IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning Algorithms For Sign Language Recognition

Authors: -

1. Sahil Sharma

Student BCA

(Shri Guru Teg Bahadur Institute of Management and Information Technology)

2. Vishal Rana

Student BCA

(Shri Guru Teg Bahadur Institute of Management and Information Technology)

3. Dr. Supreet Kaur Sahi

Assistant Professor

(Shri Guru Teg Bahadur Institute of Management and Information Technology)

Abstract

Sign language is used by deaf and mute people to communicate with the others. Using machine learning algorithms to recognize different signs and gestures the communication process can be made easy. Sign languages are not standard or universal which makes communication through sign language even harder. But machine learning algorithms with help of computer vision can automatically improve through the data available and help people to understand the signs and gestures. Vision based sign language recognition is an area of active research in computer vision and machine learning technology. Signs and gestures being a natural way of human interaction it is an area where many researchers are working on with the goal of making human computer interaction easier and natural without the need of any extra devices. So the primary goal of gesture recognition is to create systems which can identify specific human gestures or signs and use them to convey information. Here we have used computer vision and machine learning to create a

user interface system to recognize various hand gestures & sign language to help people understand it without any difficulty.

Introduction

More than 360 million of world population suffers from hearing and speech impairments.[7] Sign language is mainly used by deaf and mute people as deaf people have either hearing impairments or completely unable to hear and more people are unable to talk so they use different signs to communicate to people for their day to day life.[1]

While sign language is very important to deaf and mute people to communicate both with normal people and themselves it still gets little attention from the normal people we as a normal person tend to ignore the importance of sign language unless there are loved ones who are deaf-mute. Generally the solution to communicate with deaf and mute people is by using the services of a sign language interpreter but the use of sign language interpreter can be costly chief solutions is required so that the deaf and mute people can communicate with normal people easily. Researchers want to find a way for the deaf mute people so that they can communicate easily with normal people.{3} The breakthrough for this is the sign language recognition system. The system aims to recognize the sign language and translate it to the local language via text or speech.

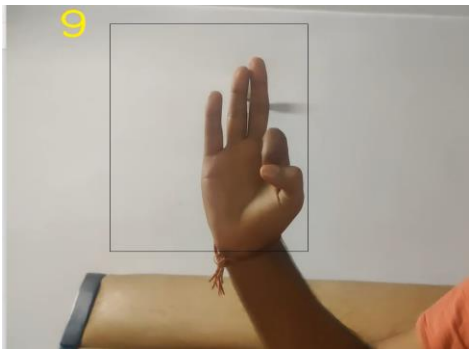
Sign language is not universal which means the different regions of the world can have different meanings for a sign which can be complex and hard to understand and remember different meanings of the same sign as well as science can be both static and dynamic. For this there are mainly two options[4] the first one is using gloves with sensors to analyze sign language and recognizing it using the data collected from the gloves this method is very expensive and this method can be effectively used to recognize dynamic signs. This method can be very expensive to be commercialized.

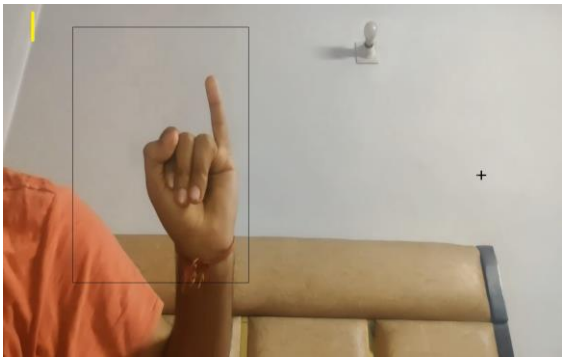
[4]The second method is using machine learning and computer vision here we use computer vision to collect data and machine algorithms that can use the data to gain experience and become efficient over time algorithms run continuously to understand the data and then predicting the answers then checking it whether it is correct and then repeating it again and again. [2]The careful feature selection and extraction are very important aspects to consider. Since visual features provide a description of the image content. This method is time consuming but the desired accuracy can be achieved without using much equipment. This method is very efficient for recognizing static signs but dynamic signs can be difficult to recognize.[6] The existing Indian Sign Language Recognition system was developed using machine learning algorithms with MATLAB.

Working

The sign language recognition system has two different parts. The first ones to detect hand gestures for this the the system uses a library called media pipe. In this library we do not need to train our system as it comes pre trained.[5] Using media pipe the detection of the hand is done and it plots 21 different points with each having a unique name to represent its position on the hand. With these points the system is able to detect whether the finger is in up position or down position. When a gesture is made in front of the camera

the system checks which fingers are in up position or in down position and is able to tell the meaning of the gesture or sign through text flashing on the screen. The second part of the system recognizes sign language, for now it is only able to detect the walls as it takes lot of time to train. For this the library OpenCV is collecting the images to create a database. It takes number of images and save them in the directory specified in the program. We have to use label image software to highlight the sign that is in the image and be label-led to specify what sign is in the image it also creates a XML file. Then we divide the images and XML files into groups that is training and testing. After that we create a label map for the signs we are training for and use TensorFlow to train the system. We have to specify how many steps we have to run the command to get the most accuracy. TensorFlow generates the code which we have to run in the CMD to train the model. Then we run the program to test the accuracy of the system.





Algorithms Used

- Convolutional Neural Network:

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning system that can take an input picture and assign importance (learnable weights and biases) to various aspects/objects in the image, as well as differentiate between them. The amount of pre-processing required by a ConvNet is much less than that required by other classification techniques. ConvNets can learn these filters/characteristics with adequate training,whereas simple techniques need hand-engineering of filters. ConvNets are multilayer artificial neural networks designed to handle 2D or 3D data as input. Every layer in the network is made up of several planes that may be 2D or 3D, and each plane is made up of numerous independent neurons composition,where nearby layer neurons are linked but same layer neurons are not.[8] A ConvNet can capture the Spatial and Temporal aspects of an image by applying appropriate filters. Furthermore, reducing the number of parameters involved and reusing weights resulted in the architecture performing better fitting to the picture collection. ConvNet's major goal is to make image processing easier by extracting relevant characteristics from images while preserving crucial information that is must for making accurate predictions. This is highly useful for developing an architecture that is not just capable of collecting and learning characteristics but also capable of handling massive volumes of data.[8]

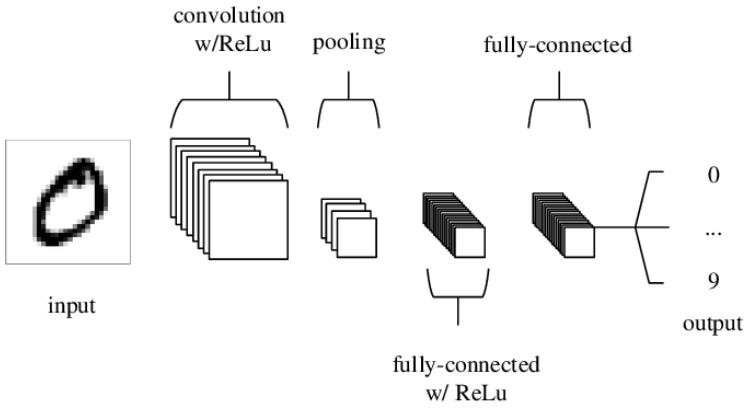


Figure-1 CNN Architecture

- SSD Mobile net V2:

The Mobile Net SSD model is a single-shot multibox detection (SSD) network that scans the pixels of an image that are inside the bounding box coordinates and class probabilities to conduct object detection. In contrast to standard residual models, the model's architecture is built on the notion of inverted residual structure, in which the residual block's input and output are narrow bottleneck layers. In addition, nonlinearities in intermediate layers are reduced, and lightweight depthwise convolution is applied. The TensorFlow object detection API includes this model.[9]

Result

The program is able to detect and tell us about the static signs and gestures. In the tables given below we can see the accuracy we are getting after the training of the model is done.

Table-1 Sign Language

Sign	Accuracy
a	90.1%
e	89.7%
i	87.8%
1	88.0%
9	89.5%

Table-2 Gestures

Gestures	Accuracy
Hello	90.2%
Yes	85.6%
No	88.3%
Thank You	91.0%

Conclusion

Sign language recognition system has been developed for classified only static signs and gestures. The development of the sign language recognition system is done using small vocabulary and self-made database. A large database for sign language recognition system is still not available for some of the signs and gestures used in different countries. Sign language and hand gestures are a powerful way for human communication with lots of potential application in the area of human computer interaction. Vision based hand gesture recognition techniques have been proven advantages compared to the traditional methods.

Future Work

This can be used as an interpreter by the disabled to interact with other people easily and it can also help them in interacting with the computer system more easily and in an efficient way. We can create a message box for the deaf and mute people in which the system will recognize the signs for alphabets and will store and merge them into a message to convey it to the other people. In future we can use bigger data sets and better testing technique for good accuracy. We can also use sensor gloves for detecting signs which will make the system to detect signs very easily and more accurately.

References

1. Wikipedia sign languages https://en.wikipedia.org/wiki/Sign_language

2. Vijay, P.K., et al., Recent Developments in Sign Language Recognition : A Review. International Journal on Advanced Computer Engineering and Communication Technology, 2012. 1(2): p. 21-26.

3. https://www.sciencedirect.com/science/article/pii/S1877050917320720

4. K. Bantupalli and Y. Xie, "American Sign Language Recognition using Deep Learning and Computer Vision," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4896-4899, doi: 10.1109/BigData.2018.8622141.

5. Mediapipe hands https://google.github.io/mediapipe/solutions/hands

6. Konstantinidis, D., Dimitropoulos, K., Daras, P.: Sign language recognition based on hand and body skeletal data. 3DTV-Conference. 2018-June, (2018). https://doi.org/10.1109/3DTV.2018.8478467.

7. https://www.who.int/en/news-room/fact-sheets/detail/deafness-and-hearing-loss

8. Al-Saffar, A. A. M., Tao, H., & Talab, M. A. (2017, October).Review of deep convolution neural network in image classification. In 2017 International Conference on Radar,37 International Journal for Modern Trends in Science and Technology Antenna, Microwave, Electronics, and Telecommunications (ICRAMET) (pp. 26-31). IEEE.

9. https://machinethink.net/blog/mobilenet-v2/ by Matthijs Hollemans [22 April 2018]