International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

# Facial recognition using Convolutional Neural Networks (CNN)

PRADEEP J

Department of Computer Science and Engineering,

Adhi College of Engineering and Technology, Kanchipuram – 631 605.

RADHIKA R

Assistant Professor, Department of Computer Science and Engineering,

Adhi College of Engineering and Technology, Kanchipuram- 631 605.

NANDA KUMAR D

Department of Computer Science and Engineering,

Adhi College of Engineering and Technology, Kanchipuram – 631 605.

MAMILLAPALLI AJIT KUMAR

Department of Computer Science and Engineering,

Adhi College of Engineering and Technology, Kanchipuram – 631 605.

GUTLAPALLI SANDEEP CHOWDARY

Department of Computer Science and Engineering,

Adhi College of Engineering and Technology, Kanchipuram – 631 605.

## ABSTRACT

Face recognition is a common biometric authentication technique used to analyse the face images and extract useful recognition information from them, which are always called as a feature vector that is used to distinguish the biological features. Face Recognition process begins with extracting the coordinates of features such as width of mouth, width of eyes, pupil, and compare it with a stored face template. The aim of the proposed system is to design a system that performs face recognition based on detecting the faces with a system to identify the accuracy of a person in the system while detecting faces. OpenCV library and image processing algorithms is used to detect and extract the face from an image thereby storing samples in order to train the system.

**Keywords** - Haarcascade algorithm, OpenCV, Convolutional Neural Networks, Deep Learning.

## INTRODUCTION

Biometrics authentication (or realistic authentication) is a form of authentication which is used for identification and access control. Biometric authentication is mainly based on physiological and behavioural characteristics. The traits such as uniqueness, permanence, measurability, performance, acceptability and circumvention is checked in an individual for biometric verification. There are various types of biometric authentication like fingerprint identification, Iris scan, retina scan, face recognition, voice analysis, etc. Fingerprint identification is most commonly used form of authentication in biometrics. But the disadvantage is that a person's fingerprint's pattern or form may change over time and fingerprint scanner does not take this into consideration. In the current scenario, there are lot of face recognition techniques and algorithm found and developed around the world. Face recognition therefore, has received a great deal of attention in various applications in the field of image processing, computer vision, etc due to several advantages it has over other biometric method. For example, in public security system, it can identify the identity of the suspect; in the bank and customs control system, it can identify and prove the identity; it also helps users safeguard its own

confidential information and experience more secure financial transaction. The proposed system is trained to recognise a set of authorized person. Haarcascade system is used to create dataset of authorized person dynamically by identifying and extracting the facial features of face helping the system to recognise the face. All others who enter the guarded area are considered strangers. Neural network is used to train the system in order to identify the stranger by comparing the dataset of all authorized person.

## RELATED WORKS

The algorithms commonly used for face recognition are active contour model and deformable template model . This model is based on the geometrical characteristic, which is first applied to the face recognition problem. Its basic idea is the difference of everyone's face because of difference in components of every face, like the eyes, noses, mouths and jaws are different. Thus the system uses the set of architectures and shapes of these components to be the features for the face recognition problem. There are five useful methods for face recognition developed in the past study.

The sub-space analysis method is often used in face recognition, which contains two methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). PCA is a technique used for identification of a smaller number of uncorrelated variables known as principal components from a larger set of data. The technique is widely used to emphasize variation and capture strong patterns in a data set. Principal component analysis is considered as a useful statistical method and used in predictive model and exploratory data analysis. The most classic method is PCA based Eigenface. This method take the face images as random variables, which turns the N× N vector of a face image to a N2 × 1 vector, and after minuses the mean data vector, uses the K-L transformation to get a set of orthogonal basis, then after keeps part of the principal components, the reduced dimension vector space of face images is obtained. It tries to find a projection direction, which can make the distance of within-class, is small and the distance of between-class is large based on the training samples projection to that direction. Compared to the PCA method, only if the training sample is large, LDA can get a better result.

Researchers employed the Hidden markov model (HMM) to solve the problem that the different appearance of facial features and the connection of each other. Based on this model, the features observed are treated as a sequence of unobserved states. Different people use different HMM parameters, and for the same person, system uses the model with same parameters to represent the observed sequence of gestures and facial expressions. Samaria first proposed the face model, which used a rectangular window sampling face images from top to bottom. Another commonly practised method for face recognition is Neural network (NN). Neural network uses its ability of learning and classifying to extract and recognize face features. Lin, etc. use the positive and negative samples for reinforcing learning to get an ideal probability result. And then increase the learning speed by applying a modular network. Proposed system was inspired by Ya Wang's Deep learning method for Face Recognition in Real-world. This system automatically generates dataset from real world surveillance video. This helps in generating dataset in various light illuminations, with different facial expressions, etc.

Improving the face recognition performance of Convolutional neural network (CNN's) by using non-CNN features. The non-CNN features showcase the characteristics from a different perspective of the targeted face images. In terms of results, Face recognition based on deep neural network works the best. The system uses CNN which is a neural network capable of handling image data. It comprises of three layers, one convolution layer, one pooling layer and one fully connected layer. CNN can learn the variations of data without prior knowledge. This method also helps in identifying a person using additional features. The system uses Labled Faces in the dataset for its implementation. A dataset of face photographs designed for studying the problem of unconstrained face recognition, known as haarcascade_frontalface_default contains more than 13,000 images of faces collected from the web.

## CNN (Convolutional Neural Networks)

A convolutional neural network (CNN or ConvNet) is **a network architecture for deep learning that learns directly from data**. CNNs are particularly useful for finding patterns in images to recognize objects, classes, and categories. They can also be quite effective for classifying audio, time-series, and signal data.

There are four types of layers for a convolutional neural network: the **convolutional layer, the pooling layer, the ReLU correction layer and the fully-connected layer.**

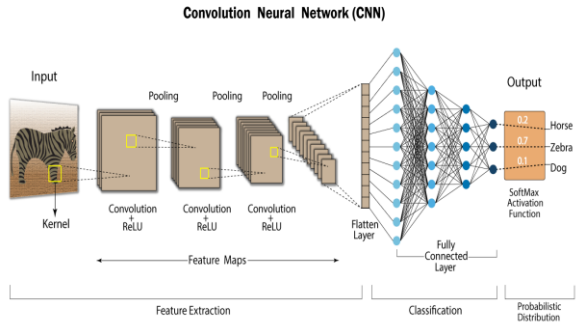


**Fig.1 Layers of Convolutional Neural Networks**

## CNN in FACE RECOGNITION

In deep learning, a convolutional neural network (CNN) is a special type of neural network that is designed to process data through multiple layers of arrays. A CNN is well-suited for applications like image recognition and is often used in face recognition software.

In CNN, convolutional layers are the fundamental building blocks that make all the magic happen. In a typical image recognition application, a convolutional layer is made up of several filters to detect the various features of the image. Understanding how this work is best illustrated with an analogy.

Suppose you saw someone walking toward you from a distance. From afar, your eyes will try to detect the edges of the figure, and you try to differentiate that figure from other objects, such as buildings or cars, etc. As the person walks closer toward you, you try to focus on the outline of the person, trying to deduce if the person is male or female, slim or fat, etc. As the person gets nearer, your focus shifts toward other features of that person, such as his facial features, if his is wearing specs, etc. In general, your focus shifts from broad features to specific features.

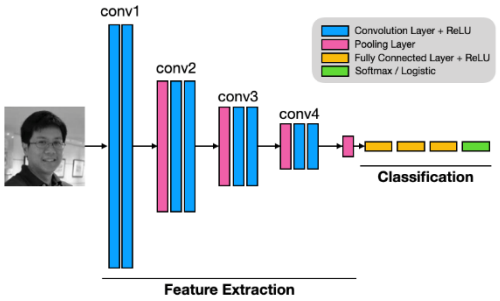


**Fig. 2 CNN in Face Recognition**

## DIGITAL IMAGE PROCESSING IN FACE RECOGNITION

Image Processing in face recognition is a computerized technique that uses an algorithm to locate and recognize a face in an image, and this technology has several uses. While there are many different facial recognition algorithms available, most programs use edge or eye detection to locate a face. Not only does image processing face recognition find a face, but most algorithms also show the user the exact pixel location of the face within the image, making it easier to find a certain face in a crowded or blurred image. In the entertainment industry, facial recognition is used for face tracking in motion capture, which is used for more realistic animation. Another common use for this technique is in facial recognition security; the algorithm is typically specialized in this case to recognize only certain faces.

## STEPS IN FACE RECOGNITION

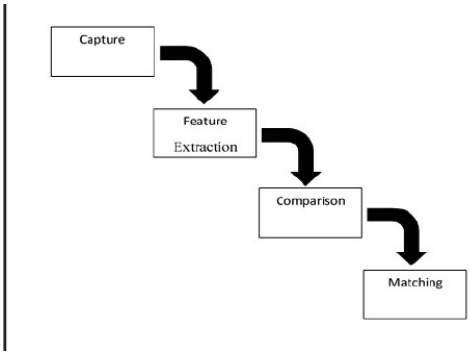


**Fig.3 Steps in Face Recognition**

There are four Steps in Face Recognition:

- Capture
- Feature Extraction
- Comparison
- Matching

## IMAGE PROCESSING ALGORITHMS

The recognition process involves a system which detect the face using algorithms PCA, LDA, LBPH which is an inbuilt algorithm in openCV library for face recognition. The camera will move a capture the images on a real time basis and again perform the face detection process.

## HAARCASCADE CLASSIFIER

Identifying a custom object in an image is known as object detection. This task can be done using several techniques, but we will use the haarcascade, the simplest method to perform object detection in this article.

Haarcascade were first introduced in 2001, and it was one of the most popular object detection algorithms in OpenCV.

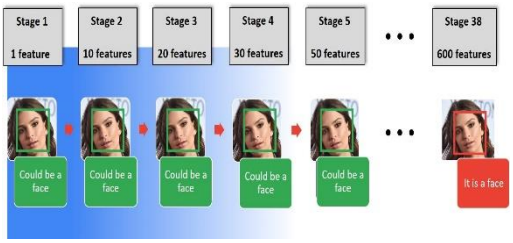


**Fig. 4 Haarcascade Classifier**

Haar cascade is an algorithm that can detect objects in images, irrespective of their scale in image and location.

This algorithm is not so complex and can run in real-time. We can train a haar-cascade detector to detect various objects like cars, bikes, buildings, fruits, etc.

Haar cascade uses the cascading window, and it tries to compute features in every window and classify whether it could be an object.

## IMPLEMENTING HAARCASCADE

The OpenCV library manages a repository containing all popular haar cascades that can be used for:

- Human face detection
- Eye detection
- Nose / Mouth detection
- Vehicle detection

It can be implemented by using OpenCV by installing OpenCV in the System for capturing and processing the image.

## IMAGE PREPROCESSING

**1)Labelled faces in the dataset**

The selection of a appropriate dataset plays a very important role in the proposed system. The dataset should consist of valid labelled images in each class in order that the neural network can learn every label. Having no restriction to the number of images per class, better results can be obtained through a wide range of images in the Haarcascade dataset. The main purpose of the Haarcascade dataset is to verify whether two

images are of the same individual or not as well as facial verification.

**2)Elimination of classes which contains less images**

Using all of the classes (individuals) would result in a useless model as many of the individuals have only a single image. It
is visionary for even the most powerful convolution neural network to learn to identify an individual from a single image.
Limiting the data to only 10 individuals with the most images in the dataset, is considered by the proposed system in order to give the model a chance to learn all the classes resulting in 10 classes with at least 50 images per class. It is indeed considered a feasible idea to avoid using all the images from the Haarcascade dataset. It is quite baffling for a neural network trained on such a dataset where 4096 individuals have only a single images of themselves.

**3) Training using Inception module**

Considering the latest and updated Inception V3 model which comprises of the parameters learned through training on the FaceNet dataset, Google's pre-trained Inception Convolution Neural Network is selected to perform image recognition as building and training the CNN is not needed.

## IMAGE PROCESSING

**1)Processing the images**

Images has to be provided to the FaceNet CNN in the form of arrays [batch_size,image_height,image_width,colour_channels]. The number of images in the training set or testing batch is represented by the batch_size, for Inception V3, the image of 299*299 is required, whereas the number of colour channels for Red-Green-Blue should be 3. For a colour channel, (x,y) denotes the pixel value between 0-255 which represents the intensity of the colour. ImageNet requires normalized pixel values between 0 and 1, which can be obtained by dividing the array by 255. Although there are built-in functions for processing the image to the required size and format,spicy and numpy method can also be used to fulfil the need. The Haarcascade data has images of 255*255*3 which will be further transformed to 299*299*3 along with normalized pixel values between 0 and 1.

**2)Define layer of Inception CNN to train**

In order to classify the objects, or to carry out image related process like facial verification, training of the parameters connection weights and biases of at least one layer of the network is said to be mandatory. As the lower layers of the convolution neural network are well skilled at identifying lower level features like shapes, colours or textures as well as top layer can differentiate the higher level features like number of appendages or eyes on a human face. It is infeasible to train the entire network on a personal laptop but if the size of the dataset is limited and a Google Cloud GPU or a consumer-grade GPU is used then the last layer of the network can be trained in a reasonable amount of time. Although achievement of record results on the task is doubtful, but the principles involved in adapting an existing model to a new dataset can be visible. Training of one layer of the network and a defined output layer is required so that the CNN can learn new classes. The parameters used by Inception V3 which is learned from FaceNet for all the layers excluding the one before the predictions hoping that all the weights and biases that are useful in differentiating objects can be applied to the face recognition task. In order to determine the trainable layer, having a view towards the structure of the network is advisable. About 12 million layers are comprised in Inception,but all of them are not involved in this process for training. Early stopping is used for training which is used to reduce over-fitting on the training set. It requires periodic testing of the network on a validation set to check the score of the cost
function (average cross entropy). The training is halted if there is no decrease in the loss for specific number of epochs. For retaining the optimal model, whenever there is some improvement in the loss, the model has to be saved. Out of all the models achieved, the one which has the best loss on the validation set is then restored at the end of the training set.
New instances are generalized as the model continues to learn the set in a much better way with each epoch, if early stopping method is avoided. Although this method has a variety of

implementations, a single validation set and stop training is preferred if there is no improvement in the loss for 20 epochs. A new Tensor Board file is created and then the model parameters are saved to restore the evaluation every time a different model is implemented.

## FACE EXPRESSION DETECTION

Haarcascade Dataset can be used to detect the face as well as the face expressions using camera. It can detect various emotions like Happy, Sad, Surprise, Disguist and Neutral.

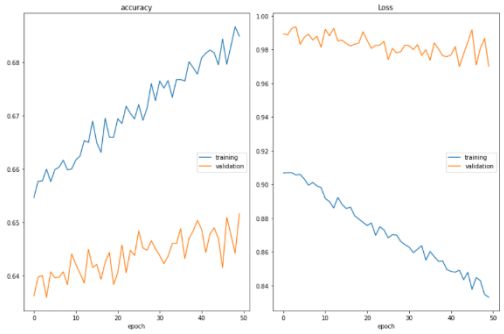


**Fig. 5 Accuracy of Facial Expression**

**Performance measure:**

**1) Loss function**

To measure the performance of a classification model whose output has a probability value between 0 and 1, Cross-entropy loss or log loss is used. It increases if the predicted value diverges form the actual labels. If the actual observation label is 1 and a probability of 0.012 is predicted the it is not assumed to be good and may result in high loss value. A perfect model has log loss log 0. For binary classification, the Cross-entropy can be calculated as follows [number of classes (M) equals to 2].

**(ylog(p)+(1-y)log(1-p))**

For multi-classification [number of classes(M) > 2], a separate loss for class label per observation is carried out followed by summing the results:

**1Myo,clog(po,c)**

**2)Optimization Function**

It is considered as one of the best optimizer till date. With a learning rate of 0.01, the proposed system uses Adam Optimization function as it is able to measure top-1 accuracy. In order to save the model during training and restore it for later, an initialize is used followed by a saver.

**3)Result**

After training the neural network, checking for usefulness is the next step. It can be done by carrying out evaluation against the test set, which is completely unseen by the model. The training will be done by going through the test set one batch at a time. 82% of individuals are correctly identified at the first attempt in this proposed system. It is successful in displaying the fact that the network has learned during training.

**4)Training on augmented data**

The new dataset created by augmentation is used again for training the system in order to increase the accuracy of face recognition. The table below shows comparison in accuracy percentage obtained before and after augmentation. In this system it is found that in order to increase accuracy shifting, flipping and illumination can be performed on the images of stored dataset. This in turn helps in increasing the classes in dataset, thereby increasing the accuracy. Thus in the graph obtained after augmentation straight line moving continuously upwards is seen in comparison to the line graph before augmentation.

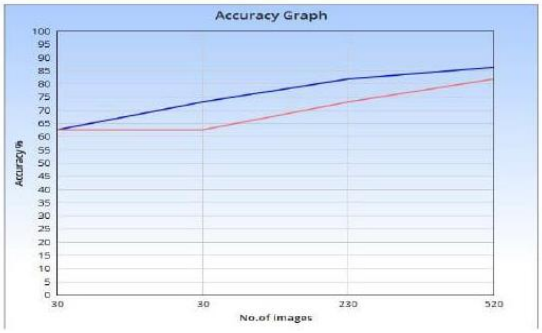


**Fig. 6 Accuracy Graph**

## CONCLUSION AND FUTURE WORKS

By using this project we can enable to detect the faces for identification and expression with the help of Haarcascade_frontalface_default dataset. In future it can be optimized to work with the new types of dataset to implement the recognition system very optimized with more accuracy and to detect using Face Mask which is to be worked out in future.

**References**:

1) F. Samaria, Face recognition using hidden Markov model [D],
Cambridge, University of Cambridge, (1994).

2) S H. Lin, S Y. Kung and L J. Lin, Face recognition/detection by
probabilistic decision based neural network [J], IEEE Trans. on
Neural Networks, vol. 8, no. 1, (1997), pp. 114-132.

3)www.codemag.com/Article/2205081/Implementing-Face-Recognition-Using-Deep-Learning-and-Support-Vector-Machines#:~:text=Deep%20Learning%20%2D%20Convolutional%20Neural%20Network,used%20in%20face%20recognition%20software.

4) Ya Wang, Ming Zhu, Face recognition in real-world surveillance
videos with deep learning method, International conference on
image,vision and computing.

5) Ze Lu, Alex Kot, Enhance deep learning performance in face
recognition, International conference on image,vision and
computing.

6)www.easytechjunkie.com/what-is-image-processing-face-recognition.htm#:~:text=Image%20processing%20face%20recognition%20is,detection%20to%20locate%20a%20face.

7)Li Xinhua, Yu Qian, Face recognition based on deep neural
network, International journal of signal processing,Vol.8,pp 29-38.

8)www.medium.com/@williamkoehrsen/facial-recognition-usinggoogles-convolution-neural-network-aa752b4240e.