SIGN LANGUAGE TO TEXT CONVERTER

[1]Yatharth Khosla, [2]Saurabh Mehanti, [3]Harshal Goyal, [4]Priyanka Garg, [4]Ms. Charul Dewan
[1,2,3,4]students, [5]Assistant Professor
Department of Information Technology
ADGITM, New Delhi, India

Abstract— Hand gesture recognition is a field of study which is important for computer and human interactions. In this paper, we present a real time recognition of hand gestures which a human makes. In our work, the hand portion is extracted from the background with the background subtraction method. Then, the palm and fingers are segmented so as to detect and recognize the fingers. Finally, a rule classifier is applied to predict the labels of hand gestures. Moreover, our method shows better performance than a state-of-art method on another data set of hand gestures. This research paper shows the sign identity of human hand shape with fingers and one thumb. The objective of this paper is to style a convenient sys-tem which is beneficial to negate any sort of barrier between a traditional person and a speech impaired person. The given gesture then dictates into the further text. The allowed system helps in performing dimensionality reduction. Nowadays researchers are paying more attention to form an outsized vocabulary for signing recognition systems.

Keywords—Tensor flow; Keras; Machine learning; Data Science;

1. INTRODUCTION

The use of physical controllers like keyboards, mouses, and joysticks for human-computer interactions hinders natural interface as there is a barrier be-tween the human and interface. In this framework, we designed a robust marker less hand gesture recognition systems which can efficiently track both the dynamic and static hand gestures. The process detects and translates the hand gestures into words we assigned them. Different gestures for alphabetic letter are used to recognize them. This could be used for interactions with dumb people who do not have ability to speak but to communicate with sign language.

Machine Learning: Machine learning is an application of artificial intelligence (AI) that provides sys-tem the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. Machine Learning algorithms are classified into 2 major categories Super-vised and Unsupervised. Supervised machine learning algorithms can apply what has been learned in the past by the machine to new data using labeled examples to predict the future outcomes. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The sys-tem is able to provide targets for any new input after sufficient training. Unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets. This project is based on supervised learning algorithm.

Image Processing: Digital Image Processing means processing digital image by means of a digital computer. We can also say that it is a use of computer algorithms, in order to get enhanced image either to extract some useful information. Different types of images are:

- Binary image: contains only two pixel elements 0 – black & 1-white.
- Black and white image: consists of only black and white colour.
- 8-bit colour: It has 256 different shades of col-ours in it and commonly known as Gray scale Image (0- Black, 255- white 127- gray).
- 16-bit colour: It is a color image format. It has 65,536 different colors in it.

2. LITERATURE REVIEW

In the past, many techniques have been used to convert the hand gesture to text. However, they were limited in terms of their functionalities. Many techniques required gloves with sensors which not only made the application more complex but also expensive. In the other version, the system was limited to a particular background without any noise or disturbance. There were some projects which were heavily dependent on heavy GPUs making it difficult for common man to use the system. Additionally, there were some systems for detections which required the object to be of a particular skin colour. Although, there have been various techniques for converting the hand gesture to text but a very few focus on converting the gesture to both text and speech with that too with limited properties. If training data is not linearly separable then it is difficult to determine optimal parameters in SVM, which appears as one of its drawbacks. Authors in pa-per, describe that there are mainly four steps in developed processing scheme, out of which, first one is, for the input RGB image, a colour transformation structure is created, because this RGB is used for col-our generation and transformed or converted image of RGB, that is, HSI is used for colour descriptor. The approach given in this for feature

set extraction is the colour co-occurrence method.

To remove these research gaps a new methodology for automatic detection image segmentation has been proposed. The advantages of proposed algorithm are as follows:

1. Use of estimators for automatic Initialization of cluster centers so there is no need of user input at the time of segmentation.
2. The detection accuracy is enhanced with proposed algorithm.
3. Proposed method is fully automatic while existing methods require user input to select the best segmentation of input image.

3. PROPOSED APPROACH

Digital camera or similar devices are use to take images of hands of different types, and then those are used to identify the correct proposed meaning of the gesture. Then different types of image processing techniques are applied on them, to process those images, to get different and useful features needed for the purpose of analyzing later. Algorithm written below illustrated the step by step approach for the proposed image recognition and segmentation processes:

- Image acquisition is the very first step that requires capturing an image with the help of a digital camera
- Preprocessing of input image to improve the quality of image and to remove the un-desired distortion from the image. To in-crease the contrast Image enhancement is also done.

Image Acquisition:

- The images of the hand gesture are captured through the camera.
- This image is in RGB(Red, Green and Blue) form.
- Colour transformation structure for the RGB hand images is created, and then, a device in-dependent of colour space transformation for the colour transformation structure is applied.

Image Preprocessing:

- To remove noise in image or other object re-moval, different pre-processing techniques is considered.
- Image clipping i.e. cropping of the hand im-age to get the interested image region. Image smoothening is done using the smoothening filter.
- Image enhancement is carried out for increasing the contrast and converting the RGB images into grey images using colour conversion.
- Then, the histogram equalization which distributes the intensities of the image is applied on the image to enhance the gestured hand image.

Before training the image is preprocessed and converted into Black and White and stored in Mat for more optimized and accurate results.
Mathematical representation of Adaptive Threshold is:

$$\mathtt{dst}(x,y) = \begin{cases} 0 & \text{if } \mathtt{src}(x,y) > T(x,y) \\ \mathtt{maxValue} & \text{otherwise} \end{cases}$$

Where T (x, y) = x (block size*block size) - c and block size is size of pixel neighbourhood that is used to calculate a threshold value for the pixel and c is constant subtracted from the mean

Mathematical representation of Dilation is:

$$\mathtt{dst}(x,y) = \max_{(x',y'):\,\mathtt{element}(x',y')\neq 0} \mathtt{src}(x+x', y+y')$$

Where src = input image; x and y are coordinates.

Image Segmentation:

Segmentation means partitioning of image into various part of same features. The segmentation can be done using various methods like otsu method, k-means clustering, converting RGB image into HIS model etc.

1. Segmentation using boundary and spot detection algorithm :-

The RGB image is converted into HIS model for segmenting. Boundary detection and spot detection helps to find the accurate gesture of hand. For bound-ary detection the 8 connectivity of pixels is considered and boundary detection algorithm is applied.

2. K-Means Clustering:-

The K-means clustering is used for classification of object based on a set of features into K number of classes. The classification of object is done by minimizing the sum of the squares of the distance between the object and the corresponding cluster.

3. Otsu Threshold algorithm:-

Thresholding creates binary images from grey level images by setting all pixels below some threshold to zero and all pixel above that threshold to one. The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

Feature Extraction:

Feature extraction plays an important role for identification of an object. In many application of image processing feature extraction is used. Color, texture, morphology, edges etc. are the features which can be used in hand gesture detection. It considers color, texture and morphology as a feature for disease detection. They have found that morphological result gives better result than the other features.

Classification:

i) Using ANN

After feature extraction is done, the learning database images are classified by using neural network. These feature vectors are considered as neurons in ANN. The output of the neuron is the function of weighted sum of the inputs. The back propagation algorithm modified SOM; Multiclass Support vector machines can be used.

ii) Back Propagation

BPNN algorithm is used in a recurrent network. Once trained, the neural network weights are fixed and can be used to compute output values for new query images which are not present in the learning database.

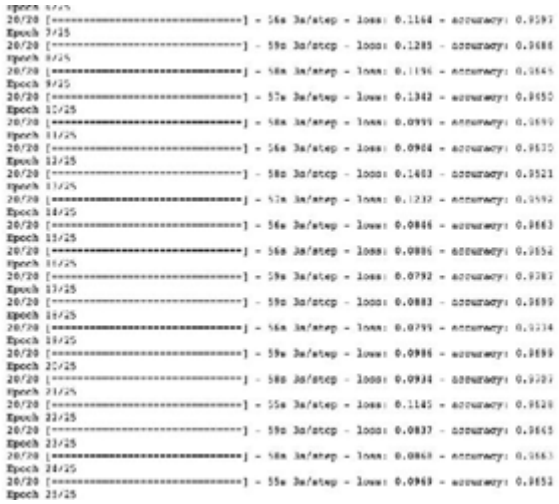

Fig 1.1 – Accuracy and loss

Recognizer Model

4. RESULTS AND ANALYSIS

An accuracy of 97.5 percent was achieved after train-ing the model.

5. CONCLUSION

The project has been successfully completed according to the initial requirements and specifications. With room for improvement in the application, the following are the functionalities and modules to be implemented during the major phase of the project:

1. A neural network trained which will will give bet-ter accuracy thus helping in giving more accurate re-sults.

2. To deploy the machine learning model to a REST-ful API for smoother updates.

3. Hosting the database to a cloud-based service such as mongodb.

4. Developing an android/Flutter app to upload the image to the software as part of the image acquisition part.

REFERENCES

1. V.PADMANABHAN, M.SORNALATHA, HAND GESTURE RECOGNITION AND VOICE CONVERSION SYSTEM FOR DUMB PEOPLE, INTERNATIONAL JOURNAL OF SCIENTIFIC & ENGINEERING RESEARCH, VOLUME 5, ISSUE 5, MAY-2014 427 ISSN 2229-5518.

2. Z. LEI, Z. H. GAN, M. JIANG, AND K. DONG, "ARTIFICIAL ROBOT NAVIGATION BASED ON GESTURE AND SPEECH RECOGNITION," PROCEEDINGS 2014 IEEE INTERNATIONAL CONFERENCE ON SECURITY, PATTERN ANALYSIS, AND CYBERNETICS (SPAC), WUHAN, 2014, PP. 323-327.

3. H. S. K., RAI S, S., PAL, S., SULTHANA K, U., & CHAKMA, S. DEVELOPMENT OF DEVICE FOR GESTURE TO SPEECH CONVERSION FOR THE MUTE COMMUNITY. 2018 INTERNATIONAL CONFERENCE ON

DESIGN INNOVATIONS FOR 3CS COMPUTE COMMUNICATE CONTROL (ICDI3C), IEEE EXPLORE (2018).TENSORFLOW TUTORIALS – OFFICIAL GUIDS.

4. S. C. MUKHOPADHYAY, "WEARABLE SENSORS FOR HUMAN ACTIVITY MONITORING: A REVIEW," IEEE SENSORS JOURNAL, VOL. 15, and NO. 3, PP. 1321-1330, MARCH 2015.DOI: 10.1109/JSEN.2014.2370945.

