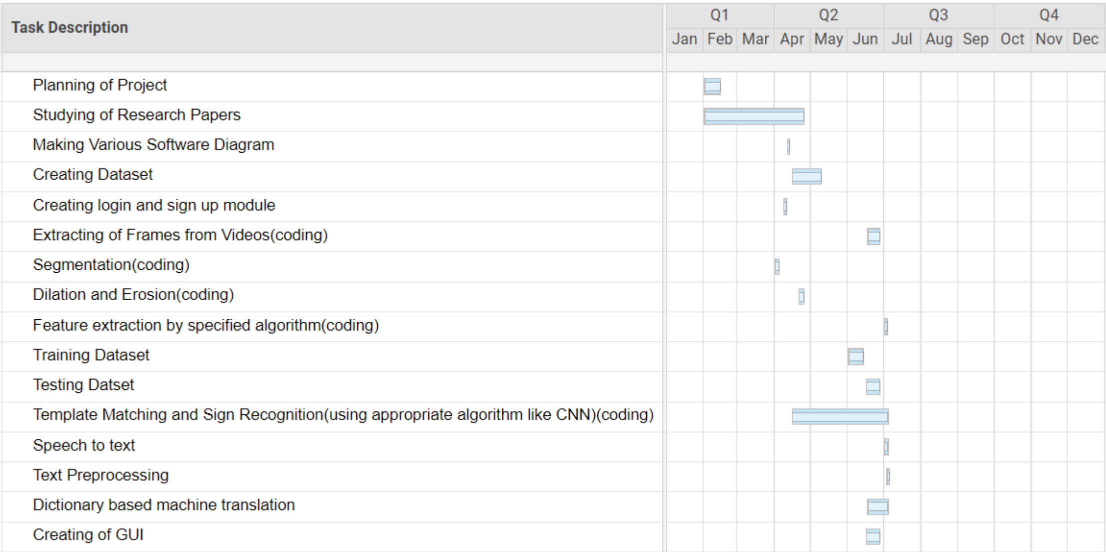


Figure 6: Gantt Chart

Gantt Chart Table

Table 12 below shows the gantt chart table in detail.

Table 12: Gantt Chart

Task Description	Duration	Start	Finish	% Complete	Status
Planning of Project	3d	02-01-20	02-14-20	100%	Completed
Studying of Research Papers	13d	02-01-20	04-25-20	100%	Completed
Making Various Software Diagram	1d	04-11-20	04-11-20	100%	Completed
Creating Dataset	5d	04-15-20	05-09-20	50%	Completed
Creating login and sign up module	1d	04-08-20	04-08-20	100%	Completed
Extracting of Frames from Videos(coding)	3d	06-17-20	06-27-20	100%	Completed
Segmentation(coding)	2d	04-01-20	04-04-20	100%	Completed
Dilation and Erosion(coding)	2d	04-21-20	04-25-20	100%	Completed
Feature extraction by specified algorithm(coding)	1d	07-01-20	07-01-20	100%	Completed

Training Dataset	3d	06-01-20	06-13-20	100%	Completed
Testing Dataset	3d	06-16-20	06-27-20	100%	Completed
Template Matching and Sign Recognition(using appropriate algorithm like CNN) (coding)	13d	04-15-20	07-04-20	100%	Completed
Speech to text	2d	07-01-20	07-04-20	100%	Completed
Text Preprocessing	2d	07-03-20	07-04-20	100%	Completed
Dictionary based machine translation	4d	06-17-20	07-04-20	100%	Completed
Creating of GUI	3d	06-16-20	06-27-20	100%	Completed

3.4 Tools and Technology

Following is the tentative list of tools and technology to be used while making the Sign Language Identifier (SLI). Some of the tools could be modified based on outcomes. As of now, given below is the list.

- Numpy
  - Working with arrays
- SpeechRecognition
  - Speech Recognition into Python application
- matplotlib.pyplot
  - Plotting graphs
- OpenCV
  - CV (Computer Vision) library for images and videos
- OS
  - For manipulating files
- Pillow
  - Image manipulation
- String
  - For punctuations
- Itertools
  - For count
- PyAudio
  - Audio I/O library
- Anaconda Navigator and PyCharm
  - Python offline coding
- Google Colab
  - Cloud coding
- Neural Networks (Keras, Tensorflow, CNN)
  - For classification and training

DESIGN SPECIFICATIONS

4.1 System Architecture

Figure 7 shows the layered architecture of the product, constituting client, server, business layer, persistence layer, database.

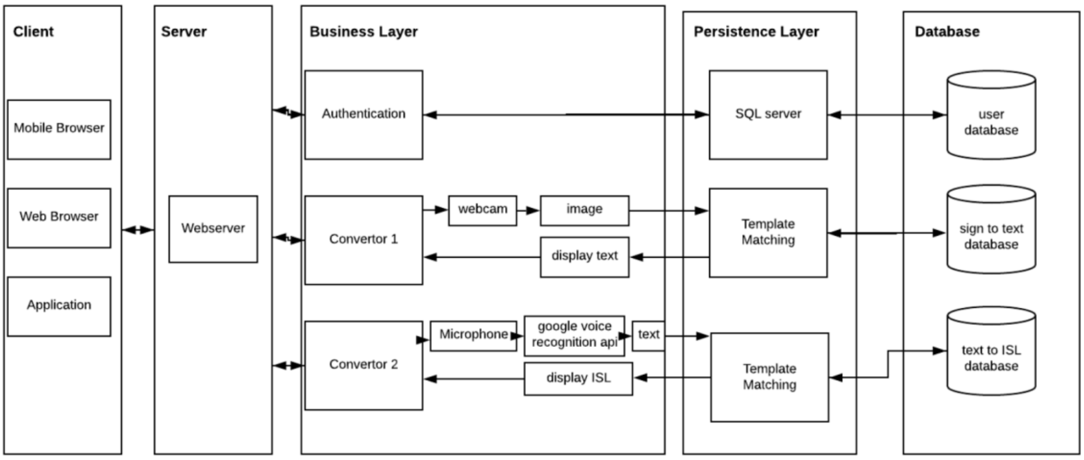


Figure 7: Layered Architecture of the product

“Sign Language Identifier” has various modules with major modules being Data Acquisition, Image preprocessing, Feature Extraction, Deep Learning Model, Template Matching and Text Output. They further have submodules as shown in Figure 10 given below. Along with that, for speech to ISL, the modules are shown in the same figure i.e. Figure 8.

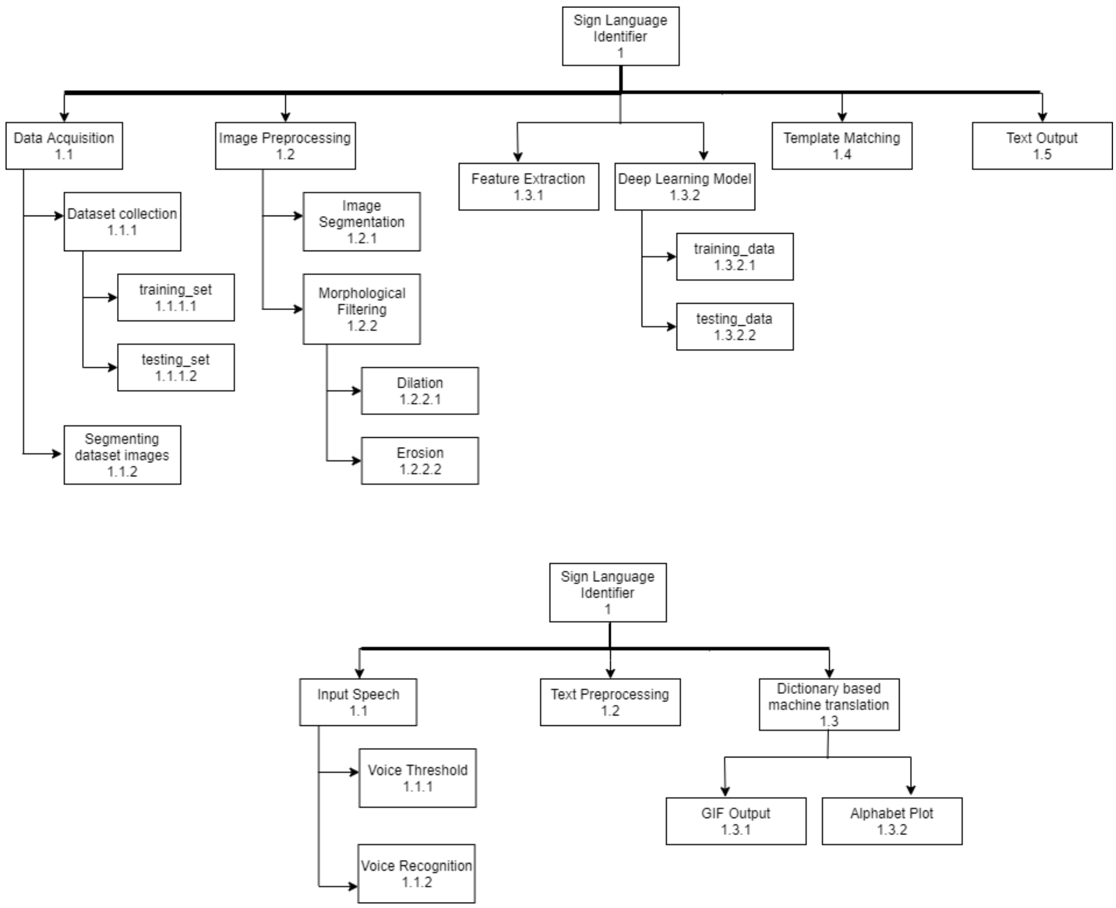


Figure 8: Work Breakdown Structure

4.2 Design Level Diagrams

Design Level Diagrams include Component Design, Data Design, Interface Design given in section 4.2.1, 4.2.2, 4.2.3 respectively.

4.2.1 Component Design

Component design is the definition and design of components after the architectural design phase. It defines the data structures, algorithms, interface characteristics and communication mechanisms allocated to each component for the system development.

Figures 9,10 serve the following purposes for the SLI project

1. It visualizes the components of the SLI system.
2. It describes the organization and relationships of the components.
3. It represents the dependencies of the components on each other.

Figure 9 shows the visualization of components for the conversion of ISL to text and figure 10 shows the same for conversion of speech to ISL.

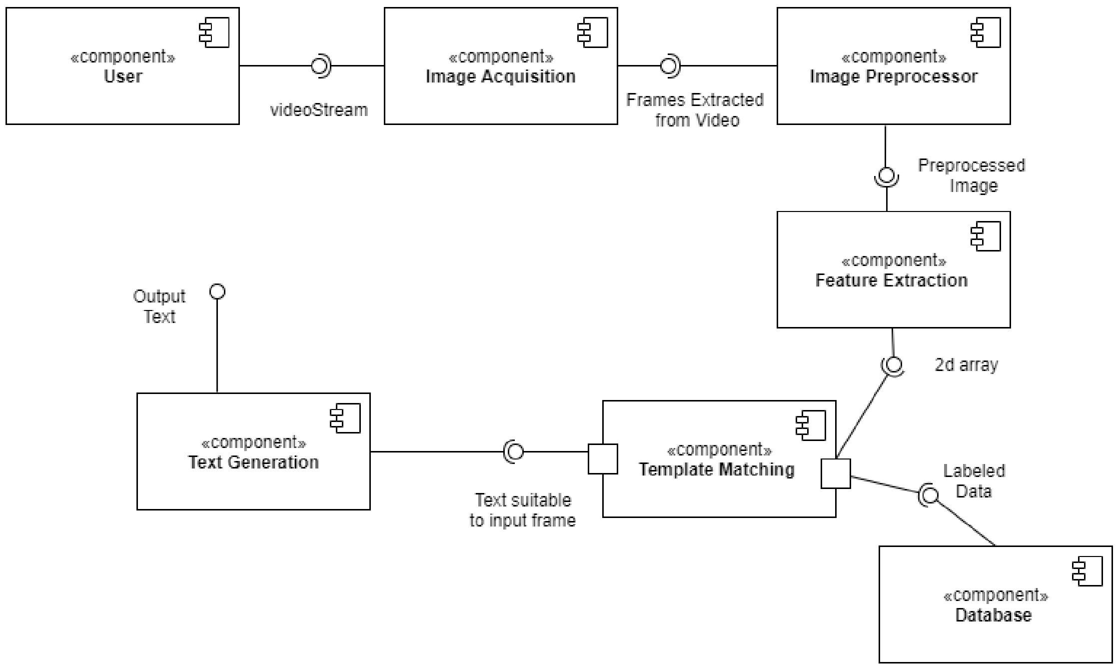


Figure 9 : Component Design for ISL to English text

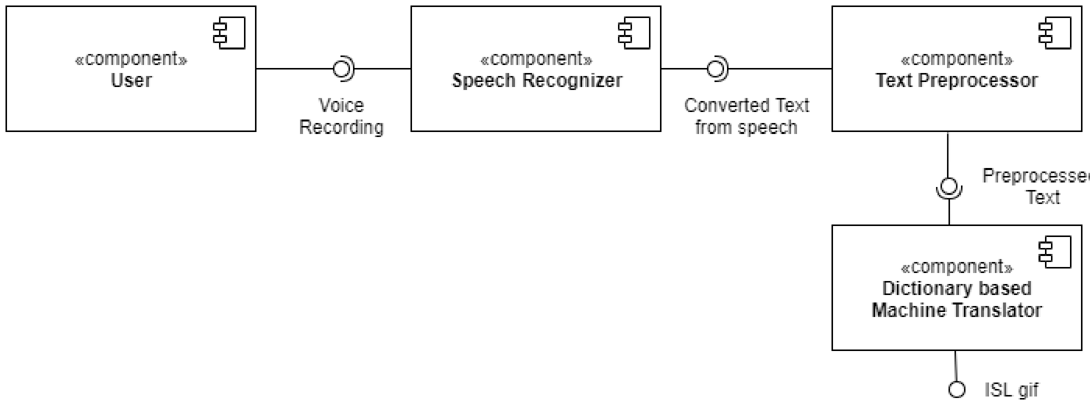


Figure 10: Component Design for Speech to ISL

4.2.2 Data Design

In data design, there will be three databases for storing user information and image to text information.

The first database will have three tables as follow:

1. User table: It will have three attributes namely username (which will also serve as primary key), password and email_id. The datatype of all the three attributes is string. Basically it will store the information of the user, provided during sign up. This information will be used for verification during log in action.
2. Saved video table: It will have two attributes namely File_no (which will also act as foreign key) and MP4_file. The datatype of the above-mentioned attributes are number and BLOG respectively. This table will save the recorded videos provided by the user doing the indian sign language during the initial stage of the sign language conversion activity.
3. Image sequence table: This table will have two attributes namely Seq_no (which will also act as foreign key) and images. Their respective data types are number and BLOG. This table will contain the sequence of images created during video preprocessing.

The second database will have one table.

1. Image to text table: This table will have two attributes namely Image (will also act as primary key) and Text. The datatype of the above-mentioned attributes are BLOG and string. It will save the english translation of the indian sign language words. This table will be prepared beforehand and it will be used in template matching.

The third database will have one table.

1. Speech to ISL table: This table will have three attributes namely Label, GIF_file and JPG_file. The datatype of the attributes are string, BLOG and BLOG respectively. It will have the gif and images of ISL corresponding to the text for the conversation to Indian Sign Language.

There are basically four relationships established between different tables. They are as follows.

1. Provides: The relationship between the user table and saved image table is provided as the user will capture the image and it will be stored corresponding to the respective user. The relation between these tables is weak as the saved image table is a weak entity. The saved image table will use its foreign key and the primary key of the user table to uniquely identify its tuple.
2. Preprocessing: The relation between saved image table and image sequence table is preprocessing as the image from the saved image table will be preprocessed and converted into image sequence of segmented image and it is also a weak relation as both the tables are weak entities. The tables will use their foreign key and primary key of the user table to uniquely identify its tuple.
3. Conversion: The relation between image sequence and image to text table is conversion as the image taken from the image sequence will serve as primary key and will give the corresponding text, which is basically the english translation of the image containing sign language words.
4. Record speech: This is the relation between user table and speech recognition table. The speech input by the user is converted to ISL using this relation. The speech is converted into text using google api and then using this table we convert it to ISL gif or ISL images.

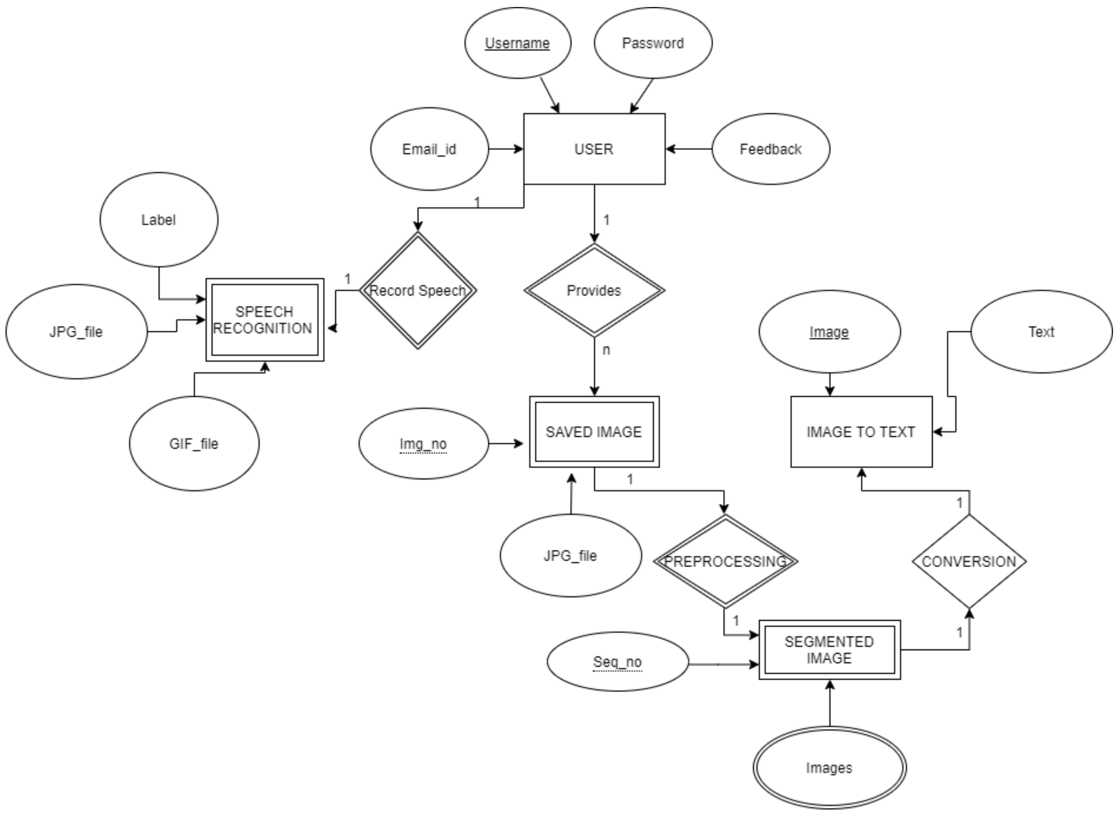


Figure 11: ER Diagram

Figure 11 represents the ER diagram depicting the relationships between different entities.

4.3 Analysis Diagrams

Analysis Model is used to describe the model, information and structure of the system and they are converted to architecture, interface and component level design in the 'design modeling'. It is a technical representation of the system. It acts as a link between system description and design model. In Analysis Modelling, information, behavior and functions of the system are defined and translated into the architecture, component and interface level design in the design modeling.

Objectives of Analysis Modelling are

1. It must establish a way of creation of software design.
2. It must describe requirements of the customer.
3. It must define a set of requirements which can be validated, once the software is built.

The model can be analysed through various diagrams which are shown below.

4.3.1 Use Case Diagram

A use case diagram is the primary form of system/software requirements for a new software program underdeveloped. Use cases specify the expected behavior (what), and not the exact method of making it happen (how).

As shown in Figure 12, there are two users in the SLI project. One is the deaf and dumb person and the other user is the system/admin. The user (specially abled person) has to login to the app before using it. Then the webcam will be switched on and the person has to perform some actions according to the Indian Sign Language that will be converted to Text Output through a series of steps which are performed in sequence as shown below. At last, there is optional feedback.

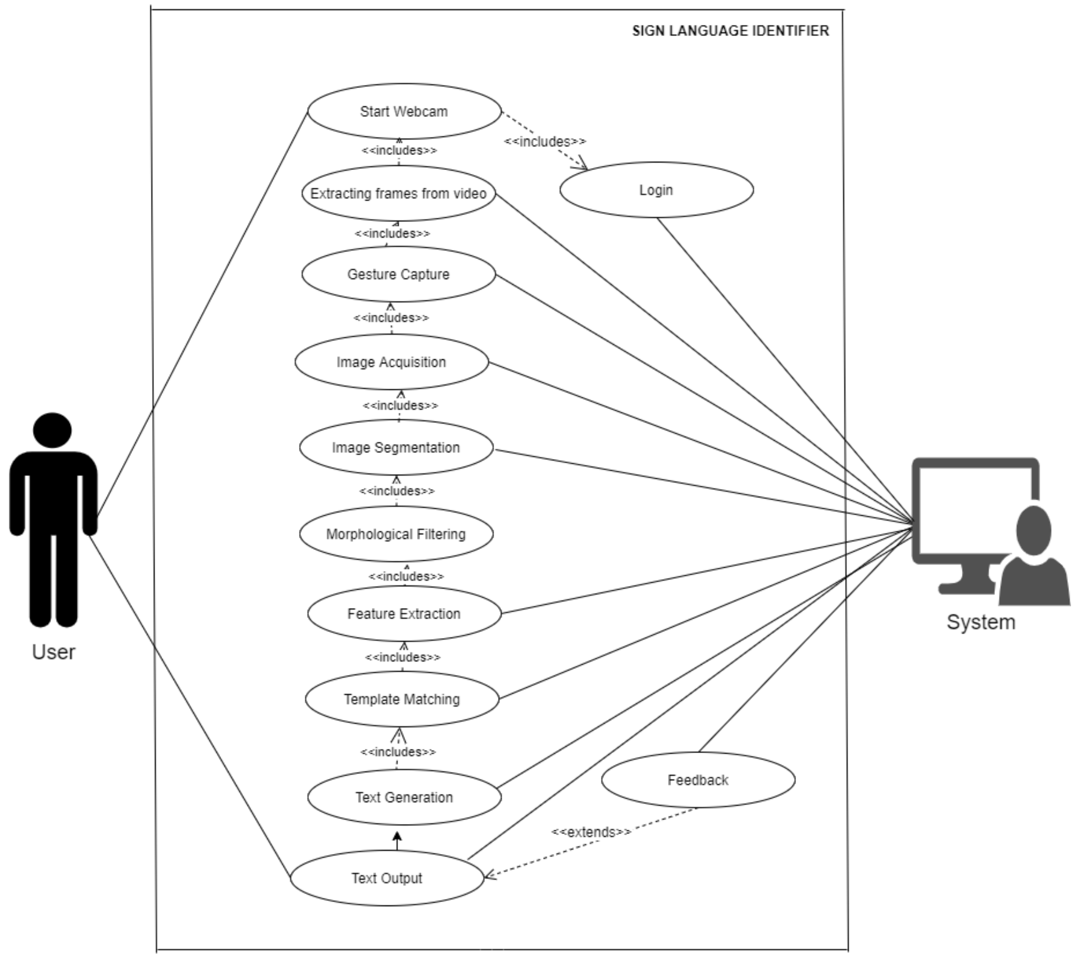


Figure 12: Use Case Diagram of ISL to text

Figure 13 is Use Case Diagram for speech to ISL.

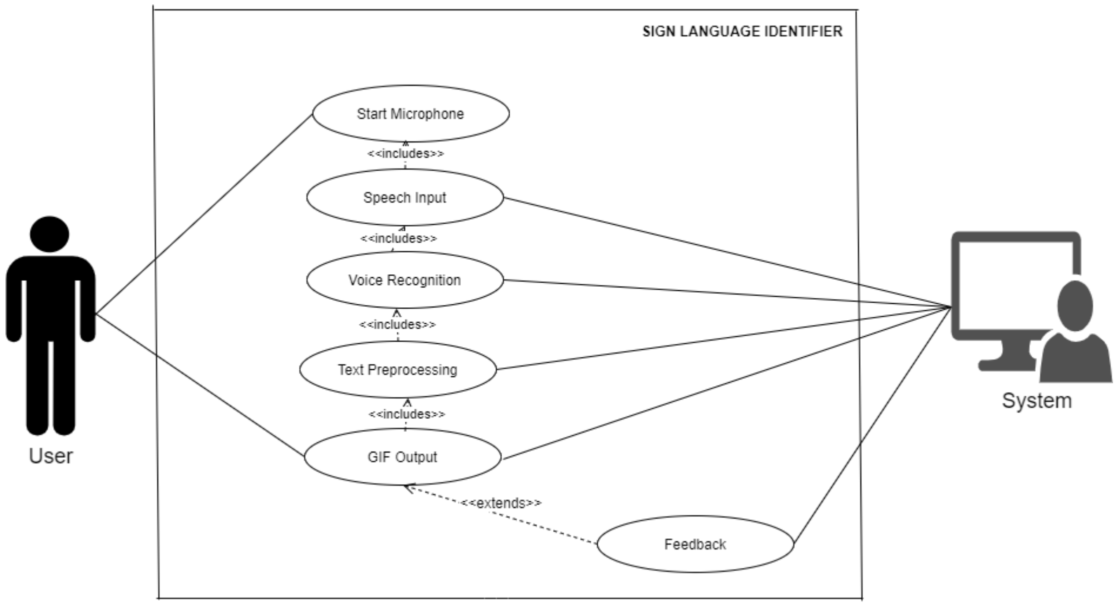


Figure 13: Use Case Diagram of speech to ISL

4.3.2 Use Case Template

Table 13 shows Use Case Template for ISL to text and Table 14 is Use Case Template for speech to ISL.

Table 13: Use Case Template for ISL to text

Use Case Title	Sign Language Identifier 1.1
Use Case Abbreviation	SLI 1.1
Description	SLI 1.1 is a web app to convert ISL to Text. This app will help the deaf and dumb people to communicate with the rest of the world. The app captures the gestures made by the deaf and dumb people through webcam and then using different technologies like image acquisition, image preprocessing, feature extraction, template matching and applying DL model. The video captured will have several frames and the corresponding sign will be converted to text on screen.
Goal	The goal of this system is to fill the gap between deaf society and the rest of the world. Using this app, the deaf and dumb people will be able to explain their thoughts to the rest of the world and will make them feel less isolated from the society and brings in sense of equality.
Pre Conditions	1. The camera used to capture images should not be blurred. 2. Users should log in before using the application. 3. A plain background must be maintained.
Basic Course	1. User logs into the system. 2. The user will grant permission to the application to use its camera.

	3. The user will perform the sign language actions in front of the camera. 4. The system will convert the speech to ISL. 5. Feedback page will be displayed.
Alternate Flow	At step 1 if login fails 1. Error messages will be displayed to enter the valid credentials and the login page will be displayed again. At step 2 if permissions are not granted 2. System will not go to the next screen and will again ask to grant the permission. At step 3 if actions are not performed according to indian sign language 3. Error messages will be displayed that actions are not performed correctly and nothing will be audible.
Post Conditions	Customers can provide feedback after using the web app post login only.
Actors	● Users ● System
Use Cases	● Start webcam ● Video capture ● Preparing image sequence ● Image acquisition ● Image preprocessing ● Feature extraction ● Template matching ● Text Generation ● Text output
Includes	● Login ● Start Webcam ● Gesture capture

	• Preparing image sequence • Image acquisition • Image preprocessing • Feature extraction • Template matching • Text Generation • Text Output
Extends	• Feedback
Modification History	10th April, 2020
Author	Samiksha Kapoor, Sanjana Vashisth, Sargunpreet Kaur, Simran

Table 14: Use Case Template for speech to ISL

Use Case Title	Sign Language Identifier 1.2
Use Case Abbreviation	SLI 1.2
Description	SLI 1.2 is a system to convert speech to ISL. This app will help the deaf and dumb people to communicate with the rest of the world. The system captures the voice and then displays the gif or sequence of plots of alphabets corresponding to the sign so performed.
Goal	The goal of this system is to fill the gap between deaf society and the rest of the world. Using this system, the deaf and dumb people will be able to explain their thoughts to the rest of the world and will make them feel less isolated from the society and brings in sense of equality.

Pre Conditions	Microphone must be in working condition.
Basic Course	1. User clicks on the “Live Voice” button. 2. A message of “Say Something” is displayed asking for voice input. 3. User speaks. 4. Said word is in the dataset. 5. A gif is displayed.
Alternate Flow	At step 1 1. User clicks on “All Done!” button. 2. System quits. At step 4 4. Said word is not in the dataset. 5. Plots of alphabets from which word is composed of is displayed.
Post Conditions	User can provide feedback.
Actors	● Users ● System
Use Cases	● Starts microphone ● Speech input ● Voice Recognition ● Text Preprocessing ● GIF Output
Includes	● Starts microphone ● Speech input ● Voice Recognition ● Text Preprocessing ● GIF Output
Extends	● Feedback

Modification History	10th April, 2020
Author	Samiksha Kapoor, Sanjana Vashisth, Sargunpreet Kaur, Simran

4.3.3 Activity Diagram

Activity diagram is basically a flowchart to represent the flow from one activity to another activity, just the difference being that parallel activities going on can be shown. The activity can be described as an operation of the system. Figures 24 and 25 represent the activity diagram for the “Sign Language Identifier” project . It serves the following purposes:

1. It draws the activity flow of the system.
2. It describes the sequence from one activity to another.

Figure 14 shows the activity diagram for conversion of ISL to text and figure 15 shows the activity diagram for conversion of speech to ISL.

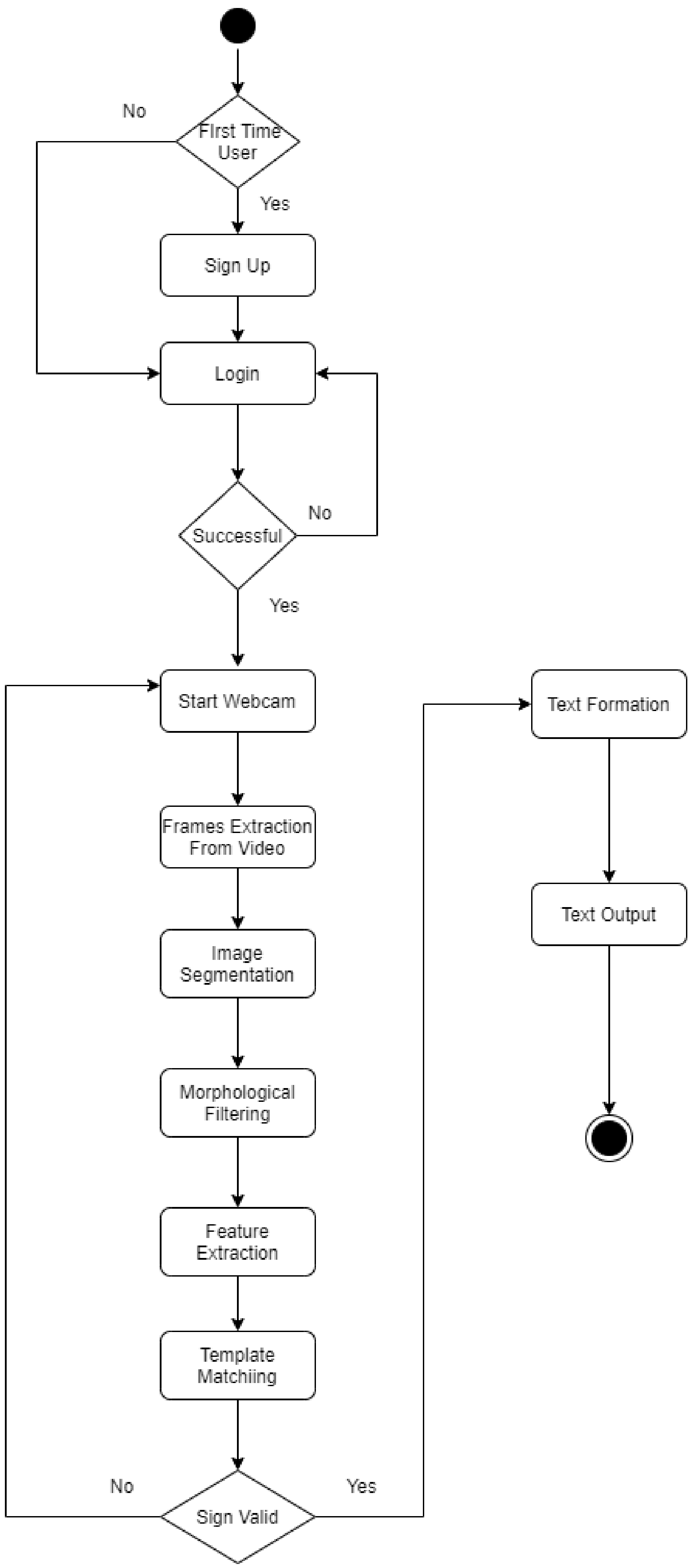


Figure 14: Activity Diagram of ISL to text

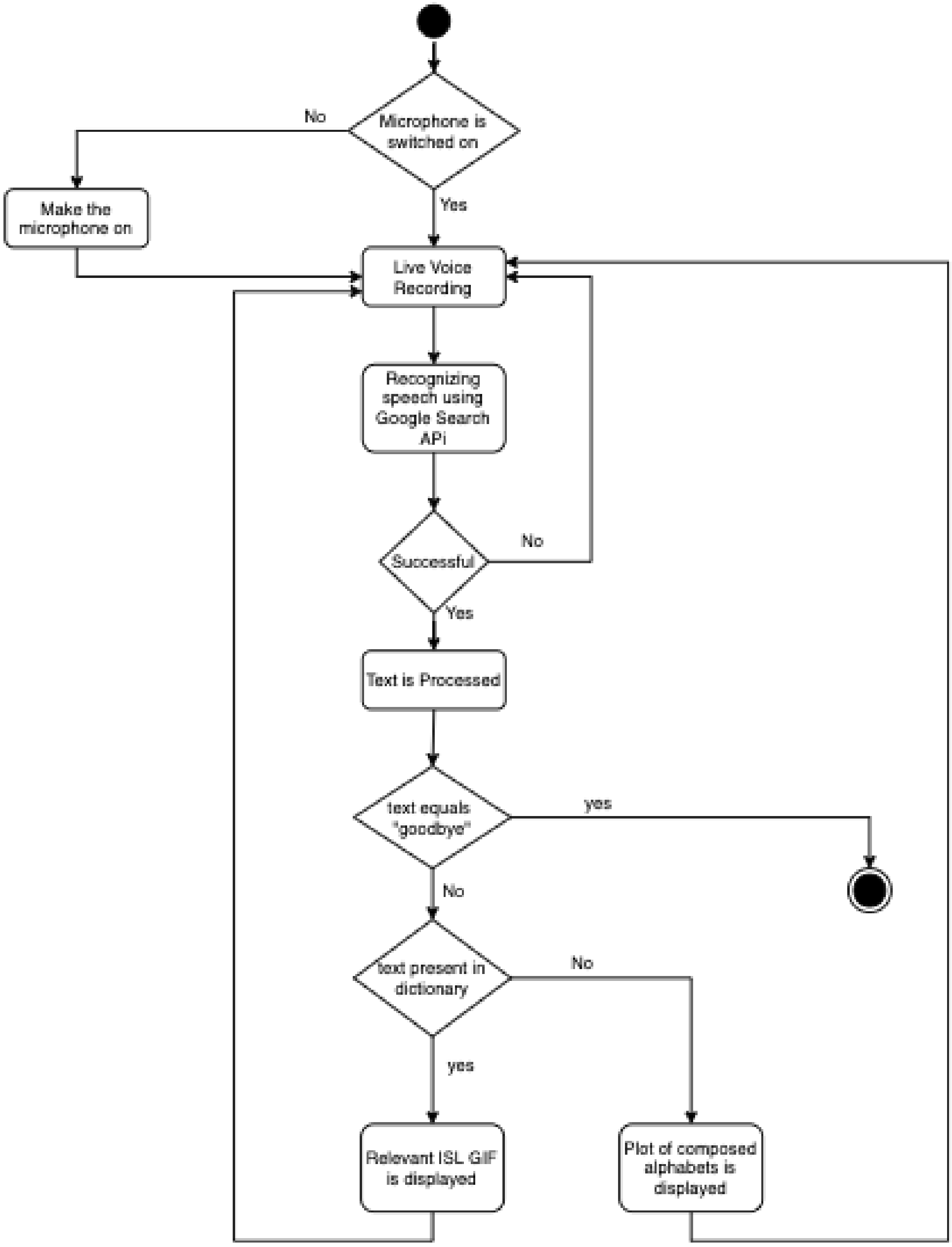


Figure 15: Activity Diagram of speech to ISL

4.3.4 Sequence Diagram

Figure 16 depicts interaction between the objects in a sequential order. There are 6 objects named User, System, Webcam, Image Preprocessor, Database and Template Matching. The diagram below shows the order in which the interaction between the objects takes place.

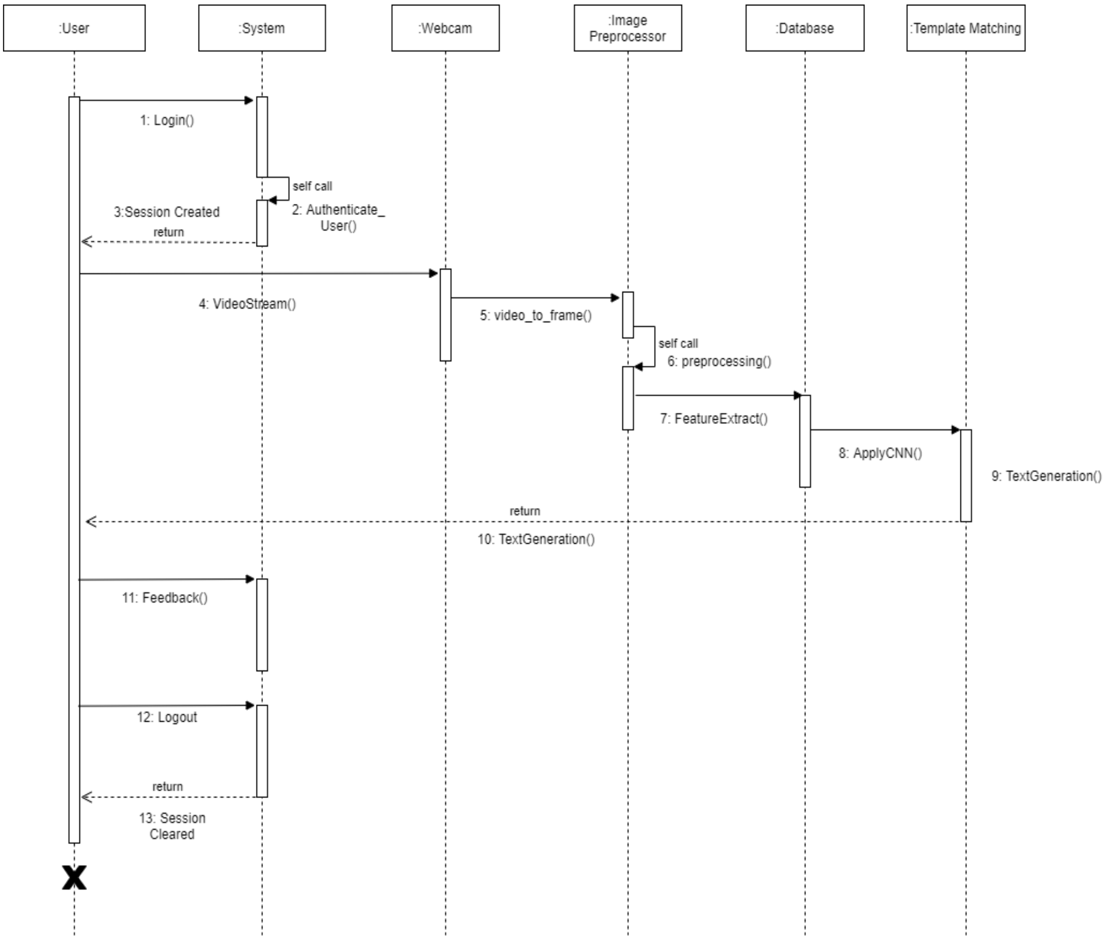


Figure 16: Sequence Diagram

One more feature has been added to the project, i.e. Converting Speech to Indian Sign language (ISL). This feature will help to communicate two-way, which means, normal people can also talk to the deaf and dumb. They just have to speak (a dictionary has been created with some words) and whatever they speak will be converted to ISL , which will be of more help for deaf people. Following are the steps followed for converting speech to ISL

If said word is present in the dataset

1: Live Voice Recording

2: Recognition of Speech using the Google Speech API

3: Text Preprocessing

4: Dictionary based Machine Translation

5: A gif is displayed performing the ISL.

If the said word is not in the dataset, then

5: Plots of alphabets from which word is composed of is displayed.

4.3.5 Class Diagram

Following Class Diagram (Figure 17) shows various classes, attributes and functions and Table 15 also shows the same.

Table 15: Classes and their functions

S.No.	Class Name	Function
1	User	Allows the user to login (if already registered), sign up (for first time registration) and feedback. It also allows the user to open the webcam and start communicating via sign language.
2	DataAcquisition	This includes storing images and converting them to HSV format.
3	ImageSegmentation	Applies automatic thresholding by using Otsu Algorithm
4	MorphologicalFiltering	Performs dilation and erosion so get a perfect image for feature extraction.
5	FeatureExtraction	Extracts the alphabet where ‘1’ is assigned in 2d array.
6	TemplateMatching	Matches signs performed by users with signs stored in the dataset, with the help of TRAIN_MODEL.h file formed by applying CNN.
7	Text_Generation	Generates text on the window as per sign made.

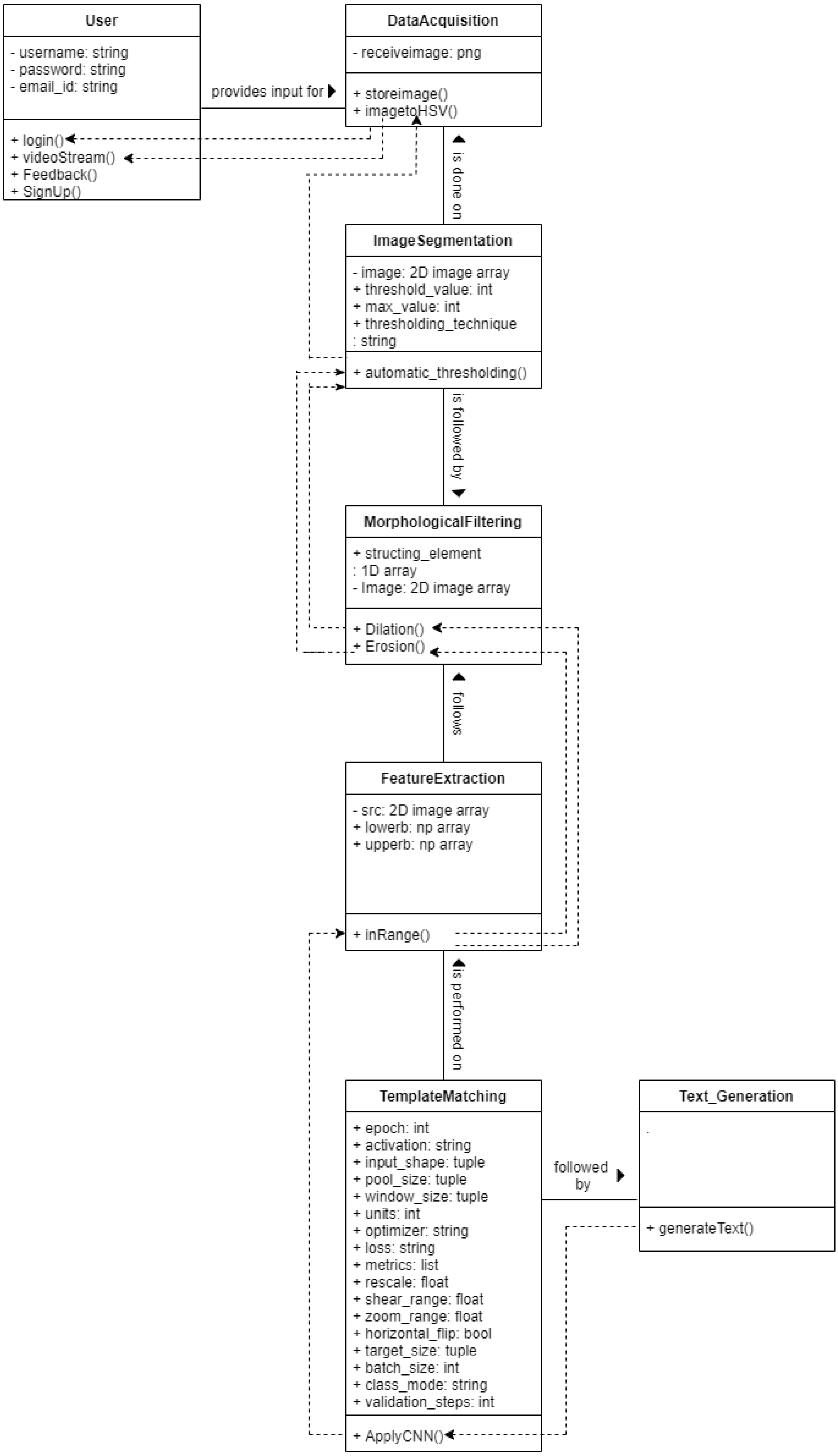


Figure 17: Class Diagram

4.3.6 State Chart Diagram

A state chart diagram is used to represent the condition of the system at finite instances of time. The following diagram i.e. Figure 18 is serving the purpose for the SLI project. There are nine states and occurrence of a particular event causes transition from one state to another.

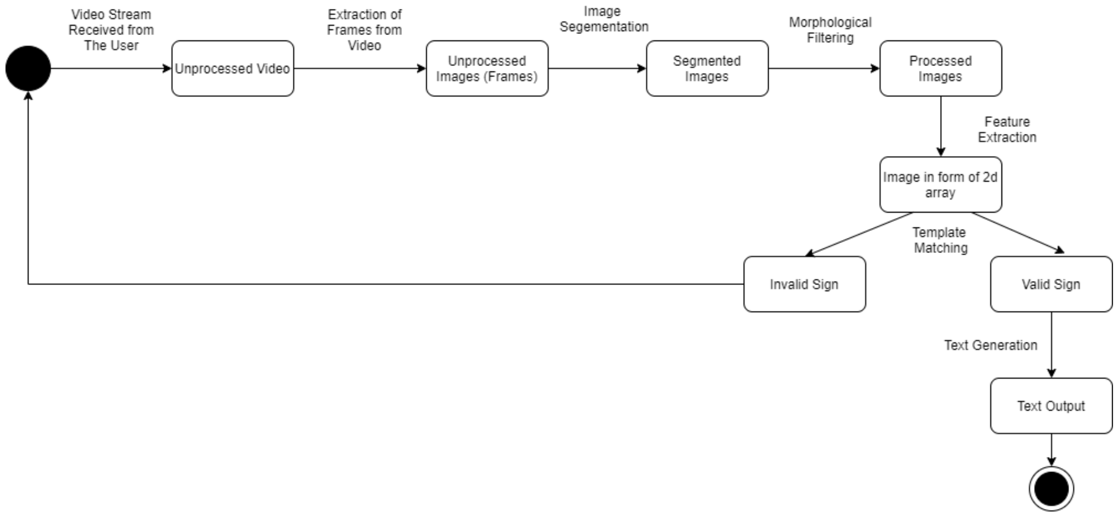


Figure 18: State Chart Diagram of ISL to text

Figure 19 shows the State Chart Diagram for conversion of speech to ISL.

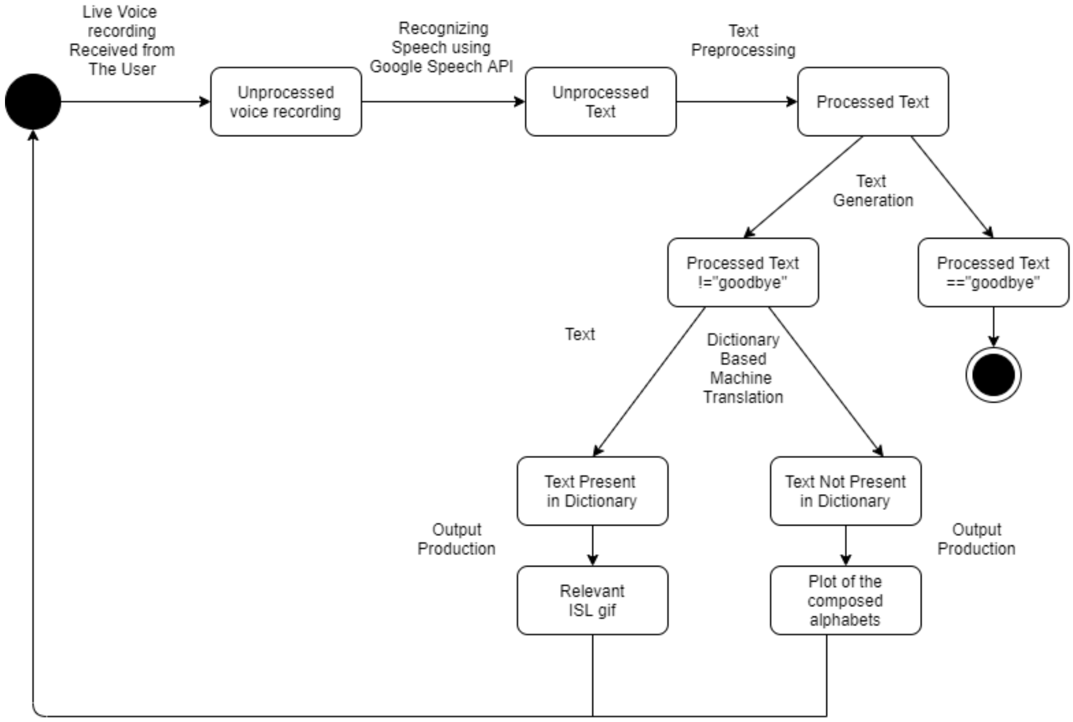


Figure 19: State Chart Diagram of speech to ISL

IMPLEMENTATION AND EXPERIMENTAL RESULTS

This section deals with discussion of implementation and experimentation with regards to the project. It also mentions all the test plans including the features to be tested, the test cases and discusses the inference drawn from the results. Furthermore, the procedural workflow and algorithm used are mentioned in it.

5.1 Experimental Setup

This involves two parts: one for ISL to Text and other for Speech to ISL.

- For both the parts, firstly the user needs to signup and create an account, then post login, he/she can avail services of "Indian Sign Language Identifier".
- For the first part (ISL to text), the user needs to have a webcam working and preferably a plain background. As he/she performs signs in ISL, the corresponding alphabet will be displayed on screen on the left hand side at the bottom.
- For the other part (Speech to ISL), the user will be asked to say something, he/she needs to speak a word, if word is in the bag of words then corresponding gif will be returned as output, else the graphs of the alphabets of which word is composed of, will be returned.
- Post both activities, the user can give optional feedback. Later on, he/she can again go to home and look for services or logout.

5.2 Experimental Analysis

The experimental analysis involves looking into data sources, how features were extracted and how hyperparameters were tuned and analyzing them. This also involves discussion about performance parameters.

5.2.1 Data

- Data Sources: Data sources include collection of dataset partly by creation of dataset on own and partly by using dataset from various github repositories.

- Dataset Preprocessing: Since, the algorithm requires the dataset to have segmented images, so a script was developed which can automatically perform this daunting task.
- Overall dataset information: Dataset includes training_set which has folders of A, B, C, …., Z which each folder contains 1200 images and testing_set also has similar subfolder arrangement, just the difference being the no. of images which are 251.
- For captured data during service avail by user: The image of only the hand portion was captured and segmentation was performed which was matched with templates (in form of images) by the trained model (CNN).

5.2.2 Performance Parameters

The proposed model for SLI could be judged on the basis of accuracy for static images and for dynamic images (taken in real time). The idea of performance can be seen with the help of Training and Testing Accuracy and also Training and Testing loss, which is discussed in detail in 5.5 section. More measures could be precision, recall for all classes A-Z. The area under the ROC curve can also give good ideas. This is with respect to ISL to text. For other part i.e. Speech to ISL, accuracy is the metric.

5.3 Working of the Project

5.3.1 Procedural Workflow

Sign language Identifier is extremely intuitive to use and operate. The first-time user has to first sign up and then login to use. There are two modules :- sign language to text and speech to sign language. The user has the choice to invoke the necessary module as and when required.

On choosing sign language to text, the webcam automatically gets started. The webcam detects the sign made by the signer and correspondingly converts the same to text. The session can be terminated by clicking the "Close Session" button on the interface.

On choosing speech to sign language, the microphone gets activated. The live voice is recorded through the microphone and is recognized using google speech API, resulting in production of the text. If the user says "goodbye", then the session is terminated else the relevant ISL gif is displayed.

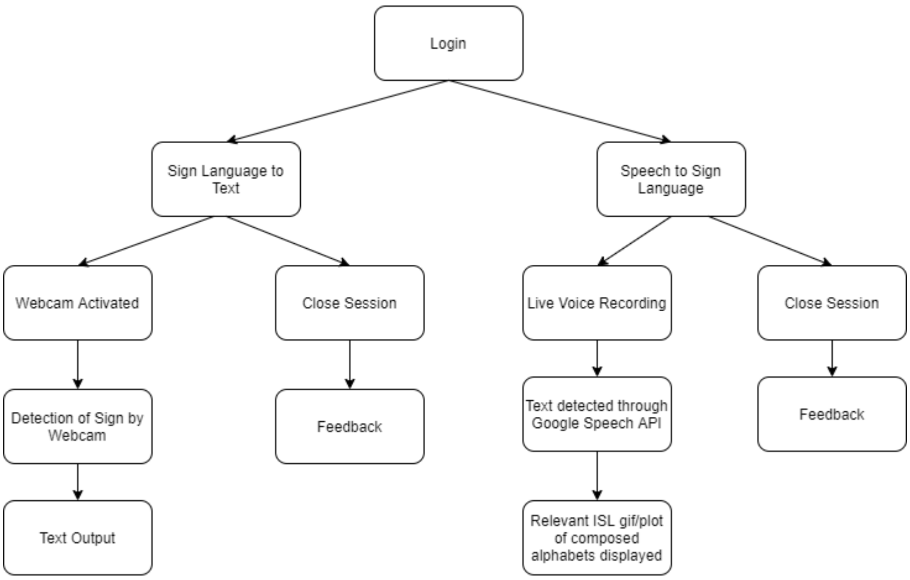


Figure 20: Procedural Workflow of the Project

Figure 20 depicts the procedural workflow of the project.

5.3.2 Algorithmic Approaches Used

5.3.2.1 ISL to Text

5.3.2.1.1 Preprocessing the image

- Capturing the image through webcam
- Since the image captured will be of BGR format, it will be converted to HSV format.
- After converting the image to HSV format, the image is converted to black and white by converting the hands in the image to white and the rest of the image to black.

Figure 21 below shows the code for the same.

```
cam = cv2.VideoCapture(0)
while True:
   ret, frame = cam.read()
   frame = cv2.flip(frame,1)
   img = cv2.rectangle(frame, (425,100),(625,300), (0,255,0), thickness=2,
                       lineType=8, shift=0)

   lower_blue = np.array([l_h, l_s, l_v])
   upper_blue = np.array([u_h, u_s, u_v])
   imcrop = img[102:298, 427:623]
   hsv = cv2.cvtColor(imcrop, cv2.COLOR_BGR2HSV)
   mask = cv2.inRange(hsv, lower_blue, upper_blue)
   mask = cv2.GaussianBlur(mask, (3, 3), 0)
```

Figure 21 : code 1

The values of l_h, l_s, l_v, u_h, u_s, u_v in Figure 37 are respective positions of the window “Slider”. This window is created with the help of code shown in Figure 22.

```
cv2.namedWindow("Slider")
cv2.createTrackbar("L - H", "Slider", 0, 179, nothing)
cv2.createTrackbar("L - S", "Slider", 58, 255, nothing)
cv2.createTrackbar("L - V", "Slider", 30, 255, nothing)
cv2.createTrackbar("U - H", "Slider", 33, 179, nothing)
cv2.createTrackbar("U - S", "Slider", 255, 255, nothing)
cv2.createTrackbar("U - V", "Slider", 255, 255, nothing)
   l_h = cv2.getTrackbarPos("L - H", "Slider")
   l_s = cv2.getTrackbarPos("L - S", "Slider")
   l_v = cv2.getTrackbarPos("L - V", "Slider")
   u_h = cv2.getTrackbarPos("U - H", "Slider")
   u_s = cv2.getTrackbarPos("U - S", "Slider")
   u_v = cv2.getTrackbarPos("U - V", "Slider")
```

Figure 22 : code 2

5.3.2.1.2 Building the CNN model

Figure 23 shows the code for building the CNN.

```
# Step 1 - Convolution Layer
classifier.add(Convolution2D(32, 3,  3, input_shape = (64, 64, 3), activation = 'relu'))

#Step 2 - Pooling
classifier.add(MaxPooling2D(pool_size =(2,2)))

#Adding second convolution layer
classifier.add(Convolution2D(32, 3,  3, activation = 'relu'))
classifier.add(MaxPooling2D(pool_size =(2,2)))

#Adding 3rd Concolution Layer
classifier.add(Convolution2D(64, 3,  3, activation = 'relu'))
classifier.add(MaxPooling2D(pool_size =(2,2)))


#Step 3 - Flattening
classifier.add(Flatten())

#Step 4 - Full Connection
classifier.add(Dense(256, activation = 'relu'))
classifier.add(Dropout(0.5))
classifier.add(Dense(26, activation = 'softmax'))
```

Figure 23 : code 3

Figure 24 shows the code for compiling the CNN.

```
#Compiling The CNN
classifier.compile(
                optimizer = optimizers.SGD(lr = 0.01),
                loss = 'categorical_crossentropy',
                metrics = ['accuracy'])
```

Figure 24: code 4

5.3.2.1.3 Recognizing image and predicting the result

Figure 25 and Figure 26 show the code for predicting the result.

```
def predictor():
       import numpy as np
       from keras.preprocessing import image
       test_image = image.load_img('1.png', target_size=(64,64))
       test_image = image.img_to_array(test_image)
       test_image = np.expand_dims(test_image, axis = 0)
       result = classifier.predict(test_image)
       if result[0][0] == 1:
              return 'A'
       elif result[0][1] == 1:
              return 'B'
       elif result[0][2] == 1:
              return 'C'
       elif result[0][3] == 1:
              return 'D'
       elif result[0][4] == 1:
              return 'E'
       elif result[0][5] == 1:
              return 'F'
       elif result[0][6] == 1:
              return 'G'
       elif result[0][7] == 1:
              return 'H'
```

Figure 25 : code 5

```
elif result[0][8] == 1:
       return 'I'
elif result[0][9] == 1:
       return 'J'
elif result[0][10] == 1:
       return 'K'
elif result[0][11] == 1:
       return 'L'
elif result[0][12] == 1:
       return 'M'
elif result[0][13] == 1:
       return 'N'
elif result[0][14] == 1:
       return 'O'
elif result[0][15] == 1:
       return 'P'
```

Figure 26 : code 6

5.3.2.1 Speech to ISL

- Firstly, live voice is recorded i.e. audio is obtained from the microphone.
- Then speech is recognized using Google Speech Recognition and corresponding text is obtained.
- After this, with the help of dictionary based machine translation, relevant ISL gif or plot of composed alphabets is displayed.

The code for the same is shown in the Figure 27-30 below.

```
def func():
       r = sr.Recognizer()
with sr.Microphone() as source:
           r.adjust_for_ambient_noise(source)
           i = 0
           while True:
               print('Say something')
               audio = r.listen(source)
```

Figure 27 : code 7

```
# recognizing speech
try:
    a = r.recognize_google(audio)
    print("You said " + a.lower())

    for c in string.punctuation:
        a = a.replace(c, "")

    if (a.lower() == 'goodbye'):
        print("OOPs!Time to say goodbye!!")
        break

    elif (a.lower() in isl_gif):

        class ImageLabel(tk.Label):


            def load(self, im):
                if isinstance(im, str):
```

Figure 28 : code 8

```
    im = Image.open(im)
self.loc = 0
self.frames = []

try:
    for i in count(1):
        self.frames.append(ImageTk.PhotoImage(im.copy()))
        im.seek(i)
except EOFError:
    pass

try:
    self.delay = im.info['duration']
except:
    self.delay = 100

if len(self.frames) == 1:
    self.config(image=self.frames[0])
else:
    self.next_frame()
```

Figure 29 : code 9

```
def unload(self):
    self.config(image=None)
    self.frames = None

def next_frame(self):
    if self.frames:
        self.loc += 1
        self.loc %= len(self.frames)
        self.config(image=self.frames[self.loc])
        self.after(self.delay, self.next_frame)
```

Figure 30 : code 10

5.3.3 Project Deployment

Our project, Sign Language Identifier (SLI) basically consists of two modules -(i) ISL to text and (ii) Speech to ISL.

It has four main menus : HOME, LOGIN, SIGNUP, ABOUT. The already registered users can directly login by filling in the username and password. The new user first needs to sign up by filling in : email id, username and password. There is an ABOUT menu which displays quick information about our project.

After successful login, the user has a choice between the two modules: sign language to text and speech to sign language. First module (sign language to text) is implemented using CNN (Convolutional Neural Network) and various image processing techniques. The second module (speech to sign language) is implemented using Google audio api for speech recognition, and dictionary based machine translation. After this the final step in the project was to deploy the model. Deployment of the model was done using Flask. Flask is a web application framework written in Python. The deployed model worked perfectly when running on the flask server.

Process of deployment:

The existing functions were embedded into html and css webpage and to have python as a backend, flask provides a facility of routing to the particular webpage.

1. Model Building: Deep Learning Model (CNN) pipeline was built to classify Signs performed into A-Z in English.
2. Webpage template: Here, the website was designed where users can avail services of ISL to text; and speech to ISL. Website design uses following technology stack:
    a. Front-end: HTML, CSS
    b. Backend: Python
3. Predict class and send results: Next, use the saved model to predict the class of the signs performed and send the results back to the webpage.

Below given is sample deployment code (not complete code)

```
@app.route('/')
@app.route('/index.html')
def index():
   return render_template("index.html")

@app.route("/login.html", methods=['POST','GET'])
def login():
   return render_template("login.html")

@app.route("/signup.html")
def signup():
   return render_template("signup.html")

@app.route("/about.html")
def about():
   return render_template("about.html")
```

5.3.4 System Screenshots

SLI has 4 main menus: HOME, LOGIN, SIGNUP, ABOUT.

Home menu looks as follows (Figure 31).

Figure 31: Welcome Screen

The second menu is Login (Figure 32) for already registered users. It asks for Username and Password. Click on the LOGIN button to complete the task.

Figure 32: Login Screen

The third menu is SignUp (Figure 33) for new users. It asks for Email Id, Username and Password. Click on the SignUp button to complete the task.

Figure 33: Sign Up Screen

Figure 34 displays the ABOUT menu.

Figure 34: About Menu

On successful login, the following screen (as shown in figure 35) is displayed from where you can choose between services available which are: Sign Language to Text and other one is Speech to Sign Language.

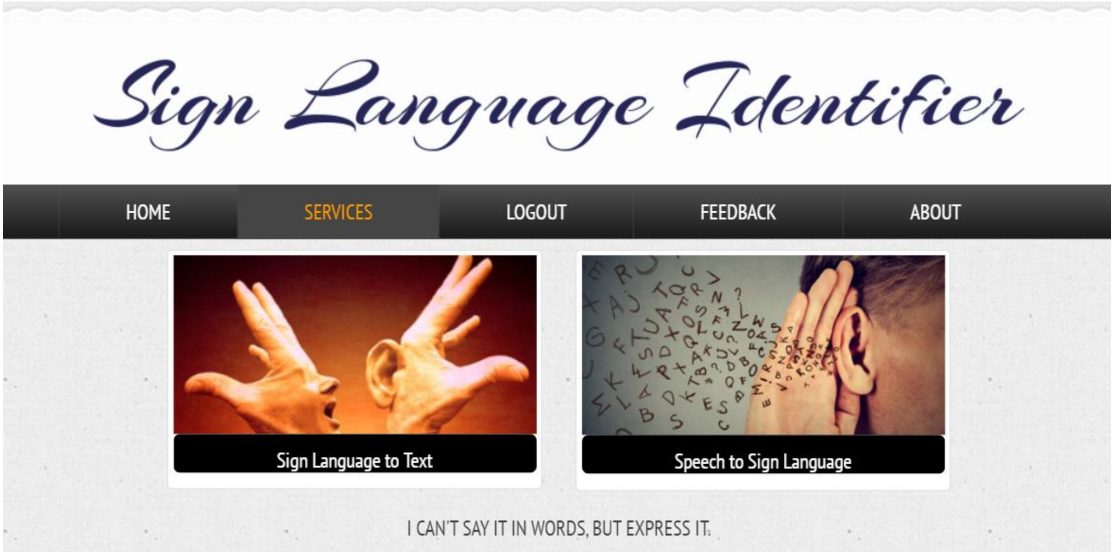


Figure 35: About Menu

If the user chooses Sign Language to Text option, then the following screen appears (Figure 36) and then the user performs sign language and corresponding text output is displayed.

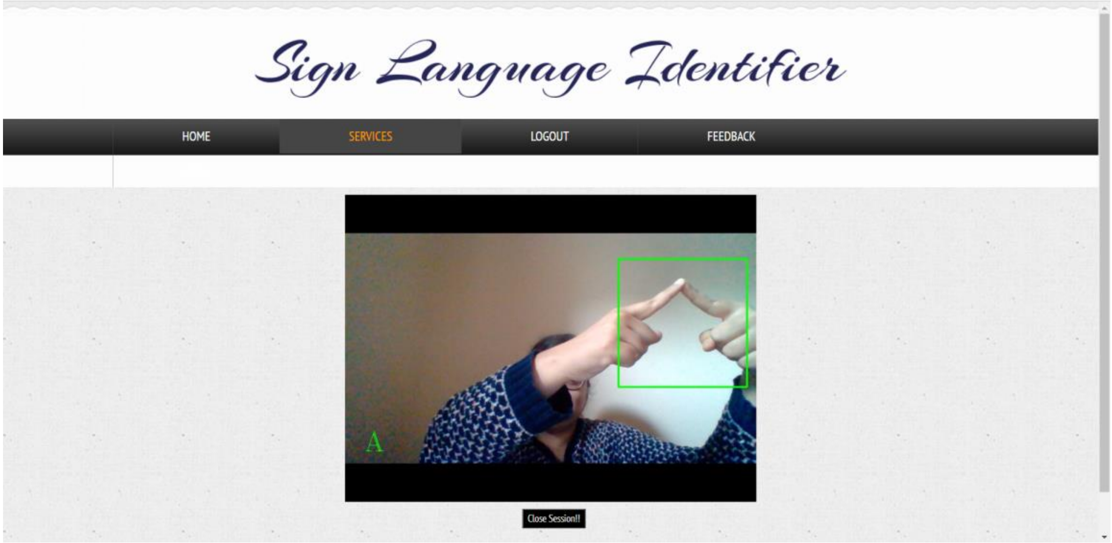


Figure 36: Indian Sign Language to Text Conversion

After the user is done with the task, he/she can click on LOGOUT which ends the session and brings the user to the homepage but if the user clicks on the "Close Session!!" button (as shown in Figure 36), then SLI asks for Feedback (Figure 37) which is optional. The user can simply log out then if he/she isn't interested or can write feedback and click on "Submit!!" button.

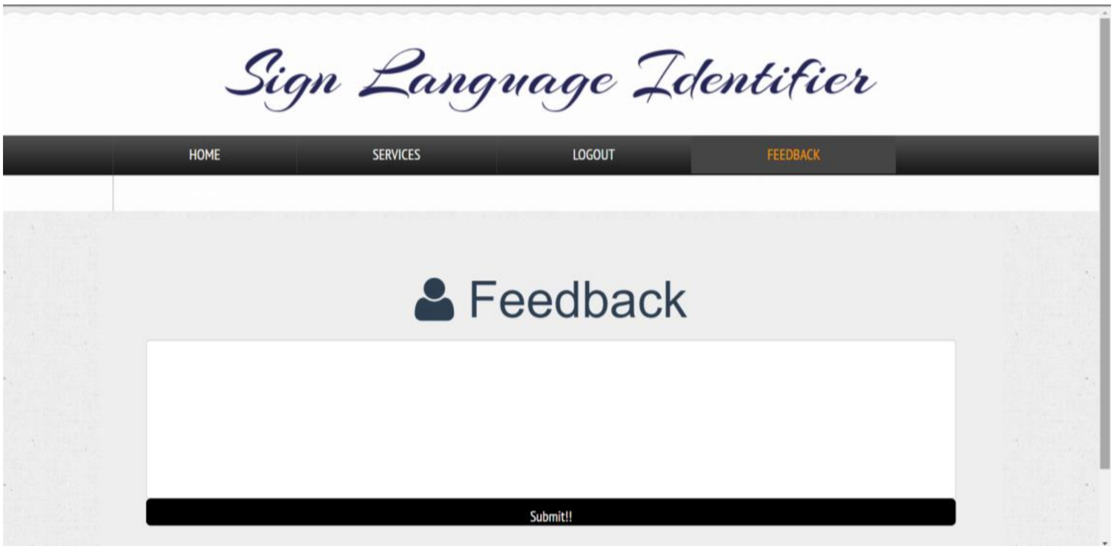


Figure 37: Feedback

If the user avails the facility of speech to ISL, then the following screen (Figure 38) appears where the user needs to click on the “Activate Microphone” button.

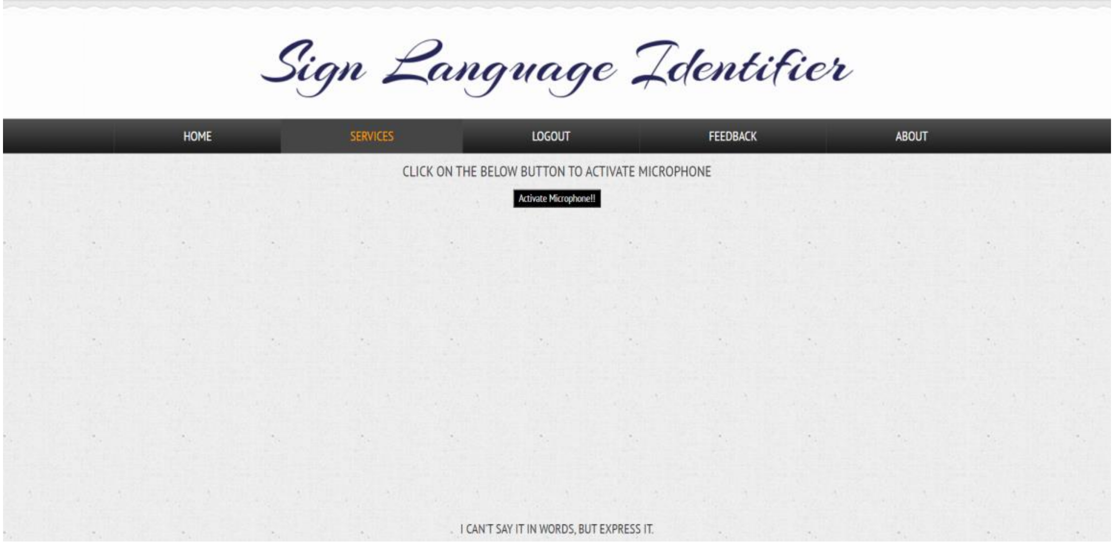


Figure 38: Screen after speech to ISL service selection

After the “Activate Microphone!!” button is clicked, SLI prompts to “SAY SOMETHING TO CONVERT TO ISL!!” as given in Figure 39.

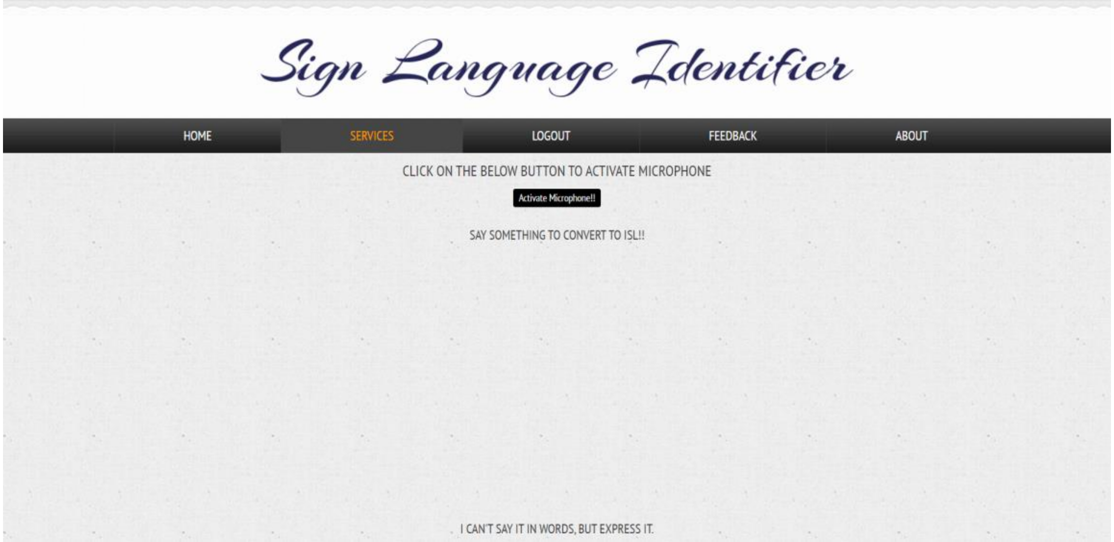


Figure 39: Prompt onclick of “Activate Microphone!!” button

The user speaks “FLOWER IS BEAUTIFUL” and the corresponding GIF is loaded as given in Figure 40.

Figure 40: GIF for spoken line

5.4 Testing Process

Testing is the most important phase where normal and corner cases are tested and that's where the developer gets an idea, how the product is performing in real time.

5.4.1 Test Plan

The following plan was adopted to test efficiency of VisionaryAid and to analyse it for future improvements:

5.4.1.1 Features to be selected

a. Sign recognition efficiency: The sign recognition efficiency has to be tested on various parameters as the sign that is to be captured can be containing any english alphabet in Indian Sign Language, thus making it necessary for such features to be tested. Following are such parameters which are tested to:

   I. ISL Alphabet signs: Sign Language Identifier on standard ISL alphabet signs without any background noises or blurs was tested.

   II. Signs with background noise: There can be signs performed with moving and colored backgrounds that need to be taken care of.g

   III. Signs with simple gesture: There can be signs with different color backgrounds which need to be ignored.

   IV. No signs performed: There can be case where no signs will be performed

   V. Alphabets under good lighting conditions: ISL Alphabet Signs performed under proper lightening

   VI. ISL Alphabet under bad Lighting Conditions: ISL Alphabet Signs performed under bad lightening

b. Speech to sign conversion efficiency: The speech to sign conversion efficiency has to be tested on various parameters as the voice that is to be heard can be containing any type of character, thus making it necessary for such features to be tested. Following are such parameters which are tested to:

   I. Different pitch voices: Voice to be heard can be of different pitch, it should have no effect on the system.

II. Different frequency voices: Voice to be heard can be of different frequencies, it should have no impact on the conversion.

III. Standard English Speech: Speech to sign conversion on standard english speech was tested.

IV. Interrogative speech: Interrogative speeches were tested.

V. Greeting speech:Greeting Speeches were tested.

VI. Imperative Speech: Imperative Speeches were tested

5.4.2 Features to be tested

Following are the features to be tested:

1. Capturing images
2. Identifying images+
3. Recording speech
4. Converting speech to Indian Sign Language

5.4.3 Test Strategy

Following are the two testing strategies for the sign language identifier:

1. To test the sign recognition efficiency we need to perform different signs in the front of the camera with plain, colored and moving backgrounds and check the output with the actual sign.
2. To test the speech to sign language we need to test it against the speeches of different pitches, frequencies and using different types of speeches.

5.4.4 Test Techniques

ISL to Text

1. Static Images: While making the neural network we do the testing of the model on the dataset and calculate its accuracy.
2. Real time Images: Capturing the images with camera, identifying the images with the help of the model and checking them against the actual signs.

Speech to Text

1. Computer recorded speech: Checking the result with the computer voice.
2. Real time speech: Live testing the result with real time speech.

5.4.5 Test Cases

There are two modules in SLI:

1. ISL to text
2. Speech to ISL

The test cases for ISL to text are given in table below (Table 16)

Table 16: Test Cases for ISL to text

Test Case	Test Case Description
1	ISL Alphabet
2	Blank
3	ISL Alphabet under good lighting conditions
4	ISL Alphabet under bad lighting conditions
5	Similar ISL Alphabets
6	Under different backgrounds

The test cases for speech to ISL are given in table below (Table 17)

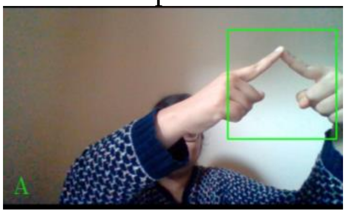
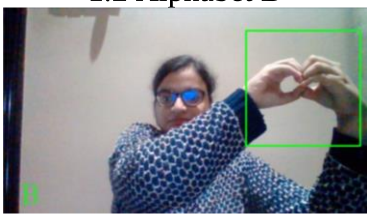
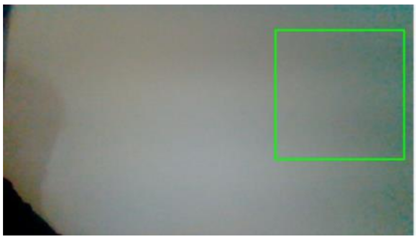
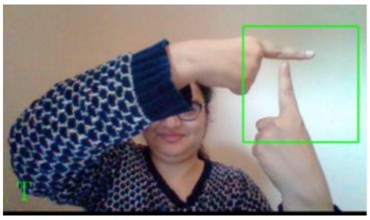
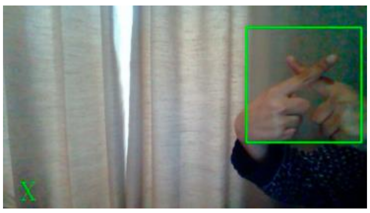
Table 17: Test Cases for speech to ISL

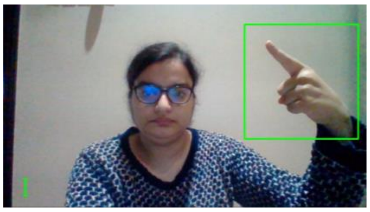
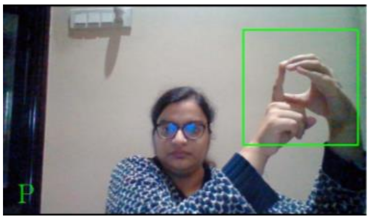
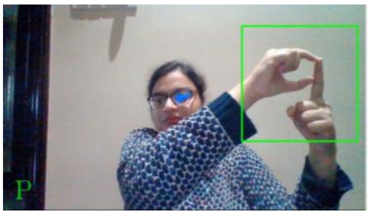
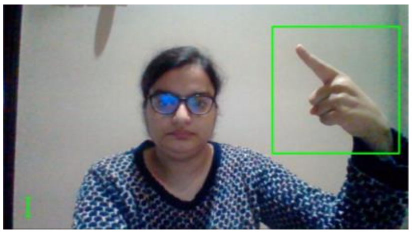
Test Case	Test Case Description
1	Spoken alphabet is ISL Alphabet
2	Declarative Sentence
3	Interrogative Sentence
4	Imperative Sentence
5	Proper Nouns

5.4.6 Test Results

On performing the test cases for ISL to text, following results were obtained and are given in table below (Table 18)

Table 18: Test Cases Results for ISL to text

Test Case	Test Case Description	Output	Status (Pass/ Fail)
1	ISL Alphabet 1.1 Alphabet A 1.2 Alphabet B	1.1 Alphabet A  1.2 Alphabet B 	1.1 Alphabet A Pass 1.2 Alphabet B Pass
2	Blank		Pass
3	ISL Alphabet under good lighting conditions		Pass
4	ISL Alphabet under bad lighting conditions		Pass
5	Similar ISL	5.1 ‘I’ and ‘L’	5.1 ‘I’ and ‘L’

	Alphabets 5.1 ‘I’ and ‘L’ 5.2 ‘D’ and ‘P’	  5.2 ‘D’ and ‘P’  	Pass 5.2 ‘D’ and ‘P’ Fail
6	Under different backgrounds	 	Pass

The test cases for speech to ISL are given in table below (Table 19)

Table 19: Test Cases for speech to ISL

Test Case	Test Case Description	Status (Pass/ Fail)
1	Spoken alphabet is ISL Alphabet	Pass

2	Declarative Sentence	Pass

3	Interrogative Sentence	Pass

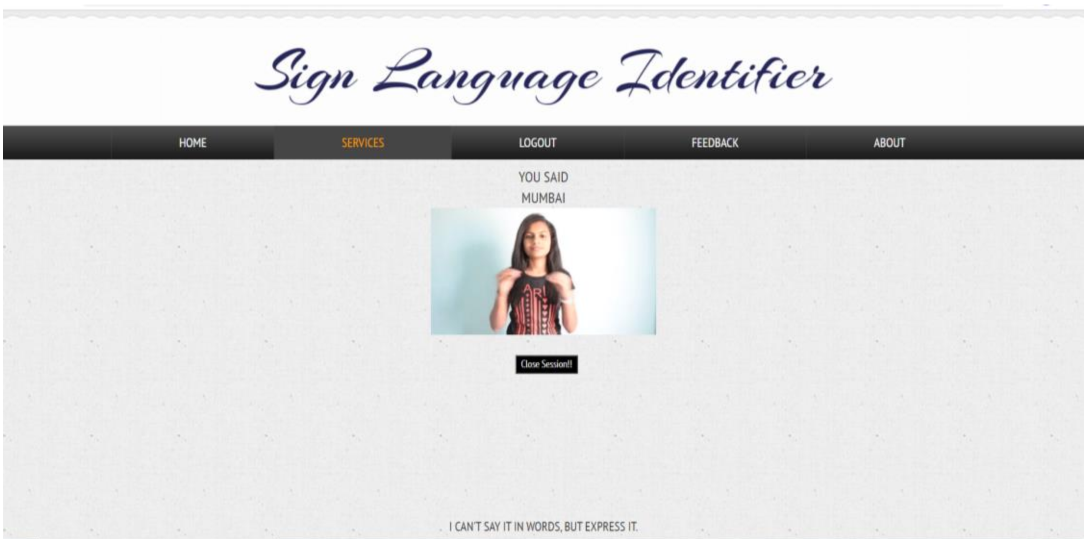
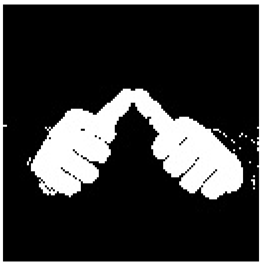
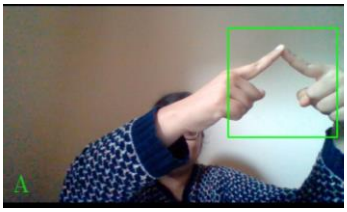
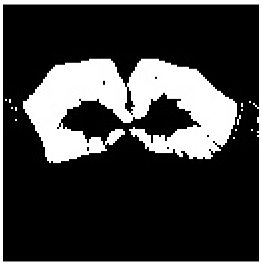
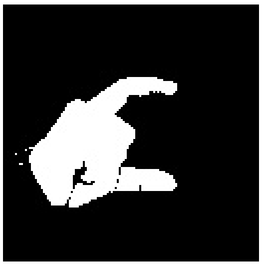
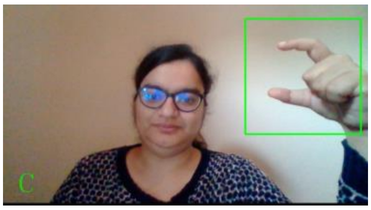
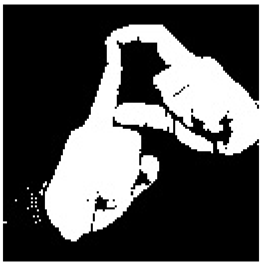
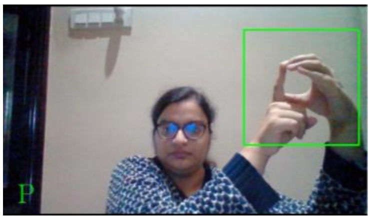
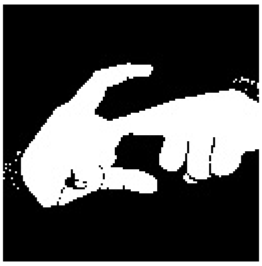
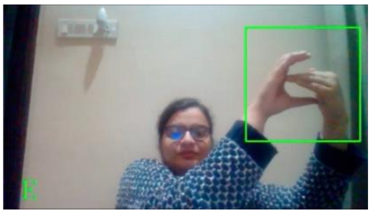
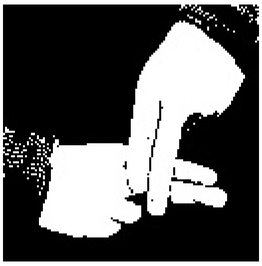
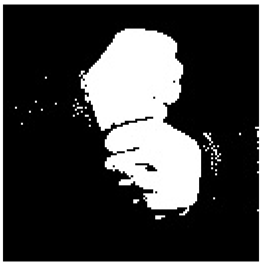
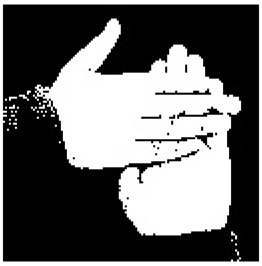
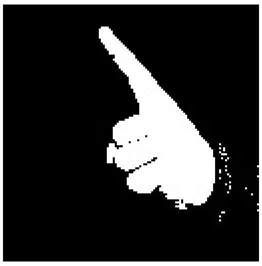
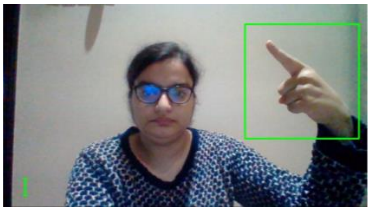
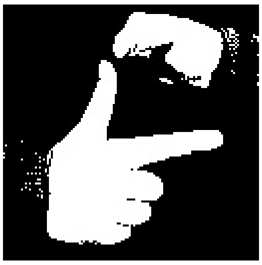
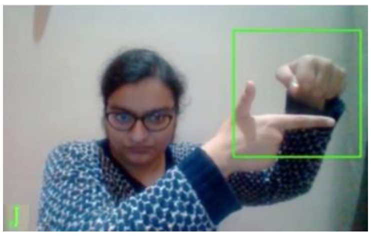
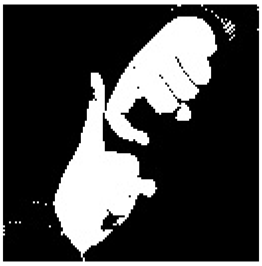
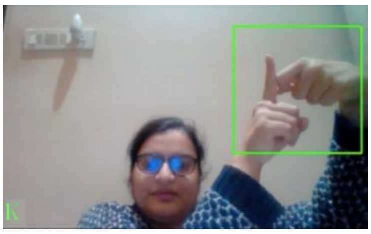
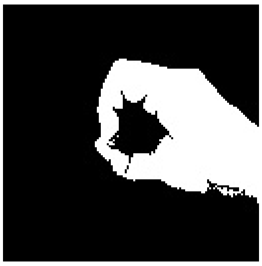
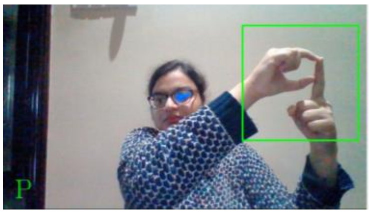
		
4	Imperative Sentence	Pass
		
5	Proper Nouns	Pass
		

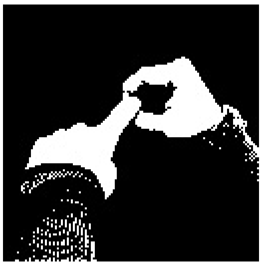
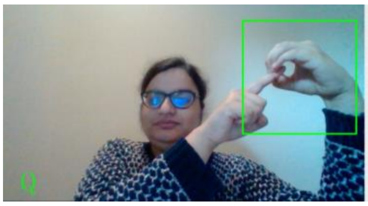
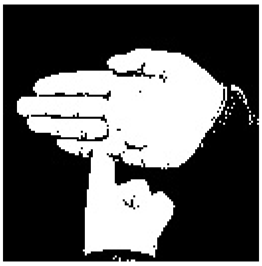
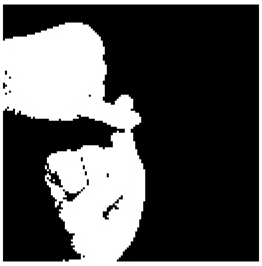
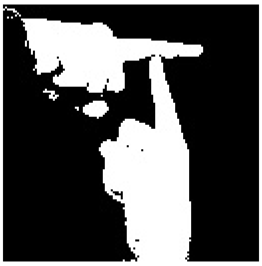
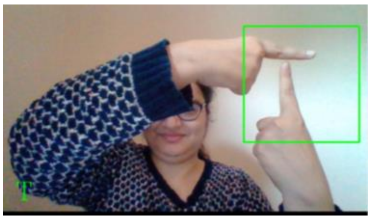
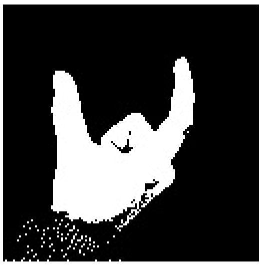
Table 20 shows the detailed result of all the ISL alphabets

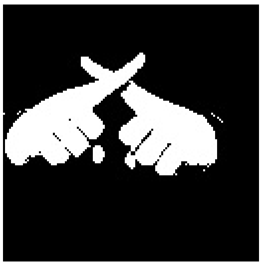
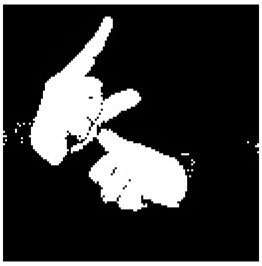
Table 20: Predicted output as per SLI

Alphabets	Symbols in ISL	Predicted Output	Correctly Identified
A			Yes
B			Yes
C			Yes
D			No

E			Yes
F			Yes
G			Yes
H			Yes
I			Yes
J			Yes

K			Yes
L			Yes
M			Yes
N			Yes
O			Yes
P			Yes

Q			Yes
R			Yes
S			Yes
T			Yes
U			Yes
V			Yes

W			Yes
X			Yes
Y			Yes
Z			Yes

5.5 Results and Discussions

The proposed model for SLI could be judged on the basis of accuracy for static images as discussed in section 5.2.2, it came out to be quite good and for dynamic images (taken in real time) there was fluctuation in result, but with proper placement of hands and plain background, it gave accuracy equivalent to static images only. The idea of performance can be seen with the help of Table 16 where it’s given regarding the alphabets, corresponding symbols and output predicted by SLI and the last column denotes if the symbol was correctly identified or not. For all the letters it gave correct results (during testing of dynamic images - real time), but for ‘D’ it mistakenly identified it as ‘P’. Another parameter could be Training and Testing Accuracy and also Training and Testing loss, which is discussed in detail in graphs given below namely Figure 41 and Figure 42. Figure 41 summarizes history for model accuracy and it can be noted that as the number of epochs increase, the model accuracy initially shot up and then later on after some epochs it remained nearly constant giving final accuracy of 0.9895 for training data and 0.9765 for test data. Figure 42 summarizes history for loss and it can be seen that the loss was huge in the very beginning of epoch and as the number of epochs increased, loss decreased to 0.0115 for training data.

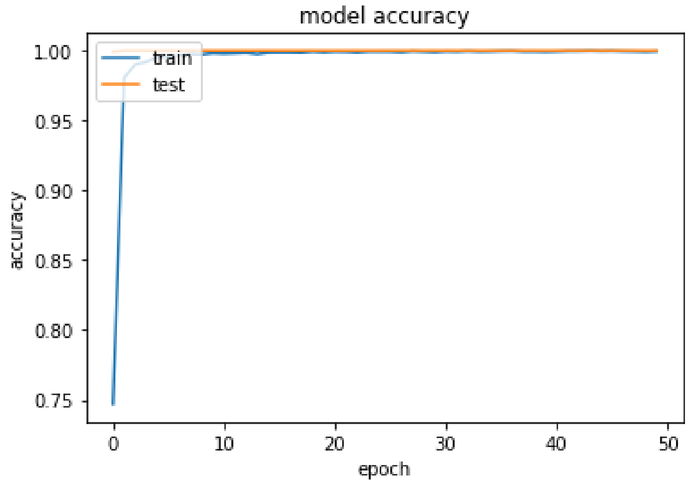


Figure 41: Model Accuracy

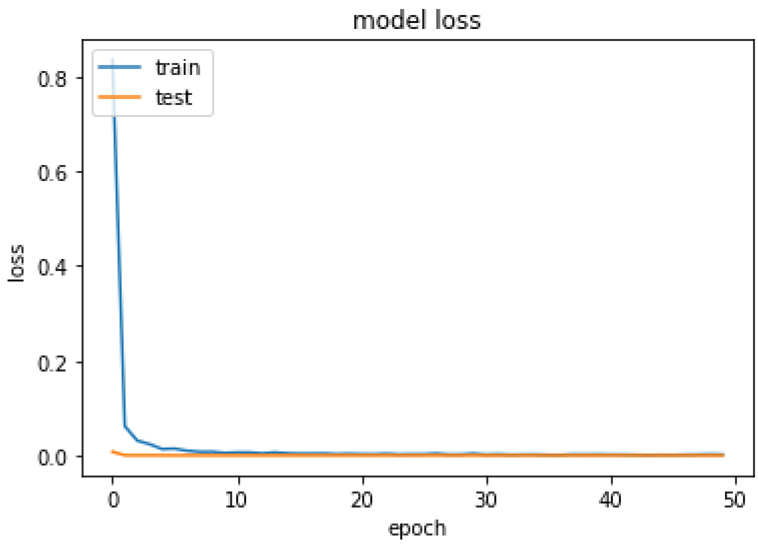


Figure 42: Model Loss

5.6 Inferences Drawn

Following are the inferences drawn from testing performed on Sign Language Identifier:

1. Sign Language Identifier is suitable for all the alphabets (A-Z) performed according to Indian Sign Language.
2. Sign Language Identifier gives results with very good accuracy of 97.25 for static images but in case of dynamic images, the output comes with some fluctuations and the user has to do some movements for getting the result.
3. Sign Language Identifier is majorly suitable with white plain background while converting Indian Sign Language to Text.
4. Sign Language Identifier successfully converts all the words which are there in our system to gif while performing speech to text. Apart from the words which are not in the system, it successfully shows the letter by letter image of the word spoken by the user.
5. Sign Language Identifier is suitable if there is no background noise while performing speech to ISL.

5.7 Validation Of Objectives

Table 21 shows the status (successful/unsuccessful) of the objectives for the ISL identifier.

Table 21: Validation of Objectives

S.No.	Objectives	Status
1.	To create a website to enable Indian Sign Language to be converted into English Text; and speech to Indian Sign Language.	Successful
2.	To perform tasks for conversion of Indian Sign Language to English. 2.1 Capture Video 2.2 Generate Text 2.3 Display Text	Successful
3.	To perform tasks for conversion speech to ISL. 3.1 Record the voice 3.2 Generate the ISL gif 3.3 Display the ISL gif	Successful
4.	To design and implement the website.	Successful
5.	To test and validate the website as per the functionality and user experience.	Successful

CONCLUSIONS AND FUTURE SCOPE

6.1 Work Accomplished (w.r.t. the Objectives)

Table 22 shows the work accomplished w.r.t cited objectives.

Table 22: Work Accomplished (w.r.t.the Objectives)

Objectives	Discussion
1. To survey and analyze the current status of usage and effectiveness of Indian Sign Language (ISL) solutions.	Sign Language Recognition is an extensive research area in the field of human computer interaction. Such recognition systems are meant to replace sign language interpreters. With the development of image processing and artificial intelligence techniques, many techniques have been recently developed in this area. Most of the signs in Indian Sign Language (ISL) are double handed and hence it is more complex compared to single handed American Sign Language (ASL). So, most of the researchers use ASL signs for creating their database. Recently, researchers from India have started working on ISL to develop automatic Indian sign language recognition systems. Mainly three steps are involved in sign language recognition-preprocessing, feature extraction and classification. The important classification methods used for recognition are Artificial Neural Networks (ANN), Support Vector

	Machine (SVM), Hidden Markov Models (HMM) etc.
2. To propose and implement an interactive solution that will convert sign language to text and speech to ISL. a.To collect ISL dataset and create ISL dictionary for commonly used words and to preprocess them. b. To apply Deep Learning Model for training on collected dataset.	The working model of sign language identifier is made, which is able to recognize fifty different image words of indian sign language. a.A dataset for english alphabets is collected. This dataset is converted into a dictionary in order to use the tensorflow module. b. CNN model is applied to the training dataset.
3. To test and demonstrate the usability and effectiveness of the proposed solution for targeted user groups (Deaf and dumb persons).	The SLI prototype is being trained on the dataset collected. After training efficiency of the model will be evaluated.

6.2 Conclusions

In the Prototype of Sign Language Identifier the working model of the “Sign Language Identifier" is made, which works on the images captured by the camera. For ISL to text conversion, the user performs the sign in the rectangular region defined. Then corresponding text appears on the screen. Internally, firstly from video, frames are extracted and then based on HSV color model and inRange() function of opencv, the thresholding takes place. Different thresholding techniques like Simple Thresholding, Adaptive Thresholding, Otsu Thresholding were used but inRange() function gave better results, then the model was trained by CNN model.

For Speech to ISL conversion, simply speech is converted to text and then the label is checked.

There are many research gaps in glove based recognition systems like the VPL dataglove or other glove systems that use wired networks makes it importable and there is a lot of noise and delay in the system, that's why the vision based recognition system is made. HMM model is used by many researchers and it has high efficiency but it is only used by those researchers who aimed at recognizing very few words or alphabets. Aim is to make the real time conversation using “Sign Language Identifier” between hearing impaired and other people, that's why CNN model for image recognition is used as its efficiency is high in recognizing large numbers of images. For the same reason it will be modified to an interface that will recognize the gestures in the video and for easy access a web app will be made.

6.3 Environmental (/Economic/ Social) Benefits

This app is of no harm to the environment and has lots of social and economic benefits. Social benefits include that it will be helping that section of the society which feels completely isolated because no one is able to understand what they want to say/express. The deaf and dumb people always feel deprived of society. They are not given equal opportunity in society. Hence, by using this app they just need to perform Indian sign language in front of a webcam and hence the sign language will be converted to text; and speech will be converted to ISL as fast as they perform the actions. Along with that, it has a lot of economic benefits too. As the app will be made by using the Indian Sign Language, it will be of great benefit to the economy as it will reduce the use of the apps and datasets made by other countries.

6.4 Reflections

The whole journey of building “Sign Language Identifier” has been a valuable experience, starting with the discovery of possible opportunities to think of the idea to the phase where the same idea was actually deployed. The team gained insight into the field of software development and now in the future, members shall feel more confident in the process of project development. Furthermore, it was learnt how to analyze the existing frameworks and perform literature surveys and utilize that analysis to identify the problem statement, research gaps and come up with the solution ideas. It was a learning of how to incorporate and take care of the user requirements. It was the time when the importance of documentation was realized and what are techniques involved in being organized about it. One of the takeaways was how to manage the resources in an efficient manner and most importantly to use common sense and build a viable and efficient model, but best takeaway was development of analytical skills while working in the team and discussing each point of the assigned task in hand in detail. The whole project helped us in exploring the skills as a computer engineer and improved confidence levels, ability to work under pressure and helped in learning project management techniques. It aided the members to be familiarized with the working and delivering of projects and how to build an entire product from just an idea.

6.5 Future Work Plan

The system currently converts sign language to text; and speech to ISL, not sign language to speech, which can be included later on. Other feature extraction algorithms can be added in conducting experiments for more accurate results. Same goes for more classifiers like Support Vector Machine (SVM), Principal Component Analysis (PCA) and many more. Their combination could be used for improvement in recognition rate. The Natural Language Processing (NLP) tasks could be focussed on, to facilitate proper context. For the web app, the user login and sign up just requires an email id, username and password, hence no personal details are asked for, ensuring no pose to security threat. Currently, limited attempts are permitted for wrong username or password. Passwords could be stored securely, by firstly modifying password with a unique way and then applying hashing method to make it more secure. Hashing and salting passwords could be used. For more sensitive information, asymmetric encryption could be employed. As of now, dataset is also limited, more dataset can be added for better results, which can be prepared by taking the pictures of different humans performing signs, different backgrounds. As of now, it is limited to the machine on which it runs, future work includes a chrome extension for SLI and a progressive web application. A digital avatar could be made for better interaction and real time experience.

PROJECT METRICS

7.1 Challenges Faced

Every new learning and trying unexplored areas always come with challenges. Initial challenge was the dataset. Since, the concerned language was ISL, so it was quite difficult to find appropriate images on which model can be trained. There are good datasets for ASL, but ISL finds it harder to make place in research, but extensive searching helped in finding suitable dataset. Later on, the next problem that everyone faced was coordination among team members during covid-19 time period when all were at their places, anyhow work from home was well-versed due to technology like Team Viewer to help get access of PC where the main project is made, Zoom and Google Meet for sharing and discussion of work progress. Amidst the project, yet another problem was to capture images from various backgrounds and then removing noise and applying filters for better output. Since, the model used CNN algorithm, so setting hyperparameters was another task. With a lot of hits and trials, some thumb rules, hyperparameters were tuned to get the best out of the algorithm. The model suffered from overfitting which was tackled by Dropout method. Conversion of speech to ISL seemed another daunting and tough task, but some knowledge and googling whenever required helped a lot. Despite challenges, it was really amazing and quite a learning journey starting from scratch and then making things work. Team is the core part which makes everything work.

7.2 Relevant Subjects

The Capstone Project requires the knowledge of multiple subjects. Some of them are direct learnings from courses (taught in Institute), and some are the skills to be learnt by the developers themselves. Table 23 provides information about the relevant subjects used to successfully complete the project.

Table 23: Subject Code and Subject Name

Subject Code	Subject Name	Description
UML501	Machine Learning	To get better insight about how "sign language identification" problems can be looked at and hence, solved.
UCS742	Deep Learning	To apply the most suitable deep learning algorithm (CNN) for sign language identification, and also knowledge of optimizers, loss functions.
UML602	Natural Language Processing	For matching gif or image corresponding to speech from a bag of words.
UCS615	Image Processing	To capture the image from the camera and further removal of noise using noise reduction filters and further enhancing the input for better output and ease of prediction for the trained model.
UCS503	Software Engineering	Software Development Lifecycle, Preparation of SRS, Working as per Scrum model, Shaping functional and non-functional requirements, Understanding and communication of ideas through UML Diagrams.

7.3 Interdisciplinary Knowledge Sharing

From doing this project, the members gained knowledge with regards to computer engineering aspects like image processing, machine learning, deep learning, natural language processing, web development and also about some aspects of biology and computer vision. Learning was also with regards to target users- the deaf and the dumb people and also got some knowledge about the Indian Sign language. Major revelation was about their everyday problems, how they face such challenges and what new ways have been brought into their lives through modernization. Note was made on the various devices available to them and how much they cost and how much efficient they are in changing their lives and helping them.

7.4 Peer Assessment Matrix

Table 24 gives view on Peer Assessment Score and Table 25 provides Student information w.r.t. Student Codes.

Table 24: Peer Assessment Marks

		Evaluation of			
		S1	**S2**	**S3**	**S4**
Evaluation By	**S1**	5	5	5	5
	S2	5	5	5	5
	S3	5	5	5	5
	S4	5	5	5	5

Table 25: Student Information

Student Code	Student Roll No.	Student Name
S1	101703476	Samiksha Kapoor
S2	101703482	Sanjana Vashisth
S3	101703488	Sargunpreet Kaur
S4	101703545	Simran

7.5 Role Playing and Work Schedule

Although all members have equally contributed to the project, the major portion of the work was done by dividing them among team members. Table 26 shows the individual roles and contributions of the team members.

Table 26: Individual Roles and Contributions

Sr.	Member Name	Member Roll no.	Contributions
1.	Samiksha Kapoor	101703476	• Studying of research papers • Creating dataset • Creating login and sign up module • Extraction of frames from videos (coding) • Segmentation (coding)
2.	Sanjana Vashisth	101703482	• Studying of research papers • Dilation and Erosion (coding) • Training dataset • Feature Extraction by specified algorithm(coding) • Making various software diagrams
3.	Sargunpreet Kaur	101703488	• Studying of research papers

			● Text preprocessing ● Template Matching and Sign Recognition using appropriate algorithm (coding) ● Conversion of Speech to Text ● Creation of GUI
4.	Simran	101703545	● Studying of research papers ● Template Matching and Sign Recognition using appropriate algorithm (coding) ● Training model ● Dictionary based machine translation

7.6 Student Outcomes Description and Performance Indicators (A-K Mapping)

Table 27: AK Mapping of various concepts

SO	Description	Outcome
A1	Applying mathematical concepts to obtain analytical and numerical solutions.	Used mathematical concepts like segmentation, dilation and erosion in image preprocessing.
A2	Applying basic principles of science towards solving engineering problems.	Used basic principles of sound to detect and convert speech to text.
A3	Applying engineering techniques for solving computing problems.	Used google speech API for speech recognition and various algorithms detect and preprocess images captured from webcam.
B1	Identify the constraints, assumptions and models for the problems.	Signer must be at a fixed distance from the webcam. Indian Sign Language is used.
B2	Use appropriate methods, tools and techniques for data collection.	Proper research on ISL was done in order to prepare an accurate dataset.
B3	Analyze and interpret results with respect to assumptions, constraints and theory.	Made the model portable and affordable.
C1	Design software system to address desired needs in different problem domains.	Helped in specialisation and better understanding.
C2	Can understand scope and constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability.	Reducing the communication barrier between deaf and dumb people and the rest. Affordable as compared to other alternatives available today.
D1	Fulfill assigned responsibility in	Learned the virtue of teamwork,

	multidisciplinary teams.	multitasking and punctuality.
D2	Can play different roles as a team player.	Increases efficiency and interdisciplinary knowledge.
E1	Identify engineering problems.	Learned to think logically, critically and practically.
E2	Develop appropriate models to formulate solutions.	Built several prototypes to achieve the final one.
E3	Use analytical and computational methods to obtain solutions.	Used segmentation, dilation, erosion, CNN and various other algorithms.
F1	Showcase professional responsibility while interacting with peers and professional communities.	Boosted confidence, listening and speaking skills.
F2	Able to evaluate the ethical dimensions of a problem.	Reviewed stats in order to come up to this final project.
G1	Produce a variety of documents such as laboratory or project reports using appropriate formats.	Taking notes and maintaining a weekly diary helped in reviewing work from time to time.
G2	Deliver well-organized and effective oral presentation.	Helped the panel to understand and give suggestions.
H1	Aware of environmental and societal impact of engineering solutions.	Less consumption of power, boosting self confidence of hearing impaired individuals.
H2	Examine economic tradeoffs in computing systems.	Using high computational yet economical components was the main objective.
I1	Able to explore and utilize resources to	Reading literature surveys of different

	enhance self-learning.	products, analyzing their drawbacks and removing them from our model was the goal.
I2	Recognize the importance of life-long learning.	Practicality always increases the learnability than reading theory.
J1	Comprehend the importance of contemporary issues.	Understanding the importance and responsibility to work in this field in order to work fro hearing impaired people.
K1	Writing code in different programming languages.	Helped in learning different programming languages, their advantages over others.
K2	Apply different data structures and algorithmic techniques.	Helped in optimization.
K3	Use software tools necessary for computer engineering domain.	Understanding working to increase the accuracy and speed rather than doing it manually.

Table 27 provides insight of AK Mapping of various concepts as shown above.

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 different sources used were first going through research papers and taking out keywords. Then, we discussed our approach with our mentor. We studied different types of sign language such as American Sign language etc to understand the language in a detailed manner. We explored already available solutions and their drawbacks. Then after seeing the pros and cons of different modules, we finalized the project.

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

Ans: The analytical methods to reach solutions were taking the lead generation survey, performing literature survey to analyse existing frameworks and identifying research gaps and studying existing solutions. Design diagrams and UML diagrams along with software requirement specification were equally very important to think through the solution ideas and make necessary changes. Experimental methods like prototyping and then it's analysis, checking efficiency of the product though testing also helped in reaching the final solution.

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

Ans: We used basic principles of science and engineering in our project. We used various algorithms for image preprocessing such as segmentation, dilation, erosion etc. We used Google Speech API to detect human speech and to convert it to text. Further algorithms like CNN were used to train the model to convert Sign language to text.

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

Ans: Our team had good communication throughout the project. We had one official whatsapp group with our mentor and one was of all the team members. Our team leader used to assign tasks to all group members and we were asked to finish it well in time.

Whatever problems came across were discussed and if we were not able to solve it on our own then we used to ask our mentor. Also, there were regular meetings with ma'am and among ourselves.

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

Ans: We used various resources such as the internet, books in the library. We went through a lot of research papers, understanding of various products already made in this field, and understanding their drawbacks. We also consulted our mentor whenever in doubt. This project wouldn't have been possible without the guidance of our mentor.

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: This one-year project taught us a lot of things. Not only the working of the project is important but also it taught us how the documentation should go side by side which keeps us in track of what part is left, what is done. Also, it taught us how real-life problems are much different than the theory one. We learn better when we are doing things and seeing it rather than reading from the book. We learnt how to use various new languages, software and tools.

APPENDIX A: REFERENCES

[1] Christopher Lee and Yangsheng Xu, “Online, interactive learning of gestures for human robot interfaces” Carnegie Mellon University, the Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996

[2] Etsuko Ueda, Yoshio Matsumoto, Masakazu Imai, Tsukasa Ogasawara, ”Hand Pose Estimation for Vision Based Human Interface”, IEEE Transactions on Industrial Electronics,Vol.50,No.4,pp.676- 684,2003.

[3] G. AnanthRao P.V.V .Kishore Ain Shams Engineering Journal Volume 9, Issue 4, December 2018, Pages 1929-1939 Selfie video based continuous Indian sign language recognition system

[4] InfoWorld 8 great Python libraries for natural language processing

[5] Mahesh Kumar N B Assistant Professor (Senior Grade), Bannari Amman Institute of Technology, Sathyamangalam, Erode, India. Conversion of Sign Language into Text

[6] Mandeep Kaur Ahuja, Amardeep Singh, “Hand Gesture Recognition Using PCA”, International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.

[7] Mohandes, Mohamed. “Arabic sign language recognition.” International conference of imaging science, systems, and technology, Las Vegas, Nevada, USA. Vol. 1. 2001.

[8] Naoum, Reyadh, Hussein H. Owaied, and Shaimaa Joudeh. Development of a new Arabic sign language recognition using k-nearest neighbor algorithm.“ (2012).

[9] Pattern Recognition Letters Volume 78, 15 July 2016, Pages 28-35 Continuous sign language recognition using level building based on fast hidden Markov model

[10] P.V.V.Kishore P.Rajesh Kumar Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text

[11] Richard Watson, “Gesture recognition techniques”, Technical report, Trinity College, Department of Computer Science, Dublin, July, Technical Report No. TCD-CS-93-11, 1993

APPENDIX B: PLAGIARISM REPORT

URKUND

Document Information

Analyzed document	SLI FINAL EVALUATION CAPSTONE.pdf (D90463842)
Submitted	12/22/2020 11:51:00 AM
Submitted by	
Submitter email	ssimran1_be17@thapar.edu
Similarity	7%
Analysis address	ssimran1_be17.thapar@analysis.urkund.com

For a detailed analysis report, click here.

APPENDIX C: TECHNICAL WRITING COURSE PROOF

Name: Samiksha Kapoor

Roll No.: 101703476

Name: Sanjana Vashisth

Roll No. : 101703482

RICE

COURSE CERTIFICATE

07/29/2020

Sanjana Vashisth

has successfully completed

Writing Skills for Engineering Leaders

an online non-credit course authorized by Rice University and offered through Coursera

EDUCATION FOR EVERYONE · coursera · COURSE CERTIFICATE

Gayle Moran
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Beata Krupa
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Verify at coursera.org/verify/BC4D8GJZWNA7

Coursera has confirmed the identity of this individual and their participation in the course.

Name: Sargunpreet Kaur

Roll No.:101703488


RICE

COURSE CERTIFICATE

07/28/2020

Sargunpreet Kaur

has successfully completed

Writing Skills for Engineering Leaders

an online non-credit course authorized by Rice University and offered through Coursera

EDUCATION FOR EVERYONE
coursera
COURSE CERTIFICATE

Gayle Moran
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Beata Krupa
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Verify at coursera.org/verify/QUBLWVVHYQKS

Coursera has confirmed the identity of this individual and their participation in the course.

Name: Simran

Roll No.: 101703545

RICE

COURSE CERTIFICATE

07/27/2020

Simran

has successfully completed

Writing Skills for Engineering Leaders

an online non-credit course authorized by Rice University and offered through Coursera

EDUCATION FOR EVERYONE

coursera

COURSE CERTIFICATE

Gayle Moran
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Beata Krupa
Lecturer in Professional and Engineering Communication
Rice Center for Engineering Leadership

Verify at coursera.org/verify/L2BJKMBJRGNA

Coursera has confirmed the identity of this individual and their participation in the course.