International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 01 | Jan 2022 www.irjet.net p-ISSN: 2395-0072

# Sign Language Recognition using Mediapipe

**Ketan Gomase[1], Akshata Dhanawade[2], Prasad Gurav[3], Sandesh Lokare[4]**

[1,2,3,4]*Student, Department of Electronics & Telecommunication Engineering, Atharva College of Engineering, Mumbai, Maharashtra, India*

---

**Abstract -** *Lack of speech is considered to be a real disability. People with these disabilities use a variety of methods to communicate with others, there are many forms available for their communication, one of the most common forms of communication as well as sign language. Sign language is used by deaf and hard of hearing people to share information with their community and others. Electronic recognition of sign language deals from signalling to touch and continues until text / speech production. Touch gestures can be classified as permanent and flexible. Steps in recognizing sign language are described in this study. Data acquisition, pre-processing and modification of data, feature extraction, segmentation and obtained results are assessed. Future research guides in this area are also recommended.*

***Key Words***: *Mediapipe, Sign language recognition [SLR], KNN, Hand Solution, Computer Interaction with Humans.*

## 1.INTRODUCTION

Sign language using sign language is designed for the deaf community, which can be used as a means of communication between friends and family of the deaf and the deaf. Sign Language Recognition is one of the fastest growing and challenging areas of research today. Many new strategies have been developed recently in this field. In this project, we will develop a program to translate sign language into OpenCV. It outlines a method that recognizes American Sign Language (ASL) and translates it into standard text.

### 1.1 Motivation and Background

Sign Language Recognition strive to develop algorithms and methods for accurately identifying the sequences of symbols produced and understanding their meaning. Many SLR methods mistreat the problem as Gesture Recognition (GR). Therefore, research has so far focused on identifying the positive characteristics and methods of differentiation in order to properly label a given signal from a set of potential indicators. However, sign language is more than just a collection of well-articulated gestures.

### 1.2 What is Gesture recognition?

Gesture recognition is a subject in computer science as well language technology for the purpose of translating a person touch with mathematical algorithms. The subdiscipline of computer vision. Gestures can come from any body movement or position but usually appears on the face or hand. The current focus in the field includes emotional recognition from facial and hand touch recognition. Users can use simple touch to control or interact with devices without the touch touching them. Many methods have been developed using cameras and computer vision algorithms to translate the signal language.

### 1.3 Sign Language

Sign languages [also known as sign languages] are languages that use visual cues to convey meaning. Sign languages are expressed in sign language as well as non-sign language objects. Sign languages are complete natural languages with their own grammar and dictionary. Sign languages are not universal and are not widely understood, although there are some striking similarities between sign languages. Linguists consider both spoken and signed communications to be natural forms of language, meaning that they both evolved into a vague aging process, one that lasted longer and evolved over time without careful planning. Sign language should not be confused with body language, a form of communication without voice.

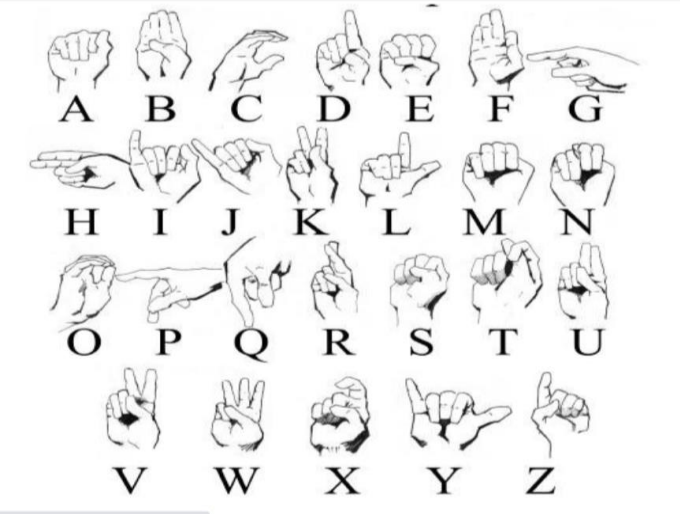


Figure 1: Sign Languages

### 1.4 Mediapipe Framework

Mediapipe Hands is a reliable hand and finger tracking device solution. It uses machine learning (ML) to understand 21 3D local hand marks from just one frame. Although modern methods depend largely on the powerful desktop locations for discovery, our approach benefits real-time performance on mobile phones, even scales to many hands. We hope to give you this handy idea working on extensive research and development society will result in cases of

misuse, to promote new applications and new research methods. Mediapipe Hands uses an integrated ML pipe of the many models working together: The palm detection model which works on the full image and returns the direct-directed hand binding box. Hand gesture model applicable to image-cut region defined by a palm detector once returns 3D hand key points with high reliability. This strategy is similar to the one used in our Mediapipe Face Mesh solution, using a face detector and a face detector a landmark model.

### 1.5 Objective

The objective of this project was to create a neural network that could distinguish between the American Sign Language (ASL) alphabet characters, if a handwritten signature is provided. This project is the first step in creating a potential sign language translator, which can take communication in sign language and translate it into written and oral language. Such a translator can greatly reduce the barrier between many deaf and hard of hearing people so that they can better communicate with others in their daily activities.

### 1.6 Summary

Improving sign language application for the deaf can be it is very important, as they will be able to easily communicate with them and those who do not understand sign language. Our program aims to take a basic step to close the connection the gap between the common people and the deaf and dumb using sign language. The main focus of this work is creativity a vision-based system for identifying spelled characters ASL. Reason for choosing a vision-based system has to do with the fact that it provides a simple and accurate way how to communicate between a person and a computer. In this report, the various stages of 36:26 were considered sections of the English Alphabets (a to z) We have used Google's Mediapipe Framework Mediapipe Solutions has improved hand recognition model and can find 21 3D Landmarks of Palm

## 2. LITERATURE SURVEY

In recent years, much research has been done on sign language recognition. This recognition technology is divided into two categories: -

### 2.1 Vision Based Approach

This method takes pictures on camera as touch data. The vision-based approach focuses heavily on touch-captured images and brings out the main and recognizable feature. Colour belts were used at the beginning of the vision-based approach. The main disadvantage of this method was the standard colour to be applied to the fingers. Then use bare hands instead of coloured ribbons. This creates a challenging problem as these systems require background, uninterrupted lighting, personal frames and a camera to achieve real-time performance. In addition, such systems must be developed to meet the requirements, including accuracy and robustness.

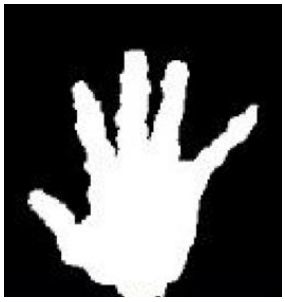

Figure 2: Sample of Vision Based Technique

Theoretical analysis is based on how people perceive information about their environment, yet it is probably the most difficult to use effectively. Several different methods have been tested so far. The first is to build a three-dimensional human hand model. The model is compared to hand images with one or more cameras, and the parameters corresponding to the shape of the palm and the combined angles are estimated. These parameters are then used to create the touch phase. The second is to take a picture using the camera and extract certain features and those features are used as input in the partition algorithm to separate.

## 3. IMPLEMENTATION METHODOLOGY

Mediapipe Hands uses a Machine Learning Pipeline that integrates multiple co-working models: A palm-type acquisition model that works in a complete image and returns a fixed hand-held binding box. A handwriting model that works with a cropped image location defined by a palm detector and restores 3D reliable key points.

So, to make a Web site we have to photograph at least 25-30 images per mark and with this model we can get 21 hand points. i.e., links [x, y, z]. x and y are common to say [0.0, 1.0] the width and height of the image respectively. The z represents the depth of the landmark and the depth of the arm at the root, and the smaller the value the closer the camera becomes. After making the Website can predict the sign with the help of the Appropriate Model. We will use the KNN algorithm.

### 3.1 Hardware & Software Requirement

1) Windows computer or Linux, Python installed and Libraries.

2) CMOS sensor (Webcam)

3) Hand Touch for Visibility

Computer Software We Used to Recognize Project Signature Recognition:

1) Python Installed Windows Os or Linux Os Machine.

2) CPU - Intel core i5 9th Gen.

3) GPU - Nvidia GTX 1050 Ti.

4) 720p60 Web Camera

5) Python 3.8.6 and IDE like VS, Spyder etc.

6) Libraries: OpenCV, TensorFlow, Keras, Mediapipe and many more basic *

7) KNN (The closest neighbours) from the Sklearn Library of Python.

### 3.2 Result

This sign language receiver can detect hand and produce co-ordinators and will be able to recognize letters (A-Z). All signs will appear in real time.

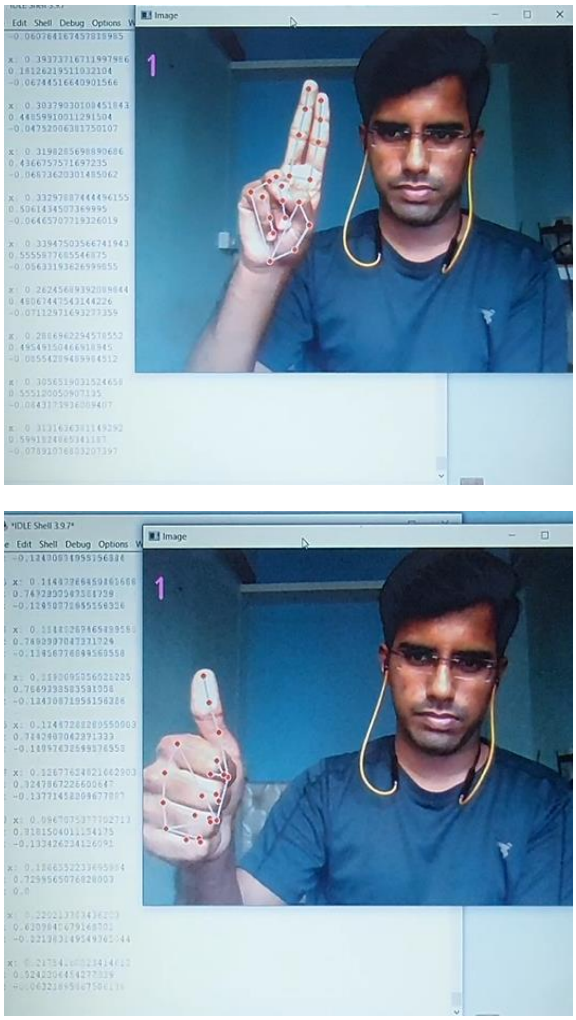

Figure 4 & 5: Co-Ordinates Generated

## 4. CONCLUSIONS

Sign Language using Mediapipe and recognition through Computer vision was partially successful and accurate an average of 17 FPS with an average accuracy of 86 to 91%. The question of perfection is another attempt to deal with it in the days to come. One hand touch detection recognition was the theme and the biggest problem with which it worked with. Mediapipe achieves an average accuracy of 95.7% palm discovery. Using the normal loss of cross entropy and not. The decoder provides an 86.22% base. So the Scope of the Future of this study will be Good to improve Human Computer Interoperability (HCI) using a very powerful and fast algorithm.

## REFERENCES

[1] IJTRS-V2-I7-005 Volume 2 Issue VII, August 2017 survey on hand gesture techniques forsign language recognition.

[2] Umang Patel & Aarti G. Ambekar “Moment based sign language recognition for IndianLanguages “2017 Third International Conference on Computing, Communication, Controland Automation (ICCUBEA).

[3] Lih-Jen kau, Wan-Lin Su, Pei-Ju Yu, Sin-Jhan Wei “A Realtime Portable Sign LanguageTranslation System'' Department of Electronic Engineering, National Taipei University ofTechnology No.1, Sec. 3, Chung-Hsiao E. Rd., Taipei 10608, Taiwan, R.O.C.

[4] Obtaining hand gesture parameters using Image Processing by Alisha Pradhan andB.B.V.L. Deepak ,2015 International Conference on Smart Technology and Management(ICSTM).

[5] Vision based sign language translation device (conference paper: February 2013 byYellapu Madhuri and Anburajan Mariamichael).

[6] K-nearest correlated neighbour classification for Indian sign language gesture recognition using feature extraction(by Bhumika Gupta, Pushkar Shukla and Ankush Mittal).

[7] Vikram Sharma M, Virtual Talk for deaf, mute, blind and normal humans, Texas instruments India Educator’s conference,2013.

[8] Hand in Hand: Automatic Sign Language to English Translation, by Daniel Stein, Philippe Dreuw, Hermann Ney and Sara Morrissey, Andy Way.

[9] Er. Aditi Kalsh, Dr. N.S. Garewal “Sign Language Recognition System'' International Journal of Computational Engineering Research 2013

[10] Prakash B Gaikwad, Dr. V.K.Bairagi,” Hand Gesture Recognition for Dumb People using Indian Sign Language” , International Journal of Advanced Research in computer Science and Software Engineering, pp:193-194, 2014.

[11] https://en.wikipedia.org/wiki/Sign_language