IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

Sign Language Translator Using Deep Learning

Jotdeep Singh[1], Kiran Devkar[2], Gaurav Manjwani[3], Archit Chandra[4], Prof. S. P. Bholane[5]
Students, Department of Computer Engineering[1,2,3,4]
Faculty, Department of Computer Engineering[5]
Sinhgad College of Engineering, Pune, Maharashtra, India

Abstract: *People suffering from speech impairment can't communicate using hearing and speech methods, they believe signing for communication. Sign language is employed among everybody who is speech impaired, but they find a tough time in communicating with people which are non-signers (people aren't proficient in sign language). So, requirement of a symbol language interpreter may be a must for speech impaired people. There has been favourable progress within the field of gesture recognition and motion recognition with current advancements in deep learning. There has been quite a significant development in computer vision which would enable us to easily track the hand gestures. The proposed system tries to try to a true time translation of hand gestures into equivalent English text. This system takes hand gestures as input through video and translates it text which might be understood by a non-signer. There will be use of CNN for classification of hand gestures. By deploying this technique, the communication gap between signers and non-signers will be reduced and they will be easily able to communicate with normal people.*

Keywords: CNN (Convolutional Neural Network), Gesture Recognition, Motion Recognition.

I. INTRODUCTION

1.1 Detailed Problem Definition

To Design and develop a system that provides real time translation of sign language using Deep Learning, considering the challenges of speech and hearing-impaired people while communicating while absence of an interpreter in order to bridge the communication gap.

1.2 Justification of Problem

Computer vision is an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do. Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from real world in order to produce numerical or symbolic information. Mainly computer vision with the theory behind artificial systems that extract information from images.

1.3 Need for the New System

According to a survey conducted by Ministry of Statistics and Programme Implementation in India there are approximately 4.2 million people suffer from either speech or hearing impairments. Those speech or hearing-impaired people face a lot of difficult in their day-to-day life. Speech impaired people heavily rely on speech interpreters for medical, legal, educational and training sessions. However, they face problems when speech interpreters are unavailable.

II. LITERATURE SURVEY

Natural Human Computer Interaction (HCI) is the demand of today's technology-oriented world. Detecting and tracking of face and hands are important for gesture recognition. Skin detection is a very popular and useful technique for detecting and tracking human-body parts. It has been much attention mainly because of its vast range of applications such as, face detection and tracking, hand detection and tracking, people retrieval in databases and Internet, etc. Many models and algorithms are being used for detection of face, hand and its gesture. Hand detection using model or classification is to build a decision rule that will discriminate between skin and non-skin pixels. Identifying skin color pixels involves finding the range of values for which most skin pixels would fall in a given color space. All external factors will be eliminated to detect the hand and its color in the image in complex background.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

There are finite number of elements which has a particular location and value which forms a digital image. These elements are known as pixel and picture element. These elements take participation to find out the hand. After this the color data plays very vital role for locating hand in vision-based system. Visions based techniques offer interaction system in dynamic environment by using cameras for capturing gesture and convert it in to meaningful command. Techniques used in hand tracking are divided into two parametric and non-parametric group. Parametric techniques focus on color concentration and non- parametric technique focuses on using statistical probability models for calculating color dispersion in training dataset. In this paper randomized lists are used for hand color classification for two stages of work: training of data and tracking of hand. In this paper, they actualize and test a continuous hand following method in view of hand shading division and conveyance utilizing randomized records. The method is easy to keep moving so as to run the client's hand openly for intuitive applications. Hand signals acknowledgment is the normal method for connection of human and machine. There are different types of methods used to locate hand in the obtained image after pre-processing. Morphology based method where image of hand is rebuilt using the image properties and extraction. Model supported method in which various models are utilized to characterize image using different models to represent in Computers.

III. System Design

Sequence Diagram

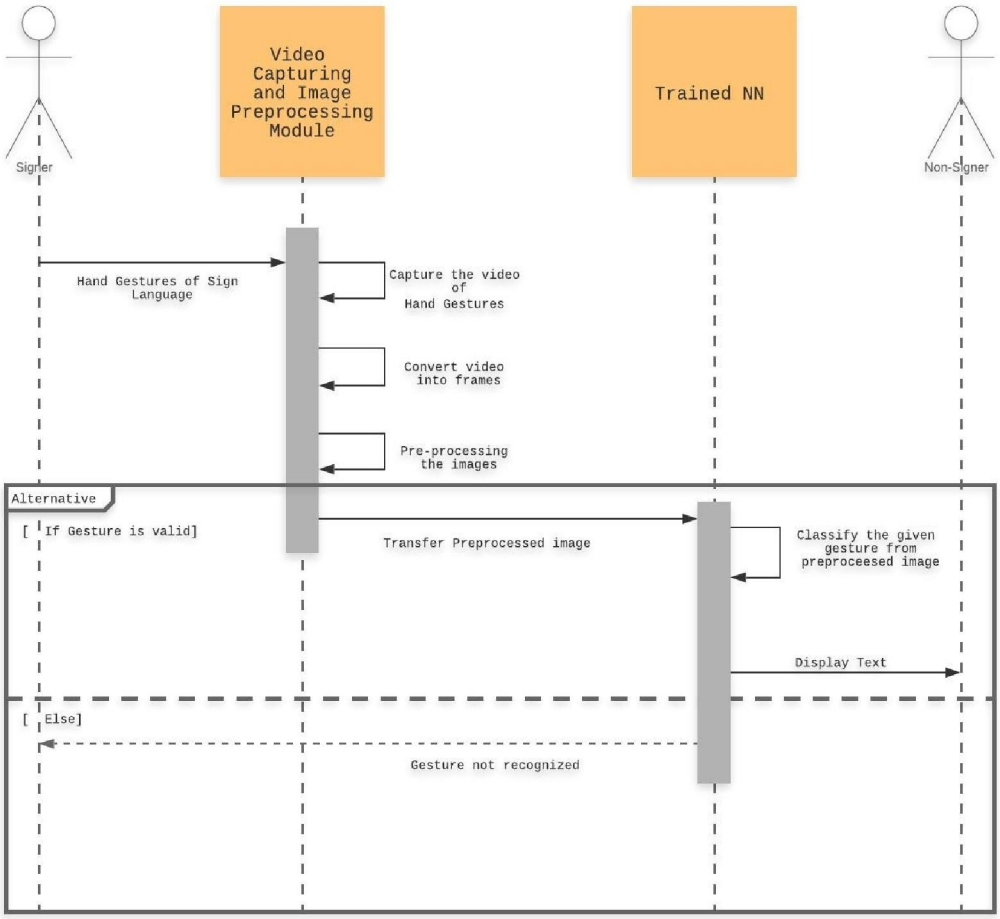

Figure 1: Sequence Diagram

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

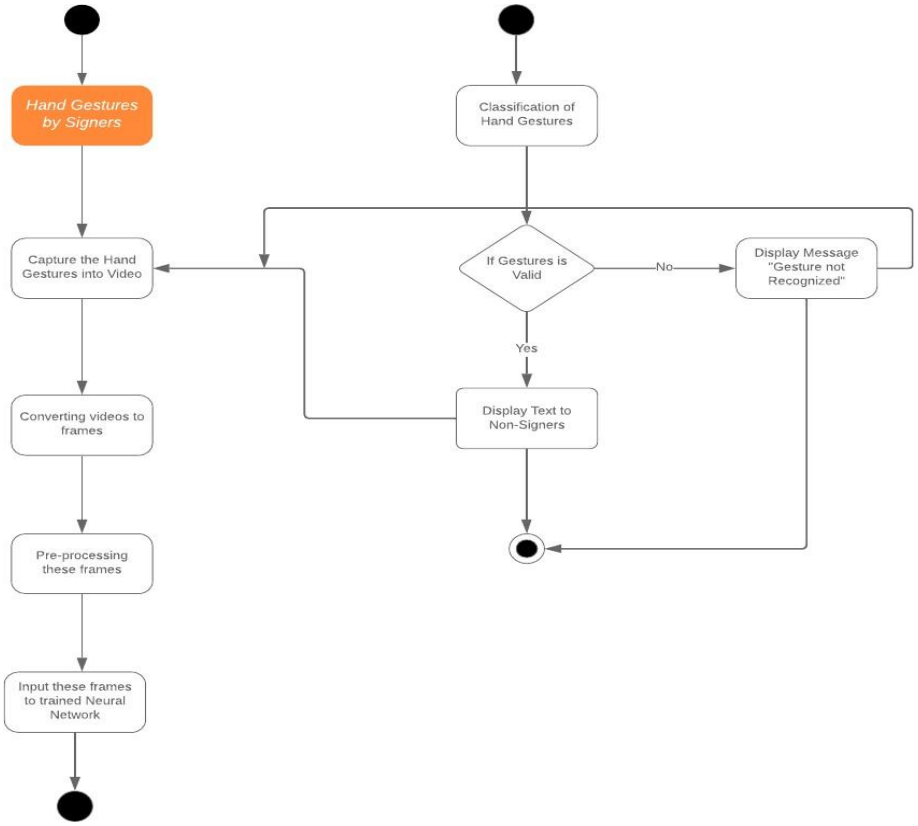

Figure 2: Activity Diagram

IV. OUTPUT

- First activity is Signer makes hand gestures in front of webcam.
- Then the next activity is to convert video into frames.
- After this, activity of pre-processing of frames is carried out.
- Next activity is to classify the hand gestures by the trained neural network and display it into English text.

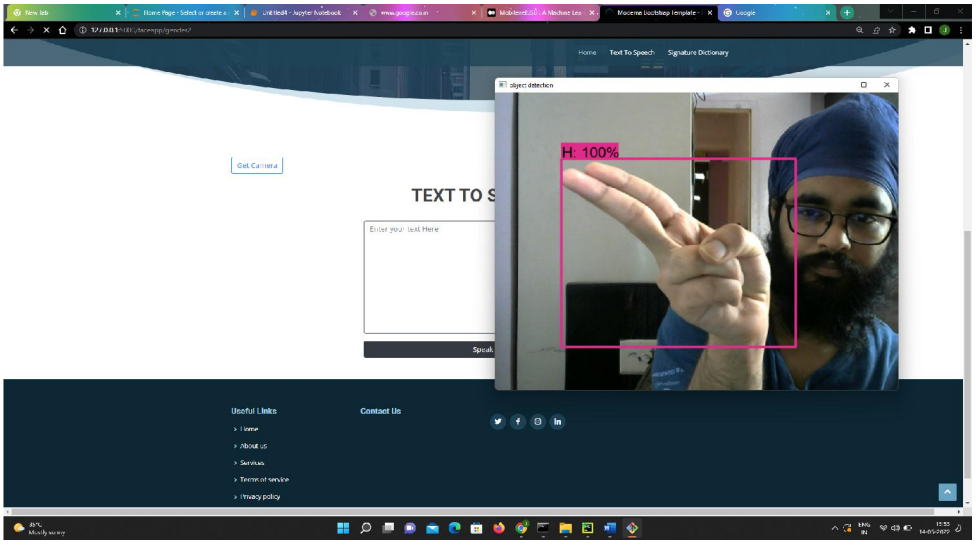

Figure 3: User Showing Sign of Letter "H"

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

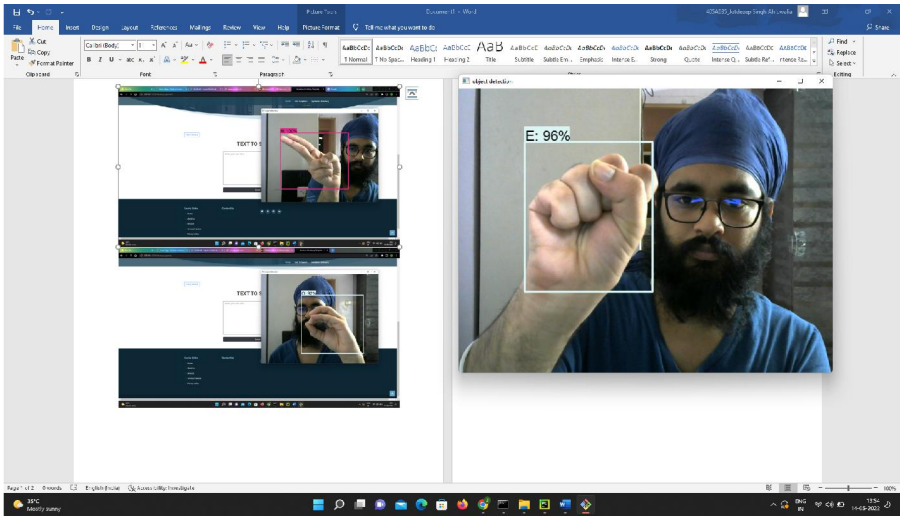

Figure 4: User Showing Sign of Letter "E"

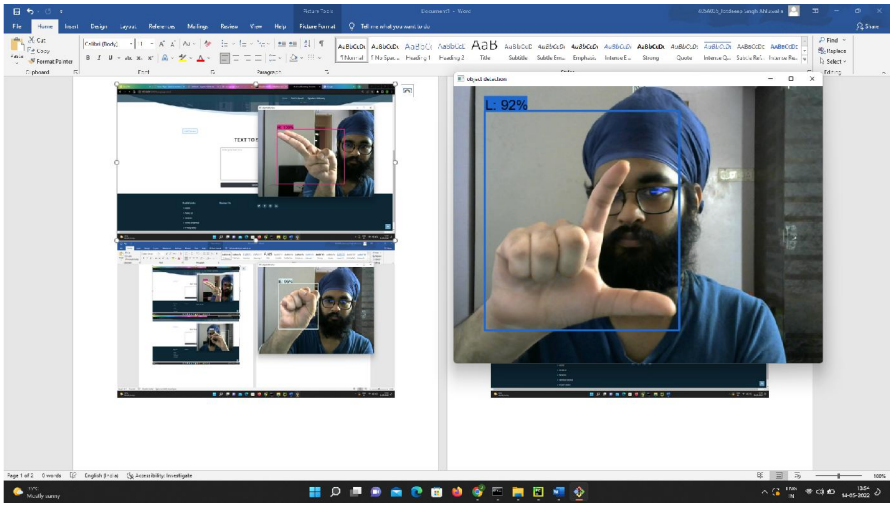

Figure 5: User Showing Sign of Letter "L"

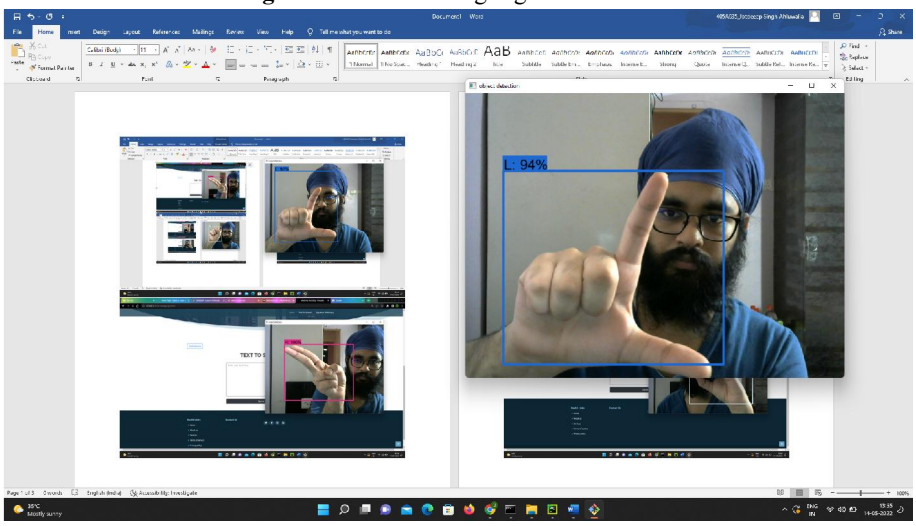

Figure 6: User Showing Sign of Letter "L"

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

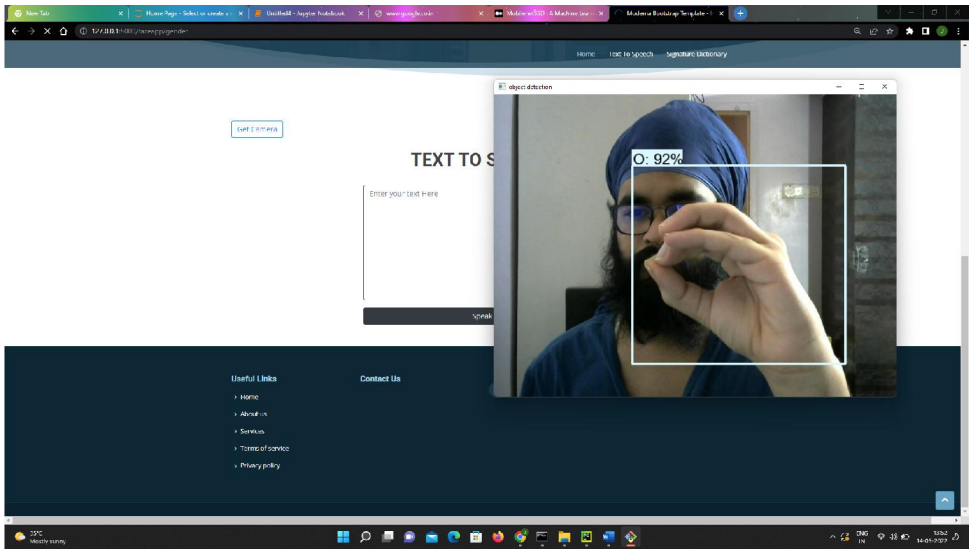

Figure 7: User Showing Sign of Letter "O"

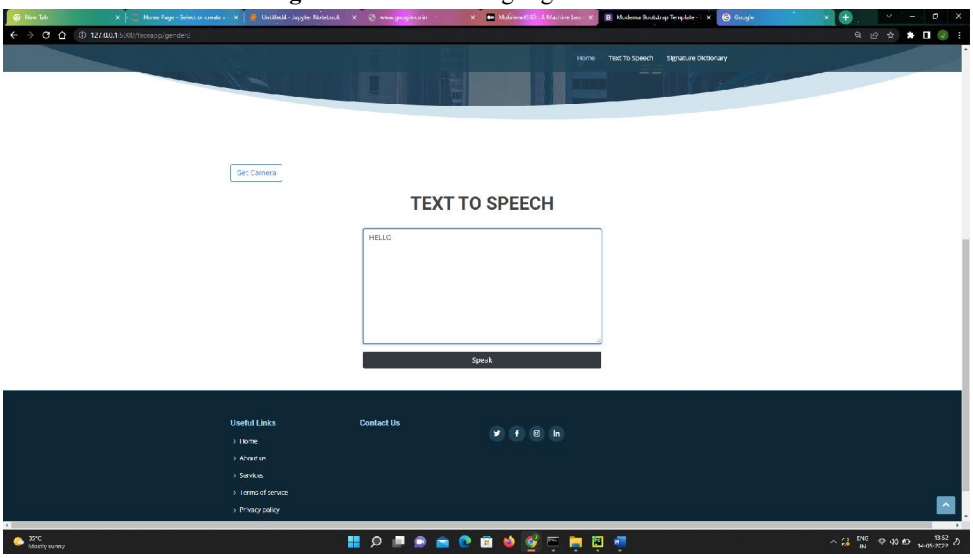

Figure 8: Final Output

V. FUTURE SCOPE

First potential improvement is that we can make words by alphabet gesture classification by adding a gesture for word break. This can work by giving a definite gesture for it then making that gesture when a signer wants to end a word. So, then the signer by making gestures for alphabets while using word break hand gesture to make words and then sentences. This would be quite beneficial for the signer to communicate with the non- signers.

Second potential improvement would be to classify hand gestures for words. Which is a daunting task as it requires the use of RNN (Recurrent Neural Network). The RNN is similar to other neural network but with a major advantage which is memory. So, it feeds the outputs a result back into itself. The classification of words requires movement of hand gestures which could be classified by RNN. RNN requires lot of processing power and large storage as it stores the output for the next epoch.

One of the important potential improvements would be to experiment with different RNN architectures for the output of the pool layer. Including GRU and Independent RNN's. In terms of CNN improvements, using Capsule Networks instead of Inception may yield better results than Inception.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, May 2022
Impact Factor: 6.252

REFERENCES

[1]. Karen Simonyan & Andrew Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition
[2]. LeCun, Y., Bengio, Y., & Hinton, G." Deep learning." vol 521 Nature, 521 pp 436-444 May 2015
[3]. Daniel Svozil, Vladimir KvasniEka, JiE Pospichal "Introduction to multi-layer feed-forward neural networks",
[4]. Alper Yilmaz, Omar Javed, Mubarak Shah, "Object Tracking: A Survey", ACM Computing Survey, Vol 38 No.4,
[5]. Deshmukh K.S., Shinde G. N, "An Adaptive Color Image Segmentation", Electronic letters on Computer Vision and Image
[6]. Soeb Hussain, Rupal Saxena, Xie Han, Jameel Ahmed Khan, "Hand Gesture Recognition Using Deep Learning",