Sign Language Hand Gesture Detection Using Machine Learning

Raghul PV[1], Ranjith T[2], Rajesh Kumaran A[3], Jainulabudeen SAK[4]
[1, 2, 3] Dept of Computer Science Engineering
[4]Assistant Professor, Dept of Computer Science Engineering
[1, 2, 3, 4] Panimalar Engineering College, Chennai

Abstract- *Sign Language and gestures is an emerging topic in today's technologies. The main focus of this is to recognize the human gestures using mathematical algorithms for human computer interaction. With the help of computers, we can communicate with the people who are having hearing impairment. Each of these communicating software has their own limitations when it comes to adapting more versatile hardware in computers. Gesture recognition is one of the essential techniques to build user-friendly interfaces. Usually gestures can be originated from any bodily motion or state, but commonly originate from the face or hand. Gesture recognition enables users to interact with the devices by identifying the actions performed by the user. This project describes how hand gestures are trained to perform certain actions to reflect the actions of the output given by the user. Gestures are used to give the input and the output is calculated.*

Keywords- Machine Learning , Convolution Neural Network, user interface , Capture and Detection

I. INTRODUCTION

Humans are able to recognize body and sign language easily. This is possible due to the combination of vision and synaptic interactions that were formed along brain development. The requirements for hand detection involve the input image from the webcam. The image should be fetched with a speed of 20 frames per second. Distance should also be maintained between the hand and the camera. Approximate distance that should be between hands the camera is around 30 to 100 cm. The video input is stored frame by frame into a matrix after preprocessing. In order to replicate this skill in computers, some problems need to be solved: how to separate objects of interest in images and which image capture technology and classification technique are more appropriate, among others. In this way, these devices are capable of capturing human gestures, developing a new medium of human-machine interaction. The uses of these devices are present in the most diverse areas, such as robotics, medicine, sign language translation, computer graphics, and augmented reality. Gesture recognition methodologies are usually divided into two categories: static or dynamic. Static gestures are those that only require the processing of a single image at the input of the classifier, the advantage of this approach is the lower computational cost. Dynamic gestures require the processing of image sequences and more complex gesture recognition approaches. In the literature, we can find several recognition methodologies based on supervised and unsupervised learning. The method of neural networks to classify the data extracted from the images which a special type of neural network is used, called learning vector quantization.

II. LITERATURE SURVEY

Author Plouffe, Guillaume &Cretu, Ana-Maria[1] developed a static gesture based application for real time system which is based on contour detection. These Authors have used K-curvature algorithm for contour detection and tracking over fingertips and Dynamic time wrapping (DTW). Dynamic time wrapping (DTW) uses Euclidian distance to find shortest path of singer's curvature.

In paper [2] author Meenakshi Pawar proposed an algorithm for Hand gesture recognition method which is purely based on Image processing and they have not used any training model. The whole process of Gesture recognition starts with image enhancement and segmentation process. Firstly, RGB image is converted YCbCr and then image is converted into binary image on which segmentation is applied. After Segmentation vertical or horizontal orientation is determined and then centroid of finger is obtained. Finally Finger region is identified using finger tips.

Hanwen Huang[3], used a different algorithm which is based on skin tone after that contour detection and segmentation process is carried out followed by gesture recognition. Author Guillaume Devineauetal [4] introduces a CNN based hand gesture recognition using deep learning for Skelton data using parallel convolution methods. They have used two separate channels based on multi perceptron for feature extraction methods. They have carefully chosen GLOROT uniform initialization for training purpose and negative log-likelihood for loss estimation[9].

Author Ali A. Alani[5] et al also proposed hand gesture recognition system using CNN like authors [4],[9] and [13], but he uses adaptive CNN(ACNN) to detect hand gesture accurately regardless of hand size or hand position. To increase the accuracy before applying Adaptive CNN, the hand images are moved horizontal and vertical direction randomly up to 20%, which increases the size of data size images and this ACNN algorithm also helps in fine tuning of parameters. Author Duhart and his team[6] deduced that the research and implementation of ANN had been a revolution in further improvement. It has given rise to Distributed Artificial Neural Controller (ANC) which can be easily used in distributed environment. Thus, we can build a global Artificial Neural Network which is composed of multiple ANC's and different mechanisms which can synchronize them. The results of testing ANN come to the conclusion that the learning rate increases significantly with the number of layers and neurons. Thus making it much faster and productive.

In [7], the authors performed a research on the flow of the hand gesture recognition and made use of the AdaBoost classifier based on the Haar feature extraction to extract the features for the gestures in a unsuitable environment. The authors also made use of the CamShift algorithm for tracking the hand gesture and identifying the hand gesture area in real-time[10].

In [8], the authors proposed a model which consisted of two parts, back-end and front-end. The back-end system consists of three modules: Camera module, Detection module and Interface module. The camera module was responsible for connecting and collecting data from different picture detector styles and transmitting this information to the frame detection module. The detection module was responsible for the image processing and finally the interface module is responsible for mapping the detected hand gestures to their associated actions.

In [9], the authors studied the Convolutional neural networks (CNNs) and stacked denoising autoencoders (SDAEs) for the recognition of 24 ASL (American Sign Language) hand gestures. They trained their model on the public databases on which recognition rates of 91.33 and 92.83 %.

The authors made use of the rectified linear activations in the hidden layers of the network, to overcome the difficulties in the learning of CNN due to the increased depths. Authors Dardas and Georganas [10], stated that the gesture detection is divided into three modules namely hand detection, tracking of hand and multiclass SVM. Here, multiple algorithm and techniques are also explained. For feature extraction they have used SIFT algorithm. K-means

clustering is used to map unified dimensional bag-of-words vector, which is building block for SVM classifier. As per the results the accuracy rate is 96.23%. In the paper of Nikam and Ambekar [11], convexity hull algorithm is used for detection of gesture. But the interesting part is instead of using RGB images which are by default captured by webcam, they are converted to HSV. So it helps to tackle with lightning condition requirements. Here multiple steps are performed to filter image and remove noise.

As per the results the accuracy rate is above 95%. Siddharth S. Rautaray [12], in 2012 found out that evaluations were not proper while studying 3D recognition patterns and proposed his study to preview existing systems and publish a detailed survey dependent on the standard set by organizations. He found out that appearance-based hand gesture representations were more prioritized than the 3D based gesture representations in the hand gesture recognition systems. As industrial applications required accurate advances and troubleshooting systems rather than error prone systems in the man to machine and machine to machine interactions, more interactive means of communication was prioritized done by hand gesture systems.

Alani and his team [5] in 2018 proposed Adapted Deep Convolutional Neural Network (ADCNN) which was adaptable for classifying hand gestures as static datasets which has varied properties like noise, light, granularity. Data augmentation was used for obtaining desired image from the real input image and consequently in turn increased dataset size. The augmented images that was produced for better recognition improved model training. ADCNN was enhanced by use of dropout regularization and L2 Regularization to overcome the difficulty of data overflowing and fitting and enhancing the training of model with best suitable input.

Hung-Yuan Chung [13], in 2019, productively combined the old image processing method alongside with the tracking method and the deep CNN that has been popular in recent years in hand gesture recognition research, achieving good recognition results given a reasonable computational load. The proposed overall hand gesture recognition system effectively combines three key components, namely, hand detection, hand tracking, and hand recognition. For hand detection, we use skin segmentation, noise processing, and background subtraction to detect the ROI of the first frame entering the webcam. For hand tracking, we use this ROI as the initial position of the tracking, and train the initial model by using the KCF algorithm.

III. WORKING OF PROPOSED SYSTEM

The proposed hand gesture detection and recognition methodology using CNN classification approach with enhancement technique stated in this paper achieves high performance with state-of-the-art methods.

We are inputting with the help of a webcam. The webcam will process the gestures and the actions which are made by the user will reflect in the system. Then the action will be converted into a command which the user is trying to convey.
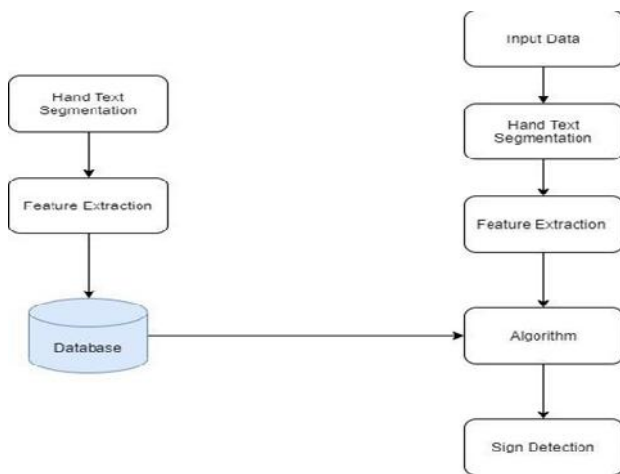

Figure 1: Gesture Detection Modules

IV. MODULE DESCRIPTION

Capture scene:

Captures the images through a web camera, which is used as an input to the system that is built. Prprocessing: Images that are captured through the webcam are compared with the dataset to recognize the valid hand movements that are needed to perform the required actions.

Hand Detection:

The requirements for hand detection involve the input image from the webcam. The image should be fetched with a speed of 20 frames per second. Distance should also be maintained between the hand and the camera. Approximate distance that should be between hand the camera is around 30 to 100 cm. The video input is stored frame by frame into a matrix after pre-processing.

Recognition:

This step involves Gesture Recognition: The number of fingers present in the hand gesture is determined by making use of defect points present in the gesture. The resultant gesture obtained is fed through a 3Dimensional Convolutional Neural Network consecutively to recognize the current gesture. Performing action: The recognized gesture is used as an input to perform the actions required by the user. These actions include zooming in, zooming out and swiping the page left or right.

Working:

1. Detection of Hand: This module detects the hand gesture by capturing an image through the web camera.
2. Detection and Defect Calculation: Finger tips are detected and the background space around its eliminated.

V. IMPLEMENTATION

With the implantation of TensorFlow Framework and Keras model for deep learning and training themodel, gesture recognition accuracy has inevitably increased. As a new image get inputted into the system, it is compared with the already trained model and processed thereafter. Once gesture data are processed, matched and evaluated, if the gesture is recognized and properly classified by the model, it carries out the respective operation by issuing back response to the gesture system.

We used the hand gesture recognition database that Kaggle provided for our project. It includes 200 pictures of different hands and hand gestures. The data collection includes a minimum of 10 hand movements involving 10 separate individuals. We have designed an Web UI to interact with the software. The Webcam identifies our action and it will convert the gestures and it will show what intended output is shown with accuracy. With the help of gestures, we can understand what command the user is trying to show.
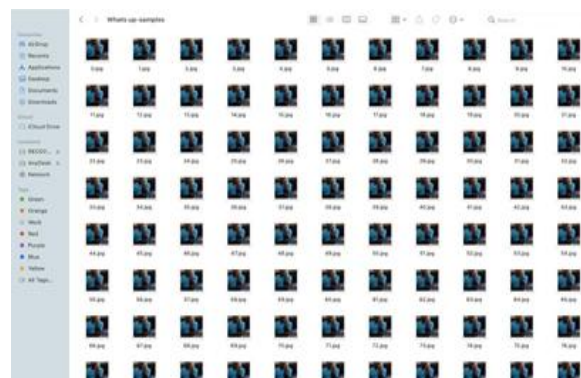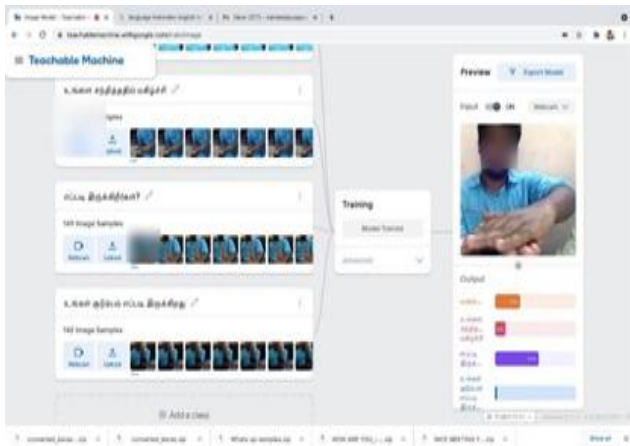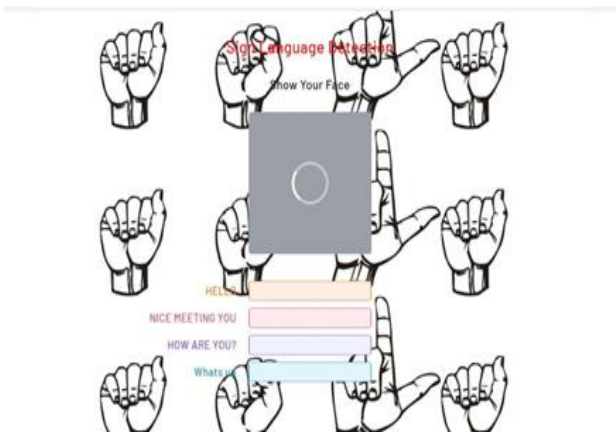

Figure 2: Dataset Images

Figure 3:Training Images
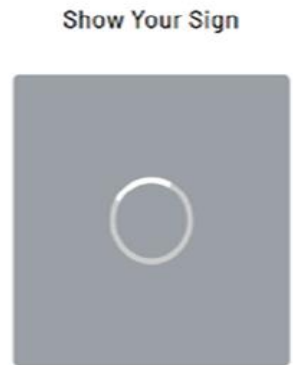

Figure 4: English WebUI
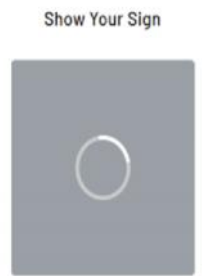

Figure 5: English WebUI

Figure 6: Hindi WebUI

Figure 7: FrenchWebUI

VI. CONCLUSION

The results also showed that the gesture recognition application was quite robust for static images. Sometimes, the video version was enormously affected by the amount of illumination, such that was necessary to check and adjust the HSV values for skin colour when starting the program to get the proper output. The adjustment was difficult to do because of the lighting conditions and the amount of objects in the background. Slight hand movements could affect gesture recognition. This gesture recognition system not only gravities

technological development but also eases human interfacing with computer, in case of person with disabilities it eases accessibility and makes particular individual independent of for performing any activity.   Nevertheless, if the hand is steady enough for long enough; the program outputs the correct command. It was also observed that while the program was executing there were memory.leaks. Attempts to remedy the problem were made by using the OpenCV functions to release memory.

REFERENCES

[1] Plouffe, Guillaume & Cretu, Ana-Maria. (2015), "Static and Dynamic H and Gesture Recognition in Depth Data Using Dynamic Time Warping," IEEE Transactions on Instrumentation and Measurement. 10.1109/TIM.2015.2498560.

[2] Panwar, Meenakshi & Mehra, Pawan. (2011), "Hand gesture recognition for human computer interaction," 1-7.10.1109/ICIIP.2011.6108940. International Journal of Advanced Science and Technology Vol.29, No. 7,(2020), pp.14186-14197 ISSN: 2005-4238 IJAST Copyright ©2020 SERSC 14197

[3] Hanwen Huangetal, "Hand Gesture Recognition with Skin Detection and Deep Learning Method" ,J.Phys.: Conf. Ser. 1213 022001

[4] Guillaume Devineau,WangXi, Fabien Moutarde, Jie Yang, "Deep Learning for Hand Gesture Recognition on Skeletal Data," 13th IEEE Conference on Automatic Face and Gesture Recognition (FG'2018), May2018,Xi'An, China. 10.1109/FG.2018.00025.hal01737771

[5] A.A.Alani,G.Cosma, A.Taherkhani and T.M.McGinnity, "Hand gesture recognition using an adapted convolutional neural network with data augmentation, "2018 4th International Conference on Information Management (ICIM), Oxford, 2018, pp.5-12.doi:10.1109/INFOMAN.2018.8392660

[6] Duhart, Clément & Bertelle, Cyrille & Paris, Ece & Lacsc, France & Fr,.(2014), "Methodology for Artificial Neural Controllers on Wireless Sensor Network," 10.1109/ICWISE.2014.7042663.

[7] J.Sun,T.Ji,S.Zhang,J.YangandG.Ji," Research on the Hand Gesture Recognition Based on Deep Learning," 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Hangzhou, China, 2018, pp. 1-4.

[8] A.Haria, A.Subramanian, N.Asokkumar, S.Poddar, J.S.Nayak, "Hand Gesture Recognition for Human Computer Interaction", 7th International Conference on Advances in Computing & Communications, ICACC-2017, 22- 24 August 2017, Cochin,India.

[9] AtulBKathole,Dr.DineshN.Chaudhari,"Pros&ConsofMachinelearningandSecurityMethods", 019.http://gujaratresearchsociety.in/index.php/JGRS,ISSN:0374-8588,Volume21Issue4

[10] Atul B Kathole, Dr.Prasad S Halgaonkar, AshviniNikhade, " Machine Learning & its Classification Techniques ",InternationalJournal of Innovative Technology and Exploring Engineering (IJITEE)ISSN:22783075,Volume-8Issue-9S3,July2019.

[11] O.K.Oyedotun,A.Khashman,"Deeplearninginvision-basedstatichandgesturerecognition",TheNatural Computing Applications Forum2016.

[12] N. H. Dardas and N. D. Georganas, "RealTime Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques," in IEEE Transactions on Instrumentation and Measurement,vol.60,no.11,pp.3592-3607,Nov.2011,doi:10.1109/TIM.2011.2161140.

[13] A.S.NikamandA.G.Ambekar,"Signlanguagerecognitionusingimagebasedhandgesturerecognition techniques," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, 2016, pp. 1-5, doi:10.1109/GET.2016.7916786