

Indian Institute of Information Technology, Vadodara

Summer Internship Report

Hand Gesture Recognition

Name: Vikas singh

ID:201452047

Aditya Raj

ID:201451032

Abstract—This report is consist of various Mathematical Algorithms used for hand gesture detection system. Hand gesture detection is boon for those patients who can not move their body due to paralysis. So basically hand gesture detection system is a human language interpreter. We implemented various Mathematical algorithms (Convolutional neural network, Fine tune, Bottleneck feature) to make hand gesture detection system.

Index Terms—Opencv, CNN, Fine tune network, Bottleneck feature

I. INTRODUCTION

Hand gesture detection can include combining hand shapes, orientation of hands to fluidly express speaker's thought. So it is basically language interpreter for person with disabilities. This project focuses on gesture recognition and it uses computer vision and machine learning techniques to achieve this goal. We used three mathematical algorithms to train the parameter for different gesture.

there are three step to accomplish hand gesture detection system.

1-Hand detection

2-Feature extraction

3-Recognition

1-Hand detection— As we had the dataset of hand gesture so we directly extract it's feature from the image.

2-Feature extraction—This step deals with extracting important features that represent important characteristics of the gesture throughout the image and then storing these features in numpy array so that we can load it from numpy array when we ll train

the model.

3-Recognition— This step deals with recognizing and classifying the performed gesture.

It has two phases

a)-the training phase which involves training the system on datasets.

b)-classification phase which involves classifying the performed gestures.

II. OBJECTIVES

The objective of this summer internship was to get knowledge and experience about how to work as a team and accomplish given task within time.

The ultimate aim of the project was to create a software that recognize pre-defined hand gestures using computer vision and machine learning algorithms.

As we divided our project into three major steps which represent the major objectives in the project. Those steps are as below

1-Hand detection

2-Feature extraction

3-Recognition

III. IMPLEMENTATION

A. Tracking

As we have downloaded data for the each gesture for training and testing so there was no need of tracking hand from the image .image itself was a hand gesture.

in this phase of project we just convert our image to grayscale and resize each image using opencv.

B. Feature extraction

This phase of implementation was extracting feature for image and store feature vector in appropriate manner so that we can deal with the feature vector while training of the model.

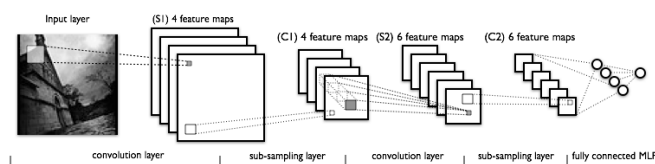
After converting image into grayscale we convert each image as a feature vector and save feature vector in numpy array.

C. Recognition

This phase is the most sensitive phase as it contains the training of model with appropriate parameter so that it can easily predict the appropriate gesture.

As we use many mathematical algorithms to implement hand gesture detection system

1-Convnet Neural network



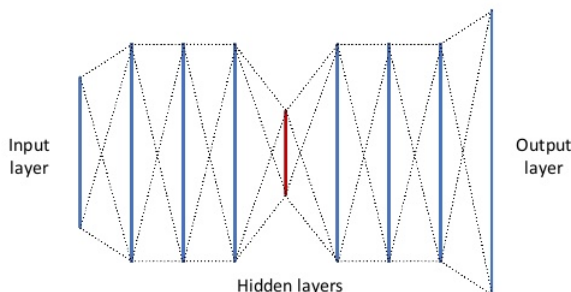
In our case we will use a very small convnet with few layers and few filters per layer, alongside data augmentation and dropout. Dropout also helps reduce overfitting, by preventing a layer from seeing twice the exact same pattern, thus acting in a way analogous to data augmentation

2-Bottleneck features



Bottleneck Features

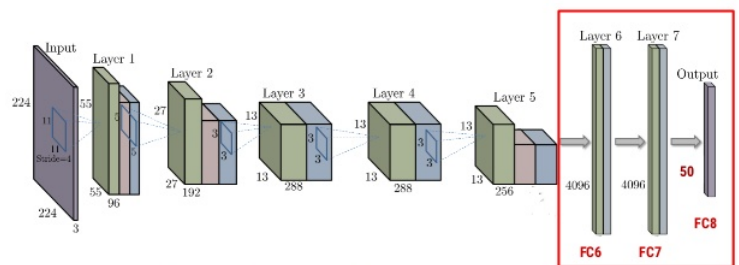
- Use one narrow hidden layer. Supervised or unsupervised training (autoencoder)



A more refined approach would be to leverage a network pre-trained on a large dataset. We will use the VGG16 architecture, pre-trained on the ImageNet dataset. Our strategy will be as follows: we will only instantiate the convolutional part of the model, everything up to the fully-connected layers. We will then run this model on our training and validation data once, recording the output (the "bottleneck features" from the VGG16 model: the last activation maps before the fully-connected layers) in two numpy arrays. Then we will train a small fully-connected model on top of the stored features.

3-Fine tune

FINE-TUNING DEEPER LAYERS ONLY



Andrei Karpathy: [Convolutional neural networks for visual recognition](#). In Stanford CS class CS231n.

40

To further improve our previous result, we can try to "fine-tune" the last convolutional block of the VGG16 model alongside the top-level classifier. Fine-tuning consists in starting from a trained network, then re-training it on a new dataset using very small weight updates. In our case, this can be done in 3 steps:

Instantiate the convolutional base of VGG16 and load its weights

Add our previously defined fully-connected model on top, and load its weights

Freeze the layers of the VGG16 model up to the last convolutional block

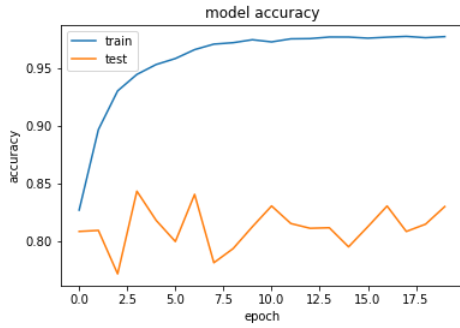
IV. EVALUATION

This section discusses the accuracy of hand gesture detection system that has been produced on different mathematical algorithms for the dataset.

A. Accuracy of convnet

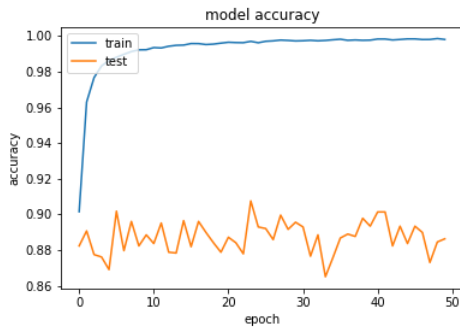
After training convolutional neural network we found accuracy about 83 so we decided to imple-

ment more appropriate model that could achieve more accuracy than convnet.



B. Accuracy of Bottleneck feature

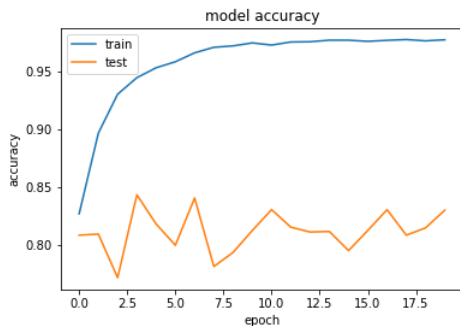
When we trained this model's parameters with the hand gesture image dataset and tested it with the testing dataset we found a better accuracy than previous one. This model was much appropriate than previous one with accuracy 90.



C. Accuracy of Fine tune

After training this model's parameter when we tested our dataset with the trained parameter we found a much better accuracy than both of the previous case.

As this model was much accurate than previous trained model with accuracy 93.



V. CONCLUSION

A. Project conclusion

Hand Gesture Recognition is very broad and difficult problem, in this project we have implemented a hand gesture recognition system that recognizes hand gesture of pre-defined sign with the implementation of Convolutional neural network model and computer vision.

B. Future work

As we had the dataset of training image and testing image. And image itself consists of hand gesture so there was no need of hand tracking and detection; it was just a feature extraction from the given dataset and train that data with the help of mathematical model.

But we can expand on the system in the future to take image in real time and predict the instruction for that gesture in real time. Basically we can work on making the recognition phase more autonomous and recognize the gestures in real time.