Tool Development for Detection of Sign Language

Nikhil Sanjay Shetiya[1], Ishwar Vikas Lonkar[2], Ayara Lovlesh Rangari[3]

Department of E&TC, SKNCOE, SPPU, Pune.

[1]nikhilsethiya03@gmail.com

[2]ishwarlonkar2206@gmail.com

[3]informtoayara@gmail.com

Abstract:

A system was developed which may function a learning tool for starters in language that involves hand detection. This system relies on a skin-color modelling technique, i.e., explicit skin-color space thresholding. The skin-color range is predetermined which may extract pixels (hand) from non–pixels (background). The pictures were fed into the model called the Convolutional Neural Network (CNN) for classification of images. In supervised learning algorithm is trained with well labelled data. It simply meaning data is already tagged with correct answer. The algorithm learns from the labelled data and predicts outcomes of unforeseen data. From given information machine learning algorithm is trained. The training algorithm is additionally a relation which learns from information then helps to predict the results of unforeseen data. It also learns from the knowledge it's predicted for. In machine learning the prediction gets better and better with the more use of it because it learns from the predictions.

Keywords: *CNN, Sign Language Detection, Deep Learning*

I .INTRODUCTION

Communication is crucial in building a nation. Good communication finishes up in better understanding, and it encompasses all the members of the community, including the deaf. within the Philippines, 1.23% of the whole population is either deaf, mute or hearing impaired Sign language bridges the gap of communication with people. However, most hearing people don't understand signing and learning it isn't an easy process. As a result, there's still an undeniable barrier between the hearing impaired and hearing majority. Over the past few decades, many efforts are made in creating an indication language recognition (SLR) system. There are two main categories in SLR, namely isolated language recognition and continuous sign classification. There are other approaches to capturing signs by using motion sensors, like electromyography (EMG) sensors, RGB cameras, Kinect sensors, leap motion controllers or their combinations .The advantage of this approach has higher accuracy, then the weakness is that it's limited movement. In recent years, the involvement of vision-based techniques for detection has become more popular, which takes input from camera (web camera, stereo camera, or 3D camera). Sandjaja and Marcos used color-coded gloves to create hand detection easier. a mix of both architectures is additionally possible, which is termed the hybrid architecture . While these are more cost effective and fewer constraining than data gloves, the weakness of this approach is lower accuracy and high computing power consumption. Most of the studies mentioned above consider translating the signs typically made by the hearing- impaired person or the signer to word(s) that the hearing majority or non-signer can understand. Although these studies proved that technology is helpful in numerous ways, their proponents think that these are intrusive to some hearing–impaired individuals. Instead, the proponents proposed a system which is ready to assist those non-signers who want to travel looking basic static signing and not being intrusive at the identical time.

II. LITERATURE SURVEY

For the past decades, research on SLR has been explored. Many studies used sensor-based devices such as Sign Speak. There are other means of capturing signs by using motion sensors, such as electromyography (EMG) sensors , RGB cameras , Kinect sensors , and leap motion controller or their combinations. Although these sensors provide accurate parameters in measurement of information, they even have limitations; first is their cost, as they require large-size datasets with diverse sign motion they going toned a high-end computers with powerful specifications; next is aesthetics, because the sensors are attached to the fingers and palms of a user, the user can encounter difficulties in setting up the device; ambient lighting conditions or backgrounds in real-world settings may also affect the popularity. Therefore, many researchers jumped from sensor-based to visual-based SLR. Several methods are developed in visual-based SLR. Because signing includes static and dynamic movements, image, and video processing was explored by many. Wang et al. used colour spaces to identify hand gestures and purchased segment images by setting a variety of the colour threshold. Hand gesture segmentation is solely done by using the hand skin threshold method. The system would not produce good results because of lighting conditions, complexion interference, and complicated backgrounds that increased noise. There are three styles of colouring detection. Another approach was taken by Balbin et al. , who used colored gloves for the hands to be identified easily by setting a precise range of the hand colouring threshold (colour of the gloves).To recognize the hand gesture, input images underwent various image processing methods or steps. First is pre-processing wherein images were converted into grayscale, and median filter is accustomed denoise the image. Next, is feature extraction wherein the colour of the hand gloves was detected and isolated from the background. Then, the image had undergone pattern recognition. The system used Kohen self-organizing maps, which are the sort of a neural network which will learn to spot patterns and group datasets in an unsupervised manner. The system was tested by five persons, and it achieved an accuracy of 97.6%. These studies propose a fancy yet manageable process of colouring thresholding; it is seen that when only the bare hands of the signer are used, it's difficult for the system to acknowledge the gesture due to different hindrances like noise. Other studies used colored gloves to solve the matter, whereas this study proposed a system that may recognize static linguistic communication without the help of gloves or hand markings but still produce acceptable results.

Existing Scenario

Real-Time Hand Gesture Detection and Recognition Using Bag Of Features. This project is finished by Nicolas D. Georganas and Nasser H. Dardas . This paper presents a novel and real-time system for interaction with an application or video game via hand gestures. Our system includes detecting and tracking bare hand in cluttered background using skin detection and hand posture contour comparison algorithm after face subtraction, recognizing hand gestures via bag-of-features and multiclass support vector machine (SVM) and building a grammar that generates gesture commands to control an application. This histogram is treated as an input vector for a multiclass SVM to make the training classifier. In proposed system for every frame captured from a webcam, the hand is detected using the algorithm.

HAND GESTURE TECHNIQUES FOR LANGUAGE RECOGNITION.This project is completed by Ms Kamal Preet Kour, Dr. (Mrs) Lini Mathew. Hand gesture recognition provides an intelligent, natural, and convenient way of human–computer interaction (HCI). signing recognition (SLR) and gesture-based control are two major applications for hand gesture recognition technologies. SLR aims to interpret sign languages automatically by a computer so on assist the deaf communicate with hearing society conveniently. Since signing could also be a fairly highly structured and largely symbolic human gesture set, SLR is also an honest basic for the event of general gesture-based HCI. This paper deals with the varied algorithm and techniques used for recognizing the hand gesture. Sign language is one among the tool of communication for physically impaired, deaf and dumb people. Real-time American English recognition using desk and wearable computer-based video.This project is finished by Starner T, Weaver J, Pentland IEEE. We present two real-time hidden Markov model-based systems for recognizing sentence-level continuous American English (ASL) employing one camera to trace the user's unadorned hands. the first system observes the user from a desk mounted camera and achieves 92 percent word accuracy. The second system mounts the camera during a cap worn by the user and achieves 98 percent accuracy

III. METHODOLOGY

In this paper we are proposing the tool development for detection of static sign language. There are ways other ways for image classification, pattern detection but the most popular as well as accurate one is Convolutional Neural Networks (CNN).

Data Set Generation:

We decided to create our own dataset for this project as we couldn't find a suitable raw image dataset in the form of RGB values which matched all our requirements.

In order to do that we used OpenCV to make our dataset. We almost captured 2000 images per sign. Then we used 80% for model training that is 1600 and 20% that is 400 for testing.

We defined a Region of Interest(ROI) then we converted that image in RGB as shown below.

Fig 1 a Region of Interest(ROI) then we converted that image in RGB

Finally, with gaussian blur filter we extracted the various features of hand/gesture. The above image after applying gaussian blur filter looks like shown below

Fig 2 a Region of Interest(ROI) then we converted that image in RGB

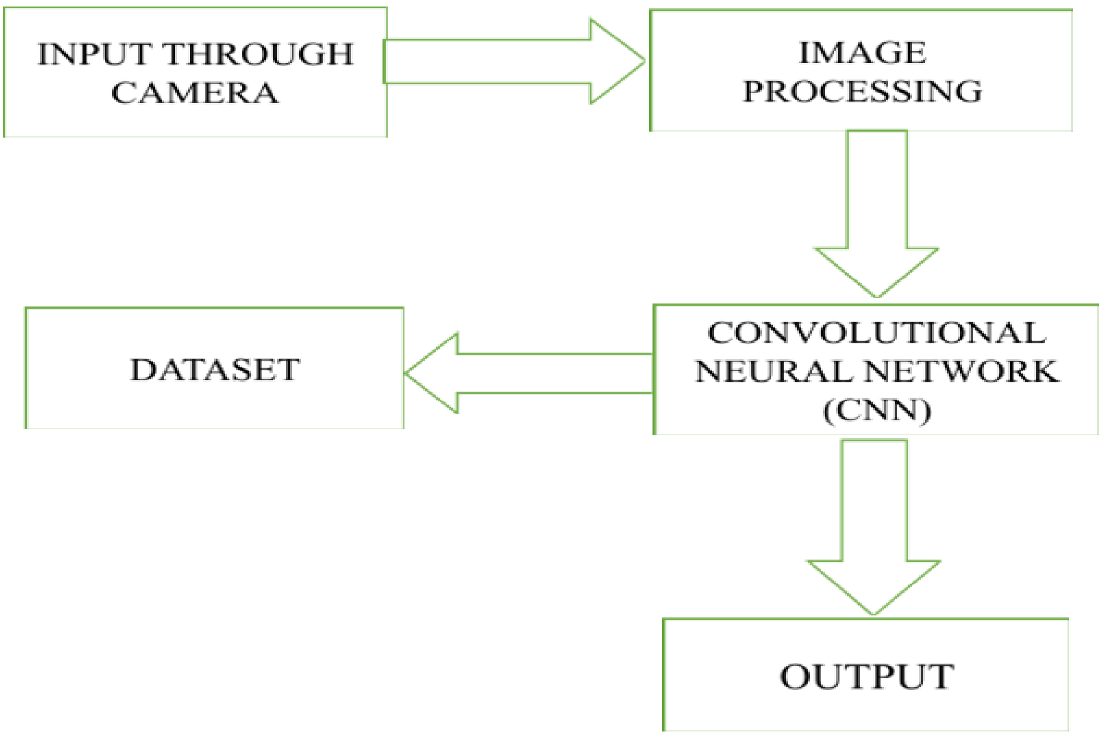

Fig 3 Block Diagram

Artificial Neural Networks:

Artificial Neural Network is a connection of neurons, replicating the structure of human brain. Each connection of neuron transfers information to another neuron. Inputs are fed into first layer of neurons which processes it and transfers to another layer of neurons called as hidden layers. After processing of information through multiple layers of hidden layers, information is passed to final output layer.

CNN:

CNN is a type of deep learning neural network algorithm. It uses randomly defined patches for input and changes itself while training, as the training is completed these patches will help to predict the results. About 80% of the images are used for training purpose and 20% of all the images will be used for testing. Convolution networks has solved the problem of classification as it corresponds to the data distribution.

In CNN, with the help of matrix filters pixels are convoluted this also shrinks the total size of image that later on helps with all the computations. We have used ReLu (Rectified Linear Unit) in each of the layers (convolutional as well as fully connected neurons).ReLu calculates max(x,0) for each input pixel. This adds nonlinearity to the formula and helps to learn more complicated features. It helps in removing the vanishing gradient problem and speeding up the training by reducing the computation time. Second most important transformation in this process is pooling. There are types of pooling e.g., max pooling, min pooling, avg pooling and these pooling techniques are used as required. At last, the grouping layer comes at work. It is responsible for reducing the dimensionality of the data and very useful for avoiding over fitting. After using folding and grouping layers, the output is directed to a fully connected.

IV. RESULTS

The developed tool with the help of CNN will distinguish between hand gestures to give outputs.

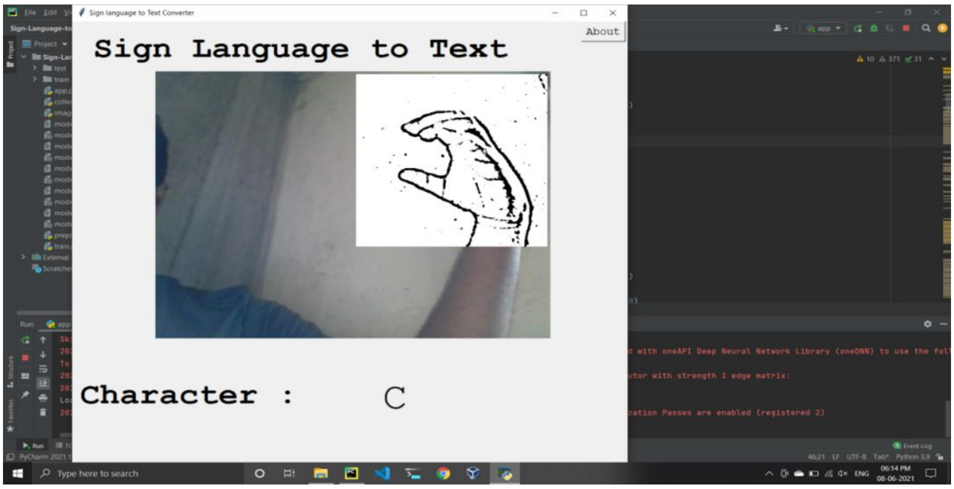

Fig 5 Gesture outline

In the screen we have adjusted another window just to show the gesture outlines. That way the user will get to know where the hand should actually to be held for accuracy of the prediction of corresponding hand gesture. we almost got the accuracy of prediction for every sign but there are still some signs e.g. T,K,D and I whose gestures are almost as same which is slightly mispredicted sometimes.

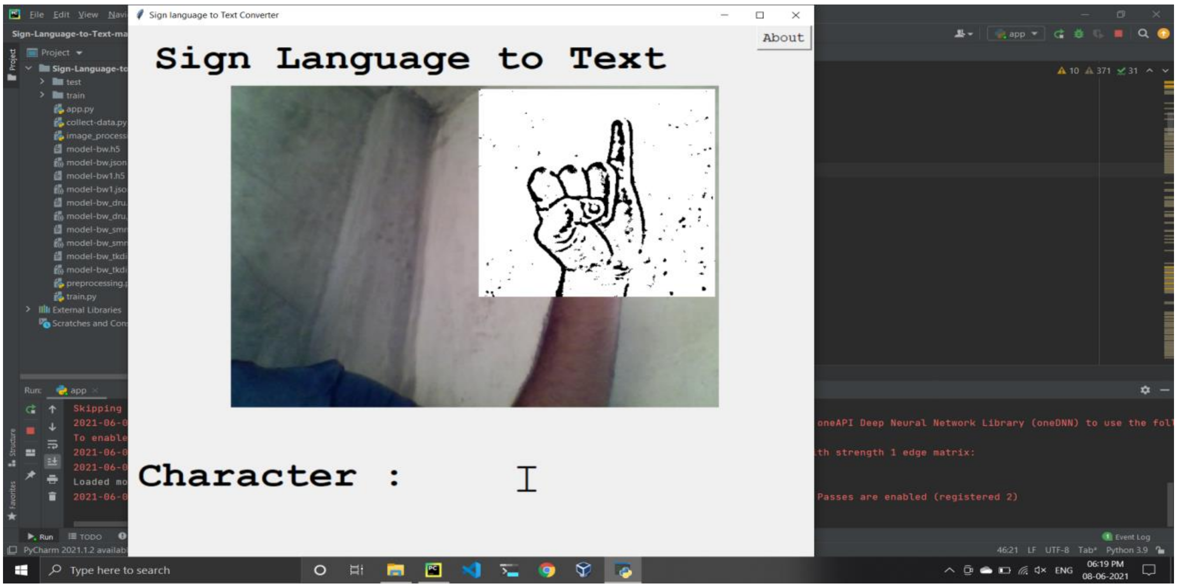

Fig 6 Sign language to text

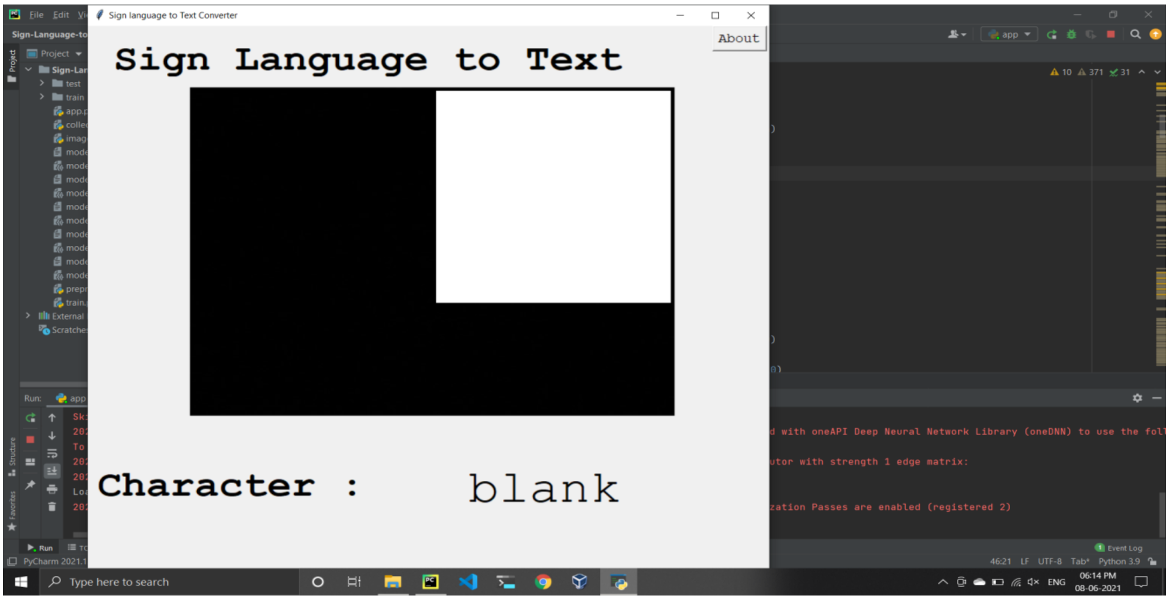

Fig 7 Sign language to text

V. CONCLUSION

There has been incredible research and advancement in this field. Aiding the cause to help the specially (deaf and mute) abled community with the help of Deep Learning and Computer Vision to be able to communicate with maximum ease. There are still some challenges for them. In this project we have made a Sign language detector which can detect 26 alphabets easily and which can be extended to detect numbers as well. The objective behind this project was to developed a tool that can detect hand gestures into its corresponding characters to familiarize users with fundamentals of sign language.

VI. REFERENCES

[1]T. Starner, J. Weaver, and A. Pentland, "Realtime american sign language recognition using desk and wearable computer-based video," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 12, pp. 1371-1375, 1998.

[2] C. Vogler and D. Metaxas, "Handshapes and movements: Multiple-channel American sign language handshapes and movements: Multiple-channel ASL recognition," in Proc. International Gesture Workshop, 2003, pp. 247-258.

[3] M. Süzgün, H. Özdemir, N. Camgöz, A. Kındıroğlu, D. Başaran, C. Togay, and L. Akarun, "HospiSign: An interactive sign language platform for hearing impaired," Journal of Naval Sciences and Engineering, vol. 11, no. 3, pp. 75-92, 2015.

[4] J. A. Deja, P. Arceo, D. G. David, P. Lawrence, and R. C. Roque, "MyoSL: A Framework for measuring usability of two arm gestural electromyography for sign language." in Proc. International Conference on Universal Access in Human-Computer interaction, 2018.

[5] Ms. Kamal Preet Kour, Dr. (Mrs.) Lini Mathew. "Sign Language Recognition Using Image Processing", in International Journal of Advance Research in Computer Science and Software Engineering.

[6]P. Mekala, Y. Gao, J. Fan, and A. Davari, "Real-time sign language recognition based on neural network architecture," in Proc. 2011 IEEE 43rd Southeastern Symposium on System Theory, 2011, pp. 195–199.

[7]J. P. Rivera and C. Ong, "Facial expression recognition in filipino sign language: Classification using 3D Animation units," in Proc. the 18th Philippine Computing Science Congress (PCSC 2018), 2018, pp. 1-8

[8]C. Ong, I. Lim, J. Lu, C. Ng, and T. Ong, "Sign-language recognition through Gesture & Movement Analysis (SIGMA)," Mechatronics and Machine Vision in Practice, vol. 3, pp. 232-245, 2018.

[9]I. N. Sandjaja and N. Marcos, "Sign language number recognition," in Proc. 2009 Fifth International Joint Conference on INC, IMS and IDC, 2009, pp. 1503-1508.