Sign Language Translator using OpenCV

Aanya Chourasiya
Computer Science and Engineering
Acropolis Institute of Technology and Research
Indore (M.P.)
aanyachourasiya20542@acropolis.In

Akshay Keswani
Computer Science and Engineering
Acropolis Institute of Technology and Research
Indore (M.P.)
akshaykeswani20194@acropolis.in

Amit Kumar Yadav
Computer Science and Engineering
Acropolis Institute of Technology and Research
Indore (M.P.)
amityadav20456@acropolis.in

Aditya Sharma
Computer Science and Engineering
Acropolis Institute of Technology and Research
Indore (M.P.)
adityasharma20340@acropolis.in

Abstract: Sign Language is the only means of communication for the deaf and mute people. But many of the normal people do not know sign language. Thus, it is difficult for the people who speak in sign language to communicate with those who don't speak that language without an interpreter. This Project focuses on a review of the literature on hand gesture techniques and introduces their merits and limitations under different circumstances. The theories of hand segmentation and the hand detection system, may be used to construct hand gesture recognition using Python and OpenCV. The use of hand gestures as a natural interface motivates research in gesture taxonomies, representations, and recognition algorithms, as well as software platforms and frameworks.

Keywords: Open-Source Computer Vision Library (OpenCV).

I. INTRODUCTION

The increased public acceptance and funding for international projects emphasizes the necessity of sign language. The desire for a computer-based solution is significant in these age of technology for the deaf people. However, re- searchers have been attacking the problem for quite sometimes now and the results are showing some promises. Although interesting technologies become available for voice recognition[2], but there is currently no commercial solution for sign recognition[2] on the market. Since it's not possible to be accompanied by a translator always, it becomes difficult for these people to communicate with normal people without a translator. Thus, a system that's able to translate sign language into words without a human translator would be extremely helpful to the deaf and mute community. The set of various technology used in the Translation model are – OpenCV [1], Tensor flow and Media pipe.

II. EXISTING SYSTEM

Based on our review, [3]HMM-based approaches have beenextensively explored in prior research, including its modifications. Deep Learning, such as Convolutional Neural Networks [1], has been popular in the past five years. [4]Hybrid CNN-HMM [1][3] and completely deep learning systems have yielded encouraging results and provide avenues for additional research.

The output of the sign language will be displayed in the text form in real time. This makes the system more efficient and hence communication of the hearing and speech impaired people easier. The images captured through web cam are compared and the result of comparison is displayed at the same time. Thus, this feature of the system makes communication very simple and delay free.

III. OBJECTIVE

The goal is to make computers to understand human language and develop a user-friendly human computer interface (HCI) [3]. Some steps toward this goal include teaching a computer to recognize speech, facial emotions, and human gestures. [1]Gestures are the nonverbally exchanged information. Gesture recognition[2] is an aspect of [3]human-computer interactionthat demonstrates an academic treatise and is a vital to popularizing the notion of a human-to-human connection, open dialogue, that must imply the correlation between the user and the machine. [1]Gesture analysis is a scientific field that can recognize gestures such as hand, arm, head, and even structural motions that usually entail a certain posture and/or motion.

Using hand gestures, the individual may send out more information in a shorter amount of time. Several approaches were explored to apply computer-vision [3] ideas to the real time processing of gesture outputs. The Computer Vision[3] study concentrates on gesture recognition[2] in the open CV [1] framework using the Python language. Language is a huge part in communication.

Languages are useless to a person with a disability.

Gesture is a vital and meaningful mode of

communication for the visually impaired person. So here is the computer-based method for regular people to understand what the differently abled individual is trying to say. For monitoring, there are various similar algorithms and object recognition systems. This allows the identification of gestures, which overcomes the boundaries and limitations of earlier systems.

IV. SCOPE

1. We can develop a model for ISL word and sentence level recognition. This will require a system that can detect changes with respect to the temporal space.

2. We can develop a complete product that will help the speech and hearing-impaired people, and thereby reduce the communication gap.

3. This allows the identification of gestures and structural motions, which overcomes the boundaries and limitations of earlier systems.

4. The purpose of Sign Language Recognition (SLR)[2][3] systems is to provide an efficient and accurate way to convert sign language into text or voice has aids for the hearing impaired for example, or enabling very young children to interact with computers (recognizing sign language), among others.

5. It Provides Real Time Functioning, Portability and has no need of Calibration.

6. In future work, proposed system can be developed and implemented using Python. Image processing part should be improved so that system will be able to communicate in both directions i.e.it should be capable of converting normal language to sign language and vice versa. We try to recognize signs that include motion. Moreover, we focus on converting the sequence of gestures into text i.e., word and sentences and then converting it into the speech which can be heard.

V. REQUIREMENT ANALYSIS

- Windows computer or Linux.
- Python Version 3.6 or above.
- CMOS sensor (Webcam).
- Python Installed Windows Operating system or Linux OS Machine.
- Software Requirements: OpenCV [1], Tensor flow, Media pipe, NumPy.

VI. DESIGN / FLOW DIAGRAMS

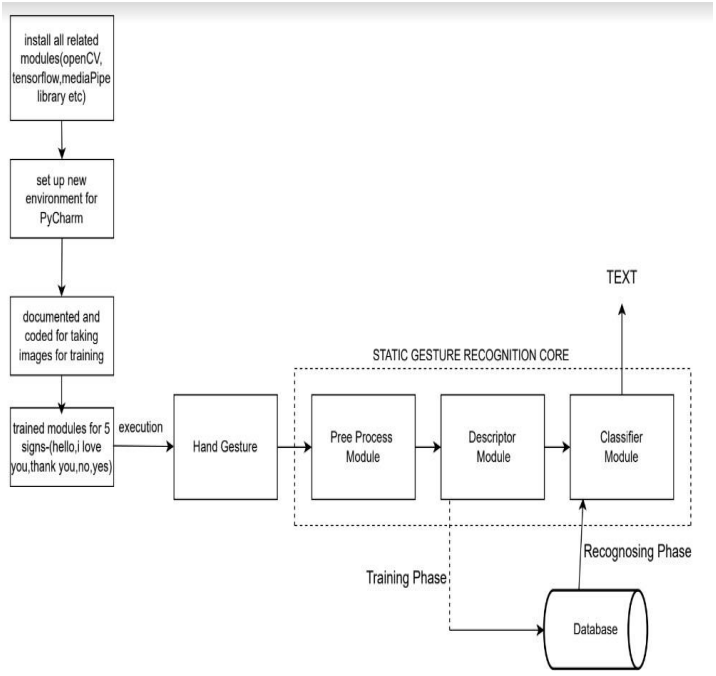

Fig. 1 Work flow Model of SLT

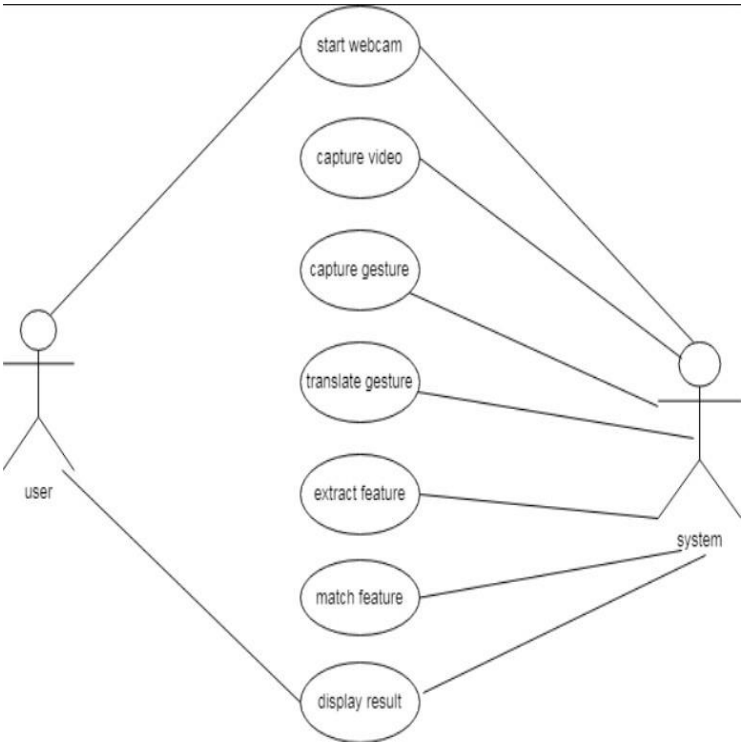

Fig.2 Use Case Diagram of SLT

VII. PROJECT DESCRIPTION

This method takes pictures on camera as touch data. The vision-based approach focuses heavily ontouch captured images and brings out the main and recognizable feature. Color belts were used at the beginning of the vision-based approach. The main disadvantage of this method was the standard color to be applied to the fingers. Then use bare hands instead of colored 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. The Model will be trainedby several set of Gestures of sign language using OpenCV [1] and Python API. We expanded several gestures of hands by our own to train the model. The gestures will be taken by the camera from multiple directions. Several different methods have been tested so far. The first is to build a 3D 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. To use the Model, the user shows some hand gestures in front of the camera. Upon encountering similar gestures in database, the model will show text/sentence as output representing the translation of the sign language. Upon encountering any gesture that is not present in the database, the model will display an Error message.

VIII. METHODOLOGY

1. Hardware & Software Requirement - Windows computer or Linux, Python installed and Libraries, CMOS sensor (Webcam), Python Installed Windows OS or Linux OS Machine.

2. Data Collection – The first step is the data-set collection. We collect the full frame (RGB) image and the corresponding depth map from the camera. It actually understands the gesture and finger positions and updates it in the Database where all relevant Collection of images are stored.

3. Extraction of Hand Gesture – The Set of relevant images will be processed and assigned a Message that the particular sign language will be conveying. All the images from multiple angles having similar Messages will be grouped together and will be assigned the Text message to be displayed in the Model.

4. Collection of Gestures in Database – The set of all Images conveying the same message have been grouped together as an entity in the database, thus the Database will serve as a collection of multiple number of entities, each of which represents a set of actions / gestures of hand conveying a message.

5. Hand-Gesture Recognition using Deep Learning – Upon User encounter with the model, the user will show hand gestures in front of the camera. Using OpenCV [1], those images will be processed and compared with a pretrained and pre-processed set of images.

6. Upon any image in the database getting matched with the Gesture, the Mapped Text message linked with the gesture will be displayed on the screen in the form of Text/sentence.

7. If the gesture performed by the user does not matches any linked image in the database, in that scenario, and Error message will be displayed on the screen of the user.

IX. EXPECTED OUTCOME

Upon training the image dataset without any augmentation, the training accuracy achieved was very high (around 99%) but, the real time performance was not up to the mark. It was predicting incorrectly most of the times because in real time hand-gestures were not placed exactly at the center and aligned vertically. In order to overcome this shortcoming, we trained our model by augmenting our dataset. The training accuracy was reduced to 89% but the real-time predictions were predominantly correct. The Model can provide an efficient and accurate way to convert sign language into text or voice has aids forthe hearing impaired for example, or enabling very young children to interact with computers (recognizing sign language), among others. The Model is capable to Provide Real Time Functioning, Portability and has no need of Calibration. In future work, large scale proposed system can be developed and implemented using Python. Image processing part should be improved so that system will be able to communicate in both directions i.e.it should be capable of converting normal language to sign language and vice versa. We will try to recognize signs that include motion. Moreover, we will focus on converting the sequence of gestures into text i.e., word and sentences and then converting it into the speech which can be heard. Also, when we increase our training image size the accuracy betters itself both during training and real time implementation. However, training images with large size is directlyproportional to the computational power of the system.

X. CONCLUSION

The aim of this project is to predict the ISL alphanumeric hand-gestures in real time. The above work shows that it can be solved with better accuracy when we actually consider the segmented RGB hand-gestures. By applying depth- based segmentation we remove the overheads of dynamic background. The segmented RGB hand- gestures were fed to 3 layered CNN [1][4] for training and testing in real time. We were able to achieve training accuracy of 89.30% and testing accuracy of 98.5%. Ourmodel may show good accuracy while predicting results both offline and online. Our project aims to make communication simpler between deaf and dumb people by introducing Computer in communication path so that sign language can be automatically captured, recognized, translated to text and displayed it on Screen. There are various methods for sign language conversion. Some of them use wired electronic glove and others use visual based approach. Moreover, we will focus on converting the sequence of gestures into text i.e., word and sentences and then converting it into the speech which can be heard.

XI. REFERENCES

1. Vinayak S Kunder, Aakash A Bhardwaj, and Vipul D Tank. Sign language recognition system with speech output.
2. Malavika Suresh, Avagyan Sinha, and RP Aneesh. Real-time hand gesturerecognition using deep learning. International Journal of Innovations and Implementations in Engineering.
3. GRS Murthy and RS Jadon. A review ofvision-based hand gestures recognition. International Journal of Information Technology and KnowledgeManagement
4. Mukesh Kumar Makwana, " Sign Language Recognition", MTech thesis, Indian Institute ofScience, Bangalore.
5. S. Hussain, R. Saxena, X. Han, J. A. Khan, and H. Shin, “Hand gesture recognition using deep learning”, 2017 International SoC DesignConference (ISOCC).
6. J. -. Lementec and P. Bajcsy, "Recognition ofarm gestures using multiple orientation sensors: gesture classification",Proceedings.
7. Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane,Comparative Study of Sign Language Recognition Systems, International Journal of Scientific and Research Publications.
8. Mukesh Kumar Makwana, " Sign Language Recognition", MTech thesis, Indian Institute of Science, Bangalore.
9. S. Hussain, R. Saxena, X. Han, J. A. Khan, and H. Shin, “Hand gesture recognition using deep learning”, 2017 International SoC Design Conference (ISOCC).
10. J. -. Lementec and P. Bajcsy, "Recognition of arm gestures using multiple orientation sensors: gesture classification", Proceedings.
11. Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane, Comparative Study of Sign Language Recognition Systems, International Journal of Scientific and Research Publications.