International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET23103107

Virtual Personal Assistant with Sign Language

Dr. D. S. Hirolikar, Hari Shelar, Prajwal Berad, Mohini Gawade, Yogesh Pawar
Department of Information Technology, PDEA's College of Engineering (Manjari B.K.), Pune, India

ARTICLEINFO

Article History:

Accepted: 10 May 2023
Published: 30 May 2023

Publication Issue
Volume 10, Issue 3
May-June-2023

Page Number
297-301

ABSTRACT

This topic is all about developing a successful Sign language recognition and translation system which is helpful for people with certain disabilities who can't use present voice enabled Virtual Assistants. Communication is important with people as it helps us understand them and gain knowledge. In case of Deaf & mute people, this becomes a problem. It is solved by creating a virtual assistant built to recognize hand gestures and translate them in normal language to perform tasks ordered by the user. This certainly helps other normal people to successfully communicate with deaf & mute people

Without worrying about any misunderstanding at the listener's end. There are more than 300 sign languages used by various cultural groups worldwide. In this article, we provide a technique for generating a sign language dataset using a camera, followed by the training of a TensorFlow model using transfer learning to produce a real-time sign language recognition system. Despite the small amount of the dataset, the system still performs well.

Keywords : Deep Learning, Virtual Assistants, Tensor Flow, Convolutional Neural Network Hand Gestures, Sign Languages.

I. INTRODUCTION

Nowadays, Communication is very important to people for connecting with each other. Today People use virtual assistants to make their day-to-day work easier. There are various virtual assistants available like Google Assistant, Microsoft Cortana, Apple Siri, Amazon Alexa, etc. but it is only used by the people who are able to speak on a major basis. A Successful Communication is done when the message sent by speaker(sender) is understood by the Listener(receiver). But if the sender is not able to speak then the voice assistant cannot understand the message and will not be able to perform the task as expected. This is where the sign language comes in. It is a language specifically made for deaf & mute people. It consists of special alphabets made of various hand movements and gestures.

There is a much larger need to make such technologies usable to the specially-abled people. Deaf & dumb people can communicate with other similar people in sign language, but normal people can't understand what they are trying to say due to lack of knowledge of their language. Such issue can easily be solved with the help of voice assistant specifically designed Digital assistant which can translate the hand gestures into commonly speaking language, English.

The suggested system has succeeded in switching from voice recognition to hand gesture recognition. The following technologies can be used by the proposed system: Convolution Neural Networks are Deep Learning algorithms that have been used for Image Recognition, transforming an image into a matrix that can be comprehended by the system model, and Tensorflow and PyTorch are the most significant libraries used for constructing these systems. In the system, CNN serves as a classifier, and OpenCV, which serves as the system's eye, will capture and process real-time hand gestures and provide outputs with the aid of a classifier.

II. Related Works

Since every virtual assistant currently in use is voice automated, Deaf-mutes and people with other disabilities cannot use them. These days, personal assistants and virtual assistants play a crucial role in our daily lives. Every business or person is converting to these technologies since they help them complete their tasks more quickly and easily. A virtual assistant is a component of this system that can take user input, comprehend it, analyse it, and carry out tasks. Users can save a tonne of time by doing this.

Understanding Deaf users is essential for designing successful sign language processing systems that satisfy user demands. Here, we provide a summary of the background information and talk about recent studies on sign language processing, which aid in our comprehension of the system's problems.

III. Literature Survey

1. “Design and Implementation of A Sign-to-Speech/Text System for Deaf and Dumb People.”

This paper presents an approach for designing and implementing a smart glove for deaf and dumb people. This research aims to develop a sign to Arabic language translator based on smart glove interfaced wirelessly with microcontroller and text/voice presenting devices. An approach has been developed and programmed to display Arabic text. The whole system has been implemented, programmed, cased and tested with very good results.

2. “A Review on Smart Gloves to Convert Sign to Speech for Mute Community”

The normal and mute people can communicate only in one way i.e. sign language, but many times communicating with normal persons they noticed difficulty. We are implementing this project to reduce the barrier between dumb and normal person. This device design is based on the embedded system.

3. “On Design and Implementation of A Sign-to Speech/Text Syst”

In this paper two approaches have been investigated for designing and implementing a sign to speech/text translator. The approaches have been developed and implemented to display a dual language (Arabic and English) text and voice. In the first part of this paper, a vision-based system is developed and demonstrated. In the second part, a glove-based system is designed and implemented. The second system is based on wireless-interfaced glove, microcontroller and presenting devices to translate the sign to Arabic/English language. The two developed systems have been tested with very good results

IV. METHODS AND MATERIAL

A. Dataset Description :

We used the publicly available Sign Language MINST dataset for American sign language on Kaggle to test our model's ability to recognise signs.

The American Sign Language dataset includes 24 alphabetical gestures out of a total of 26 alphabets. J and Z were left out of the dataset since they are motion-based alphabets and because we thought about working with just static photos. There are respectively 27,455 and 7,172 instances of 24 alphabetical gestures in the sample. To fine-tune the model, 30% of the data from the training pictures were selected as validation samples.

B. Proposed Deep CNN Architecture:

One of the oldest and most powerful neural systems in the field of images is the convolution neural network, or CNN recognition. After successfully completing the ImageNet challenge, CNN became well-known . In light of its intrinsic construction of layers, a convolutional neural system provides a few benefits over other types of neural systems.

Convolutional, pooling, and fully linked layers make up the bulk of CNN's layers. The information lattice is replicated with a few convolutional sections or channels in the convolutional layer to produce an element map that specifies the kind of highlight that persists in the image. The amount of highlight maps is reduced by the interpretation and scaling variation provided in the pooling layer. The fully connected layer, which comes after all previous convolutional and pooling layers, clarifies what kinds of traits are present in the picture and what kinds aren't. The Convolutional Neural Network's core layers are these three categories of layers. In comparison to other sophisticated learning system designs, CNN also has fewer limits.

Three pooling layers and a total of six convolutional layers were used in this investigation. Convolution layer ReLU activation function was used. To get more in-depth features, the subsequently added two convolutional layers increased the number of convolutional filters from 64 to 128 in the initial two convolutional layers. In order to extract even more in-depth features from the picture, 256 convolutional filters were used in the last two convolutional layers. Every convolution layer has a 3x3 size filter in place, while the pooling layer used Max pooling. The feature that had been removed from the image was finally moved in the direction of the completely linked layer. The buried layer also used the ReLU activation algorithm. The hidden layer included 256 neurons that encoded the mapping between the inputs from the fully connected layer and the outputs from the output layer. The output layer included a total of 24 nodes with a softmax activation function because the classification challenge involves 24 American sign language symbols.

V. ALGORITHM

CNN :

Convolutional Neural Networks specialized for applications in image amp; video recognition. CNN is mainly used in image analysis tasks like Image recognition, Object detection amp; Segmentation. There are Four types of layers in Convolutional Neural Networks:

1) Convolutional Layer: In a typical neural network each input neuron is connected to the next hidden

layer. In CNN, only a small region of the input layer neurons connect to the neuron hidden layer.

2) Pooling Layer: The pooling layer is used to reduce the dimensionality of the feature map. There will be multiple activation amp; pooling layers inside the hidden layer of the CNN.
3) Flatten: - Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector.
4) Fully-Connected layer: Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer

Methodology

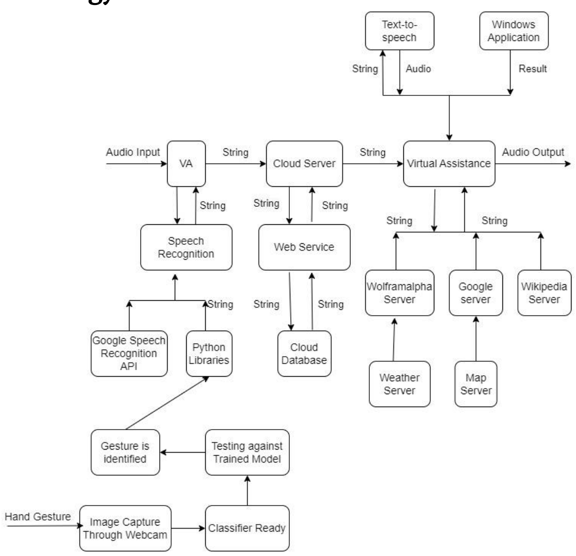
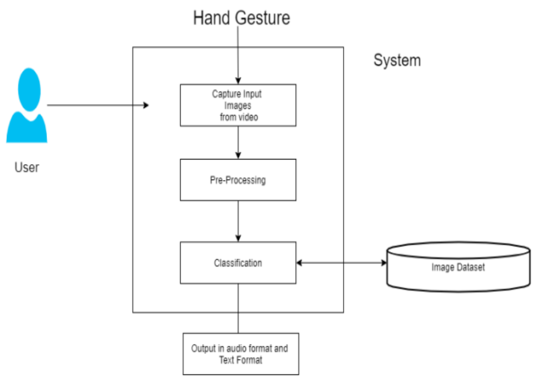


Fig. System Architecture



Following is a simple summary of the system's workflow:

In front of the webcam, a hand motion is made as seen in Fig. This sign motion is translated into text, and the written output is then spoken aloud to the assistant as an input. The assistant considers the query and provides an audio response. The output of this audio format is text. The text output will then be seen on the monitor.

The following is a quick technological description of the system: First, a sign language database is made available to the system in order to train it with a variety of hand gestures labelled with their respective labels. After training is completed, the system enters identification mode. It now utilises the camera input picture and runs it through the classifier to discover its neighbours based on the training samples or labels supplied in the previous phase. If the Gesture does not match the specified dataset, it appends a label to the frame, indicating that the system recognises the next hand motions accordingly. Once the gesture is recognised, it converts text to speech format using Python libraries, and the assistant can recognise the voice and provide output to the system in audio as well as on the display screen. The entire procedure can be performed several times. However, hand gestures will be recognised only if they match the trained dataset standards.

VI. Conclusion

Sign Language is a tool to reduce the communication gap between deaf-mute people and normal person. This system which is proposed above gives the methodology which aims to do the same as the two-way communication is possible. This method proposed here facilitates the conversion on the sign into speech. This overcomes the requirement of a translator since real time conversion is used. The system acts a voice of the person who is deaf-mute. This project is a step towards helping a specially challenged people. This can be further enhanced by making it more user friendly, efficient, portable, compatible for more signs and as well as dynamic signs. This can be further improvised so as to making it compatible for the mobile phones using the built-in camera of the phone. We can increase the distance at which it can be used by using a longer trans-receiver module or over Wi-Fi.

VII. Future Work

In future work, proposed system can be developed and implemented using Raspberry Pi. Image Processing part should be improved so that System would 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 which 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.

VI. REFERENCES

[1]. Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak (Feb 2015), "Study of Sign Language Translation us ing Gesture Recognition", International Journal of Advanced Research in Computer and Communication Engineering

[2]. Jayshree R. Pansare, Maya Ingle (2016)," Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram", at 2016 International Conference on Image, Vision and Computing.

[3]. Arslan Arif, Syed Tahir Hussain Rizvi, Iqra Jawaid, Muhammad Adam Waleed, Techno-Talk: An American Sign Language (ASL) Translator, at CoDIT'16 - April 6-8, 2016, Malta

[4]. Justin K. Chen, Debabrata Sengupta, Rukmani Ravi Sundaram," Sign Language Gesture Recognition with Unsupervised Feature Learning"

[5]. Matheesha Fernando, Janaka Wijayanayaka,"Low cost approach for Real Time Sign Language Recognition" at 2013 IEEE 8th International Conference on Industrial and Information Systems, ICIIS 2013, Aug. 18-20, 2013, Sri

[6]. M Vijayalakshmi, "Sign Language To Speech Conversion", At 2016 Fifth International Conference On Recent Trends In Information

[7]. xia Wu, Chong Pan,YufengJin, Shengli Sun, and Guangyi Shi Shaoxing, "Improvement of Chinese Sign Language Translation System based on Collaboration of Arm and Finger Sensing Nodes", At The 6th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems June 19-22, 2016,

[8]. Poonam Chavan1, Prof. Tushar Ghorpade2, Prof. Puja Padiya, "Indian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier" at 2016 Symposium on Colossal Data Analysis and Networking (CDAN)

[9]. Ms.Manisha D. Raut, Ms. Pallavi Dhok, Mr. Ketan Machhale, Ms.Jaspreet Manjeet Hora, "A System forRecognition of Indian SignLanguage for Deaf People using Otsu's Algorithm", International Research Journal of Engineering and Technology (IRJET)

[10]. Swapna, Shabnam Shaikh "Literature survey on hand gesture recognition for video processing" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)

Cite this article as :

Dr. D. S. Hirolikar, Hari Shelar, Prajwal Berad, Mohini Gawade, Yogesh Pawar, "Virtual Assistant With Sign Language", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 3, pp. 297-301, May-June 2023. Available at doi : https://doi.org/10.32628/IJSRSET23103107

Journal URL : https://ijsrset.com/IJSRSET23103107