

Hand Gesture Recognition for Sign Language using CNN

Sayali Samgiskar¹ Dhanashree Sakharkar² Priyanka Sonawane³ Rahul Jadhav⁴

Prof. Yogesh Shahare⁵

^{1,2,3,4,5}MGM College of Engineering and Technology, India

Abstract— In this Hand Gesture recognition for sign language project a real time vision based system is proposed to monitor objects (hand fingers). It is built based on the Raspberry Pi with camera module and programmed with Python programming Language supported by Open Source Computer Vision (OpenCV) library. The Raspberry Pi embeds with an image-processing algorithm called hand gesture, which monitors an object (hand fingers) with its extracted features. The essential aim of hand gesture recognition system is to establish a communication between human and computerized systems for the sake of control. The mobile system is built and tested to prove the effectiveness of the proposed. It has many applications in traffic control, human computer interaction, gesture recognition, augmented reality and surveillance. It led to a system that has the ability of surveillance and applications in detecting and monitoring a known object. Raspberry pi is a small sized PC board suitable for real-time projects. The main purpose of the work presented in this project is to make a system capable of detecting and monitoring some features for objects that specified according to an image processing algorithms using Raspberry Pi and camera module. The feature extraction algorithm programmed with Python supported by OpenCV libraries, and executed with the Raspberry Pi attached with an external camera.

Keywords: Convolutional Neural Network (CNN), Python, Raspberry Pi 3B+, Gesture Recognition

I. INTRODUCTION

Hand Gesture Recognition system is widely used technology for helping the dumb and deaf people. Human hand has remained a popular choice to convey information and messages in the situation where other forms like speech cannot be used. Gesture recognition is a type of perceptual computing user interfaces that allows computers to capture and interpret human gestures as commands. The general definition of gesture recognition is the ability of a computer to understand gestures and execute commands based on those gestures. The proposed approach, we will first look at the related works done in this field. The general purpose of the Hand gesture recognition System is to make a system capable of detecting and monitoring some features for objects that specified according to an image processing algorithms using Raspberry Pi and camera module. The feature extraction algorithm programmed with Python supported by OpenCV libraries, and executed with the Raspberry Pi attached with an external camera. This system is working well even in poor illumination conditions. Hand gesture algorithm that embeds in the Raspberry Pi is used to detect and monitoring hand gesture with the image thresholding. Images thresholding is type of images processing. This image analysis technique is a type of image segmentation that isolates objects by converting grayscale images into binary images. Image thresholding is most effective in images with high levels of contrast.

II. LITERATURE REVIEW

Several methods are proposed for both dynamic and static hand gestures. [1] “A method for hand gesture recognition”, In this paper we present a method for hand gesture recognition using Microsoft kinect allow capturing dense, and three dimensional scans of an object in real time.[2] “Hand Gesture Recognition using K-means clustering and SVM”, This paper discusses hand gesture recognition as input command for biologic premium robot using k means clustering and support vector machine. [3] “Hand Gesture Recognition for sign language using KNN and svm”. This proposes two new approaches of hand gesture recognition which will recognize sign language gestures in real time environment.

A. Existing System

There are many systems available today that have implemented gesture recognition system. Some of these systems have special requirements and working conditions with proper accuracy levels depending upon the technology used. Hand gesture is used for communicating between two different people with the help of sign language. Many of the system have hand gesture conversion system form image processing to text. Which help to communicate two people with each other with help of this system. Hand gesture may be of static type or dynamic type. In static type of hand gesture, certain position and posture is used for communication purpose. Dynamic hand gesture is more suitable for real time application.

III. PROBLEM DEFINITION

We aim for developing a deaf and dumb gesture recognition system for establishing communication with the deaf and dumb people. Communicating with the dumb and deaf becomes very complicated as normal people cannot understand the sign language used by them. So for better communication with this people we are using hand gesture recognition for sign language. The Hand gesture recognition system for sign language is gesture recognition system which detect and monitoring hand moment with the help of camera module. By identifying the gesture, it will show the respective output for the same.

IV. EXITING METHODOLOGY

There are different sign languages according to countries. Most popular and widely used amongst them is the American Sign Language. There are many systems in existence today that have implemented gesture recognition system of one or the other kind. Many of these systems have special requirements and working conditions and accuracy levels depending upon the technology used and the algorithms implemented to run the system. Hand gesture is used for communicating between human and machine as well as between two different people with the help of sign language.

Hand gesture may be of static type or dynamic type. In static type of hand gesture, certain position and posture is used for communication purpose. Dynamic hand gesture is more suitable for real time application.

In Existing system KNN and SVM algorithm are used but its accuracy level for image processing was low so our system is using CNN. We use better accuracy while image processing.

A. Processing the Input Image:

The input image needs to be processed to extract the required features from it which can be compared with the database to obtain results. The Figure 2 shows an overview of the steps to be performed in processing the input image

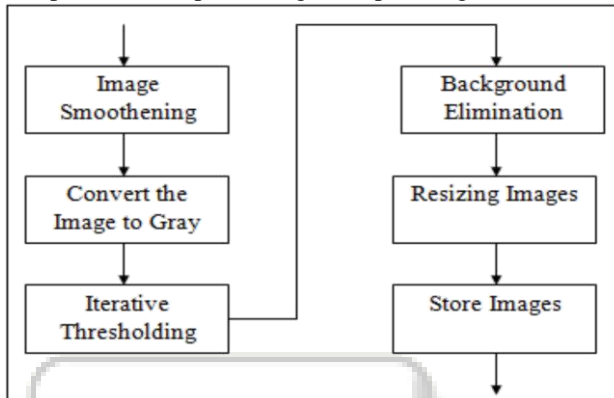


Fig. 1: Overview of steps in processing input image

Firstly, according to the proposed method the image need to be smoothed by applying the mean filter over the complete image. This will blur the image and balance the level of colors to certain extent in the image. Then convert this RGB image to grayscale. To do this, for every pixel we multiply the R component to 0.2125, G component to 0.7154 and B component to 0.0721, and then replace the corresponding pixel in the grayscale image with the sum of these products. This has to be done for every pixel in the input image. Now, this image needs to be converted to Gray image format. We use iterative threshold for eliminating background image and focus on foreground image. For this first capture 30 frames for background image after that consider hand which is foreground image captured. Images thresholding help to eliminate background and focus on foreground images. We are concerned about the white area in this image.

Thresholding=Foreground Images-Background Image

We have successfully eliminated the background image after that all capture image is resizing and store into respective folder. To implement the proposed method, we have used the tensor flow library for implementing the image processing tasks.

B. Image Thresholding Method

Image thresholding is a simple, yet effective, way of partitioning an image into a foreground and background. This image analysis technique is a type of image segmentation that isolates objects by converting RGB images into Gray images. Image thresholding is most effective in images with high levels of contrast.

While developing the method for automatic processing many issues related to images need to be taken

care of. One such problem is separating the background of the image from its foreground to make foreground image distinctly visible for future processing. That's way we used Image thresholding.

When the background of an image is removed, it is necessary that rest of the information must suffer less degradation as for as possible. The intensity of the character can be used to separate the character from background using thresholding technique by comparing the gray level or other feature to a reference value. Thresholding can be of two types: Global thresholding: where single threshold value is used for the entire image. Local thresholding method to compute the threshold for each pixel based on a neighborhood of the pixel.

C. CNN Algorithm

Convolutional neural network, like neural network, are made up of neurons with learnable weights and biases. Each neuron receives several inputs, takes a weighted sum over them, pass it through an activation function and responds with an output. The whole network has a loss function and all the tips n tricks that we developed for neural networks still applied on convolutional neural network. Neural network as it name suggest, is a machine learning technique which is model after the brain structure. It comprises of a network of leaning units called neurons. This neurons learn how to convert the input signal into corresponding output signal.

How Do convolutional Neural Networks Work?

There are four layered concepts we should understand in Convolutional Neural Networks:

- 1) Convolution,
- 2) ReLu,
- 3) Pooling and
- 4) Full Connectedness (Fully Connected Layer).

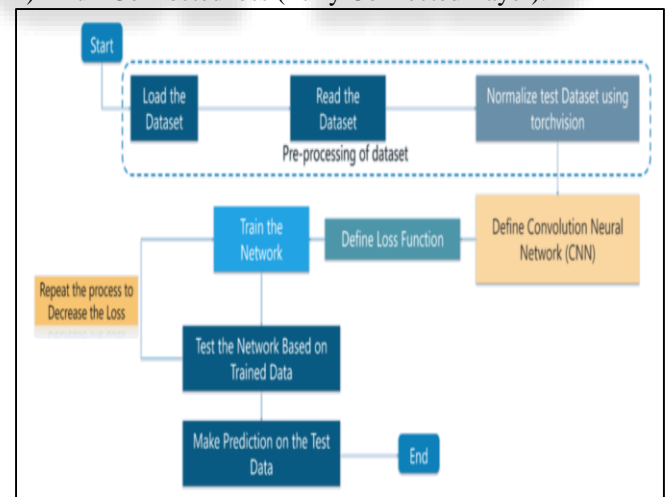


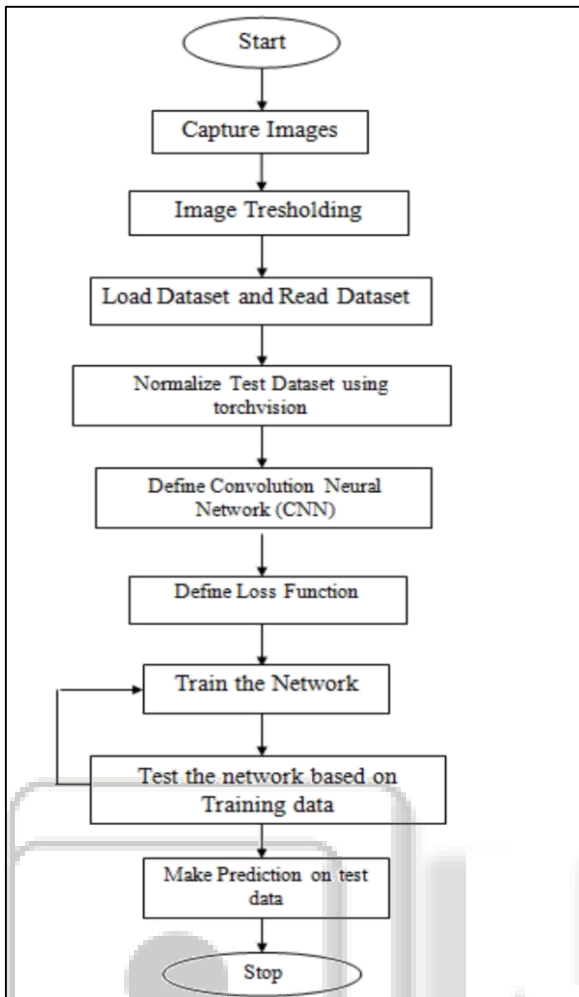
Fig. 2: Working on CNN

V. ARCHITECTURE

Hand gesture recognition system consists of the following steps.

- 1) Pre-processing and hand segmentations
- 2) Hand detection and tracking
- 3) Hand posture recognition
- 4) Hand gesture classification

There is CNN (Convolution neural Network) algorithm used.



VI. CONCLUSION

With the proposed method, we found that the probability of obtaining desired output is around 0.96 which is sufficient and can be enough to make it suitable to be used on a larger scale for the intended purpose. As compared to other existing systems, our approach uses a simpler mechanism to detect the Hand sign and compare it with the database. It can be easily implemented using the Tensor flow library as observed in our experiment. We can, therefore conclude that our proposed approach for sign language translation is better and give best results with least possible processing overhead. The proposed system does not require the use of any gloves or extra camera to capture complete images. It simply uses the computer webcam with capability of capture sharp images for our purpose. We expect that more research will be performed in this field so that the proposed system can be extended to support more gestures that make use of motion and maintain its performance results. There are great number of applications for this system that can help improve human life in different walks of life. Future scope is in our proposed system we are using American Sign Language. In future we can use on Indian language by using same proposed system. Instead of American language we can create our own language image data set for sign language

REFERENCES

- [1] Hand Gesture Recognition for HumanMachine Interaction Elena Sánchez-Nielsen, Luis Antón-Canalís and Mario Hernández-Tejera, Journal of WSCG, Vol.12, No.1-3, ISSN 1213- 6972.
- [2] Chance M. Glenn, DivyaMandloi, KanthiSarella, and MuhammedLonon, An Image Processing Technique for the Translation of ASL Finger-Spelling to Digital Audio or Text. The Laboratory for Advanced Communications Technology / CASCI ECTET Department / CAST Rochester Institute of Technology Rochester, New York 14623.
- [3] Real time Hand Gesture Recognition using a Range Camera Zhi Li, Ray Jarvis. Monash University, Wellington Road Clayton, Victoria AUSTRALIA.
- [4] Gonzalez/Woods, Digital Image Processing, Ch.10 page:599 [5] <http://www.aforge.net/>, Used for AForge.net library and framework details.
- [5] Ali A. Abed, Sarah A. Rahman, October 11, 12 2016. Computer Vision for Object Recognition and Tracking Based on Raspberry Pi. International Conference on Change, Innovation, Informatics and Disruptive Technology ICCIIDT'16, London- U.K.
- [6] John G. Allen, Richard Y. D. Xu, Jesse S. Jin, 2004. Object tracking using camshift algorithm and multiple quantized feature spaces. Proceedings of the Pan-Sydney area workshop on Visual information processing. Australian Computer Society, Inc., Vol. 36, pp. 3-7.