Synopsis
On
Real-Time Sign Language Detection

Submitted By : Aastha Shreya
Deeksha Goyal

Branch/Sec : B. Tech (CS) / H

University Roll NO. : 191500006
191500236

Submitted To : Mr. Amir Khan
Faculty Name : Mr. Amir Khan

INTRODUCTION

Name of our project is Real-Time Sign Language Detection, which basically will convert Sign Language to Text and speech in real-time. Main motive to make this project is to make communication flow smoothly between differently-abled and abled people.

Sign Language is a visual-gestural language used by deaf or hard hearing people to exchange information between their own community and with other people. It involves hand movements and other parts of the body. As spoken languages use the oratory faculties(involves tongue, larynx, lips etc.) to produce frequency against specific words and grammatical combinations to convey meaningful information, and received by auditory faculties(Ear) and processed accordingly. In the same way sign language also works with the combination of complex grammar, hand gestures and position of head and body parts.

There was no formal Indian Sign Language until 1978, observing the advantage of sign language in aiding the deaf people for communication in public places and latest gadgets, Indian Sign Language started in 1978 in India only.

Although Sign Language is issued, many well hearing people do not bother to learn it and those who can not see need translation of the same. So to create smooth communication between these groups of people we need some mechanism, and that mechanism can be Real-Time Sign Language to speech assuming the person is not as well.

Further, In this project we will be implementing different Deep Learning and Machine Learning models to effectively convert the sign language to text/speech. For the time being our project works only on Indian Sign Language, It can be extended to different languages as well, so that our motive of making communication smooth doesn't get affected.

EXISTING SYSTEM

Over the past 20 years, researchers have used classifiers from a variety of categories that can be classified roughly into linear classifiers, neural networks and Bayesian networks. Brandon Garcia and Sigberto Alarcon Viesca were able to implement a robust model for the letters a-e and a modest one for letters a -k.

The Surrey and Massey datasets on ASL along with the GoogleNet architecture The system takes the input of the user's sign language video, classifies the actions for each letter and then tries to come up with the most accurate words. The factors such as lighting conditions and sign language border detection are considered while designing the project. They attained a validation accuracy of nearly 98% with five letters and 74% with ten letters.

CNNs have seen a rapid improvement in image classification with many proposed models like GoogleNet, AlexNet giving an accuracy almost near to human perception. The main cause of the recent improvement in CNNs has been due to the ImageNet Large Scale Visual Recognition Competition(ILSVRC). For image processing it is proposed to use the OpenCV library along with TensorFlow and Keras which will be used for training the classifier. For other mathematical calculations the NumPy Array in python can be used. Steps include:

1. Softmax Regression
2. Recognizing Hand Gestures
3. Neural Networks
4. Haar Cascade
5. Deep Learning
6. Keras and Tensorflow Python Libraries

USE OF THE PROJECT

This project has been decided keeping in view the differently abled people of the society to equally get involved in general conversation with those who are abled. For instance, one who can not understand Hindi language movies can use a translator to convert into their preferred language, in the same way this project will convert Sign Language to Spoken Language.

This would help a large section of people convey their message smoothly to those people who don't understand Sign Language. In this way, such a section of people can get involved in regular conversations. Also, there are section of people who can not speak but they have great skill to work at certain job posts and there is possibility that large number of people don't understand Sign Language in that case it would make them aloof from others and may lead to loss of their job, but using this application can help them express themselves in easy and effectively.

There are several people who are visually impared but not deaf or hard of hearing, in that case speech conversion will help them understand what the person in front is conveying to them.

FEASIBILITY OF PROJECT

To convert Sign Language to text most effective methods involves are explained in the further sections :

Softmax Regression : Softmax regression allows us to handle y(i) ϵ {1,...,N} where N is the number of classes.

Recognizing Hand Gestures : Hand gestures can be recognized using several methods. Two of the most widely used methods are: Haar Cascades and Neural Networks.

Neural Networks : Neural networks are more like brain neurons which help in solving complex problems of pattern recognition better than procedural algorithms.

Haar Cascade : we train a cascade function on many positive and negative images. It can therefore be used to detect objects from other images.

Deep Learning : Deep learning is used to mask the levels of abstraction in a machine learning algorithm. It contains a set of hidden layers each using the output of the previous layer for better feature extraction and pattern recognition. We have implemented deep learning using Keras deep learning which contains Tensorflow.

Also, we will try using YOLOV5 to train a custom video dataset by applying different lighting and features and see what will be the result, the efficient way would be selected for completing the project.

FUNCTIONAL SPECIFICATION

The working or the use of the project is very simple, the user only has to input their sign language gestures through their devices' camera then this input would go through the complex processing used and will provide the user with text and audio conversion seen on screen and heard through speaker of the device respectively, in real-time. Points to be considered while passing input is to make sure that camera covers the user completely from head to chest.

SOFTWARE SPECIFICATION

Technology Used : Deep Learning and Machine Learning
Language Used : Python

Hardware Requirement

Processor : Intel Pentium 4 or higher
Operating System : Windows, Mac, Linux
Ram : Minimum 8GB or higher
Hardware Devices : Camera (good quality 3MP)
Hard Disk : 10GB or higher
Display : 10" or 17" Monitor

FUTURE SCOPE

The project now has time latency and if further research and work put in will make it more vastly used among differently-abled and abled persons to make smooth flow of communication.
This only works only for Indian Sign Language further it can be enhanced and extended for different countries sign language to make it versatile. And if possible I would be more than pleased to convert this as an application for easy and portable use.