Face Detection and Emotion Analysis Using Convolution Neural Network

Rakshanda Mulay[1], Srushti Yerawar[2] and Mohammed Abdul Mujeeb[3]

[1]Computer Engineering, MIT Academy Of Engineering, Pune, India

[2]Computer Engineering, MIT Academy Of Engineering, Pune, India

[3]Computer Engineering, MIT Academy Of Engineering, Pune, India

[1]rsmulay@mitaoe.ac.in, [2]sgyerawar@mitaoe.ac.in, [3]mamujeeb@mitaoe.ac.in

Manuscript Received: Manuscript Accepted:

Abstract

Humans have been recognizing face for far long as its key factor is to understand human behaviour and mental state. Only verbal communication is not enough, facial expression also plays a vital role in predicting human behaviour. As the world is advancing in various field it is having great importance in the technology by reading and understanding human behaviour. Many researches are going on to identify trends and pattern in human beings to understand human behaviour. A major step in this process is facial recognition. In this paper we propose easy ways of facial recognition to identify 7 types of standard emotions. By following this paper the researcher can get quite good result and can be proved revolutionary for future model of computer based emotion recognition system.

Keywords: *communication , facial expressions, human behaviour, emotions, emotion recognition.*

I. INTRODUCTION

A. Need of Face and Emotion detection

To understand the intentions of people, facial emotions act as an important factor in human communication. In general, people infer the emotional states of other people such as happy, sad, anger using facial expressions and vocal tone.

There are two types of conveying components. One is verbal components and other is non-verbal components. Facial expression is one of the main non verbal information channels in interpersonal communication. According to different surveys verbal components convey one-third of human communication, and nonverbal components convey two-thirds. Therefore, it is obvious that research of facial emotion is gaining lot of attention. It is also affective in computing and computer animation.[1]

B. What is Neural network?

The Neural networks are a form of artificial intelligence, which have been modeled after, and inspired by the processes of the human brain. Structurally, Neural network consist of highly interconnected processing elements. [2]

The learning potential of Neural Network has been demonstrated in different areas of application, such as pattern recognition, function prediction, and robot control.

The structure of the Neural network is as follows :

Each branch contains entities which are known as weights. For computation the following process is carried out by a neural network .

1) Input values go into the input layer .

2) The values are multiplied with the weights of the branch joining that node and the next hidden node and that values gets propagated forward.

3) Summation of all incoming values is done and passed through an activation function for mapping the value between 0 to 1[1]

4) This process gets repeated till the values reach the output layer and that is the desired value.

C. What is TensorFlow?

Tensor-flow is an open source library which is used for various numerical computation and for large complex machine learning problems. It integrates machine learning, deep learning models and algorithms which makes them useful by the way of a common metaphor. For building purpose this provides an efficient front-end Application.

It can train and run deep neural networks for handwritten digit classification, image recognition, word embeddings, recurrent neural networks, sequence-to-sequence models for machine translation, natural language processing, and PDE (partial differential equation) based simulations. Tensor Flow supports production prediction at scale, with the same models used for training.

II. CONVOLUTION NEURAL NETWORKS AND IMAGE CLASSIFICATION

Convolution Neural Network (CNN) is used for progressively extraction of higher and higher-level representations of the image content. A CNN takes just the image's raw pixel data as input and "learns" how to extract these features, and ultimately infer what object they constitute instead of preprocessing the data to derive features like textures and shapes. [3]

The image classification takes the input image and the definition of its class.

Here the image in the picture is an elephant. But the computer sees the pictures quite differently:

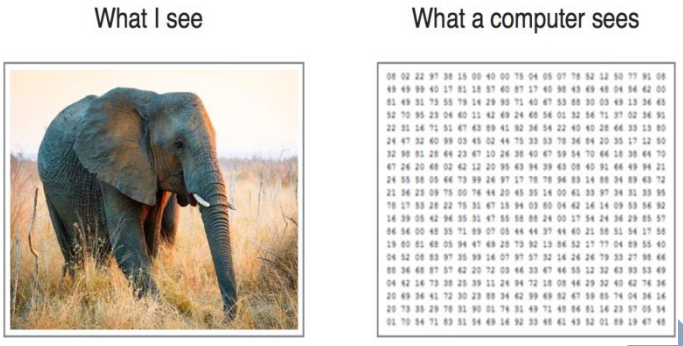

Fig1. Image recognition by computer

Instead of image, computer observe it as an array of pixel. Consider if size of image is 400 x 400. The size of the array will be 400x400x4. Where 400 x 400 is width and height respectively whereas 4 is RGB channel values. Each of these values ranges by computer to 0 to 255. This value describes the intensity of pixel

The computer looks for the characteristics of the base level. Usually, human consider these characteristics as large ears and trunks. For computers these characteristics are like boundaries. And then through the groups of convolution computer constructs the layer of the abstract.

Thus the output of the image is generated by passing the image through fully connected layers.[4]

III. ALGORITHM

Collecting the data from the scratch and then labeling it, takes lot of time to detect the image. But due to the advancement of technology we can carry out some shortcut methods. We can take a standard trained piece of model and reuse it in new model. This improves the efficiency of the result and gets the data in less time.

So the other way is to collect the set of images. Since it needs angry, sad, happy, surprised as factors in result, it has to gather images which are decided as per the project requirement. To get the result more accurately it needs nearly 200 images for each factor. This will help to retrain the model in an efficient way.

Dataset named FERC 2013 can be used efficiently to get the output data.

To get the expression of the live image, first capture the image from the web camera. Then crop the image at a specific size. This image should be saved in new location and that location path should be given in retrain code.

After completing this step, run the retrain model in python and get result.

IV. IMPLEMENTATION

A. What is haar casscade classifier?

A Haar Cascade is basically a classifier which is used to detect the object for which it has been trained for, from the source. The Haar Cascade is by superimposing the positive image over a set of negative images. The training is generally done on a server and on various stages.

B. Steps to perform:

This is a three step process.

In the first step, we load the XML file for detecting the presence of faces and then we retrain our network with our image on five different categories. After that, we import the label_image.py program.

1) Implementation of Open CV HAAR CASCADES

Use the "Frontal Face Alt" Classifier for detecting the presence of Face in the Web-Cam. These classifier needs to be included in one file .There are also some other classifiers which can be included in the file for the presence of eye, lips etc.

This file needs to be load in other file where it captures the image from the web cam and validates with the classifiers.

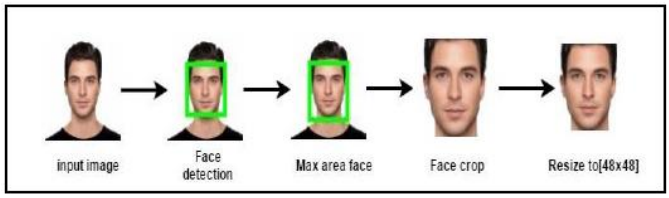

Fig2. Face detection and resizing

2) ReTraining the Network - Tensorflow Image Classifier

This step will consist of several sub steps:

First create a directory named images. In this directory, create five or six sub directories with names like Happy, Sad, Angry, Calm and Neutral. You can add more than this.

Fill these directories with respective images.

E.g., In "Happy" directory, fill only those images of person who are happy.

Resize the image in its standard size which will be convenient for recongnition. Or in simple way just download the FERC 2013 dataset and add it to the python file. This will give you the result more accurately and within less time.

Once the image is cleaned, you are ready to retrain the network.

3)Importing the ReTrained Model and Setting Everything Up

Now run the retrain model which will give the specified results.

This result can be obtained by using graph like

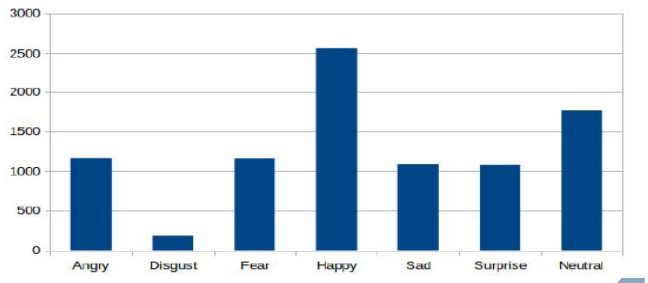

Fig3. Result Visualization

V. SYSTEM ARCHITECTURE

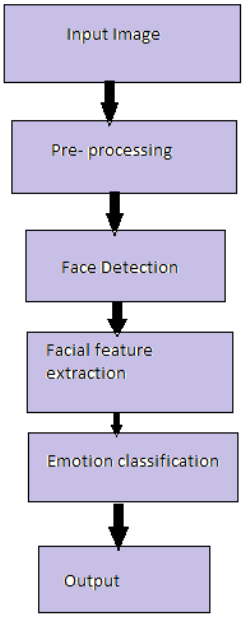

Fig 4.Image processing

The above given diagram gives the general idea of the system. Here the first need is to capture the images through web cameras. This image captured needs to be resized in standard size. This all needs to be done through openCV. After resizing the image it is retrained through various images or through dataset, from which we will get the specified result of the emotions.

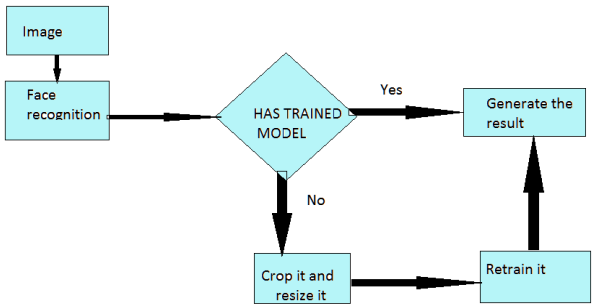

Fig5. Trained Image validation

This diagram shows the flow of system. After capturing the image from the web camera, it needs to be checked whether the image is already trained or not. If it is trained than directly, the results are generated else it needs to be resized and trained. For that firstly, it needs to be resized in the desired size. After cropping the image it needs to be retrained using various images and finally, the results are generated.

VI. APPLICATIONS

Facial Emotion detection can have a wide range of applications:

1)Making cars safer and personalized:

Imagine a car is telling you to stop the car and have a coffee break. This can be done by having the emotion analysis in car where the car can recognize the sleepy and drowsy driver.

2) Facial Emotion Detection in Interviews:

All recruiters would appreciate technology that can tell them what a candidate is feeling. This solution can classify facial expressions into seven different emotions that can be seen through various statistical tools and can analyze whether candidate fits for the job or not.

3)Testing for Video Games:

A Videogame hooks you and gets your adrenaline pumping. Facial expressions while playing is a great metric to understand if the game is successful in making your experience enjoyable.

4) Market Research:

Facial Emotion Recognition allows market research organizations to measure moment-by-moment facial expressions of emotions (facial coding) automatically and analyze the results. Further the advertisement field can be increased by knowing the expression of the user.

Fig 6. Analysis of emotions through Emoji

For example, if a brand wants to launch the limited version of a new product it can target showing ads to only those clients who have expressed the emotion of "excitement" in the previous product launched. Or alternatively, if a client is typically angry when interacting with a brand, targeted discounts can be offered to assuage them.

5)Emotion-recognition-in-Health-Care:
The industry which can take the major advantage is Health care system where with AI-powered recognition software can decide when patients necessitate medicine or to help out physicians determine who to see first.

VII. ADVANTAGES

Emotions detection is one of the challenging task. But machine learning algorithms have shown great promise.

By using Facial Emotion Recognition, businesses can process images, and videos in real-time for monitoring video feeds or automating video analytics, thus it will save costs and make life better for their users.

It can combine the science of psychology, human expressions and artificial intelligence to recognize different emotions on an individual's face automatically. This algorithm can identify seven different types of emotional states in real-time: happiness, sadness, disgust, surprise, anger, and fear.

VIII. CONCLUSION

The mental state of a person can be understood from its facial expression. Detecting facial expression can be proved very useful in field such as Testing for Video Games, Market Research, and Facial Emotion Detection in Interviews, Making cars safer and personalized to avoid accident.

Its Uses are immensely growing so its need.

In this paper we used convolution neural network to identify the types of standard emotions. We took a standard trained piece of model and reuse it in new model. This improves the efficiency of the result and gets the data in less time.

The data bank we applied in this research namely FERC-2013 database. In, FERC-2013 contains about 32000 low resolution face pictures of dissimilar age groups and having different degrees of angle are available. We have trained our system on FERC-2013 database. Since images are 'very clear' and have well define expressions, they effortlessly classified for different emotions on a face. Therefore, the convolution networks are trained with the FERC-2013 database.[1]

We have used tensor-flow open source library which can train and run deep neural networks for image recognition.Using convolution neural networks and tensor-flow a real-time emotion recognition system is proposed and developed .

REFERENCES

[1]Rajesh Kumar , Ravi Kant Kumar, “Facial Emotion Analysis using Deep Convolution Neural Network”, IEEE Trans, July 2017

[2]C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with Convolutionals", *Co RR*, vol. abs/1409, pp. 4842, 2014.

[3]M. D. Zeiler, R. Fergus, "Visualizing and understanding convolutional networks", ECCV, 2014.

[4]K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition", ICLR, 2015.

[5]O. Abdel-hamid, L. Deng, D. Yu, Exploring Convolutional Neural Network Structures and Optimization Techniques for Speech Recognition, pp. 3366-3370, August 2013.

[6]G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. arxiv:1207.0580, 2012.

[7]Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(8):1798-1828, 2013.

[8]. Dumoulin, F. Visin, A guide to convolution arithmetic for deep learning, pp. 1-28, 2016.

[9] D. H. Nguyen, B. Widrow, "Neural networks for self-learning control systems", IEEE Contr. Syst. Mag., vol. 10, no. 3, April 1990.

[10] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Int. Conf. Neural Inf. Process. Syst., pp. 1097-1105, 2012.