International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

Sign Language to Text Converter Using Hand Gestures

Akanksha Pawar, Harshada Bhosale, Mansi Diwan
Department of Computer Engineering, Zeal College of Engineering & Research, Pune, Maharashtra, India

Article Info
Volume 8, Issue 2
Page Number : 435-439

Publication Issue :
March-April-2022

Article History
Accepted: 10 April 2022
Published: 22 April 2022

ABSTRACT

Sign Language is been used by hearing impaired people who are unable to interact through speech with normal people. According to the WHO study currently around more than 1.5 billion people live with hearing loss all over the world. Without sign language it becomes hard to interact with the hearing-impaired person as common people are unaware of sign language and gestures. Hence communication becomes hard. There is a need for a intermediate to translate what they want to express. Our work aims to improve the communication with the hearing-impaired person. This paper does American sign language recognition using google's Mediapipe Hands API to detect landmarks on the hands to detect gestures and uses Support Vector Machine (SVM) classification algorithm for achieving high accuracy. The model works on web application. The proposed concept (sign language conversion to text, speech and image) is a breakthrough for helping the hearing-impaired community.

Keywords—Support Vector Machine, Image Processing, Mediapipe, Gesture Recognition, Machine Learning.

I. INTRODUCTION

Sign Language allows the deaf and hearing impaired people to communicate with each other or a normal person. There are consist of over 300 different sign languages around the world in which the top 4 used sign languages are American Sign Language (ASL), British, Australian and New Zealand Sign Language (BANZSL), Chinese Sign Language (CSL) and French Sign Language (LSF). 1.5 billion people worldwide have impaired hearing loss, according to the World Health Organization (WHO).With the growing adoption of telehealth, hearing impaired people should be able to communicate naturally with their healthcare network, colleagues and peers regardless of whether the second person knows sign language.

Why one should learn sign language? It has many benefits such as it boosts the brain functioning. It strengthens our cognitive function and gives a great impact on improving factors like reasoning, increases memory, attention span, creativity, and communication skills. Sign language can be easier to learn than any other spoken languages not only for the hearing impaired person but also for people suffering from disorders such as autism, dyslexia, Aphasia, Cerebral Palsy, etc.

In this project we will be using python and its various libraries to perform American sign language recognition system. It includes libraries such as Media

pipe Hands API for landmarks, OpenCV for image processing, Numpy to perform and store operations related to array, Pandas, SKlearn. It also includes database to store user credentials. So the database used is firebase database. We have also used Flask Framework as it is used for developing web applications, it is highly flexible and gives higher performance. Including this it also highly compatible with latest technologies. The system is compatible with Windows 7 and above. Further it uses Support Vector Machine algorithm for classification it uses .

In vision-based gesture recognition, a camera is used as input. Videos are broken down into frames before processing. Hence vision based methods are preferred over gesture-based approaches as anyone with a smartphone can convert sign language to text/speech and it is relatively cost- effective.

II. METHODOLOGY

Sign language is used by deaf and dumb people for communication. It's a bundle of hand gestures, facial expressions, and body language that helps them to communicate with the rest of society. The system is based on American Sign Language in which we have used 26 letters trained and recognized resulting in text and voice out of that particular letter. The system works on Support Vector Machine(SVM).

A. Dataset

We used supervised learning modal for training our dataset which predicts the right output for a particular letter in sign language. We added our own set of alphabets A-Z to the csv file and used them as training data. Initially we checked on the accuracy of self generated images of dataset and tested them manually. Our dataset contains total 100 images as SVM algorithm works more precisely with smaller datasets. The dataset used for this work was made by the team themselves since there was no readily available dataset for ISL that fulfilled all the requirements of the work.

A python script was written for collecting the data from all the team members. The script was capable of detecting and extracting the landmarks of one or two hands using libraries like openCV and mediapipe.

Fig. 1. American Sign Language

1) Dataset for single handed gestures:

For each time a gesture was shown in the camera, the Mediapipe API returned the collection of handedness of the detected or tracked hands i.e is it a left or right hand) and collection of detected or tracked hands, where each hand is represented as a list of 21 hand landmarks and each landmark is composed of x, y and z. Using the count of detected hands, the python script written for data collection sends the data to the function that handles one handed gestures data pre-processing. Only the x and y coordinates are used from all the 21 landmarks detected by Mediapipe Hands API.

These coordinates were inserted in a list in proper order preceded by the alphabet that is represented by the gesture. More than 15000 such records were dumped in the CSV file for each character.

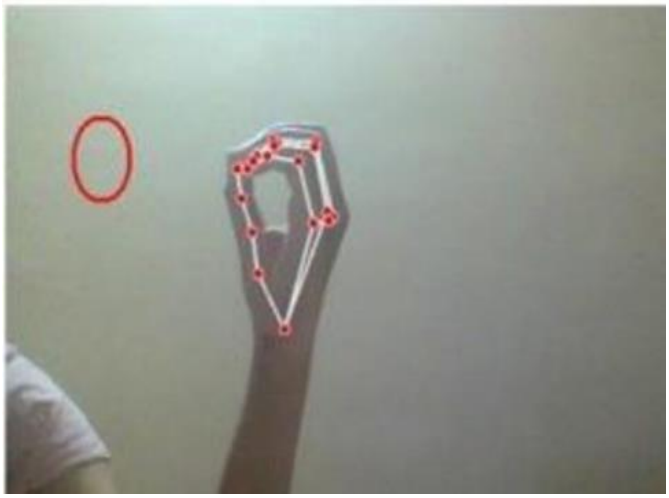

Fig. 2. Login Successful

performance evaluation, and (c) a collection of re-usable inference and processing components called calculators. Hand tracking using MediaPipe involves two stages:

Palm detection - MediaPipe works on the complete input image and provides a cropped image of the hand.
Hand landmarks identification - MediaPipe finds the 21 hand landmarks on the cropped image of the hand.
Hand tracking using MediaPipe involves two stages:

Fig. 3. Dataset

B. Hand Detection

The palm detection model from Mediapipe handles the whole hand detection process. The webcam feed is fed into the Mediapipe API, which returns a list of detected/tracked hands, each of which is defined by a list of 21 hand landmarks, each of which is made up of x, y, and z. The landmark depth is represented by z, with the origin being the depth at the thumb, and the smaller the value, the closer the landmark is to the camera. It also returns a list of the detected/tracked hands' handedness, as stated in section

1) Image processing: Image processing basically includes the following three steps:

- Importing the image via image acquisition tools.
- Analysing and manipulating the image.
- Output in which result can be altered image or report that is based on image analysis.

2) Mediapipe Hands API : MediaPipe's system achieves real-time efficiency on a mobile phone, and also scales to several hands, while existing state-of-the-art methods focus largely on strong desktop environments for inference. MediaPipe consists of three main parts: (a) a framework for inference from sensory data (b) a set of tools for

C. Classification

The image dataset's attributes serve as the learning model's training data. This knowledge can be used to train classification models such as SVM, Random Forest, K-nearest neighbours KNN, Decision Tree, and Neural Networks. The alphabets can be predicted using the trained model.

1) SVM: A support vector machine (SVM) is a type of deep learning algorithm that performs supervised learning for classification or regression of data groups.

Support vector machines are used to sort two data groups by like classification. The algorithms draw lines (hyperplanes) to separate the groups according to patterns. Like other supervised learning machines, an SVM requires labeled data to be trained. It is a great algorithm to choose when you are working with smaller datasets that have tens to many thousands of features. They typically find more accurate results when put next to other algorithms due to their ability to handle small, complex datasets.

III. IMPLEMENTATION DETAILS

This project includes real time detection of American Sign Language. It detects all alphabets from A-Z resulting in text and and speech. The hearing impaired kids will be able to learn the basics.

Fig 4. includes the first page so the user can login to the sytem in which on enering the correct credentials the login is successful it access the firebase database for data.

Fig 5 shows failed login for entering the wrong credentials.

Fig 6 includes sign up to register as a member for accessing the system services.

Fig 7 includes the main page during the real time detection of the gestures where a user performs the gesturein front of the web camera the cam accesses the image through frame and further the image goes fo the internal processing and classification which is hided from the user. It also includes the reference chart for American Sign Language for beginners.

Fig 8 includes the final output of the system gone through number of processes like image processing, working with various libraries to fulfill the needs, SVM classification and fetching and matching the data to the dataset for detection. It results in a pop window the shown UI with "A for apple" sound at the background. Similarly it works for other Alphabets.

Fig. 4. Login Successful

Fig. 5. Login Failed

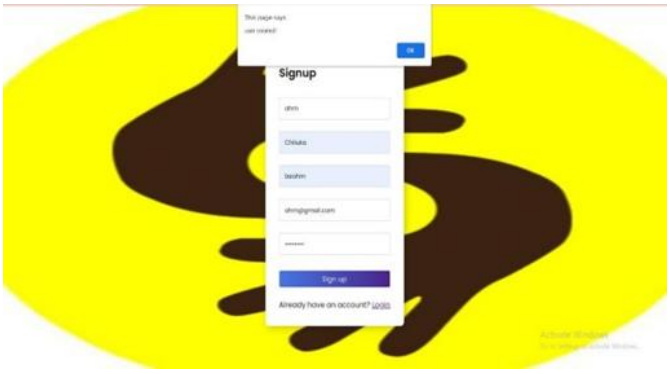

Fig. 6. Signup

IV. RESULTS AND DISCUSSION

The project demonstrates how the SVM algorithm helps us to attain the challenges required for solving the problem of communication with hearing impaired people which have no basic knowledge about sign language communicaion. The hearing impaired person can hereby learn the basics of American sign language through sign language to text converter platform.

Further the system should be able to achieve more and more accuracy with bulk of more data to detect. The system must also detect numbers, perform classification for various actions such as hello, bye, etc. usually used for greeting with the other person.

V. CONCLUSION AND FUTURE SCOPE

This paper compares different techniques and chooses the most optimal approach for creating a vision-based application for sign language to text/speech conversion for hearing imapired people. The proposed system could efficiently recognize the alphabets from images using a customized SVM model. The system will provide an interface which will easily communicate with deaf people by sign recognition. For the social use, this technique is extremely helpful for hearing impaired people. The project focuses on distinguishing among various different alphabets of English language. Also this system mainly emphasis on kids who have hearing loss, so that they can learn alphabets through sign language by themselves as the system is very user friendly.

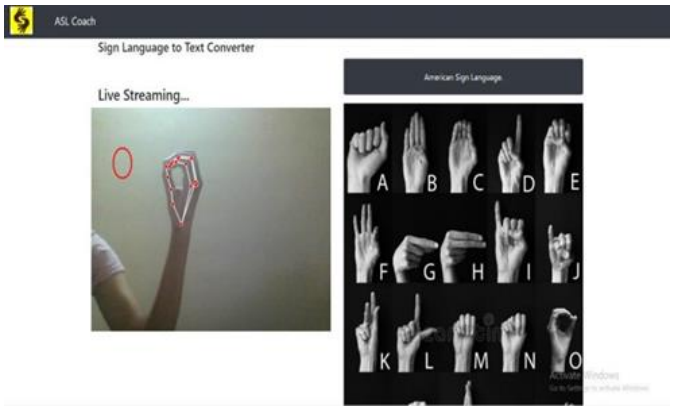

Fig. 7. Real Time Detection

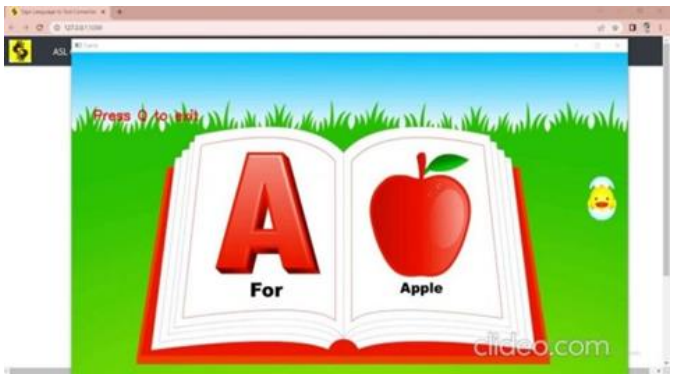

Fig. 8. Detection of 'A'

Nowadays, applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted so as to perform various applications. Sign language recognition System is a powerful tool to prepare an expert knowledge , edge detect and the combination of inaccurate information from different sources.

As of now the proposed system includes on Alphabets A-Z, but in future we are planning to add numbers, continuous action, expressions to form a sentence and works like good morning, hello, bye, etc. so that the people can learn quickly.

VI. REFERENCES

[1]. Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh and Indra R."Realtime Conversion of Sign Language to Text and Speech". Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020).

[2]. Subhalaxmi Chakraborty, Nanak Bandyopadhyay, Piyal Chakraverty, Swatilekha Banerjee, Zinnia Sarkar, Sweta Ghosh."Indian Sign Language Classification (ISL) using Machine Learning ".American Journal of Electronics and Communication, Vol. I (3), 17-21.

[3]. Ketan Gomase1, Akshata Dhanawade2, Prasad Gurav3, Sandesh Lokare4."Sign Language Recognition using Mediapipe". Volume: 09 Issue: 01 | Jan 2022.

[4]. Adarsh Vishwakarma1 Niraj Yadav2 Prajnay Yadav3 Vaibhav Singh4 "Sign LanguageRecognition Using Mediapipe Framework with Python" (Vol. 9, Issue 3, 2021