Facial emotion recognition using MATLAB

[1]Vineet Kumar Singh, [2]Shivani Kumar, [3]Monica Sehrawat
[1]Research Student, [2]Research Student, [3]Assistant Professor
Computer Science and Engineering,
ABES Institute of Technology, Ghaziabad, India

Abstract : The most important cognitive function that our brain effectively does is detecting and interpret face and facial expressions during communication or any specific scenario with minimal or no hassle. With the boom in artificial intelligence systems and machine learning in recent years has led to development of various intelligent systems which can itself learn and extract knowledge from the variety of data fed to the systems This led to the emergence of intelligent systems which can learn, identify and understand human emotion via verbal communication i.e. Speech or Text or non-verbal communication i.e. Facial expression and body language. The goal is to develop a facial emotion recognition model which can understand human facial expressions and detect the mood and mind state of a person based on input data. It uses the concept of computer vision and machine learning to identify the emotions of the person based on his/her facial expression**.**

Index Terms **– Facial, Emotion recognition, FER, Haar features, HOG vector algorithm, Viola Jones algorithm, KNN algorithm, Adaboost.**

I. INTRODUCTION

A Human emotion plays an important role in the interpersonal relationship. We possess and express emotions in everyday interactions as they are vital in effective communication to boost non- verbal cues for better understanding. We express our feelings and emotions many times in facial expressions rather than verbal communication. For instance, A simple smile can indicate our approval or scowl might express displeasure or disagreement. Facial expression includes smiling, crying, yawning, eye rolling, frowning, etc. While another facial expression can express a sudden shock or excitement, like opening one's mouth widely. People do a great deal of inferencing from the perceived facial expression for example "You seem sad" or "You look happy" and the fact we understand emotions and know how to react accordingly greatly enriches the communication between individuals or groups. As they are the most important way of expressing ourselves and it takes more priority over verbal expressions as we believe what we see which lacks in communication via email or text.

On this basis automatic recognition of emotions has been an active research topic from early eras therefore, there are several advances made in this field in recent years by researchers and tech giants leading to implementations of facial emotion recognition in prevalent and upcoming technologies. Emotions are reflected from hand, gestures of the body, speech and through facial expressions. Hence extracting and understanding of emotion has high importance of the interaction between human and machine communication.

Engineers and researchers have tried to analyze facial expressions in an attempt to understand and classify them so that this knowledge can be used to train our computers using machine learning algorithms using video inputs and images as input trained data. It is done by the first detection of a face from the input image, then extracting and processing the facial expression information and then finally classification of expression (i.e. In which category they lie in).

The aim of this project is to develop a system using machine learning which can analyze different faces provided as input images and identify the emotion of a person in the image. This project uses Hog Vector Algorithm, for feature extraction and for image processing by edge detection and blob detection in image, Viola Jones Algorithm for face detection using Haar feature for mapping similar face features like eyes, nose, ears, etc., K- nearest neighbor or KNN classifier is instance-based learning for classifying input data into different sets and store them which is used for pattern recognition and neural network classification for pattern and image analysis.

II. OBJECTIVES

The study was carried out in order to determine the following objectives:

a) To explore the present status of usage and applications of facial emotion recognition system in the industries.

b) To find out the areas in which the facial emotion recognition system is being used in the industry and day to day life.

c) To identify the merits and demerits of using facial emotion recognition system.

d) To check how the different type of facial emotions can be used to predict the emotions of the person in the picture when given as an input to the model.

e) To make the emotion recognition process faster, accurate and reliable.

III. FACIAL EMOTION RECOGNITION

We humans can expand their knowledge to adapt to the changing environment and to accomplish it we must "learn". Learning is a process of acquiring knowledge about something through study, experience, or by being taught by some external agent and is a continuous process by which a system improves performance or experience. Learning is a most important attribute of all living animals on this planet but is most well developed and prominent in human beings but this is not present in computers which work on binary data i.e. 0's and 1's. Then a big question raised "How to make computers learn?

With the advent of Machine Learning and Artificial Intelligence in recent years have brought in a new era of computing systems which provide them advanced capabilities of self-learning artificially by analyzing the data sets given as input. When we say machine learns, we actually mean that the machine is able to make predictions from past data and is able to process information on the basis of that data. Due to increase in capability of hardware in recent years has led to tremendous increase in computing power of a computer system which helped in the development of systems in which are capable of performing such tasks which speed and precision. Machine learning made computers bring one step closer to mimic human-like intelligence and giving them the ability to think and process knowledge out of different types of input data like images, videos, text, etc. Facial emotion recognition (FER) is an important topic in the fields of computer vision and artificial intelligence owing to its significant academic and commercial potential. According to different surveys, verbal components convey one-third of human communication, and nonverbal components convey two-thirds. Therefore, it has its applications not only in the perceptual and cognitive sciences but also in affective computing and computer animations. It has also been increasing recently with the rapid development of artificial intelligent techniques, including in human-computer interaction (HCI), virtual reality (VR), augment reality (AR), advanced driver assistant systems (ADASs), and entertainment. It can also be used in day to day life as in video calling mobile application for detecting emotions or in selfie applications to apply filters to our detected face like Snapchat©.

Now let us discuss how Facial Emotion Recognition works. Let's take an overview of how it works and we will discuss it in detail later when we will discuss algorithms.

Consider this example given below. In this, we have provided the face of a girl who is sad as an input image to our model. The FER model will first detect the face in the image using image processing then feature extraction process will be done on the input image. After the extraction of features of the face in the image, the classification of the features is done and the output result of this process will be the emotion depicted in the input image.

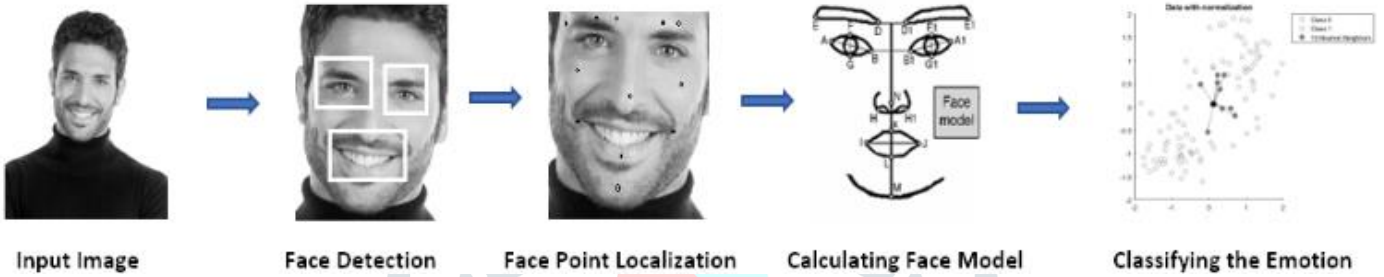

Fig 1.1: Above image shows how a Facial Emotion Recognition works.

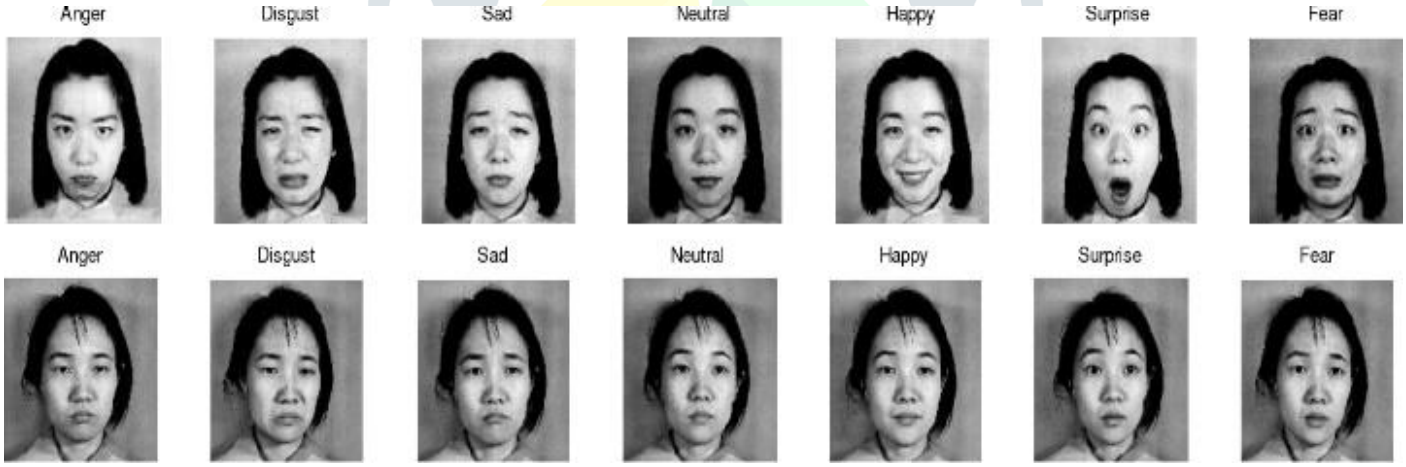

Fig 1.2: Examples of various types of emotions of a person.

IV. ALGORITHMS USED

In this project, we will be using MATLAB as the Machine Learning computational environment and we will be using the following algorithms as:

1) HOG Vector algorithm for feature extraction from an image.

2) Viola Jones for face detection

3) K-Nearest Neighbors for classification of the emotion

We will talk about these tools in detail below.

4.1 Histogram of Oriented Gradients (HOG) Descriptor

Histogram of oriented gradient which is also known as HOG feature descriptor is used to detect objects in computer vision and image processing in which data is given as an input image or video. A feature descriptor is an image representator that processes the input image by extracting useful information from it and discarding all other extraneous information. This technique counts occurrences of gradient orientation in a localized part of an image. It is widely utilized as a feature described image region for object detection such as a human face or human body detection. The HOG algorithm works by checking the local object appearance and shape that can often be characterized the distribution of the local intensity gradients or edge directions, even when we don't have any precise knowledge of the corresponding gradient or edge positions. The orientation analysis is robust to lighting changes since the histogramming of the gradients gives us its translational invariance. The HOG feature summarizes the distribution of measurements within the image regions and is particularly useful for recognition of textured objects with deformable shapes. It is simple and fast so the histogram can be calculated quickly.

4.2 Viola Jones algorithm

Paul Viola and Michael Jones created a fast and robust method for face detection. It has four basic parts in face detection: Haar feature extraction, integral image, AdaBoost, and classifier cascade. The basic principle of the Viola-Jones algorithm is to first scan the input image and then capable of detecting faces in it. The integral image is made by making each pixel equal to the entire sum of all pixels above and to the left of the selected pixel**.** The AdaBoost algorithm helps in focuses on classification problems and aims to convert a set of weak classifiers into a strong one and increase the performance of the classifier. The cascades object classifier is basically used to detect people's faces, noses, eyes, mouth.

The technique relies on the use of simple Haar features that are evaluated quickly through the use of a new integral image. All human faces share some similar properties. These regularities may be matched using Haar Features.

Basic features common to every human face are: Eyes, lips, nose, ears

The composition of properties of recognizing facial features are as follows:

- Location and size of the bridge of the nose, mouth, and eyes
- The upper region of cheeks is brighter than the eye region.
- The eye region is darker than the nose region.
- Value of gradients of intensities of individual pixels

Using integral image, it generates a large set of features from the detected face and then uses the boosting algorithm AdaBoost to reduce the over-complete set and the introduction of a degenerative tree of the boosted classifiers provides the robust and fast interferences. The detector is applied in a scanning fashion and used on gray-scale images, the scanned window that is applied can also be scaled, as well as the features evaluated.

4.3 K-Nearest Neighbor (KNN) algorithm

The *k*-nearest neighbor algorithm is a non-parametric, instance-based method which uses supervised learning used for classification and regression. When you say a technique is non-parametric, it means that it does not make any assumptions on the underlying data distribution and it is important in the real world, most of the practical data does not obey the typical theoretical assumptions made (e.g. Gaussian mixtures, linearly separable, etc.) as data in world is a mixture of multiple types of data. Instance-based learning means that algorithm we do not need to explicitly make the model learn Rather, it memorizes the training data which were subsequently used as "knowledge" for the prediction phase in training of the model. This means that only when a data is to be predicted (i.e. when we give input to the model and ask it to predict a label on the given input), will the algorithm use the training instances to show output. It is also a lazy algorithm. What this means is that it does not use the training data points to do any generalization i.e. there is no explicit training phase or it is very minimal which makes training phase pretty fast. K-NN algorithm simply relies on the distance between all feature vectors, we have the different category associated with each type of input so we can predict and return an actual category for the image. It searches the memorized training observations for the 'K' instances that most closely resemble the new instance and assigns to it their most common class.

The Euclidean distance metric is often chosen to determine the closeness between the data points in KNN. The Euclidean distance is given by:

$$d(x, x') = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \ldots + (x_n - x'_n)^2}$$

V. Conclusion

To develop and maintain a reliable facial emotion recognition system is a critical task. Therefore, proper training of the model needs to be done prior to prediction. Facial Emotion Recognition is the future of a more intrutive way of interaction when integrated with Artificially Intelligent systems or a virtual assistant. The progress is still being made to use it in real time using live or dynamic video data which is being given as an input to the model and also to improve speed and accuracy in previous models.

VI. Acknowledgment

This research was supported by my mentor Prof. Monica Sehrawat at ABES Institute of Technology. I thank my project partner Shivani Kumar from who provided insight and expertise that greatly assisted the research. I would also like to show our gratitude to the faculty at ABES Institute of Technology for sharing their pearls of wisdom with us during the course of this research.

REFERENCES

[1] Morrison RL, Bellack AS, Mueser KT: Deficits in facial-affect recognition and schizophrenia. Schizophr Bull 1988;14:67–83

[2] Mandal MK, Pandey R, Prasad AB: Facial expressions of emotions and schizophrenia: a review. Schizophr Bull 1998; 24:399–412

[3] Edwards J, Pattison PE, Jackson HJ, Wales RJ: Facial affect and affective prosody recognition in first-episode schizophrenia. Schizophr Res 2001; 48:235–253

[4] Kerr SL, Neale JM: Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? J Abnorm Psychol 1993; 102:312–318

[5] Salem JE, Kring AM, Kerr SL: More evidence for generalized poor performance in facial emotion perception in schizophrenia. J Abnorm Psychol 1996; 105:480–483

[6] Silver H, Shlomo N, Turner T, Gur RC: Perception of happy and sad facial expressions in chronic schizophrenia: evidence for two evaluative systems. Schizophr Res 2002; 55:171–1