International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 2, Mar-Apr-2020, ISSN (Online): 2395-566X

Sign Language Translation to Voice For Dumb People

Rosemarry Antony, Stephy Paul, Silji Simon C, Assistant Prof. Scaria Alex
Dept. of Computer Science
Sahrdaya College of Engineering and Technology
Thrissur, India
rosemarypallipadan@gmail.com,siljisimonc@gmail.com,stephypaulachandy@gmail.com,scariaalex@sahrdaya.ac.in

***Abstract* –Communication is the only medium by which we can share or convey our thoughts, but for a person with disability(dumb) faces difficulty in communicating. It is impossible for them to communicate with others until ordinary people learn the sign language for communication. So this system will help the dumb people to intract with other people. Some existing methods are sign language to text conversion and using smart gloves. In the already existing systems signs are identified from images. The normal outputs will be alphabets or numbers and may not handle connective words like is,are etc. It is more comfortable if it is converted to voice rather than text for dumb people. In this system a webcam will capture the hand gesture from video and perform hand feature extraction. Then predict the symbol. Finally we get the result as voice. So that it would be audible to everyone. Here we are identifying signs from live video. It is the main difference from other methods. The output will be collection of words. For handling connectives special languages want to implement. Machine learning is the domain of this system. This is implemented using python. The main objective of this system is to help dumb people in their communication.**

***Keywords*– Machine learning, Feature extraction.**

I. INTRODUCTION

the society it is very difficult for the dumb people to talk with the ordinary people. It becomes impossible for them to communicate with the normal people unless and until people like us learn the sign language for the communication.The sign language of dumb is little difficult to learn and it is not easy for everybody to learn that language. So every person cannot share their thoughts with these physically impaired people. So here is a system which help the dumb people to communicate with every one. In this system a webcam is placed in front of a physically impaired person.

The physically impaired person would place his finger in front of the web camera and thewebcam will capture the hand gesture then it will perform image processing using Principle Component Analysis (PCA) algorithm. The co-ordinates captured will be mapped with the one previously stored image and accordingly exact picture from the database will be identified. Continuing in this way physically impaired person will be able to go through the whole sentence that he wants to communicate. Later on this sentence will be converted into speech so that it would be audible to everyone. By using this system the dumb people would be benefited as they can communicate with everyone freely. Speech impairment is a disability which affects an indi- viduals ability to communicate with others using speech and hearing. People who cannot speak will use other media of communication such as sign language. Although sign language is everywhere in recent times, there remains a challenge for non-sign language speakers to communicate with sign language speakers . With recent advances in technologies there has been promising progress in the fields of motion and gesture recognition using deep learning. This proposed model takes video sequences and extracts temporal and spatial features from them.

II.PROBLEM DEFINITION

In our society there are many people who cannot communicate properly with others. It is difficult for the dumb people to express or convey their opinions to ordinary people. Normal persons may not be understand their sign language. This will create a communication gap between them. This system will help dumb people to communicate with others. Here their sign language is capturing using camera as a video. Then we do the hand feature extraction. Finally it will predict the symbol. Special languages will be implemented forrepresenting connecting words like is,are etc. Using machine learning and image processing the signs are converted to voice. Noise is removed from the image using image processing. Using this system other people can clearly understand the dumb people.

III.RELATED WORK

Language (SL) is commonly known to be the primary language of deaf people, and usually captured in the form of video. In [1] Runpeng Cui, Hu Liu, and Changshui Zhang, they developed a continuous SL recognition framework with deep neural networks, which directly coverts videos of SL sentences to sequences of ordered gloss labels. A deep convolutional neural networks with stacked temporal fusion layers as the feature extraction module, and bi-directional recurrent neural networks as the sequence learning module were using. Here, end-to-end recognition model is trained first for alignment proposal, and then use the alignment proposal as strong supervisory information to directly tune the feature extraction module.

People use sign language gestures for non-verbal communication to express their thoughts and emotions. But non-signers find it extremely difficult to understand,hence trained sign language interpreters are needed during medical and legal appointments, educational and training sessions.[2]use a CNN (Convolutional Neural Network) model named inception to ex- tract spatial features from the video stream for Sign Language Recognition (SLR).

Then, by using a LSTM (Long Short-Term Memory) [3], a RNN (Recurrent Neural Network) model, we can extract temporal features from the video sequences via two methods: Using the outputs from the Softmax and the Pool layer of the CNN respectively. A system that is capable to detect static hand signs of alphabets automatically in American Sign Language (ASL)[4]. To do that, two combined concepts AdaBoost and Haar like classifiers are used. In this work, to increase the accuracy of the system, a huge database for training process was needed, and it generates impressive results. Input of this is live video and output is the text.

The recently developed depth sensors mainly the Kinect sensor, have provided many new opportunities for Human Computer Interaction (HCI) [5]. Kinect sensor has brought great progress, in human body tracking, face and human action recognition, robust hand gesture recognition remains an open problem. Compared to the entire human body, the hand is considered as a small object with more complex articulations and more easily affected by segmentation errors.

Thus it is very difficult to recognize hand gestures. Robust part based hand gesture recognition system is focused in this paper using Kinect sensor. A novel distance metric was proposed to handle the noisy shapes from the kinect sensor, Finger Earth Mover's Distance (FEMD), to measure the dissimilarity between hand shapes. This matches only the finger parts not the whole hand, it can better distinguish the hand gestures of slight differences. Related work is with a recurrent three-dimensional convolutional neural network that performs condetection and classification of dynamic hand gestures from multi-modal data[6].

Here is an algorithm for joint segmentation and classification of dynamic hand gestures from continuous depth, color and stereo-IR data streams. Develop on the recent success of CNN classifiers for gesture recognition, so this is a network that employs a recurrent 3D-CNN with connectionist temporal classification (CTC)[7]. In another work, implementation of an American Sign Language (ASL) finger spelling translator based on skin segmen- tation and machine learning algorithms[8]. Automatic human skin segmentation algorithm based on color information was used there. The skin- color distribution as a bivariate normal distribution in CbCr plane. Then Convolutional Neural Net- work (CNN) is used to take the features from images and Deep Learning Method is used to train a classifier to recognize Sign Language.

IV.PROPOSED SYSTEM

By using this proposed system, we aim to do the following.The main area where this can be used is in public places like ticket issuing counters, hospitals etc. This can be even used to teach the sign language to normal people.The core objective of this project is to recognizing the gestures and displaying the respective word. The first step involves capturing the gesture using a webcam. The web cam captures the image then the image is processed with pose estimation algorithm in tensor-flow utility. Capturing signs from real world and translating them is the main objective of this work. The real-world signs are read using a webcam which placed infront of the impaired person ,that will captures both static and moving images of the objects in front of it.

The images captured from the camera are then stored in a testing dataset for further analysis. Before storage, it performs methods for detecting our area of interests, ie, hand detection is performed. Then a peak value analysis is performed to get the peak values of hands. Next the features from the area of interest is taken into account. It is then compared with our trained dataset to generate matched outputs. For this , gesture matching tech- niques are used. Various open platform libraries are available for gesture matching. For simplicity and accuracy opencv, one of the best among them is selected. Once the output matches, it uses text to speech conversion techniques to generate the voice.

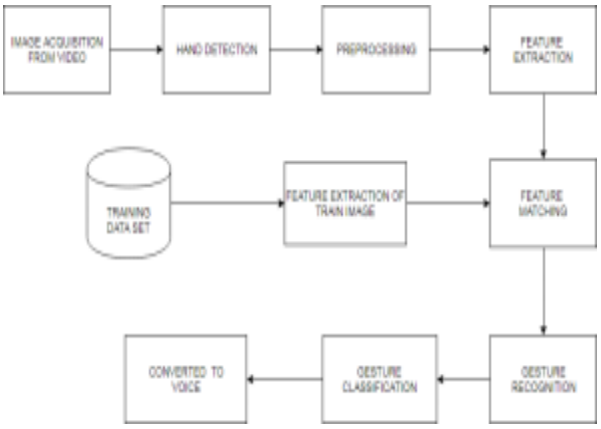

Fig. 1. A Block visualization for the proposed system.

The images captured from the camera are then stored in a testing dataset for further analysis. Before storage, it performs methods for detecting our area of interests, ie, hand detection is performed. Then a peak value analysis is performed to get the peak values of hands. Next the features from the area of interest is taken into account. It is then compared with our trained dataset to generate matched outputs. For this , gesture matching techniques are used. Various open platform libraries are available for gesture matching. For simplicity and accuracy opencv, one of the best among them is selected. Once the output matches, it uses text to speech conversion techniques to generate the voice.

V. CONCLUSION AND FUTURE WORK

Communication has a major role in our life because we all are social being. Deaf-Dumb people cannot be fully part in effective communication. This system is to improve the communication with the dumb people who are using any sign language to express themselves. In this system a webcam will capture the hand gesture from video and perform hand feature extraction. Then predict the symbol. Finally we get the result as voice. So that it would be audible to everyone. Here we are identifying signs from live video. It is the main difference from other methods. This system will be very helpful for the dumb people for expressing their opinions and for proper communication. This will lead to proper interaction between dumb people and the normal people.

In future it can be developed into a mobile application of such system that enables everyone be able to speak with dumb people. In this design webcam is used to acquire input image. The image features are extracted via a trained Convolutional Neural Network. In future we can implement an android based application for dumb people to communicate with normal peo- ple. Everyone uses mobilephones, so communication through mobile phone is considered very important in enhancing better understand in social situation. So there is a great opportunity for mankind in the communication sector of all over the world.

REFERENCES

[1]. Runpeng Cui, Hu Liu, and Changshui Zhang, “A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training”, IEEE,2018;
[2]. Kshitij Bantupalli,Ying Xie,“American Sign Language Recognition using Deep
[3]. Learning and Computer Vision”, IEEE International Conference on Big Data (Big Data),2018.
[4]. Sepp Hochreiter et al.,“Long Short-Term Memory,”, Neural Computa- tion 9(8): 1735-1780,1997.
[5]. Vi N.T. Truong ,Chuan-Kai Yang,Quoc-Viet Tran,Translator for American Sign Language to Text and Speech ”, IEEE 5th Global Conference on Consumer Electronics,2016.
[6]. Z. Ren, J. Yuan, J. Meng, and Z. Zha,“Robust Part-Based Hand Gesture Recognition Using Kinect Sensor”, ” IEEE Trans. Multimedia, vol. 15, no. 5, pp. 1110–1120,2013
[7]. P. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, and J. Kauz, “Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural network”, in Proc. IEEE Conf. Comput. Vis. Pattern Recog,2016.