Hand Gesture Image Recognitions Using CNN Algorithm for Smart Interaction

Prajwal.P.Kasture[1], Sachin.V.Suryawanshi[2], Suraj.S.Malge[3], Sushilkumar.S.Salve[4]
[1,2,3] *Student, Sinhgad College of Engineering, Pune, India*
[4]*faculty, Sinhgad College of Engineering, Pune, India*

***Abstract*-Sign language is one way to communicate with deaf people. Hand gestures are one of the typical methods used in sign language. One should learn sign language to interact with them. Learning usually takes place in peer groups. There are few study materials available for sign learning. Because of this, the process of learning sign language is very difficulttask. This project presents a solution that will not only automatically recognize hand gestures but alsoconvert it into text output so that the impaired person can easily communicate with normal people. In this paper, we propose a convolution neural network (CNN) method to recognize hand gestures. To achieve the desired level of performance, The skin model and the calibration of hand position and orientation are applied to obtain the training and testing data for the CNN. Since the light condition seriously affects the skin color, we adopt a Gaussian Mixture model (GMM) to train the skin model, which is used to robustly filter out non-skin colors in an image. Then the calibrated images are used to train the CNN.**

***Index Terms*-convolution neural network (CNN), Hand Gesture, MobileNet, Sign.**

I.INTRODUCTION

Communication involves the exchange of information, and this can only occur when it is most effective if all participants use a common language. Sign language is widely used by people who are deaf. A Communication through signing is only made up of different signals, framed by different states of hand, its developments, directions as well as looks.The basic goal of Human Computer Interaction is to improve the interaction between users and computers by making the computer more receptive to user needs.

A gesture in sign language is a particular movement of thehands with a specific shape made out of them.

Gesture recognition is a computing process that attempts to recognise and interpret human gestures through the use of mathematical algorithms. Gesture recognition is not limited to just human hand gestures. But rather, it can be used to recognise everything from head nods to different walking gaits.

Hand gesture detection can be done using a webcam. The detection ofGestures is done using the various techniques of contour analysis and feature extraction. Image Recognition of Hand Gestures Using CNN TechniqueA convolutional neural network (CNN, or ConvNet), is another class of deep neural network.CNNs are most commonly employed in computer vision. Given a series of images or videos from the real world, with the utilization of CNN, the AI system learns to automatically extract the features of these inputs to complete a specific task, e.g., image classification, face authentication, and image semantic segmentation. There are many types of CNN models. but we are using the MobileNetV2 model. MobileNets are CNNs that can be fitted on a cell phone to group pictures or identify objects with low inactivity. MobileNets have been created by Andrew G. Trillion et al. They are typically tiny CNN designs, which make them simple to run continuously, utilising inserted gadgets like cell phones and robots.

II.RELATED WORK

Sr. No	Paper Title	Authors	Methodology
1	Sign Language Recognition System using Convolutional Neural Network and Computer Vision	Mohammad Elham Walizad, Mehreen Hurroo.	Used web camera to shoot the hand gestures. The images undergo a series of processing operations whereby the backgrounds are detected and eliminated using the color extraction algorithm.
2	Sign	Mohamm	They divided their

	language Recognition Using Machine Learning Algorithm	ad Elham Walizad, Mehreen Hurroo.	approach to tackle the classification problem into three stages. 1.segment the skin part 2.extract relevant features from theskin segmented images 3.use the extracted features as input into various supervised learning models for training and then finally use the trained models for classification.
3	Wenjin Zhang, Jiacun Wang	Dynamic Hand Gesture Recogniti on Based on Short-Term Sampling Neural Networks	The organization coordinates a few all around demonstrated modules together to gain both present moment and long-haul highlights from video sources of info and in the interim keep away from concentrated Calculation.The network integrates several well provided modules from both short term and long term features.
4	Munir Oudah, Ali Al-Naji and Javaan Chahl	Hand Gesture Recogniti on Based on Computer Vision: A Review of Technique s	instrumented sensor technology and computer vision. In other words, the hand sign can be classified under many headings, such as posture and gesture, as well as dynamic and static, or a hybrid of the two.

III. METHODOLOGY

A. Figures:

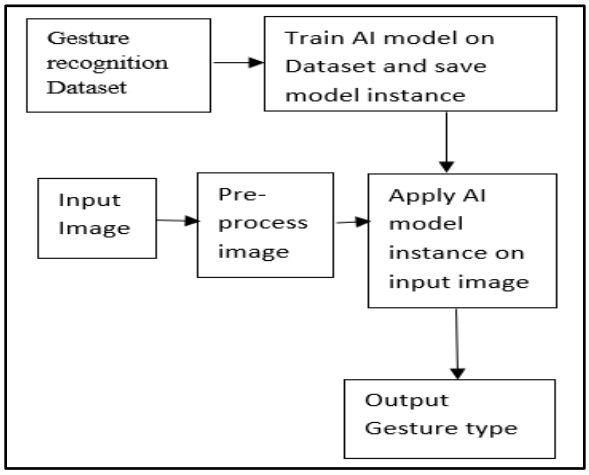

Fig.1.Block Diagram of the Proposed Model

B. Phases of development:

1.Designing a module for gesture detection and recognition using still images.

2.Designing a module for processing the gesture, extracting features and finally converting it into Text format.

3. Connecting all the modules together and testing the integrity and accuracy of the system.

C. Explore Dataset:

The biggest problem in deep learning when you want to train a good model for prediction task, is to find a large dataset. We chose to utilize open-source dataset accessible on Kaggle site for preparing AI model Dataset has motions: clench hand, open hand, two finger contains 2k pictures per signal

The proposed system presents a recognition algorithm to recognize a set of Five specific static hand gestures, namely: Open, peace, fist, three finger, Index.

D. Preprocessing:

In the pre-processing stage, some operations are applied to extract the hand gesture from its background and prepare the hand gesture image for the feature extraction stage. We are using the following image pre-processing techniques:

1.Data Augmentation

The images in the dataset are not used directly. Instead, only augmented images are provided to the model. There are different types of data augmentation techniques out there, but we are using

1.Horizontal and Vertical Shift Augmentation: A shift to an image means moving all the pixels of the image in one direction, such as horizontally or vertically, while keeping the image dimensions the same.

2.Random Rotation Augmentation: A rotation augmentation randomly rotates the image clockwise by a given number of degrees from 0 to 360.

3.Random Zoom Augmentation: A zoom augmentation randomly zooms the image in and either adds new pixel values around the image or interpolates pixel values, respectively.

2.Image standardisation

The ImageDataGenerator class in Keras provides a suite of techniques for scaling pixel values in your image dataset prior to modeling.

E. Model Creation:

1.CNN based AI model will be prepared on above dataset.

A convolutional neural network (CNN, or ConvNet) is another class of deep neural networks. CNNs are most commonly employed in computer vision. Given a series of images or videos from the real world, with the utilization of CNN, the AI system learns to automatically extract the features of these inputs to complete a specific task, e.g., image classification, face authentication, and image semantic segmentation. The CNN or convolutional neural networks are the most commonly used algorithms for image classification problems. An image classifier takes a photograph or video as an input and classifies it into one of the possible categories that it was trained to identify. They have applications in various fields like driver less cars, defense, healthcare etc. There are many types of CNN models. We are using the MobileNetV2 architecture. MobileNets are CNNs that can be fit on a mobile device to classify images or detect objects with low latency. MobileNets have been developed by Andrew G Trillion et al. They are usually very small CNN architectures, which makes them easy to run in real-time using embedded devices like smartphones and drones.

2.Tuning of AI model will be finished utilizing Keras Tuners.

Neural networks trained on smaller datasets often tend to over-fit and are therefore less likely to be accurate on new data. Theoretically, the best way to train a model could be to try out as many different combination of different parameter values and then take an average of those individual results to come up with a generalized result

F. Testing:

Once the information picture is taken, it will be handled and changed over into structure which is OK by AI model.Computer based intelligence model will examine the information picture and show anticipated yield.

G. Algorithms:

1. CNN: Convolutional Neural Network algorithm will be used to train the model for gesture recognition.

The transfer learning model of CNN used: MobileNet V2

2.Mean Square Error: An algorithm used to evaluate the accuracy of the model.

3.Stochastic Gradient Descent: An algorithm used to optimize the loss of AI model

IV.RESULT

The data set divided into two groups, one used for training and other for testing.The preparation set comprises of70% of the total information and staying 30% are utilized astesting.

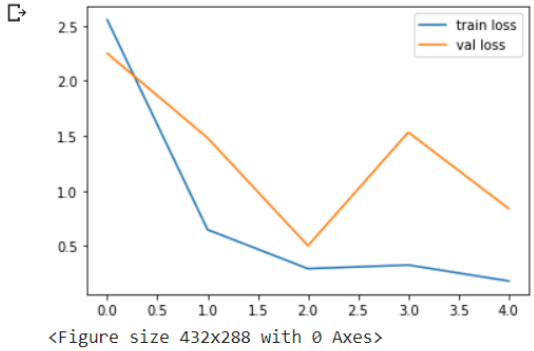

Fig.2. The graph of training loss and validation loss

The graph represents the training loss and validation loss of the model. Both the losses are gradually decreasing. Although this model may not work perfectly every single time, there is a lot we can do to improve its performance.To test the model, we are using some images of hand gestures captured. The model takes these images as an input parameter and predicts the class they belong to. Before passing the image, we need to ensure that we are using the same dimensions that we used during the training phase. For the following test images, the model predicted correctly. However, if the image contains too many other items in the background, it isn't always accurate.

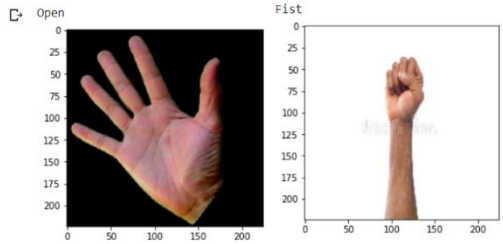

Fig.3. A snapshot of the model's predicted outcome

V.CONCLUSION

As a simple and down to earth method for accomplishing human-PC communication, in this

arrangement hand motion. Overall, the arrangement intends to give help to those in need accordingly guaranteeing social significance. Individuals can without much of a stretch speak with one another.

The easy-to-use nature of the framework guarantee that individuals can utilize it with no trouble and intricacy. The application is cost productive and dispenses with the utilization of costly innovation.

REFERENCES

[1] Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks Wenjin Zhang, Jiacun Wang, 2021

[2] Hand-Gesture Recognition-Algorithm based on Finger Counting, M. Perimal, S.N. Basah, M.J.A. Safar, M.J.A. Safar, 218.

[3] Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks Wenjin Zhang, Jiacun Wang, 2021.

[4] Static Hand Gesture Recognition Based on Convolutional Neural Networks, aimundo F. Pinto, Carlos D. B. Borges, Antônio M. A. Almeida, 2019.

[5] Andrew G. Howard et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. CoRR, abs/1704.04861, 2017.