International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

Real-Time Face Recognition Using an Optimized Neural Network

Ferdous Irtiaz Khan
Dept. of EDA, University of Kent
Canterbury, UK

Submitted: 05-02-2022 Revised: 18-02-2022 Accepted: 20-02-2022

ABSTRACT:For security and automation face recognition is a very important aspect and needs efficient results. Till now many techniques had been developed to accurately identify the correct faces. In this paper Viola-Jones algorithm is combined with the neural network method i.eAlexNet and a particular thresholding method is proposed for real time face recognition. The proposed method accurately detects faces from a live camera produced identification accuracy of 95%. The quite simple and efficient methodology can be a solution for automated real time security system.

KEYWORDS: AlexNet, Neural Network, Real-Time Face Recognition, Viola-Jones Algorithm.

I. INTRODUCTION

Facial recognition is the phenomena which typically involves feature extraction, feature reduction and recognition or classification. For security and validity checks, face recognition is widely used nowadays. Generally, the procedure followed for facial recognition starts with detection of important geometry of faces such as eyes, nose, mouths, etc. Then they are transferred to a database and stored in devices (PC, memory card, microcontroller, etc.), where they are processed, and the unique features are later used for recognition.

Numerous techniques have been developed till now to carry out facial recognition. The initial technique involves detection of faces and then recognition of the accurate face. Ban et al. [1] proposed a technique of face recognition based on skin color. During feature extraction skin color is taken under considerations and histogram built with the available information. In order to do so the RGB space is converted to YCbCr space which contains information of luminance and chrominance and will be a distinct feature for individual person. In [2], a combination of histogram of gradient (HOG) and congenital pulmonary airway malformation (CPAM) is used to recognize faces. With HOG local object appearance and edge is extracted and with CPAM human color vision is detected. In another work the energy function of face color is used track faces [3]. The energy function of color is a mixture of external and internal forces and the internal ones come in the shape of snakes. These individual shapes may be utilized as unique identification of faces. Tsai [4] suggested a scheme which uses Eigenfaces to detect faces and neural network to examine the faces. Among all the techniques developed so far Viola-Jones algorithm is widely used because of its high detection rate and ability to run in real time[5][6]. Its mechanism is described in detail in the next section.

After detection of faces and extraction of features, generally machine learning methods or deep learning methods are used in the modern systems for efficient face recognition. Machine learning technologies basically learn from given set of data and make decision based on that. On the other hand, deep learning methodology can train itself from the given data and make decision on itself, enabling it to improve its performance with the help of neural network. The advancement of pre-trained neural network such as AlexNet, GoogLeNet, etc. have increased efficiency and made life easier. Experimental results have shown that AlexNet performs better than generalized convolutional neural network (CNN) and GoogleNet[7].

In this paper, for the detection of faces, Viola-Jones algorithm is used and for recognition purpose, an improved AlexNet method is considered. A combination of both helped facial recognition to be efficient, fast and bear high accuracy.

The paper is organized as follows: after introduction,Viola-Jones algorithm and theory behind Alex Net is discussed. In part III the proposed methodology is explained. In part IV, the results and analysis are provided. In last part the conclusion is given.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

II. VIOLA-JONES ALGORITHM&ALEXNET

The model was proposed by Paul Viola and Michael Jones in 2001 which was initially tested to detect objects. It is now widely used to detect faces. As shown by Figure01, it follows 4 steps namely: Haar feature extraction, integral image creation, adaboost training and cascading classifierfor detection of face.

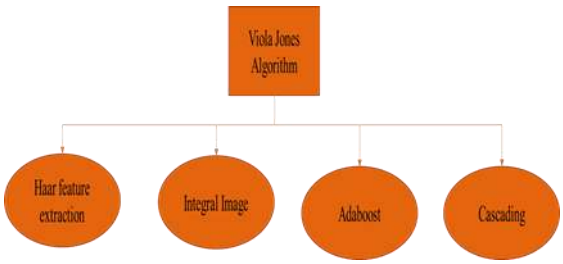

Figure 1: Steps of Viola-Jones Algorithm

A. Haar Feature Extraction

The process is in close approximation to convolution kernel. It generates different rectangles having white and black regions. The pixels of the individual regions (black/white) are added up and then subtracted from each other to give the result.

Figure 2[8] shows how Haar feature extraction works. Rectangles that resemble important structures such as eyes, nose, mouth, etc. are passed through the entire face or image. At the overlapping point of the certain features the subtraction of the white and black pixels will result in maximum value.

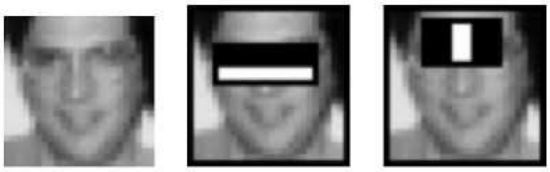

Figure 2: Haar feature extraction [8]

B. Integral Image

Integral Image is a process that help to reduce calculation and as a result the computation time is faster. In this method the value of a pixel at a point (x,y) is the sum of all the pixels above and to the left of (x,y).

In figure 3, it can be seen that by applying integral image method the sum of all the pixels at region 'D' can be calculated just by adding the diagonal pixels (1 and 4) at the edge of D and subtracting the other diagonal elements (2 and 3).

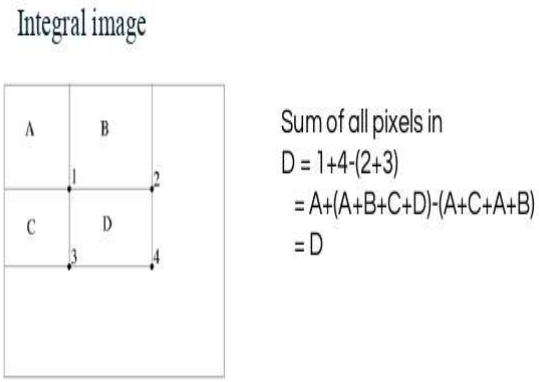

Figure 3: Integral Image Process

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

C. Adaboost

The Adaboost method helps to get rid of redundant features. It distinguishes relevant and irrelevant features. The method is better than random guessing as it sums up all the weak classifiers multiplied with a corresponding weight and forms a strong classifier.

Adaboost is used to form a strong classifier from weak classifiers. For example, if there are a number of weak classifiers such as $f_1(x), f_2(x), f_3(x)$, etc. Then a strong classifier $F(x)$ can be produced by multiplying with $\alpha_1, \alpha_2, \alpha_3$, etc. As shown by the equation (1):

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \cdots \quad (1)$$

D. Cascading

The features are split into different stages. Each stage is checked at a time and if it proves that the input is definitely not valid then further stages are not checked and the process terminates.

Figure 4 depicts how cascading is done. The features are divided into stages and input checked sequentially. If at any stage the results are mismatch it rejects the input. This method saves processing time.

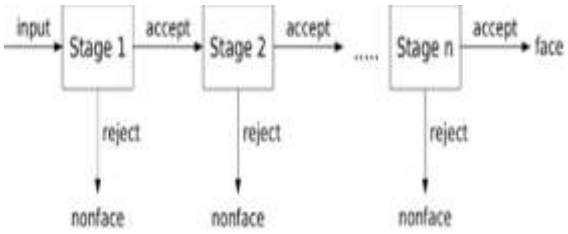

Figure 4: Cascading Method

E. AlexNet

The model was first proposed by Alex [9] and won the 2010 ImageNet competition with a wide range. The AlexNet architecture is a convolutional neural network (CNN) that consists of 8 deep layers. There are five convolution layers and three fully connected layers. The characteristics that make AlexNet special are: Rectified Linear Units (ReLU), multiple GPUs and overlapping pooling. ReLU is much faster than the conventional tanh function. Multiple GPUs enable larger model to be trained and save computation time.

Figure 5 [10]shows the AlexNet architecture. It takes on 227*227 sized color pictures. The output of the final softmax layer has 1000 classifications. Thus, it is able to classify 1000 different objects or pictures. The 5 convolution layers and 3 fully connected layers are also seen. The second, fourth- and fifth-convolution layer kernels are connected to the layers of the same GPU, whereas the third layer ones are connected to the second layer.

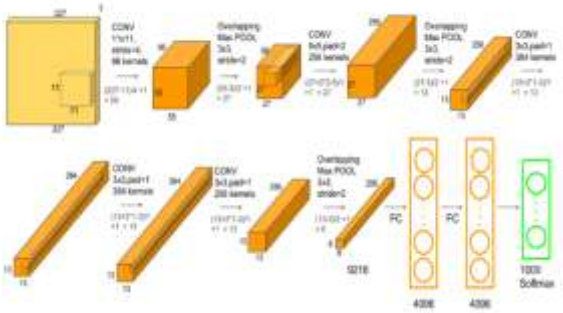

Figure 5: AlexNetArchitecture[10]

III. METHODOLOGY

The process flow chart of figure 6 shows how the image processing technique has been utilized in this paper for optimized face recognition.

Initially 'N' samples of images of each people who are to be counted as valid is captured and is the system is trained using AlexNet. Images are captured at different angles, weather conditions

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

and light effects to improve the performance of the deep learning network.

While validating results the live video camera then captures images of the person who comes in front of it. the video camera checks the frames from it capture. The initial step is the face detection part and if no face is detected the camera keeps on scanning. When the face is detected from its frame with the help of Viola-Jones algorithm, the face image is captured. Next, the features are extracted from the captured image and passed on to the system for comparison with the pre-stored verified images. The system is trained beforehand with all the verified images of the people who are allowed to enter by the help of AlexNet. The data from the trained set of images are compared with the new image to provide a decision. If the data matches, a counter adds up.

The process is repeated until 'N' number of images, that is the same number of images that the AlexNet is trained with, is captured and compared. Considering system error and thresholding at least half of the images need to match for the system to declare the person as 'Verified'. Else it is counted as an 'invalid' person.This thresholding is controlled by the user. Increasing the thresholding will increase accuracy but system operation time may be sacrificed.

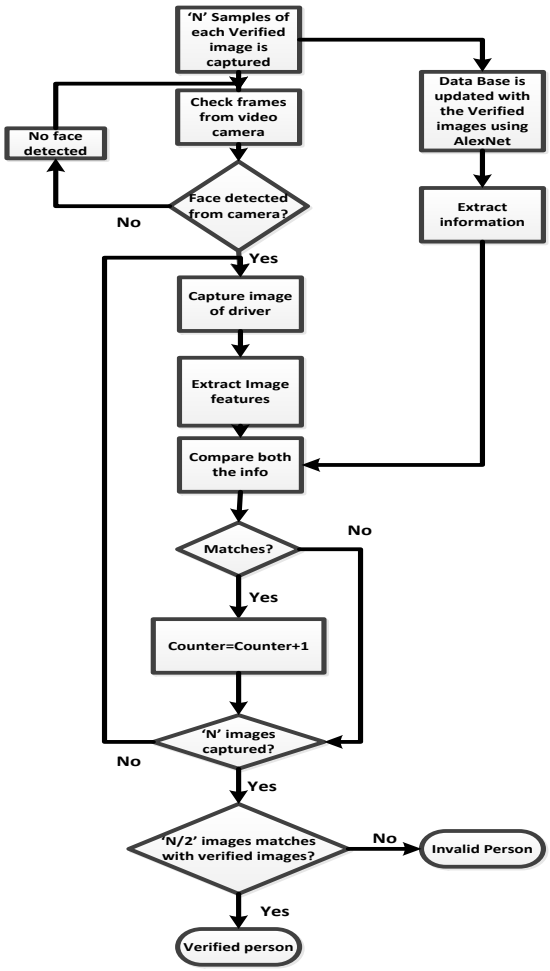

Figure 6: Flowchart of Proposed Methodology

IV. RESULTS AND ANALYSIS

Matlab R2020b has been used in this project for face recognition coding.The face recognition coding is divided into three sections:

a. Face Data collection
b. Model Training
c. Real time face recognition

a. Face Data collection

Two codes are developed for face data collection. The first one captures real images when the face is placed in front of the camera. Multiple images of the face are captured according to the wish of the user and processed. AlexNet requires the image size to be 227 by 227, so the images are

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

automatically cropped to the mentioned size. The second code is designed to facilitate data storage if some verified images are already present in the system. They are resized and only the facial portion is extracted for accurate training by AlexNet.

b. Model Training

The AlexNet is modelled with 4 individual images as verified person. The faces of Figure 7 depict those verified images. For each classification the model is trained with 10 images of a single person. A mixture of bright, dim, reflected, different exposure pictures are used for training so that even at adverse condition the model can recognize faces with accuracy. Figure 7 shows some of the samples of those images. Some of the images are collected from the dataset [11] and some used with permission from the individual.

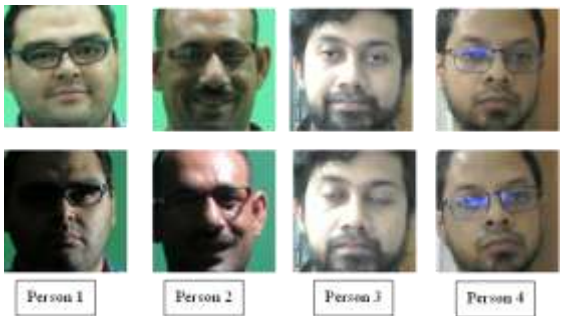

Figure 7: Images of 4 person that the Alexnet is trained with

c. Real time face recognition

Finally, for the real time recognition the code is developed to recognize verified and not verified faces. The code detects faces from the frame using Viola-Jones algorithm and then checks its feature with the stored dataset of the AlexNet. It gives its verdict based on it.

D. Testing of proposed methodology

According to the process flow, after storing of verified images the camera starts to detect faces. By the help of Viola-Jones algorithm the face is detected at every frame and after detection of face the face is recognized from the dataset already present in the AlexNet The process is repeated for 'N' number of samples which is 10 for this case. If the face is a valid result(N/2 samples matches), signals are forwarded to the system, else it is counted as intruder.

The following valid person (4) and invalid person (4) as given in Figure 8 are checked for verification purpose.

Figure 8: Dataset of valid and invalid tested sample

Initially the code was checked to work for real-time. Some of the results are given in Figure 9, the camera is able to detect the faces really quick and in real time. It can detect the verified images and non-verified images accurately. Samples were tested with real face in front of the camera and with pictures from mobile phone display. Even when there are no faces in front of the camera it is able to determine no faces are present. A box is created when a face is detected and on top of the frame it is mentioned whether the face is verified/intruder/no face.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

Figure 9: Real time face verification

One of the important parameters of the project is to work the facial recognition efficiently and with great accuracy under adverse condition. These factors are very much necessary because for automatic gate control security is a major concern. The matlab code is updated by introducing multiple samples verification method. In multiple sample verification method whenever a face is detected in front of the camera, it will capture 10 face snapshots. It then validates these 10 snapshots with the dataset provided by AlexNet. It is assumed that some of the snapshots might be erroneous and so if 50% i.e 5 of the snapshots are valid then the person is deemed a valid person. The opposite is for invalid person. Although the process time is sacrificed a little bit but accuracy is emphasized in terms of security.

In figure 10, it isshown how the test is carried out. The 8 samples (4 valid+4 invalid) are tested by varying the distance and also by continuously moving the camera or sample. Blurred images, very bright images and test at dark environment is carried out for further verification of results. Each of the samples are tested 10 times to check the efficiency of the proposed method. Figure 11& 12 shows the bar diagram of the accuracy check for valid and invalid samples. In Figure 12, it is observed that only one of the samples gives 90% accuracy whereas all the other results have 100% accuracy. From figure 13, it is visible that the accuracy are as follows: 90%, 80%, 100% and 100%. The overall accuracy of the test system came to be 95%, which is a very satisfying result.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 2 Feb 2022, pp: 1071-1078 www.ijaem.net ISSN: 2395-5252

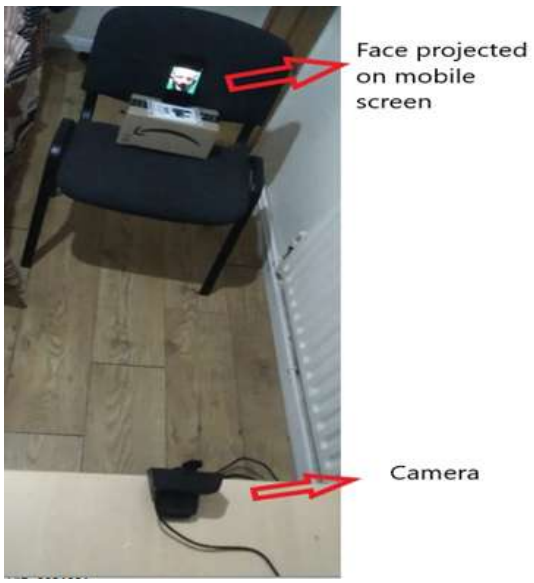

Figure 10: Real time testing of face recognition

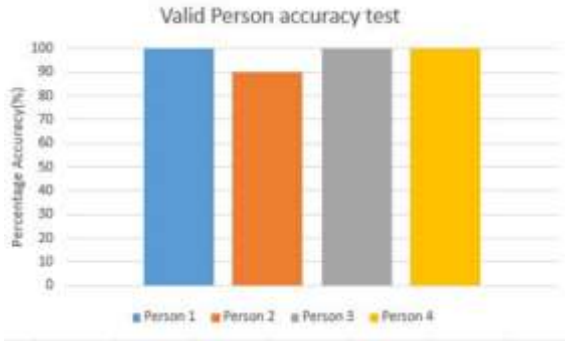

Figure 11: Accuracy of Valid samples

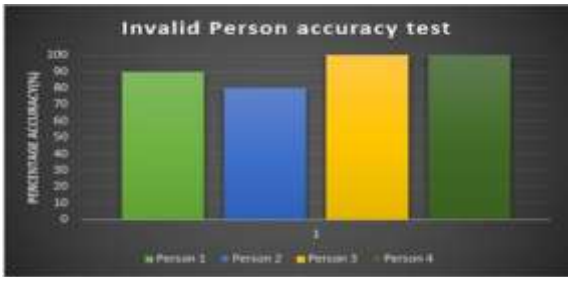

Figure 12: Accuracy of Invalid person samples

V. DISCUSSION

The modern world requires a more secured and automated system. This paper presents a Real Time Face Recognition system by utilizing Viola-Jones algorithm, ALexNet and an optimized identification process.

The proposed optimized technique is tested for different types of samples and at different conditions like dark, bright, blurred environment and by varying the distances. The proposed method is a very simple one but provides an accuracy of around 95%. The added advantage is that the user can alter its efficiency or can make it faster by changing the threshold.

REFERENCES

[1] Y. Ban, S. K. Kim, S. Kim, K. A. Toh, and S. Lee, "Face detection based on skin color likelihood," Pattern Recognit., vol. 47, no. 4, pp. 1573–1585, 2014, doi: 10.1016/j.patcog.2013.11.005.

[2] I. J. Jacob and P. Betty, "Face Recognition system using Local Texture and Spatial Colour Layout," J. Xidian Univ., vol. 14, no. 4, pp. 5208–5214, 2020, doi: 10.37896/jxu14.4/107.

[3] K. H. Seo, W. Kim, C. Oh, and J. J. Lee, "Face detection and facial feature extraction using color snake," IEEE Int. Symp. Ind. Electron., vol. 2, no. January 2015, pp. 457–462, 2002, doi: 10.1109/isie.2002.1026332.

[4] C. C. Tsai, W. C. Cheng, J. S. Taur, and C. W. Tao, "Face detection using eigenface and neural network," Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern., vol. 5, pp. 4343–4347, 2006, doi: 10.1109/ICSMC.2006.384817.

[5] G. Saadouli, M. I. Elburdani, R. M. Al-Qatouni, S. Kunhoth, and S. Al-Maadeed, "Automatic and Secure Electronic Gate System Using Fusion of License Plate, Car Make Recognition and Face Detection," 2020 IEEE Int. Conf. Informatics, IoT, Enabling Technol. ICIoT 2020, pp. 79–84, 2020, doi: 10.1109/ICIoT48696.2020.9089615.

[6] P. Viola and Michael Jones, "Robust Real-Time Face Dtection," Int. J. Comput. Vis., vol. 40, no. 2, pp. 50–62, 2004, doi: 10.1023/B.

[7] N. A. B. M. Kasim, N. H. B. A. Rahman, Z. Ibrahim, and N. N. A. Mangshor, "Celebrity face recognition using deep learning," Indones. J. Electr. Eng. Comput. Sci., vol. 12, no. 2, pp. 476–481, 2018, doi: 10.11591/ijeecs.v12.i2.pp476-481.

[8] P. Viola, "The Viola / Jones Face Detector Classifier is Learned from Labeled Data," Procedia Comput. Sci., pp. 22–23, 2001.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.

[10] M. ul Hassan, "AlexNet - ImageNet Classification with Convolutional Neural Networks." https://neurohive.io/en/popular-networks/alexnet-imagenet-classification-with-deep-convolutional-neural-networks/ (accessed Feb. 15, 2021).

[11] M. Afifi and A. Abdelhamed, "AFIF4: Deep gender classification based on AdaBoost-based fusion of isolated facial features and foggy faces," J. Vis. Commun. Image Represent., vol. 62, pp. 77–86, Jul. 2019, doi: 10.1016/j.jvcir.2019.05.001.