ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse230952020.pdf
https://doi.org/10.30534/ijatcse/2020/230952020

Design and Development of a Sign Language Gesture Recognition using Open CV

[1]S. Sai Kumar [2]Y.K. Sundara Krishna [3]Praveen Tumuluru [4]P. Ravi Kiran
[1]Assistant Professor, Dept. of IT, P.V.P.Siddhartha Institute Of Technology, India, saikumar.senagavarapu@gmail.com
[2]Professor, Dept. of Computer Science, Krishna University, India, yksk2010@gmail.com
[3]Assistant Professor, Dept. of CSE, KL University, India, praveenluru@gmail.com
[4]Assistant Professor, Dept. of CSE, Sri Kalahasteeswara Institute of Technology, India, camravikiran@gmail.com

ABSTRACT

There are different kinds of people in this world who have different types of disabilities. Being not able to speak is considered as serious disability. For helping those people, we create an application that can solve the problem of having normal conversation with others. We developed a sign language application for the easy communication between deaf and dumb people and normal people.
The purpose of our paper is to detect gestures through the webcam and display what the gesture is. Here, we have considered different signs from American Sign Language system corresponding to 26 alphabet letters. In order to make this work, a gesture is shown to the computer's webcam resulting in generation of corresponding letter from the images that were defined in dataset. There were actually some of the models, making this Paper success. The first one is CNN, that is used to differentiate gestures and the second is ROI, used to remove background images that are not necessary and also to indentify the hands having different skin tones.

Key words: CNN, ROI, Sign Language

1. INTRODUCTION

For the sign language to text conversion, there is already software exists using Arduino, which is implemented by using IOT. Here human hands are considered and some kind of gestures is made by using those hands. For detecting those gestures we use sensors. And for that purpose we wear gloves on our hands which have sensors under them. By analyzing the data taken from the input data glove, some signs and gestures are recognized [1][2][3].

The input taken from sensor gloves are analysed by Arduino board and converted to digital form i.e., result which is displayed on monitor or LCD screen[4][5].

Disadvantages
The software system that already exists(by using Arduino) results in the following drawbacks:

i. If any wire is disconnected when placing the finger on the sensor then no output will be displayed.
ii. The cost is too high as this is implemented using sensors, hand gloves and arduino's.
iii. As many wires connecting to the gloves, it restricts user's free hand movements[6].

2. PROPOSED SYSTEM

This Paper focuses on converting from sign language to text. In this system, the corresponding letter will be generated according to the gesture given to the computer's webcam using Open CV in python. The generation of such alphabet is mainly because of the three models i.e., CNN, ROI and Image pre processing we have used in the system. CNN is used mainly for differentiating gestures and ROI is used to remove background pictures, which is not necessary and also to identify the hands having different skin tones [18][2][7][8][9].

For this Paper we have considered 26 alphabet and their signs according to the American Sign Language system and the dataset for defining the gestures.

ADVANTAGES OF PROPOSED SYSTEM

i. The main advantage over existing system is that, a user can be given various gestures without any sense of restrictions that was in existing system[10].
ii. System implementation is low cost.
iii. High efficiency.

3. MODULE DESCRIPTION

The Modules present in the project are

A. Image Acquisition
B. Image Preprocessing
C. Hand Detection by colour segmentation
D. Gesture Recognition

A. Image Acquisition

In our paper, first and foremost images or videos should be captured by the webcam of the device in order detect the gestures. This process is done by using the OpenCV module which is wide spread module for the capturing of video and images in python language[11][17].

B. Image Preprocessing

After collection of the image in the image acquisition module the image needs to get preprocessed i.e, it needs to undergo the gray scale conversion to reduce the pixel density and easy processing of the image. We have used ROI(Region Of Interest) to remove the background and to identify the hands having different skin tones on the temporal features[16].

C. Hand Detection By Color Segmentation

Colour Segmentation is done by using CV tobgc method**.** The images captured in ROI are taken to detect the skin by this model. After the detection of skin(hand) from the image it is converted to binary image which recognizes the gesture[15].

D. Gesture Recognition

From the binary image that is captured, gestures are recognized from comparing to the many images in the trained dataset and gives out the specified result[14][2].

4. FUNCTIONAL REQUIREMENTS

a) Bounding Box

The Bounding Box is a rectangle drawn on the image which tightly fits the object(hand) in the image[12]. A bounding box exists for every instance of every object in the image. For the box, 4 numbers (center x, center y, width, height) are predicted. This can be trained using a distance measure between predicted and ground truth bounding box. The distance measure is a jaccard distance which computes intersection over union between the predicted and ground truth boxes [12].

b) OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products[11].

The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc[11][18].

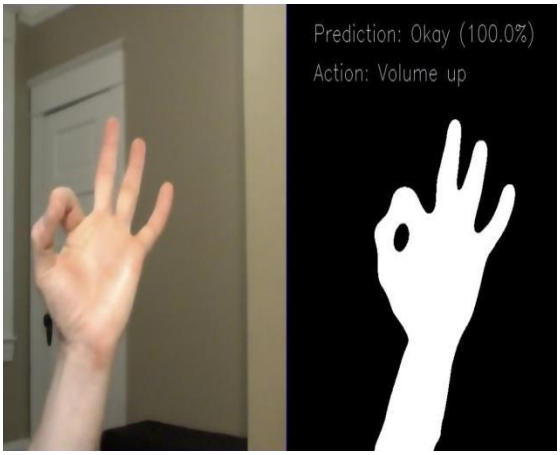


Figure 1: Generated mask image using Open CV

c) Convolution Neural Network(CNN)

The model we used is built with Keras using Convolutional Neural Networks (CNN). A convolutional neural network is a special type of deep neural network which performs extremely well for image classification purposes. A CNN basically consists of an input layer[13], an output layer and a hidden layer which can have multiple numbers of layers. A convolution operation is performed on these layers using a filter that performs 2D matrix multiplication on the layer and filter.

The CNN model architecture consists of the following layers:

- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 32 nodes, kernel size 3
- Convolutional layer; 64 nodes, kernel size 3
- Fully connected layer; 128 nodes

The final layer is also a fully connected layer with 2 nodes. In all the layers, a Relu activation function is used except the output layer in which we used Softmax [13][19]. The following figure 2 represents the different layers in CNN Model

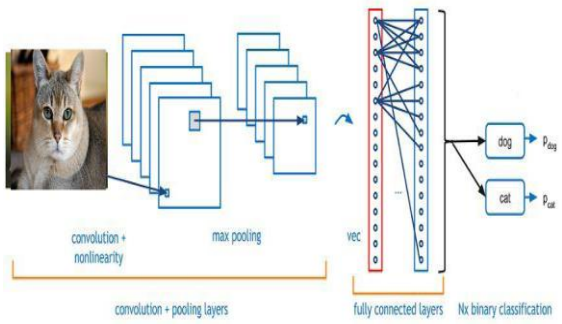


Figure 2: Different layers in CNN Model

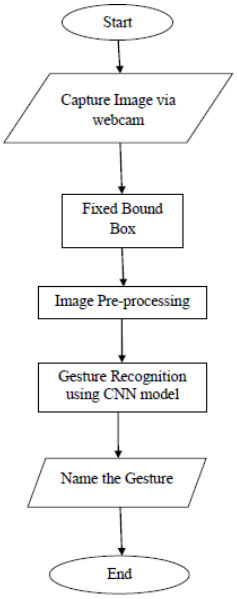
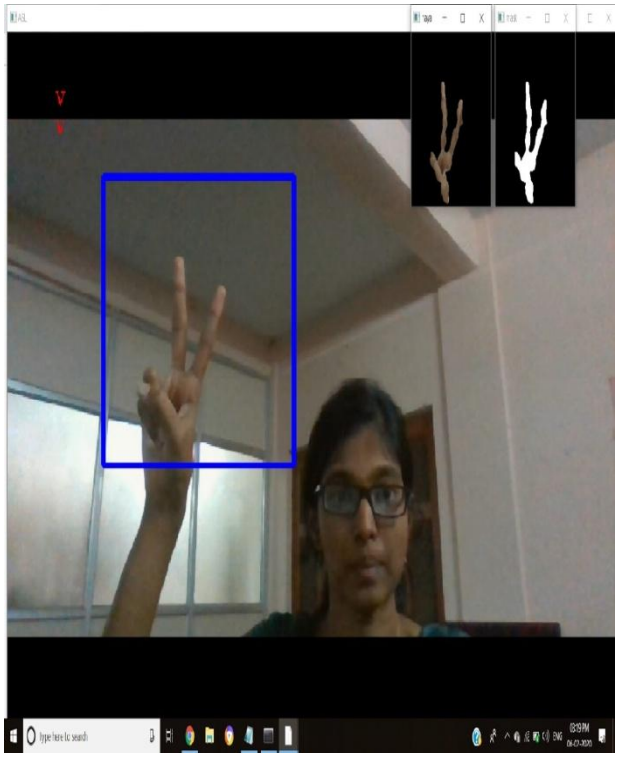


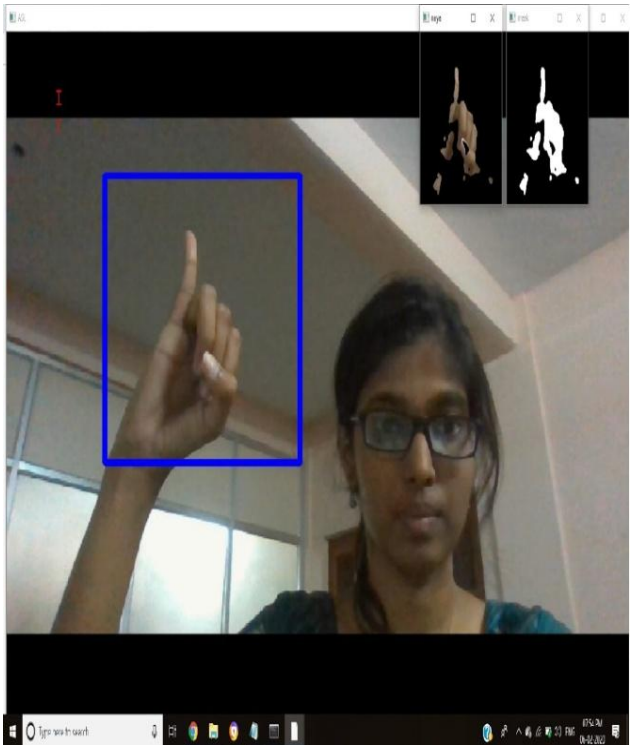
Figure 3. Flow chart for Sign Language Gesture Recognition

In the above figure 3,when started image is captured via webcam and there is a ROI (boundbox) where hand gestures are given .Image preprocessing is done on the images and by using CNN model we detect the gestures.

5. RESULTS

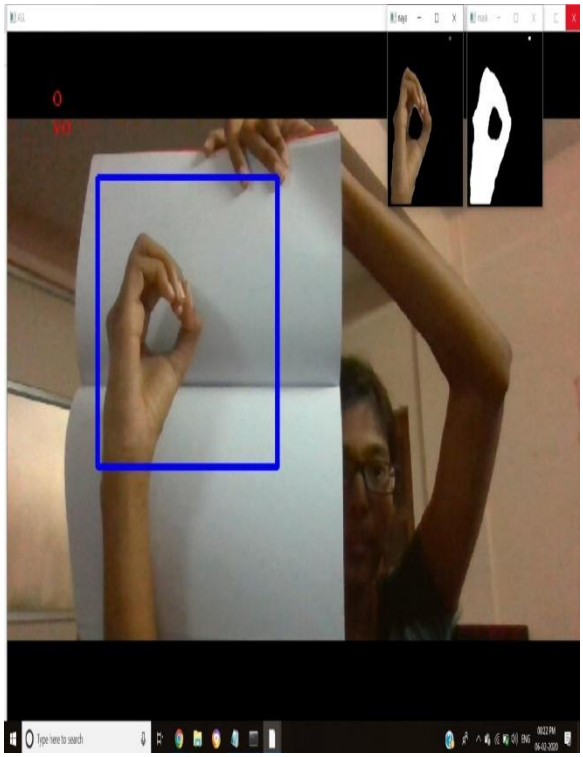


Description: Showing the alphabet V

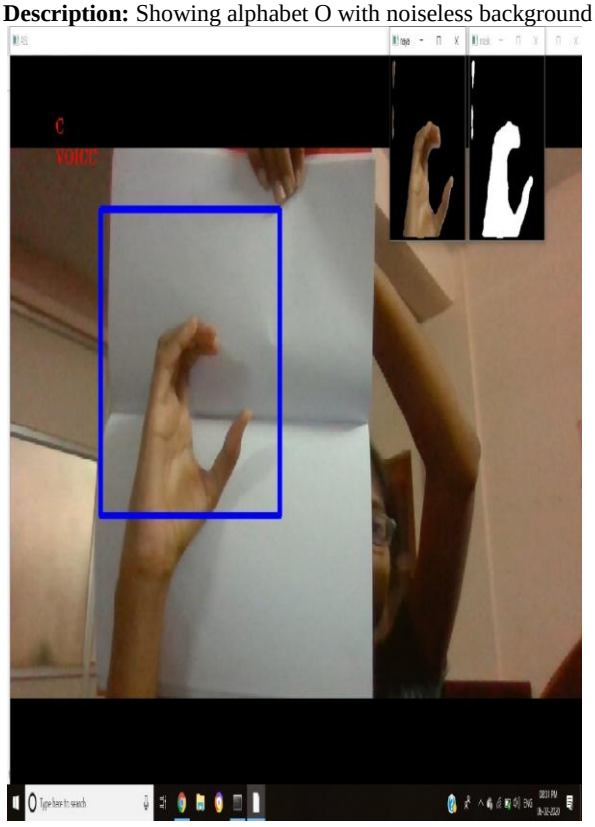


Description: Showing the alphabet I

Description: Showing the alphabet H



Description: Showing alphabet O with noiseless background



Description: Showing alphabet C with noiseless background

6.CONCLUSION

In human communication, gestures play a very prominent role especially hand gestures have lots of potential applications in the area of human computer interaction. Gesture based recognition techniques especially vision based hand gestures have many advantages compared with others. But, Current work of hand gesture recognition is only a small contribution towards achieving great results in field of sign language recognition. Here in our Paper we can only be able to recognise the gestures of alphabets from American Sign Language (ASL).

The work we have done can be further extended with better accuracy of recognition of gestures. we can move on to sentence level too in future by this method.

REFERENCES

1. Y. Zhou, G. Jiang, Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition", *Pattern Recognit.*, vol. 49, pp. 102-114, 2016.
   Show Context CrossRef Google Scholar
2. K. Liu, N. Kehtarnavaz, "Real-time robust vision-based hand gesture recognition using stereo images", *J. Real-Time Image Process.*, vol. 11, no. 1, pp. 201-209, Jan. 2016.
   Show Context CrossRef Google Scholar
3. R. Mahesh, "Gesture recognition: A survey of Gesture Recognition techniques using Neural networks", *Glob. J. Comput. Sci. Technol. Neural Artif. Intell*, vol. 13, no. 3, 2013.Show Context Google Scholar
4. L. Kulkarni, "Appearance Based Recognition of American Sign Language Using Gesture Segmentation IJCSE)", *International Journal on Computer Science and Engineering*, vol. 02, no. 03, pp. 560-565, 2010.
   Show Context Google Scholar
5. E., N. Kyriakos, A Dynamic Gesture and Posture Recognition System. DOI 10.1007/s10846-013-9983-7, 2014.
6. *Ciresan, Dan; Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber (2011). "Flexible, High Performance Convolutional Neural Networks for Image Classification" (PDF). Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence-Volume Volume Two. **2**: 1237–1242. Retrieved 17 November 2013.*
7. https://ieeexplore.ieee.org/document/7569545
8. https://github.com/hthuwal/sign-language-gesture-recognition
9. https://ieeexplore.ieee.org/document/8417200
10. https://www.geeksforgeeks.org/Paper-idea-sign-language-translator-speech-impaired/
11. https://www.researchgate.net/publication/297757591_Sign_Language_Recognition_Using_Convolutional_Neural_Networks
12. https://opencv.org/about/
13. https://www.mililink.com/upload/article/1627575595aams_vol188_june_2019_a14_p741751_shakti,_harish_saini_and_ashish_kumar.pdf
14. https://medium.com/@pallawi.ds/ai-starter-build-your-first-convolution-neural-network-in-keras-from-scratch-to-perform-a059eaa6d4ff

15. https://towardsdatascience.com/training-a-neural-network-to-detect-gestures-with-opencv-in-python-e09b0a12bdf1
16. troindia.in/journal/ijcesr/vol4iss6part3/43-49.pdf
17. https://www.analyticsvidhya.com/blog/2019/04/introduction-imagesegmentation-techniques-python/
18. https://docs.opencv.org/2.4/opencv_tutorials.pdf
19. https://worldwidescience.org/topicpages/s/sign+recognition+software.html
20. Praveen Tumuluru, Burra Lakshmi Ramani et al. "OpenCV Algorithms for facial recognition", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075,Volume-8 Issue-8 June, 2019.
21. Senapati, RK; Prasad, PMK; Swain, G; Shankar, TN, Volumetric medical image compression using 3D listless embedded block partitioning, SPRINGERPLUS, 2016, 10.1186/s40064-016-3784-y.
22. Sahu, AK; Swain, G, A Novel Multi Stego-image based Data Hiding Method for Gray Scale Image, Pertanika journal of science and technology,2019.
23. Pathak M,Srinivasu N,Bairagi V, Effective segmentation of sclera, iris and pupil in noisy eye images Telkomnika (Telecommunication Computing Electronics and Control) 2019.
24. Vallabhaneni, R. B.; Rajesh, V.On the Performance Characteristics of Embedded Techniques for Medical Image Compression,JOURNAL OF SCIENTIFIC & INDUSTRIAL RESEARCH, pp 662-665, 2017
25. Ayushree; Arora, Sandeep Kumar, comparative analysis of aodv and dsdv using machine learning approach in manet journal of engineering science and technology, 3315-3328, 2017
26. Kavitha, K.; Sandhya, B.; Rao, B. Thirumala, Evaluation of Distance Measures for Feature based Image Registration using AlexNet, International journal of advanced computer science and applications, 284-290,2018.
27. Jaya Lakshmi,R, & Subba Rao, GV 2017, "A re-constructive algorithm to improve image recovery in compressed senescing", Journal of Theoretical and Applied information technology, vol-95, no-20, pp 5443-5453.
28. Sheela Rani, C.M, Dheeraj,K,Sai Vineeth reddy,P& Satyasai,K.2018, "Image segmentation for pattern reorganization in surveillance", International Journal of Engineering and advanced technology, vol-7,no-3,pp 45-49.
29. Sampath Dakshina murthy, A., Satyanarana murthy, P,Rajesh,V., Hasane ahammed,S. & Omkar Lakshmi jagan,B. 2019, "Execution of natural random forest machine learning compression", International journal of Pharmacetical research, vol-11, n0.4,pp.1241-1255.