*International Journal For Technological Research In Engineering*
*Volume 6, Issue 7, March-2019* *ISSN (Online): 2347 - 4718*

# FACE DETECTION USING DEEP NEURAL NETWORK FOR BEHAVIOUR ANALYSIS

Ashish Dusane[1], Mrs. Vineeta Tiwari[2], Mr. Manish Nirmal[3], Prof. Bhushan Chaudhari[4]
[1]M.Tech, Department of Computer Science Engineering
[2,3]Department of ACTS CDAC ACTS – PUNE, India
[4]Department of Computer Sciences Sandip University, Nashik

***Abstract: Future Smart Classroom that we envision will significantly enhance learning experience and seamless communication among students and teachers using real-time sensing and machine intelligence. In that, we research a Smart Classroom system that consists of these components like Student Attendance using Facial Recognition algorithms, human behavior, such as hand gestures, facial expressions and body language, Sentiment Analysis based of Student behavior.***

***We base our suggested system components on existing research in deep learning-based Emotion / Behavior recognition. In this paper comprehensive study of the process of deep learning-based Emotion / Behaviour recognition has been done and determining the computational requirements of a system that incorporates these technologies. Based on these requirements, we research a study of the system like: i) Smart classroom automatic attendance system by using two deep learning facial recognition algorithms "FaceNet" and "DeepFace". ii) Behavior Analysis Model Based on Facial Recognition using Neural Network or Haar Classifier. iii) Based on Students Behavior Analysis automatic Feedback for Teacher.***

***Keywords: Face Detection, Face Recognition, Face Identification, Behaviour Analysis***

## I. INTRODUCTION

To analyse the advantages and disadvantages, problems encountered and solutions found, when using face detection and recognition in an academic environment to keep track of the attendance of students, where in this process we use a CCTV camera to be fixed at the entry point of a classroom, which automatically captures the image of the person and checks the observed image with the facial database using proposed system. It is typically used for two purposes. Initially marking the attendance for a student by comparing the face images produced recently which does recognition of human who are new or strange to the environment i.e. an unauthorized person. For identification and verification of image, a newly emerging trend 3-Diemention Face Recognition is used which claims to provide more accuracy in matching the image databases and has an ability to recognize a subject at different views.

Also from that face detection we capture the student's emotion just like Happy, Sad, Neutral, Angry, Disgusted, Surprised, etc. from that emotion we analyse it and from that analysis we get the final overall students behavior for particular lecture. So from that students behavior we get the result in form of teacher feedback also and students feedback also. Subsequently, detailed description of the three main parts Student attendance system by using two deep learning facial recognition algorithms ―FaceNet‖ and ―DeepFace‖. ii) Behavior Analysis Model Based on Facial Recognition using : Neural Network or Haar Classifier. iii) Based on Students Behavior Analysis automatic Feedback for Teacher.

MOTIVATION:

Face recognition has been a sought after problem of biometrics and it has a variety of applications in modern life. The problems of face recognition attracts researchers working in biometrics, pattern recognition field and computer vision. Several face recognition algorithms are also used in many different applications apart from biometrics, such as video compressions, indexing etc. They can also be used to classify multimedia content, to allow fast and efficient searching for material that is of interest to the user. An efficient face recognition system can be of great help in forensic sciences, identification for law enforcement, surveillance, authentication for banking and security system, and giving preferential access to authorized users i.e. access control for secured areas etc. The problem of face recognition has gained even more importance after the recent increase in the terrorism related incidents. Use of face recognition for authentication also reduces the need of remembering passwords and can provide a much greater security if face recognition is used in combination with other security measures for access control. The cost of the license for an efficient commercial Face recognition system ranges from 30,000 $ to 150,000 $ which shows the significant value of the problem. Though face recognition is considered to be a very crucial authentication system but even after two decades continuous research and evolution of many face recognition algorithms, a truly robust and efficient system that can produce good results in real time and normal conditions.

## II. RELATED WORK

Face Detection :

Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images. Face detection also refers to the psychological process by which humans locate and attend to faces in a visual scene.

Face detection can be regarded as a specific case of object-class detection. In object-class detection, the task is to find

  

the locations and sizes of all objects in an image that belong to a given class. Examples include upper torsos, pedestrians, and cars. Face-detection algorithms focus on the detection of frontal human faces. It is analogous to image detection in which the image of a person is matched bit by bit. Image matches with the image stores in database. Any facial feature changes in the database will invalidate the matching process.
Firstly, the possible human eye regions are detected by testing all the valley regions in the gray-level image. Then the genetic algorithm is used to generate all the possible face regions which include the eyebrows, the iris, the nostril and the mouth corners.

Each possible face candidate is normalized to reduce both the lightning effect, which is caused by uneven illumination; and the shirring effect, which is due to head movement. The fitness value of each candidate is measured based on its projection on the Eigen-faces. After a number of iterations, all the face candidates with a high fitness value are selected for further verification. At this stage, the face symmetry is measured and the existence of the different facial features is verified for each face candidate.

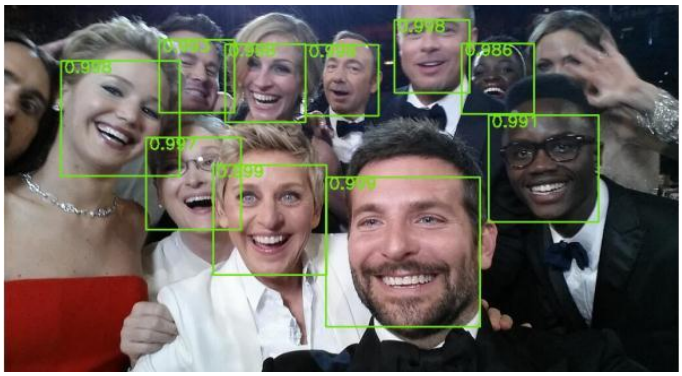

Fig. 1: Face Detection in Images

Face Recognition :

A facial recognition is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems work, but in general, they work by comparing selected features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analysing patterns based on the person's facial textures and shape. While initially a form of computer application, it has seen wider uses in
It is typically used as access control in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems. Although the accuracy of facial recognition system as a biometric technology is lower than iris recognition and fingerprint recognition, it is widely adopted due to its contactless and non-invasive process.[3]
Recently, it has also become popular as a commercial identification and marketing tool.

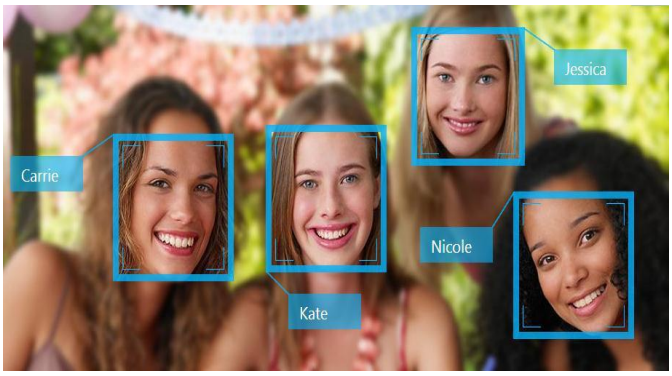


Fig 2 : Face Recognition in Image

DIGITAL IMAGE PROCESSING :
The term digital image processing generally refers to the processing of a two-dimensional picture by a digital computer. A digital image is an array of real numbers represented by a finite number of bits. The principle advantage of Digital Image Processing methods is its versatility, repeatability and the preservation of original data precision. The various Image Processing techniques are:

- Image pre-processing
- Image enhancement
- Image segmentation
- Feature extraction
- Image classification

Steps of Image Processing are :

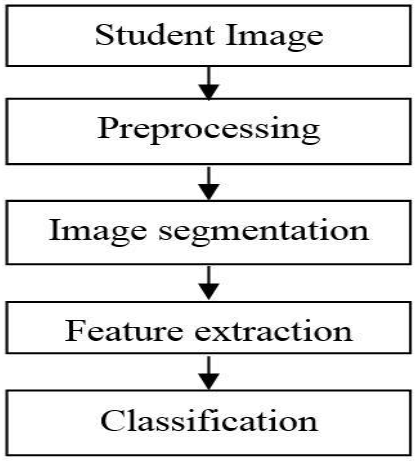


Fig 3 : Image processing process

IMAGE PREPROCESSING :
Image pre-processing is the process in which image data is recorded by sensors on a satellite restrain errors related to geometry and brightness values of the pixels integrated in picture. These errors are corrected using appropriate mathematical models which are either definite or statistical models. Image enhancement is the process of modifying image by changing the pixel brightness values to improve its visual impact. Image enhancement involves a collection of techniques that are used to improve the visual appearance of an image, or to convert the image to a form which is better suited for human or machine interpretation.
a. Contrast Stretching
b. Noise Filtering

  

c. Histogram modification

a. Contrast Stretching :
Contrast Stretching is the process of changing the impact on the images through some processing. Some images (eg. deserts, dense forests, snow, clouds and under hazy conditions over heterogeneous regions) are homogeneous i.e., they do not have much change in their levels. Let's consider in terms of histogram representation, they are characterized as the occurrence of very narrow peaks. The homogeneity can also be due to the incorrect illumination of the scene. Ultimately the images hence obtained are not easily interpretable due to poor human perceptibility. This is because there exists only a narrow range of gray-levels in the image having provision for wider range of gray-levels. Basically contrast stretching methods are designed exclusively for frequently encountered situations. In image contrast there are different stretching techniques have been developed to stretch the narrow range to the whole of the available dynamic range.

b. Noise Filtering :
Noise Filtering is used to filter the unnecessary information from an image. It is also used to remove various types of noises from the images. Mostly this feature is interactive. Various filters like low pass, high pass, mean, median etc., are available.

c. Histogram Modification :
Histogram has a lot of importance in image enhancement. It reflects the characteristics of image. By modifying the histogram, image characteristics can be modified. One such example is Histogram Equalization. Histogram equalization is a nonlinear stretch that redistributes pixel values so that there is approximately the same number of pixels with each value within a range. The result approximates a flat histogram. Therefore, contrast is increased at the peaks and lessened at the tails.

IMAGE SEGMENTATION:
Image segmentation is one of the key problems in image processing method. Generally image segmentation is the process that subdivides an image into its constituent parts or objects. The level to which this subdivision is carried out depends on the problem being solved, i.e., the segmentation should stop when the objects of interest in an application have been isolated for example in autonomous air-to-ground target acquisition let's consider suppose our interest lies in identifying vehicles on a road, the first step is to segment the road from the image and then to segment the contents of the road down to potential vehicles. Image thresholding technique is basically are used for image segmentation.

Segmentation of images involves sometimes not only the discrimination between objects and the background, but also separation between different regions. One method for such separation is known as watershed segmentation.

FEATURE EXTRACTION:
Feature extraction techniques are basically developed to extract features in synthetic aperture radar images. Usually his technique extracts high-level features needed in order to perform classification of target images. Features are those items which uniquely describe a target, such as size, shape, composition, location etc. Segmentation techniques are used to isolate the desired object from the scene so that measurements can be made on it subsequently. When the pre-processing and the desired level of segmentation has been achieved, some feature extraction technique is applied to the segments to obtain features, which is followed by application of classification and post processing techniques. Feature selection of a feature extraction method is the single most important factor in achieving high recognition performance. Feature extraction has been given as ―extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the between class pattern variability‖. Thus, the selection of a suitable feature extraction technique according to the input to be applied needs to be done with utmost care. Taking into consideration all these factors, it becomes essential to look at the various available techniques for feature extraction in a given domain, covering vast possibilities of cases.

IMAGE CLASSIFICATION :
In the image classification the proposed algorithm provides energy efficient path for data transmission and maximizes the lifetime of entire network. As the performance of the proposed algorithm is analyzed between two metrics in future with some modifications in design considerations the performance of the proposed algorithm can be compared with other energy efficient algorithm. We have used very small network of 5 nodes, as number of nodes increases the complexity will increase. Image classification is the labelling of a pixel or a group of pixels based on its grey value. Classification is one of the most often used methods of information extraction. In Classification, usually multiple features are used for a set of pixels i.e., many images of a particular object are needed. In Remote Sensing area, this procedure assumes that the imagery of a specific geographic area is collected in multiple regions of the electromagnetic spectrum and is in good registration. Most of the information extraction techniques rely on analysis of the spectral reflectance properties of such imagery and employ special algorithms designed to perform various types of 'spectral analysis'. The process of multispectral classification can be performed using either of the two methods: Supervised or Unsupervised.

Challenges in face detection and recognition :
Detecting and recognizing faces are challenging as faces have a wide variability in poses, shapes, sizes and texture. The problems or challenges in face detection and recognition are listed as follow:
Pose:
A face can vary depends on the position of the camera during the image is captured.
Presence of structural components: There may be another

  

additional components on the face such as spectacles, moustache or beard. These components may have different types, shapes, colours and textures.

Facial expression:

The facial expression resembles directly on the person's face.

Occlusion

A face may be partially obstructed by someone else or something when the image is captured among crowds.

Image orientation

It involves with the variation in rotation of the camera's optical axis.

## III. DEEP LEARNING

Deep learning uses layered architecture for building computational models which mainly contains input layer, hidden layer which extract features from image and output layer which classifies objects. It mainly working like human brain.

Recurrent Neural Network :

The recurrent neural network (RNN), unlike feedforward neural networks, can operate effectively on sequences of data with variable input length

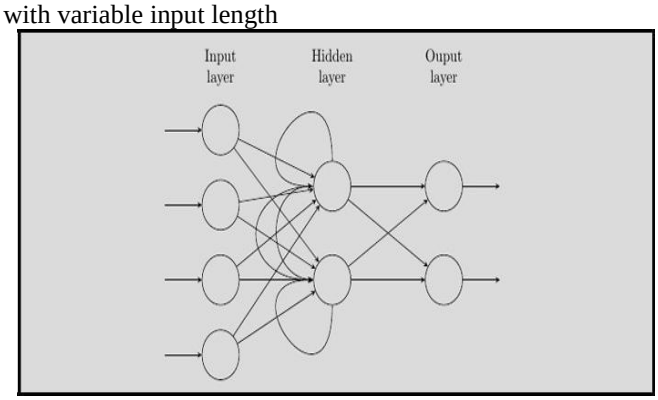


Fig. 4 : Recurrent Neural Network

RNNs uses knowledge of its previous state as an input for its current prediction, and we can repeat this process for an arbitrary number of steps allowing the network to propagate information via its hidden state through time. This is essentially like giving a neural network a short-term memory. This feature makes RNNs very effective for working with sequences of data that occur over time.

The two variants on the basic RNN architecture that help solve a common problem with training RNNs are Gated RNNs, and Long Short-Term Memory RNNs (LSTMs).

Convolutional Neural Network :

The typical use of convolutional networks is on classification tasks, where the output to an image is a single class label. However, in many visual tasks, especially in biomedical image processing, the desired output should include localization, i.e., a class label is supposed to be assigned to each pixel. Moreover, thousands of training images are usually beyond reach in biomedical tasks. One of the most popular types of deep neural networks is known as convolutional neural networks (CNN or ConvNet). A convolutional neural network (CNN) is a class of deep, feed-forward networks, composed of one or more convolutional layers with fully connected layers (matching those in typical Artificial neural networks) on top. A CNN convolves learned features with input data, and uses 2D convolutional layers, making this architecture well suited to processing 2D data, such as images.

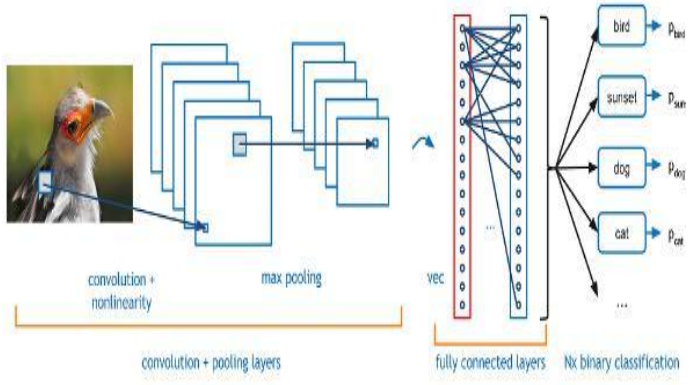


Fig 5 : CNN Architecture

CNNs eliminate the need for manual feature extraction, so you do not need to identify features used to classify images. The CNN works by extracting features directly from images. The relevant features are not pretrained; they are learned while the network trains on a collection of images. This automated feature extraction makes deep learning models highly accurate for computer vision tasks such as object classification. CNNs learn to detect different features of an image using tens or hundreds of hidden layers. Every hidden layer increases the complexity of the learned image features. For example, the first hidden layer could learn how to detect edges, and the last learns how to detect more complex shapes specifically catered to the shape of the object we are trying to recognize. Regular Neural Networks transform an input by putting it through a series of hidden layers. Every layer is made up of a set of neurons, where each layer is fully connected to all neurons in the layer before. Finally, there is a last fully-connected layer—the output layer—that represent the predictions.

CNNs have Two Components :

The Hidden layers/Feature extraction part

In this part, the network will perform a series of convolutions and pooling operations during which the features are detected. If you had a picture of a zebra, this is the part where the network would recognize its stripes, two ears, and four legs.

The Classification part:

Here, the fully connected layers will serve as a classifier on top of these extracted features. They will assign a probability for the object on the image being what the algorithm predicts it is.

Why Deep learning?

Deep learning take huge amount of data and gives output with good performance. The network automatically learn features and do not need to do it manually. It try to learn low level features first such as edges, corners and contours then

  

high level features such as recognize objects in image.

Haar:
Haar Feature-based Cascade Classifier for Object Detection. The word ―cascade‖ in the classifier name means that the resultant classifier consists of several simpler classifiers (stages) that are applied subsequently to a region of interest until at some stage the candidate is rejected or all thestages are passed.
Object Detection using Haar feature-based cascade classifiers is an effective object detectionmethod proposed by Paul Viola and Michael Jones in their paper, "Rapid Object Detection usinga Boosted Cascade of Simple Features" in 2001. It is a machine learning based approach where acascade function is trained from a lot of positive and negative images. It is then used to detectobjects in other images.
In that algorithm initially, the algorithm needs a lot of positive images(images of faces) and negative images (images without faces) to train the classifier. Then needto extract features from it. They are just like our convolutional kernel. Each feature is a single value obtained by subtracting sum of pixelsunder white rectangle from sum of pixels under black rectangle.

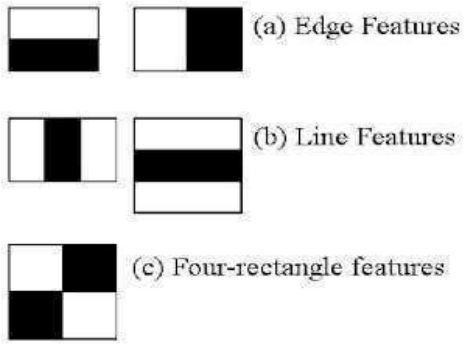


Fig 6 : HAAR Cascade

## IV. CONCLUSION

Based on the literature survey on this topic of Face Detection using Neural Network, different aspects of how face detection & recognition works are studied. Various survey papers are reviewed and analysis is done for each paper. The purpose of this research is to provide an overview of the functionality of Face Detection using Deep Neural Network for Analyzing the Emotion and Behavior of Student. This proposed work can enhance the overall performance of Face Detection process and moreover focus on students behavior thoroughly.

## REFERENCES

[1] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf. "Deep Face: Closing the Gap to Human-Level Performance in Face Verification‖,2014 IEEE Conference on Computer Vision and Pattern Recognition
[2] SanunSrisuk, SurachaiOngkittikul"Robust Face Recognition based on Weighted DeepFace", 5th International Electrical Engineering Congress, Pattaya, Thailand, 8-10 March 2017.
[3] Gurpreet Kaur, Sukhvir Kaur, Amit Walia."Face Recognition System using Deep Face and Neural Network", IJSRD - International Journal for Scientific Research & Development| Vol. 4, Issue 05, 2016
[4] Justus Schwan, EsamGhaleb, Enrique Hortal and Stylianos Asteriadis, "High-Performance and Lightweight Real-Time Deep Face Emotion Recognition", 2017 IEEE
[5] Florian Schroff, Dmitry Kalenichenko, James Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", 2015 IEEE
[6] Adnan Firoze, Tonmoay Deb. "Face Recognition Time Reduction Based on Partitioned Faces without Compromising Accuracy and aReview of state-of-the-art Face Recognition Approaches", ICIGP 2018, February 24–26, 2018, Hong Kong, Hong Kong.
[7] Enrique Correa, ArnoudJonker, Michael Ozo, Rob Stolk ―Emotion Recognition using Deep Convolutional Neural Networks", June 30, 2016.

  