Journal of Engineering Sciences
Vol 10, Issue 12, DEC/2019
ISSN NO:0377-9254

ANALYSIS OF AI BASED FACE DETECTION TECHNIQUES

Arnold Sachith A Hans[1] and Smitha Rao[2]
[1]*Department of Computer Science & Engineering, School of Engineering Presidency University, Bengaluru, India*

[2]*Professor, Department of Computer Science & Engineering, School of Engineering, Presidency University, Bengaluru, India*

Abstract: **In this fast-moving world, computer vision has been gaining lots of importance in the field of Artificial Intelligence. Face Recognition is one among the most important field which has been gaining attention in the recent days due to its application in various fields such as Security and Surveillance applications, marketing sector, computer graphics. Face detection is a technique which is used to detect the faces in a image/video frame. This is the primary and a crucial step carried out before the Face recognition process. Some of the ways to detect human faces in a digital image is by using Haar Based Face Detection, Deep learning-based Face detector, Tiny Face detector, Kernel methods etc. In our study face detection is carried out by using OpenCV and a comparison was carried out between the Haar Cascade model and a deep learning model with the pre-defined constraints in accordance with the properties of the image and the physical distance parameters of the camera position etc. The results of face detection with both the models was analysed on different factors such as accuracy, feasibility, total number of faces detected with reference to the actual number of faces existing in the image. Conclusions were drawn to suggest the best methodology well suited for a non-cooperating scenario where testing and train conditions are different such as non-intrusive based face recognition application.**

Keywords: Deep Learning, Face Detection, Machine Learning, Face recognition, Tiny Face Detector

1. INTRODUCTION

The rapid growth in the Computer Vision Technology has brought a wide amount of changes in perspective of the ability of the computer to understand the outer world through digital images or videos. One among the most trending technologies is the Face Detection process.

Some of the Face detection applications are Facial motion capture which is used to produce Computer Graphics, computer animation for movies, real-time avatars and games. Face recognition – Face detection algorithms are necessary to determine which parts of an image which is used to generate the face prints that are compared with previously stored face prints to check whether match is there or not. Marketing – A camera captures the image of face, then the system calculates the gender race, and age range of the face detected, once the information is collected a series of advertisements can be played that is specific towards the detected age/gender race. Emotional Inference – to help people with autism understand the feelings of people around them, Lip reading – help computers determine who is speaking which is needed when security is important, Face tracking, Pose estimation, Used by Face book that offers an option to the user to tag their friends on uploaded images and then will automatically tag the face whenever the same face appears in the photos. In the process of face analysis, face detection tells the face analysis algorithms which part of an image (or video) to focus on when, recognizing gender, identifying age, and analysing emotions based on facial expressions.

Face detection using an Artificial Intelligence based technology that can locate and identify the presence of human face in digital pictures and videos. This can be regarded as a special case of object-class detection, where the task is to find the location and specify the sizes the objects that belong to a given class, in this case it is the faces – within a particular image. Face detection applications use algorithms that determine whether images are positive images (Image which has face) or negative images (Image which does not have face) to perform this activity accurately the algorithms must be trained on huge datasets containing millions of face images and non-face images. The algorithms are trained it will able to predict if there are any faces in the image provided and if any face is found then the algorithm then it detects the particular location of the face and draws a bounding box around the detected face.

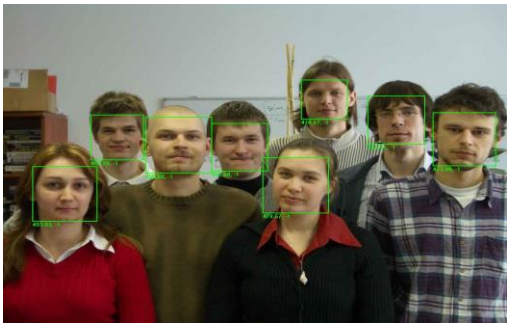


Fig 1: Face detection by drawing a square box around the detected face. Reference: https://www.neurotechnology.com/verilook.html

Starting from the Eigen Faces (Late 19880s-Early 1990s) to the Face book implementing Facial Recognition functionality (2010 - Present) the face detection has evolved from non AI to AI based technology and has seen rapid advancements in the improvement of this technology for the better accuracy. Presently open source Computer vision (OpenCV) has given the freedom for lots of researchers to conduct the research on this domain to work on this trending technology.

One of the basic classical techniques for face detection was presented by Paul Viola and Michael Jones which uses the weak classifier cascade for face detection and this feature has been embedded in opencv platform in the form of .xml files and is available to the user from the opencv distribution packages. There are four components in this Haar feature based Cascade classifier. Firstly user needs to generate features and Haar features are used for the extraction of features from the images. Second, the introduction of a new image representation called the integral image which reduces the calculations involved in the process. The third component is an efficient classifier which is built using AdaBoost learning algorithm to select the best critical feature. The fourth component is cascading classifiers which helps in the identification of the useful features and discards the background regions of the image which is not necessary for the computation spending more computation time on face-like regions.

Another technique is by using Hidden Deep Learning based Face detection which uses deep neural network for its training and prediction in real time. This is also found in the opencv repository as dnn module, in this case the networks are pre-trained with a huge amount of data with deep learning frameworks. Directly forward pass the input and get the output for the application.

This paper focuses on comparative study of two face detection techniques for Face recognition for non-intrusive attendance system. The organization of the paper is as follows: Section II discusses the literature review in the area of face detection, Section III includes the methodology of the analysis carried out, section IV presents the results of testing, and in section V conclusions are drawn based on the results obtained.

2. LITERATURE REVIEW

Humans attempt to provide the capability of vision systems to machines to understand the outer world and to act according to the input received from these systems gave rise in the rapid developments in the field of computer vision systems. The advancements in this field make computer systems more robust which helped the machines to understand humans.

The major turnover was found through the introduction of Haar feature [1] based Cascade classifier which gave rise to the concept of Integral image, the learning algorithm for the machine was based on Adaboost and to reduce the computation time a method named Cascade classifier was used. This technique served as the base for most of the object detection algorithms and has been extensively used in most of the face detection applications, this work served as the base for many further researches. In [4] the authors have mentioned about the ways to overcome the problems faced in unconstrained scenarios such as video surveillance or images captured by hand held devices. A Normalized Pixel Difference (NPD) was proposed which had the capability of extracting desirable properties such as scale invariance, boundedness and ability to reconstruct the original image. Resulting Face detection template was scaled for multiscale face detection, along with this a deep quadratic tree learning method was used to construct a single soft-cascade AdaBoost classifier to handle complex face manifolds and arbitrary pose and occlusion conditions. In [5], [9] the skin color model along with a improved AdaBoost algorithm was used for the face detection. A very interesting concept Tiny Face detector [12] which was introduced which helped the machines to detect the human faces from huge crowds. In [15] authors have shown that if the pose between the test and reference images match exactly, the best recognition performance was achievable using a particular

face recognition system. They also mentioned that a gradual decrease in recognition performance as the difference in pose between test and reference set increased.

In this paper we are going to use the technique of [1] Haar based cascade classifier and deep learning based face detector which comes with the pre-trained model for face detection.

3. EXPERIMENTAL SETUP

For Face detection creation of dataset is a humongous task. Face detection was carried out for the Application of Non-intrusive attendance marking system. Understanding and analysing testing conditions increases prediction rate and hence the experimental analysis was carried out keeping in mind the following assumptions and the results obtained are with respect to the under mentioned constraints. Moderate lightning conditions was maintained with a medium contrast level, the Resolution was– 72dpi X 72dpi, Background – Plain white background, Dimensions of the Room – 17.5ft X 32ft (Breadth X Length), Height of the Camera was 182.88 cm from the ground when the pictures were captured for the analysis, Distance of Camera from xmin position was 152.4cm and xmax was 640.08 cm, Angle of the Camera –A1 = 39.67 degree, A2 = 88.789 degree. Properties of Camera – Tabulated in Table II, Profile of the Face – Frontal Face, Total of 196 images were captured, Format of the Image – JPG format. The above mentioned parameters were kept constant and the further research was carried out for the non-intrusive based Attendance marking system.

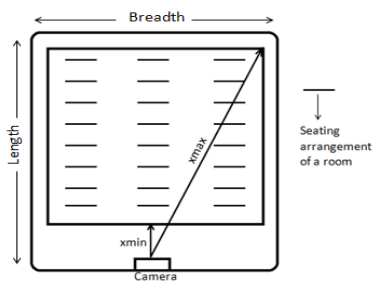


Fig 2: The top view of the camera arrangement

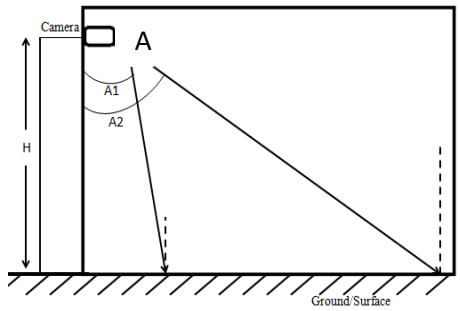


Fig 3: Th side view of the amera arrangement

From figure 2 the top view of the camera arrangement can be known. xmin refers to the minimum distance from the camera to the face which can be detected and xmax refers to the maximum distance from the camera to the face which can be detected. The horizontal line depicts the seating arrangement in a room and the same type of arrangement was maintained when the experimental analysis was carried out. From figure 3 the side view of the camera arrangement is shown. The camera is mounted on the wall at a height H from the ground/surface. A1 refers to the minimum inclined angle of the camera to detect the face at xmin while A2 refers to the maximum inclined angle of the camera to detect the face at xmax. The following formulas were used to calculate the angle of A1 and A2:

$$AC = \sqrt{(H)^2 + (xmin)^2} \quad (1)$$

$$\text{Angle A1} = \sin^{-1}(xmin/AC) \quad (2)$$

$$AD = \sqrt{(H)^2 + (xmax)^2} \quad (3)$$

$$\text{Angle A2} = \sin^{-1}(xmax/AD) \quad (4)$$

The figure 2 and 3 shows different physical parameters of the setup to standardize the overall parameters of the experiment. Parameters considered are: xmin = 152.4cm, xmax = 640.08 cm, H = 182.88 cm, A1 = 39.67 degree, A2 = 88.789 degree. The camera was kept at a height H and inclined at an angle A1 and A2 respectively so as to cover the maximum number of faces present in the room. All the above parameters are a part of the design issue rather than research issue. The details/properties of the images used for the analysis are tabulated below.

TABLE I. PROPERTIES OF THE IMAGE AND CAMERA

Parameters	Camera
Dimensions	4160x 3120

Width	4160 pixels
Height	3120 pixels
Horizontal Resolution	72 dpi
Vertical Resolution	72 dpi
Bit depth	24
Resolution unit	2
Color Representation	sRGB
F-stop	f/2
Exposure time	1/100 sec.
ISO speed	ISO-173
Brightness Range	2.1-2.74
Item type	JPG file
Original Size range (MB)	1.02-1.4

When the face detection was carried out there was no changes made to the image in terms of resolution or image size. The original images were taken to carry out the experimental analysis.

3.1 HAAR FEATURE – BASED CASCADE CLASSIFIER

The first method used was Object detection using Haar feature-based cascade classifier which is an effective object recognition method proposed by Paul Viola and Michael Jones in their paper[1]. The cascade function is trained from a lot of positive images (Image which has face) and negative images (Image which does not have face), it is then used to detect faces in an image (or video). During the training process the features are extracted from the images used for training. For this technique of feature extraction the Haar features technique is used as shown below.

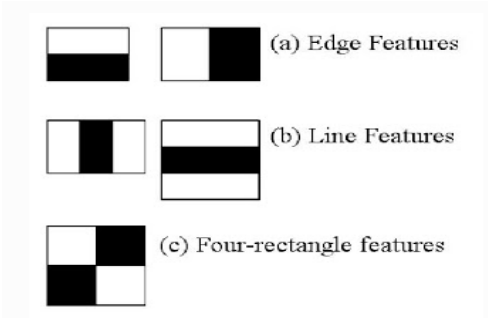
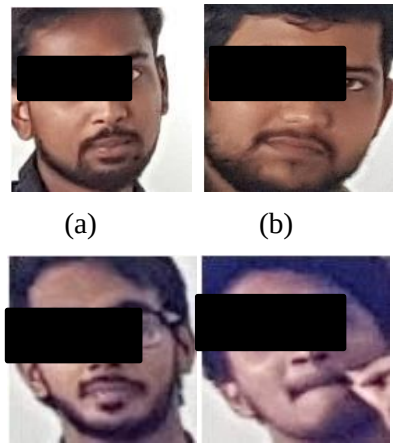


Figure 4: Haar Features, The figure was taken up from https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

Each feature is a single value which is obtained by subtracting the sum of pixels under white rectangle from the sum of pixels under black rectangle. All the possible sizes and locations of each kernel is used to calculate lot of features. When this is analysed practically it was observed that even a simple 24X24 window results over 160000 features. For each feature calculation, user needs to find sum of pixels under white and black rectangles. Later the technique of integral images was introduced which made the calculations of sum of pixels much easier. To select the best suitable features among many the technique of Adaboost was implemented. For each feature, it finds the best threshold which will classify the faces to positive and negative. The errors or misclassifications were observed during this process, but the features with minimum error rate are the features that best classifies the images which is with face and without face. Final classifier is a weighted sum of the weak classifiers. It was called weak because it alone can't classify the image, but together with other forms a strong classifier. The concept of Cascade of Classifiers was introduced to reduce the computation time by eliminating the checking of non face region in an image. For example instead of applying all the 6000 features [17] on a window, group the features into different stages of classifiers and apply one-by-one. Normally the first few stages of classifiers contain very less number of features. If a window fails the first stage, discard it. The remaining features in an image are not considered. But if it passes, apply the second stage of features and continue the process. The window which passes all stages is a face region. According to the authors, on an average, 10 features out of 6000+ are evaluated per sub-window.

OpenCV comes with a trainer as well as detector. User can train a classifier for any object like car, planes etc OpenCV can be used to create one. OpenCV already contains many pre-trained classifiers for face, eyes, smile etc.

Fig 5: User sitting at both xmin and xmax position.



(a) (b)

(c) (d)

Figure 6: Faces detected with a person at xmin (6a) and xmax (6c) position.

In figure 6 the user is sitting at xmin and xmax position as shown in figure 2 and this is the minimum and maximum distance at which the face can be detected by the camera. Figure 6 is the cropping of faces detected in Figure 5. The dimensions of the picture 6a is 276 X 276 pixels, 6b is 170 X 170 pixels, 6c is 86 X 86 pixels, 6d is 99 X 99 pixels.

Overall considering all the images the results were calculated in the form of Success Detection rate (SDR). The formula of the same has been mentioned below:

$$\text{SDR} = \frac{Correct\ Detected\ Faces}{Total\ number\ of\ faces}\ X\ 100 \quad (5)$$

The experimental analysis was carried out for 196 images.

$$\text{SDR} = \frac{1686}{1986}\ X\ 100$$

$$\text{SDR} = 83.54\ \%$$

False Detection Rate (FDR) is calculated by the formula mentioned below:

$$\text{FDR} = \frac{False\ Detection}{Total\ Number\ of\ Detections} X100 \quad (6)$$

$$\text{FDR} = \frac{20}{1686}\ X\ 100$$

$$\text{FDR} = 1.2\%$$

3.2 DEEP LEARNING BASED FACE DETECTOR

Deep learning in OpenCV has a highly improved "deep neural networks (dnn)" module. This module supports a number of deep learning frameworks, including Torch/PyTorch, Tensor Flow, and Caffe. The Caffe-based face detector can be found in the face_detector sub-directory of dnn samples. When using OpenCV's deep neural network module with Caffee models, there were two sets of files which was used. prototxt file which defines the model architecture and the. caffemodel file which contains the weights for the actual layers [18]. Both of these files were used when using models trained using Caffe for deep learning. Caffe model is widely used because of its advanced architecture, its flexibility to switch between CPU and GPU for training purpose.

OpenCV's deep learning face detector is based on the Single Shot Detector (SSD) framework with a ResNet base network [18]. SSD is used for detecting objects in images using deep neural network and this discretizes the output space of bounding boxes. The scores are generated by the network when there is a presence of each object category and produces adjustments to the box for better match of the object shape. The network combines predictions from multiple feature, maps with different resolutions to naturally handle objects of various sizes. The main advantage of using SSD is that it completely eliminates object proposal generation and subsequent pixel or feature resampling and encapsulates all computation in a single network which saves a lot of time, this makes SSD easy to train and straight forward to integrate into systems that requires detection component. It is provides a much better accuracy even with smaller input for 300X300 resolution input a 72.1% mAP on 58FPS in NVIDIA and for 500X500 resolution input a 75.1% mAP faster as compared to R-CNN model.

Initially in coding we import the required packages such as image, prototxt file, and the caffe model file from the appropriate directory. The dnn.blobFromImage takes care of pre-processing which includes setting the blob dimensions and normalization. Here the dimensions are reduced to 300X300 resolution

size since caffe model this spatial size for its computation.

Figure 7: Face detection of person sitting at xmin position while the face of the person sitting at x position was not detected.

Figure 8: The face of the person at xmin position is detected at 95.36% possibility. The dimension of the picture is 256X400 pixels.

The Out of 196 pictures with the person sitting at xmin position the model could predict 10 faces of the person sitting at xmin position.

Figure 9: Failure to detect the face of the person sitting at xmax position as indicated in the figure.

Out of 196 pictures with the person sitting at xmax position the model could not detect any face of the person sitting at xmax position. There were no False Detections observed by the model. Overall the Success Detection Rate for face detection using Deep neural network is calculated as:

$$\text{SDR} = \frac{Correct\ Detected\ Faces}{Total\ number\ of\ faces}\ X\ 100 \quad (7)$$

$$\text{SDR} = \frac{11}{1986}\ X\ 100$$

SDR = 0.55%

4. RESULT ANALYSIS

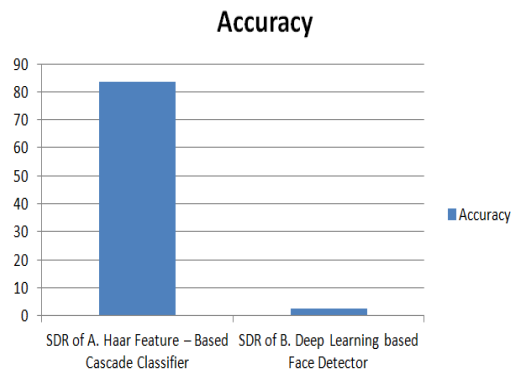


Fig 10: Accuracy Graph

This application involves a non-co-operative subject unlike biometric systems. The size of the cropped images of the detected faces ranges from maximum 276X276 pixels to minimum 99X99 pixels. The maximum and minimum pixel size is of the person sitting at xmin and xmax position respectively. This information serves as the base for the preparation of the training dataset for face recognition tasks.

Based on the analysis carried out, we can infer that between Haar feature-based Cascade classifier and Deep Learning technique, Haar feature-based Cascade classifier gave a very good accuracy in detecting the face for the given set of images with pre-determined constraints mentioned in section III, though with more data which includes the given scenario the deep learning technique might work good for the face detection purpose.

5. CONCLUSION AND FUTURE WORK

This research was carried out to adopt suitable face detection technique to find out the type of training set images required for the face recognition system for Non-intrusive attendance system. The physical parameters as shown in figure 2 and 3 were considered during the analysis.

Our study has also shown that the dimension of the images in training set for face recognition system needs to be varied. According to our study the dimension sizes vary from maximum 276X276 pixels and minimum 99X99 pixels as shown in figure 6. To have a better face recognition system one needs to have a training

set images comprising of a combination of all these dimension sizes.

Our findings in this paper are subject to three limitations. First, the image quality parameters were considered to be independent. Secondly, all the images used in this study were taken from a room whose size has been mentioned in section 3, and Finally these findings are were carried only under the frontal face profile conditions. From the calculations and graphs it was concluded that Haar feature-based Cascade classifier serves as the best face detection technique. We expect that this face detection technique would help us prepare good training set images which in turns helps in building the best model for the Non intrusive based face recognition application.

6. REFERENCES

[1] Paul Viola and Michael Jones "Rapid Object Detection using Boosted Cascade Simple Features" Conference on computer vision and pattern recognition 2001.

[2] Ofualagba Godswill, Omijie Osas, Orobor Anderson, Ibhadode Oseikhuemen and Odiete Etse "Automated Student Attendance Management System using Face Recognition" International Journal of Education Research and Information Science 2018.

[3] Momotaz Begum Panna and Md. Imdadul Islam "Human Face Detection based on Combination of Linear Regression, PCA and Fuzzy C-means Clustering" International Journal of Computer Science and information Security (IJCSIS), Vol 17, No7, July 2019.

[4] Shengcai Liao, Anil K. Jain and Stan Z. Li "A fast and accurate unconstrained Face Detector" DOI 10.1109/TPAMI.2015.2448075, IEEE transactions on Pattern Analysis and Machine Intelligence.

[5] Zhi-Zhen Li, Qing-Hua Zeng, Xiang-Dong Li and Yang-Yu "Face Detection Technology Based on Combining Skin Color Model with Improved Adaboost Algorithm" IEEE 4th International Conference on Signal and Image Processing 2019.

[6] Chen Yan, Zhengqun Wang, Chunlin Xu "Gentle Adaboost Algorithm based on multi-feature fusion for face detection" IET Journals The institution of Engineering and Technology eISSN 2051-3305 February 2019.

[7] Ariel Ruiz-Garcia, Vasile Palade, Ibrahim Almakky and Mark Elshaw "Deep Q-learning for Illumination and Rotation Invariant Face Detection" IJCNN 2019, International Joint Conference on Neural Networks, Budapest Hungary. 14-19 July 2019.

[8] Shuainan Wang, Tong XU, Wei Li and Haifeng Sun "CSSD: Cascade Single Shot Face Detector" IJCNN 2019, International Joint Conference on Neural Networks, Budapest Hungary. 14-19 July 2019.

[9] Zhengqi Xu and Tianfei Shen "Non-Positive Face Detection Based on YCgCr Color Space" 2019-11-14

[10] Mrs. Monali Nitin Chaudhari, Ms. Mrinali Deshmukh, Ms. Gayatri Ramakrishnan and Ms. Rakshita Parolkar "Face Detection using Viola Jones Algorithm and Neural Networks" 2019.

[11] Neethu A, Athi Narayanan S and Kamal Bijlani "People Count Estimation Using Hybird Face Detection Method" IEEE 2016

[12] Chan Zhang, Tao Li, Song Guo, Ning Li, Yingqi Gao and Kai Wang "Aggregation Connection Network for Tiny Face Detection" IJCNN 2019, International Joint Conference on Neural Networks, Budapest, Hungary July 2019.

[13] Chien-Yu Chen, Jian- Juin Ding, Hung-Wei Hsu and Yih-Cherng Lee "Advanced Orientation Robust Face Detection Algorithm Using Prominent Features and Hybrid Learning Techniques" 2018 IEEE.

[14] Rong Qi, Rui-Sheng Jia, Qi-Chao Mao, Hong-Mei Sun, Ling-Qun Zuo "Face Detection method based on Cascaded Convolutional Networks" DOI 10.1109/ACCESS.2019.2934563, IEEE Access.

[15] Abhishek Datta, Raymond Veldhuis, Luuk Spreeuwers "The impact of image quality on Face Recognition"

[16] Rajeev Ranjan, Ankan Bansal, Jingxiao Zheng, Hongyu Xu, Joshua Gleason, Boyu Lu, ANirudh Nanduri, Jun-Cheng Chen, Carlos D, Castillo, Rama Chellappa "A fast and accurate system for Face Detection, Identification and Verification" DOI 10.1109/TBIOM.2019.2908436, IEEE Journal of LATEX class files, VOL. 14, NO. 8, August 2015

[17] https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

[18] https://www.pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/