International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 09 | Sep 2021 www.irjet.net p-ISSN: 2395-0072

Hand Gesture Recognition System for Dumb People using Image Processing

Swaroop Y P[1], Abhishek T G[2], Kishor K[3], Ravindra Prasad Y K[4], Anisha P S[5]

[1,2,3,4]U.G. Students, Department of Electronics and Communication Engineering, Maharaja Institute of Technology Mysore, Mandya, Karnataka, India
[5]Assistant Professor, Department of Electronics and Communication Engineering, Maharaja Institute of Technology Mysore, Mandya, Karnataka, India

---

Abstract - *Communication between a dumb and normal person has been always a challenging task. This work focuses on finding a unique technique that helps the dumb people by letting them to convey their message through Hand Gestures. The hand gesture recognition system provides an innovative, natural and user-friendly way of interaction. The hand gesture recognition-based man-machine interface is being developed using Deep learning. Gestures always vary in orientation and shape from person to person. So, non-linearity exists in this problem. Recent research has proved the importance of Convolutional Neural Network (CNN) in image processing. The unique architecture of CNN allows to extract relevant information from the input images, which help us to learn complex and non-linear relationships among images. Hence in this paper a Hand gesture recognition method deploying CNN is used with Raspberry Pi which is interfaced with a webcam and speaker.*

Key Words: Hand gesture, Deep learning, Convolutional Neural Network.

1. INTRODUCTION

Communication acts as a main channel between people to communicate with each other. But there are some unfortunate ones who are deprived by this who are known as dumb people. Dumb peoples always find difficulties to communicate with normal people. These people who cannot speak make use of Sign language to communicate with other people who knows the meaning of that particular sign or there must be an interpreter who always needs to translate their sign language to the normal people. This challenge makes them uncomfortable, and they feel discriminated in the society. In the contemporary era, there are several new technologies which can be implemented for making physically disable people to lead their life in a normal way without facing any difficulties. Nowadays Hand gesture play an important role in interchanging information. By taking this into the account, a "Hand Gesture Recognition system for dumb people" is built, which localizes and track the hand gestures of dumb people and convert it into speech in order to maintain a communication channel with other people.

Deep learning is a fragment of a wide-ranging family of Artificial Intelligence. It essentially puts a light on the concept of a multi-layer perceptron learning. A Convolutional Neural Network commonly known as a Comp Net is a neural network class used in deep learning which is most applied to images and videos for their analysis. A CNN is a technique, or a machine learning model that can be applied to images to make them interpretable by machines. It can be implemented in other data analysis and classification problems as well. It is a type of artificial neural network which has a specialty of being able to deduce or distinguish patterns and understand them. It is different than other deep learning models as it has an extra set of hidden layers called the convolutional layers along with the standard hidden layers. It can have one or more than one convolutional layer followed by the fully connected layers. The system will be learning features from each gesture and then further classify it. The entire notion is to make a machine learn and to interprets the thoughts of dumb people.

2. LITERATURE REVIEW

The idea used in assistive translator for deaf and dumb people [1], involves glove-based technique comprising of flex sensors tactile sensor and accelerometer sensors. They overcome "pot and mechanical assembly technique" since it requires very precise assembly, and it is bit delicate. The whole assembly is placed on a palm and different strings are connected to the fingers, whenever position of pot varies it results in a variable voltage. But this mechanism is uncomfortable and inconvenient for daily use and precision required is more, but system does not provide enough accuracy to match the requirement. The advantage of this system is that the system-based technique offers greater mobility and reduce ambiguity among gestures. The drawback of this system is that the output of the system is an alphabet, combining these alphabets and comprising the word is a complex thing and takes more time.

The aim of this hand gesture recognition system [2], involves "finger spelling" that means spelling out words in an alphabetical language using hand gesture. They use wireless glove which is fitted with flex sensors (bend sensors) which recognize the gesture by means of resistance associated with the flex sensors. The advantage of this system is, it provides more accuracy in terms of hand gesture recognition and also it is cost efficient. The drawback of this system is, it causes

uncomfortable to wear the gloves whenever they want to communicate and also this system carries a lot of cables and flex sensors which are somewhat delicate in terms of quality and any wear and tear can cause changes in resistance which impacts the final output.

A system which uses color bands to implement the sign language is proposed in [3]. These bands are matched with color bands and the corresponding output is generated. This system overcomes the usage of electromechanical gloves for recognition. The output accuracy is improved due to usage of colour bands. The drawback of the system is that the user has to wear different color bands to represent different sign gestures.

The author has proposed a Gesture to Speech (G2S) system [4] which is developed using the skin color segmentation. Based on the values of RGB in the image frame the skin color is detected. They use Centre of Region (COR) of the hand region as well the farthest point from the COR for feature extraction process. By taking wrist segment as the reference segment, classification of the gesture is made. The system shows maximum classification accuracy of 80%.

Hand Gesture Recognition and Voice Conversion for Deaf and Dumb [5] uses Principal Component Analysis (PCA) method for feature extraction. K-Nearest neighbor method and Support Vector Machine (SVM) algorithm for classification. This method take image as input gives text and speech as an output. Accuracy of the system is up to 90%.

In [6] The idea of this work is classifying ISL static gestures captured under indistinct conditions, gestures were divided into single handed or double handed gestures. Geometric descriptors and HOG features are used to extract features. A database consisting of 260 images captured under simple and complex background for experimental purposes. By comparing KNN and SVM classifier, it is concluded that SVM was Superior to KNN algorithm in terms of accuracy on both geometric and HOG features. SVM has achieved the highest accuracy of 94.23%

3. BLOCK DIAGRAM

Fig -1 shows the block diagram of the proposed hand gesture recognition system using image processing. The first step is to create the dataset of hand gesture. The created dataset is divided into 2 parts, the 90% of the images are used for Training and remaining 10% of the images are used for Testing. In Training process, the set of images for training are fed as input to convolution neural network, and the output of the CNN is image annotation model which is saved into the disc. In the Testing process, the set of test images are taken and are checked for match with the image annotation model which is stored in the disc, and the classification of images takes place, and the annotation results are obtained. In real time hand gesture recognition process, the real time gesture is taken as input through laptop webcam, the region of interest of hand is obtained and compared with the image annotation model to check similarity and the respective output class is obtained. The prerecorded audio formats are stored in raspberry pi and the resultant class from open cv is fed as input to raspberry pi. According to the output class the particular audio track will be played through the speakers.

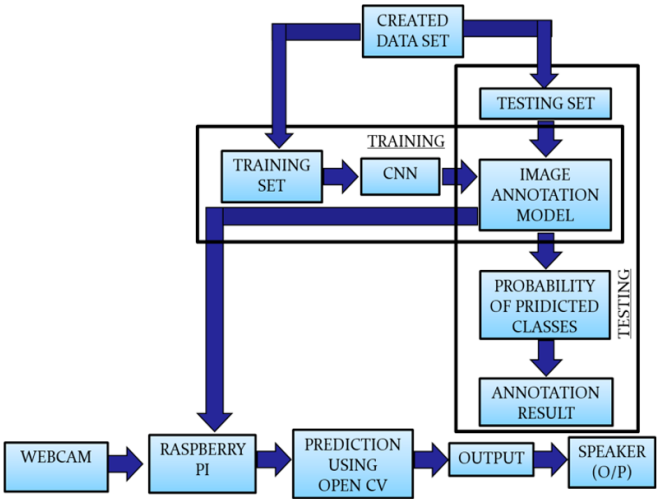

Fig -1: Hand recognition system using Image processing.

4. METHODOLOGY

The proposed system has the following steps:

1) The database of hand gestures is created.
2) Constructing a convolution neural network (CNN) architecture.
3) Training the CNN.
4) Testing the CNN.
5) Real time hand gesture recognition.

4.1 Creation of Dataset

A dataset is created using laptop camera which consists of 3600 images and stored in binary format. In the binary image the pixels of the hand region are white in color, and the background pixels will be black in color, this is done to overcome colour based prediction of CNN. These images are sub divided into 5 classes such as 1,2,3,4, and 5. Each class consists of 600 images. The dataset created is divided into training set and testing set, the 90% of the images that is 2700 images are stored in training set and the remaining 10% of the images that is 300 images are stored in testing set. Examples of dataset created is as shown in Fig-2.

Fig -2: Hand Gesture Dataset

4.2 Construction of Convolutional Neural Network (CNN)

The CNN is built using 6 different layers. The first layer is a Sequential layer which is used as a linear stack, where the instance of sequential layer is created and then the further layers will be added to it sequentially. The combination of Convolution and Pooling layer will continuously extract the features by down sampling the input images. The obtain results are fed into Flatten layer, where the input image is flattened which give rise to N-dimensional vector, where N is the number of classes from which the model selects the desired class. Once flattening of input image is completed the Dropout layer is used to prevent the model from overfitting by making the weights of some redundant neurons in a particular layer equal to zero. The Activation function such as Relu layer and Softmax layer are used to classify the images respectively to their classes.

4.3 Training of CNN

Once the neural network is constructed, the next step is to train the network by providing the training dataset. Training a neural network consist of 2 phases namely, Forward phase and Backward phase. In Forward phase input is passed completely through the network, during this phase each layer will be assigned with random weights. Preceding the Forward phase there is a Backward phase where the gradients are backpropagated and weights are updated, in this phase each layer receives a gradient and return a gradient. The number of epochs and steps for epochs are specified in this phase. After training, the extracted features of training dataset and the weights of the respective layers are stored, this trained model is called Image Annotation model.

4.4 Testing of CNN

After training, the model has to be tested to check the proper working of the system. The testing dataset is now given as input to the neural network, where the features of testing data is compared with image annotation model and the respective class is obtained as the output.

4.5 Real-time Gesture Recognition using Open CV

The real-time gesture recognition is performed by giving hand gestures as input through laptop webcam. Once the hand gestures are obtained, these hand gestures are thresholded and the contours are mapped around the hand region and the region of interest is obtained. Then the region of interest is compared with the image annotation model and the output class is recognized.

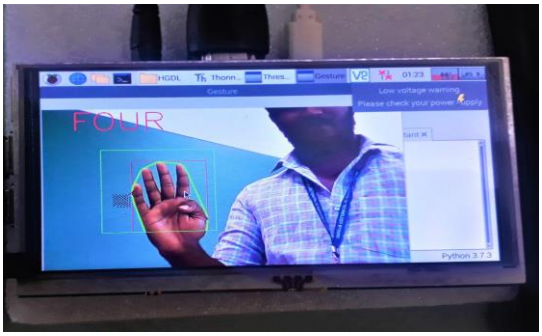

Fig -3: Real-time Gesture Recognition

The image annotation model and the real time gesture recognition code is dumped into the Raspberry Pi, and also a webcam and speaker are interfaced as shown in Fig -4. Webcam is connected to the USB port and the speaker is connected to the 3.5mm audio jack of the Raspberry Pi. The webcam is used to capture the hand gesture of dumb people, and the obtained gesture is compared with image annotation model, once the match to the input image is found, it is displayed on LCD and the respective audio track is played through the speakers.

Fig -4: Raspberry Pi interfaced with webcam and speaker

5. CONCLUSIONS

The proposed system acts as a communication bridge between dumb and normal people. There are significant number of researches are already going on in this area to improve the accuracy of real time performance of these models. The proposed model has achieved the accuracy rate up to 95%

Fig -5: Obtained results

REFERENCES

[1] S. B. Shrote, Mandar Deshpande, Prashant Deshmukh, Sanjaykumar Mathapati, "Assistive Translator for Deaf and Dumb People", International Journal of Electronics Communication and Computer Engineering, Volume 5, Issue (4) July, Technovision-2014, ISSN 2249–071X.

[2] Mr. R. Jagadish, R. Gayathri, R. Mohanapriya, R. Kalaivani and S. Keerthana, "Hand Gesture recognition System for Deaf and Dumb Persons", Indo-Iranian Journal of Scientific Research (IIJSR), Volume 2, Issue 1, Pages 139-146, January-March 2018.

[3] Channaiah Chandana K, Nikhita K, Nikitha P, Bhavani N K, Sudeep J, "Hand Gestures Recognition System for Deaf, Dumb and Blind People", International Journal of Innovative Research in Computer and Comunication Engineering, Vol. 5, Issue 5, May 2017.

[4] Prof. R. R. Itkarkar and Dr. Anil V Nandi, "Hand Gesture to Speech Conversion using Matlab", 4th ICCCNT 2013 July 4-6, 2013, IEEE-31661.

[5] Rupesh Prajapati, Vedant Pandey, Nupur Jamindar, Neeraj Yadav, Prof. Neelam Phadnis, "Hand Gesture Recognition and Voice Conversion for Deaf and Dumb", International Research Journal of Engineering and Technology (IRJET), Vol. 5, Issue 4, 4th Apr 2018.

[6] A. Singh, S. Arora, P. Shukla and A. Mittal, "Indian Sign Language gesture classification as single or double handed gestures", 2015 Third International Conference on Image Information Processing (ICIIP), pp. 378-381.

[7] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo, "Indian Sign Language Recognition using Neural Networks and KNN Classifiers", ARPN Journal of Engineering and Applied Sciences, VOL. 9, NO. 8, AUGUST 2014.

[8] Neha V. Tavari, Prof. A. V. Deorankar, "Indian Sign Langue Recognition based on Histograms of Oriented Gradient", International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 3657-3660.

[9] Miss. Juhi Ekbote and Mrs. Mahasweta Joshi, "Indian Sign Language Recognition using ANN and SVM Classifiers", 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).