

# **Sign language recognition using**

## **ML technologies**

Submitted in partial fulfillment of the requirements of the  
degree

**BACHELOR OF ENGINEERING IN COMPUTER  
ENGINEERING**

By

1. Agarwal Gunjan Hemant - 312002
2. Ishan Ahmed - 312018
3. Tamim Ahmad - 312055
4. Tarmale Sudarshan Sanjay -312056

**Prof. Saiqa Khan**



**Department of Computer Engineering**

**M.H. SABOO SIDDIK COLLEGE OF ENGINEERING**

**8, Saboo Siddik Road, Byculla, Mumbai - 400 008**

**Department of Computer Engineering**

**University of Mumbai**

**(AY 2021-22)**

# CERTIFICATE

This is to certify that the Mini Project entitled “**Sign language recognition using ML technologies**” is a bonafide work of submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of “**Bachelor of Engineering**” in “**Computer Engineering**”.

**(Prof. Saiqa Khan)**

Mentor

**(Dr. Zainab Pirani)**

Head of Department

**(Dr. Ganesh Kame)**

I/c Principal

## Mini Project Approval

This Mini Project entitled “**Sign language recognition using ML technologies**”  
by **Agarwal Gunjan Hemant (312002), Ishan Ahmed (312018), Tamim Ahmad (312055), Tarmale Sudarshan Sanjay (312056)** is approved for the degree of  
**Bachelor of Engineering in Computer Engineering.**

### Examiners

**1**.....  
(Internal Examiner Name & Sign)

**2**.....  
(External Examiner name & Sign)

Date:

Place:

# Contents

	<b>Abstract</b>	<b>05</b>
	<b>Acknowledgement</b>	<b>06</b>
	<b>List of figures</b>	<b>07</b>
	<b>List of Tables</b>	<b>08</b>
<b>1</b>	<b>Introduction</b>	<b>09</b>
	1.1 Introduction	10
	1.2 Motivation	10
	1.3 Problem Statement & Objectives	11
	1.4 Organization of the Report	11
<b>2</b>	<b>Literature Survey</b>	<b>12</b>
	2.1 Survey of Existing System	13
	2.2 Limitation Existing system or Research gap	17
	2.3 Mini Project Contribution	17
<b>3</b>	<b>Proposed System</b>	<b>18</b>
	3.1 Introduction	19
	3.2 Architecture/ Framework	19
	3.3 Algorithm and Process Design	19
	3.4 Details of Hardware & Software	20
	3.5 Experiment and Results for Validation and Verification	20
<b>4</b>	<b>References</b>	<b>21</b>

## Abstract

Sign language recognition is a leap forward for helping deaf-mute people. In sign language, every sign has a meaning assigned to it, so that it becomes easy to understand and interpret by the people. The main primary objective of our project is to bridge the gap between the deaf and dumb people and ordinary people without the need of an intermediary translator. Gesture recognition is a computer based visual technique used to detect and identify the object in an image or video. An application of this gesture recognition technique involves translating the sign language into the American language which can be further understood and interpreted by normal people. Many researchers have proposed their systems for the implementation of the ASL system.

This report is a review of some studies related to the same topic. We addressed different approaches that used Convolutional Neural Networks (CNNs), K- Nearest Neighbors algorithm (KNN), Edge detection algorithm, Deep Neural Network (DNN), fuzzy clustering machine learning algorithm, Kernelized correlation filters (KCF) algorithm along with their results and drawbacks. Furthermore, we report on research gaps while summarizing these studies. From the above research papers, we got an average accuracy of 90.24%. To overcome the drawbacks, we propose a system which collects images for deep learning using webcam and OpenCV, with the help of TensorFlow Object Detection and Python that allows you to translate sign language in real time.

## **Acknowledgments**

We would like to express our heartfelt gratitude to our director Dr. Mohiuddin Ahmed and our principal Dr. Ganesh Kame, M.H. Saboo Siddik College of Engineering for providing us all the facilities, support, and wonderful environment to meet our project requirements. We take this opportunity to express our profound gratitude and deep regards to our internal mentor Mrs. Saiqa Khan for their exemplary guidance, monitoring and constant encouragement throughout the course of this project.

To list who all have helped us is difficult because they are so numerous and the depth is so enormous. We would like to acknowledge the following as being idealistic channels and fresh dimensions in the completion of this project. We take this opportunity to thank the M.H. Saboo Siddik College of Engineering for giving us a chance to do this project and for providing the necessary facilities required for completion of this project. We take this opportunity to thank our mentor for her moral support and guidance. We would also like to express our sincere gratitude towards our project guide whose guidance and care made the project successful. We would like to thank our College Library, for having provided various reference books and papers related to our project.

Lastly, we would like to thank each and every person who directly or indirectly helped us in the completion of the project especially our classmates and our friends who supported us throughout my project.

## List of Figures

<b>Fig. No.</b>	<b>Figures</b>	<b>Page No.</b>
3.1	The architecture of the Proposed System	9
3.2	Block diagram of the Proposed System	10

## List of Tables

Table. No.	Tables	Page No.
2.1	Summary of Studies Performed	7
3.1	Software Requirements	11



# INTRODUCTION

## 1.Introduction

Sign language (SL) is the basic means by which deaf and mute people communicate with the normal people. As normal people cannot understand the sign language used by dumb people, it is difficult to understand the message conveyed by the dumb/deaf people. Gesture recognition is a technology that uses sensors to read and interpret hand movements and translate them in the form of text which can be interpreted by the normal people.

Gestures are different movements employed in the communication process. Either the hand or the body makes gestures. Sign language uses gestures that typically use visually transmitted models. Gesture recognition is a computer-assisted visual technique that makes it possible to detect and identify an object in an image or video. An application of this gesture recognition technique implies translating the sign language into the American language which can be better understood and interpreted by normal people. Recognizing sign language is a step forward in helping people who are deaf-mute. In sign language, each sign has a meaning that is attributed to it, so that it becomes easy for the people to understand and interpret.

As there is need of a system which translate the sign language into text which is helpful for the normal people to understand. There is much need of such system to convey the message of deaf and dumb people. The current system which we generated is only based on hand gestures which will detect hand gesture and with the help of different algorithm interpret the text from the given sign.

### 1.2 Motivation

A language translator is extensively utilized by the mute people for converting and giving shape to their thoughts. A system is in urgent need of recognizing and translating sign language.

Lack of efficient gesture detection system designed specifically for the differently abled, motivates us as a team to do something great in this field. The proposed work aims at converting such sign gestures into speech that can be understood by normal people. The entire model pipeline is developed by CNN architecture for the classification of 26 alphabets and one extra alphabet for null character. The proposed work has achieved an efficiency of 99.88%.

### 1.3 Problem Statement & Objective

Sign language recognition is a leap forward for helping deaf-mute people. In sign language, every sign has a meaning assigned to it, so that it becomes easy to understand and interpret by the people. The main objective of our project is to bridge the gap between speech and the deaf and ordinary people.

We propose a system using Sign language recognition which can be implemented in the following way –

- Platforms used for conducting virtual meet.  
Example- Google meetings, Zoom, etc.
- It can be implemented in an app which will perform the task of converting sign language spoken by the dumb and deaf people to language which can be understood by the normal people.

### 1.4. Organization of report

The rest of the report is structured as follows: Section 2 contains the literature survey wherein a summary of various papers for sign language recognition is mentioned along with their limitations and research gaps. In addition, we have mentioned the research gaps we will be working on, highlighting our contribution to the research area. Section 3 contains details of the proposed system as well as hardware and software requirements, results and experiments, conclusions

# LITERATURE SURVEY

## 2.Literature Survey

### 2.1 Survey of Existing Systems

Amrutha K et al. performed ML Based Sign Language Recognition System in which they use the technique of convex hull method for feature extraction and finally KNN with Euclidean Distance for classification. They made candidates sign in front of the camera in a controlled environment and with the help of the same created dataset for each gesture. This model that they created showed an accuracy of 65%. The model showed less performance when the distance between the camera and the object is not considerable. The detection and recognition were less when the hand was moved at a fast pace.

Soma Shrenika et al. performed two techniques (1) Performing pre-processing steps on the image, that is, convert the acquired image, which is in RGB model to gray scale image. (2) Track the edges by using canny edge detection algorithm. (3) Edge detection algorithm was used to detect the sign in the image. American Sign Language dataset has numbers labelled from zero to nine and alphabets from a to z. This data set has 70 samples for each of the 36 symbols. There are 70 samples for each symbol. The process includes removal of noise and other less important data and applying smoothing algorithm to image and displaying the sign alphabet for the given gesture. The features of ASL are all of right hand only. There is no provision for left hand.

Diksha Hatibaruah et al. used Convolutional Neural Networks (CNNs) to train the system and Histogram Back Projection technique for segmentation of images. To train the system, they have used Indian SL database, consisting of 26 alphabets along with 10 digits. After training they achieved testing accuracy of 99.89% and validation accuracy of 99.85%

Qinglian Yang et al. carried out a gesture recognition system based on the Deep Neural Network (DNN) and the Leap Motion controller. In this paper, a total of 2000 frames of each gesture were collected from each of the 5 volunteers. Dataset is divided in two sets Training and Testing where training set is 20000, and the testing set is 10000. The average precision rate can reach more than 98% at the end of the training. And after 10000 paces, the precision of the model is excellent with the loss function remained below 0. 1. The limitation is that here we used the Leap motion controller while the work might also have been done simply using a webcam or any other technique. The use of the controller adds to the total cost of the project.

Muthu Mariappan H et al. performed Training and prediction of hand gestures by applying fuzzy clustering machine learning algorithm. The Regions of Interest (ROI) are identified and tracked using the skin segmentation feature of OpenCV. The data samples were collected for 80 words and 50 sentences of everyday usage terms of ISL. The videos were recorded from ten volunteers of our collaborator school, using a digital camera. This sign language recognition system, for recognizing the words of Indian Sign Language has produced 75 % accuracy in gesture labelling.

Saransh Sharma et al. implemented a system in which Skin color detection has been done in YCbCr color space. They have used Haar Cascade Classifiers and LBPH recognizer for face detection and recognition. A Database named SS100 has been created in which 6-10 images for each person has been stored, in total 100 images are stored. Each image has a different expression and postures. After the testing the model has accuracy of 95.2% and 92% in facial recognition. There are some limitations, which are needed to be addressed. Recognizing a greater number of gestures would be helpful for performing more tasks.

Hung-Yuan Chung performed Kernelized correlation filters (KCF) algorithm to track the detected Region of Interest and they have done Skin segmentation is with YCbCr to remove the unwanted background. For the dataset 800 images were collected for each hand gesture so a total of 4800 training images were used for the training model. The training data set can reach a recognition rate of 99.90%, and the test data set has a recognition rate of 95.61%. The system is limited to only 6 hand gestures.

Felix Zhan implemented a system in which he has used CNN classifier for dynamic hand gesture recognition and Spatio-temporal data augmentation techniques to get an additional 4000 images. The dataset used consisted of 500 images of 9 hand gestures using webcam to evaluate the model. This CNN classifier system showed an accuracy of 98.74% on the data set. The drawback of this system is that the dataset is limited to only 9 hand gestures.

Sr. No.	Study	Techniques	Dataset	Result	Drawbacks
1]	Amrutha K et al (2021)	The convex hull method was used for feature extraction and finally KNN with Euclidean Distance for classification.	Candidate was made to sign in front of the camera in an environment and with this dataset have been created for each gesture.	The tested model showed an accuracy of 65%. The accuracy can be increased further.	The model showed less performance when the distance between the camera and the object is not considerable. The detection and recognition were less when the hand was moved at a fast pace.
2]	Soma Shrenika et al (2020)	1. Performing pre-processing steps on the image, that is, convert the acquired image, which is in RGB model to Gray scale image. 2. Track the edges by using canny edge detection algorithm. 3. Edge detection algorithm was used to detect the sign in the image.	American Sign Language dataset has numbers labelled from zero to nine and alphabets from a to z. This data set has 70 samples for each of the 36 symbols. There are 70 samples for each symbol.	The process includes removal of noise and other less important data and applying smoothing algorithm to image and displaying the sign alphabet for the given gesture.	The features of ASL are all of right hand only. There is no provision for left hand.
3]	Diksha Hatibaruah et al (2020)	1.Convolutional Neural Networks (CNNs) is used to train the system 2.Histogram Back Projection technique for segmentation of images	To train the system, we have used Indian SL database, consisting of 26 alphabets along with 10 digits	After training we find out the testing accuracy of 99.89% and validation accuracy of 99.85%	The limitation is that the background must be light with good lighting conditions. Also, Individual letters cannot be concatenated to form meaningful words. In addition, it is applicable only for static gestures.
4]	Qinglian Yang et al (2020)	Gesture recognition system based on Deep Neural Network (DNN) and Leap Motion controller	1. In this paper, a total of 2000 frames of each gesture were collected from each of the 5 volunteers. 2. Dataset is divided in two sets Training and Testing where training set is 20000, and the testing set is 10000.	The average accuracy rate can basically reach more than 98% at the end of the training. And after 10,000 steps, the accuracy of the model is excellent with the loss function remained below 0.1.	The limitation is that here we have used Leap motion controller whereas the work could have also been done by simply using a webcam or any other technique. The use of the controller increases the overall cost of the project.
5]	Muthu Mariappan H et al (2019)	1. Training and prediction of hand gestures are performed by applying fuzzy clustering machine learning algorithm 2. The Regions of Interest (ROI) are identified and tracked using the skin segmentation feature of OpenCV	The data samples are collected for 80 words and 50 sentences of everyday usage terms of ISL. The videos are recorded from ten volunteers of our collaborator school, using a digital camera.	This sign language recognition system, for recognising the words of Indian Sign Language has produced 75 % accuracy in gesture labelling.	FCM is efficient, it requires more computation time than the others.
6]	Saransh Sharma et al (2019)	1. Skin colour detection has been done in YCbCr colour space	A Database named SS100 has been created in which 6-10 images for each	The model has accuracy of 95.2% and 92% in facial recognition	There are some limitations, which are needed to be addressed. Recognizing a greater

		2.Haar Cascade Classifiers and LBPH recognizer are being used for face detection and recognition	person has been stored, in total 100 images are stored. Each image has a different expression and postures.		number of gestures would be helpful for performing more tasks
7]	Hung-Yuan Chung (2019)	1. Kernelized correlation filters (KCF) algorithm to track the detected Region of Interest. 2. Skin segmentation is done with YCbCr to remove the unwanted background.	800 images were collected for each hand gesture so a total of 4800 training images were used for the training model.	The training data set can reach a recognition rate of 99.90%, and the test data set has a recognition rate of 95.61%	The system is limited to only 6 hand gestures
8]	Felix Zhan (2019)	1. CNN classifier for dynamic hand gesture recognition 2. Spatio-temporal data augmentation techniques to get an additional 4000 images	500 images of 9 hand gestures using webcam to evaluate the model. Each image is a 50x50 pixels.	800 images were collected for each hand gesture so a total of 4800 training images were used for the training model .	The dataset is limited to 9 hand gesture only.

Table 2.1. Summary of Studies Performed



## 2.2 Limitations of Existing System and Research Gaps

The systems discussed in the earlier section proved to be fruitful in their respective requirements. However, there are research gaps that we discuss in this section. Following are the limitations of the existing systems for fruit counting and detection.

- The detection and recognition were less when the hand was moved at a fast pace.
- Individual letters cannot be concatenated to form meaningful words
- The background must be light with good lighting conditions.
- The features of ASL are all of right hand only. There is no provision for left hand.
- There are some limitations, which are needed to be addressed. Recognizing a greater number of gestures would be helpful for performing more tasks

## 2.3 Mini Project Contribution

This mini project is anticipating to contributing to the social field by proposing a solution for sign language recognition in video conference platforms. We have in view to work on implementing this system globally in video conference platforms which was not considered in earlier systems.

In addition to the above points, we also intend to train our system on images taken from different distances and every possible angle. The acquisition of images from varied distances will help us train the system on a variety of data and help us in ensuring adequate accuracy.

# PROPOSED SYSTEM

## 3. Proposed System

### 3.1 Introduction

The proposed methodology for the project is to create a dataset of images and to train the system i.e., reinforcement learning method of machine learning to provide a more accurate output the captured image will process for the region of interest of hand gesture and the decision will be taken. This Proposed system can therefore be used in real-time applications which perform the task of converting sign language spoken by the dumb and deaf people to language which can be understood by the normal people. This project can be integrated with platforms used for conducting virtual meet.

Example- Google meetings, Zoom, Microsoft teams etc.

### 3.2 Architecture/Framework

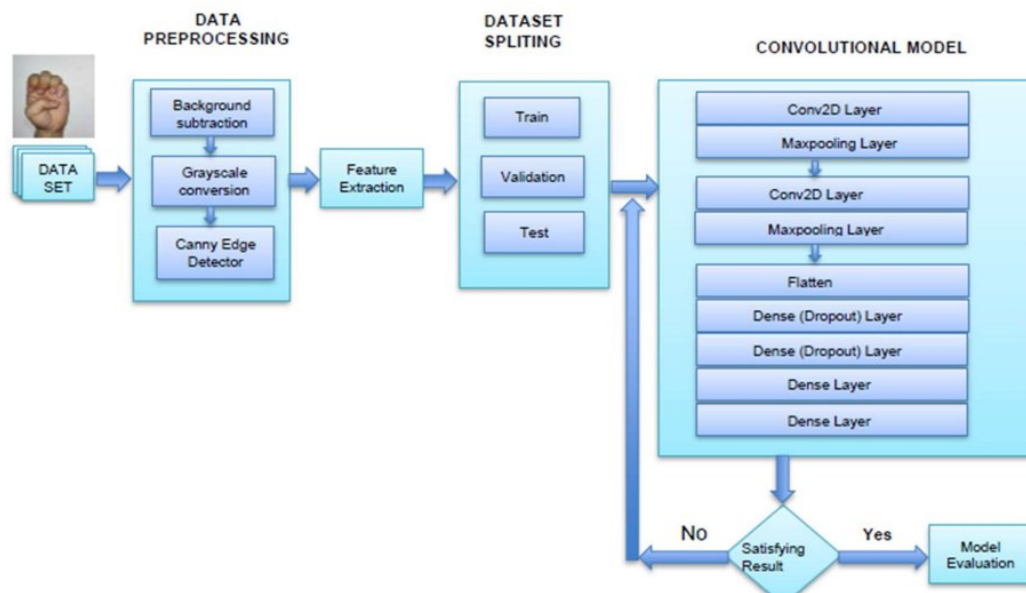


Figure 3.2. Block Diagram of the Proposed System

### 3.3 Algorithm and Process Design

1. Import pretrained model(asl\_classifier.h5).
2. Create a label dictionary.
3. Taking input from user.
4. Converting image from RGB to Grayscale.
5. Applied gaussian blur to images.
6. Separating foreground and background using Adaptive Threshold.
7. Images are then resized using resize function of OpenCV.
8. Predicting the result using loaded pretrained model.

### 3.4 Details of Hardware and Software

#### Hardware:

The input of this system will be taken by the web-cam of the user's Laptop. Thus, the hardware component in this proposed system is the web-cam for image capturing as the input.

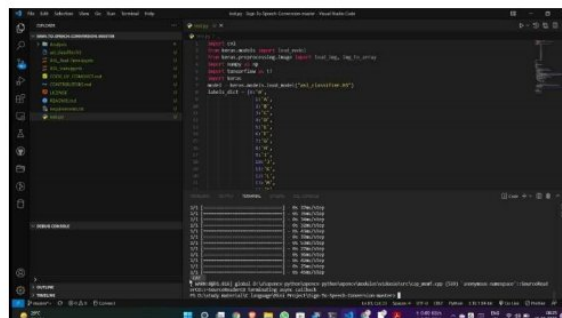
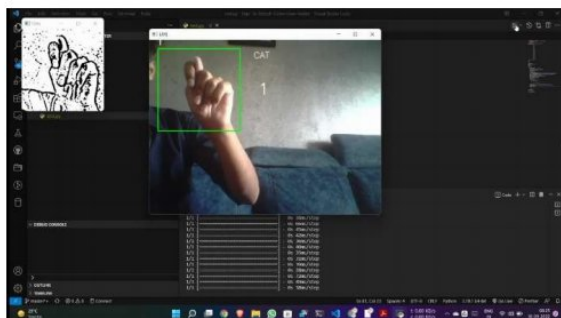
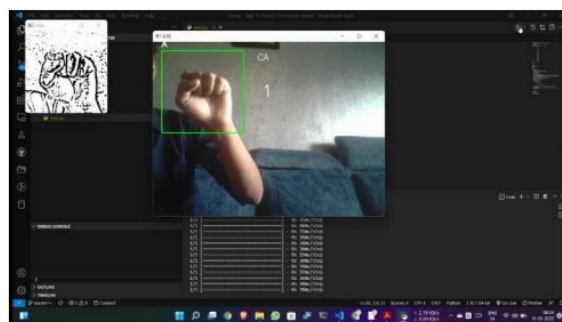
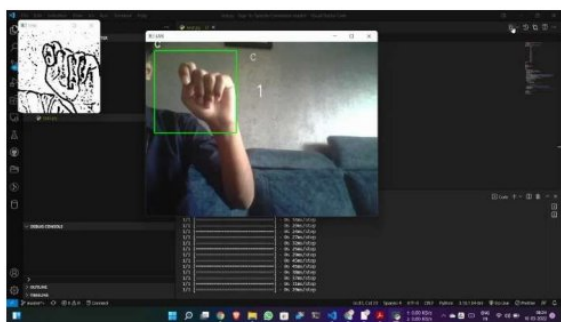
#### Software:

Table 3.1 lists the software requirements for the project.

Table 3.1. Software Requirements

<b>Hardware</b>	Web-Cam, Laptop
<b>Coding Language</b>	Python
<b>External Tools</b>	<ul style="list-style-type: none"> <li>Tensorflow-1.15.0</li> <li>OpenCV</li> <li>Keras-2.1.6</li> <li>H5py-2.10.0</li> <li>Pycocotools</li> <li>Pandas</li> </ul>
<b>System Requirement</b>	<ul style="list-style-type: none"> <li>Python 3.5-3.7</li> <li>pip 19.0 or later</li> <li>Microsoft Visual C++ Redistributable for Visual Studio 2015, 2017 and 2019</li> <li>Windows 7 or later</li> <li>GPU support requires a CUDA<sup>®</sup>-enabled card (Ubuntu and Windows)</li> </ul>

### 3.5 Experiment and Result



## 4.References

- [1] Amrutha K et al " ML Based Sign Language Recognition System" ©2021 IEEE | DOI: 10.1109/ICITIIT51526.2021.9399594
- [2] Sona Shrenika et al "Sign Language Recognition Using Template Matching Technique" ©2020 IEEE | DOI: 10.1109/ICCSEA49143.2020.9132899
- [3] Diksha Hatibaruah et al "A Static Hand Gesture Based Sign Language Recognition System using Convolutional Neural Networks" ©2020 IEEE | DOI: 10.1109/INDICON49873.2020.9342405
- [4] Qinglian Yang et al "Leap Motion Hand Gesture Recognition Based on Deep Neural Network," ©2020 Chinese Control And Decision Conference (CCDC), 2020, pp. 2089-2093, DOI: 10.1109/CCDC49329.2020.9164723.
- [5] H. Muthu Mariappan et al "Real-Time Recognition of Indian Sign Language," ©2019 International Conference on Computational Intelligence in Data Science (ICCIDS), 2019, pp. 1-6, DOI: 10.1109/ICCIDS.2019.8862125.
- [6] Saaransh Sharma et al "A Static Hand Gesture and Face Recognition System for Blind People," ©2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), 2019, pp. 534-539, DOI: 10.1109/SPIN.2019.8711706.
- [7] Hung-Yuan Chung "An Efficient Hand Gesture Recognition System Based on Deep CNN," ©2019 IEEE International Conference on Industrial Technology (ICIT), 2019, pp. 853-858, DOI: 10.1109/ICIT.2019.8755038.
- [8] Felix Zhan, "Hand Gesture Recognition with Convolution Neural Networks," ©2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), 2019, pp. 295-298, DOI: 10.1109/IRI.2019.00054.