REAL TIME FACE DETECTION AND RECOGNITION SYSTEM USING HAAR CASCADE CLASSIFIER AND NEURAL NETWORKS

Anirban Chakraborty[1], Shilpa Sharma [2]

[1] *Student,Department Of Artificial Intelligence, Lovely Professional University, Punjab, India.*
[2]*Professor, Department Of Computer Science And Engineering, Lovely Professional University, Punjab, India*
[1]*anirbanchakraborty456@gmail.com* [2]*shilpa.sharma@lpu.co.in*

Abstract: Home protection and privacy have become one of the most critical aspects in today's world. As technology progresses at an exponential pace, the times are not far ahead for each house to be fitted with sophisticated security systems to deal with regular burglary and theft. But as one side of the tech progresses, so do its detrimental counterparts. DES encryption can be an indicator of how easily an encrypted piece of information can be deciphered. Not long after its release, DES encryption was referred to as 'unsafe' and with today's modern application, anything like DES might be an open invitation to hack. With many developments in the field, the technology has, in many respects, surpassed the use of biometrics (finger prints). Face recognition, nowadays, is present in almost every smart device that has some piece of information stored that holds importance to its users. With facial recognition gaining popularity, many tech companies have come with their own patent to make a technology related to Facial Recognition on the market. This paper suggests a somewhat related concept as to how home protection can be improved by using a face detection and recognition algorithm (Haar Cascade Classifier).

Keyword: Face Recognition, Haar Cascade, KNN Algorithm, LBP.

Introduction

The method of facial recognition is not a modern area of technology, it dates back to the early 1960s when Woody Bledsoe, along with Helen Chan and Charles Bisson, worked on a device to identify human faces.

Face Recognition is a process in which the device first recognizes the presence of a face in the picture. When it is detected, the later stages help to find a match for the picture associated with those already present in the database. A traditional face detector doesn't get to the depth of the picture, it just checks for the similarities. This can also prove fatal, since it would take a lot of work to circumvent this kind of device. Face recognition systems may be required as human biometrics (fingerprints) begin to fade away after a certain age.

This project takes into account the current safety system scenario and proposes a system that can be detected and defined with greater precision and time. Every day, the technology industry is gaining new heights with ML and AI. The secret to any security system accuracy is to introduce the ML-based algorithm and to train the system accordingly, we will be able to reduce the chances of having false positives and negatives.

Now, while most face recognition systems use the LBP (Local Binary Pattern) algorithm, our system operates on the Haar Cascade Classifier. Instead of generating some graphical information from the image, such as LBP, the Haar Cascade Classifier looks directly at the related facial features based on rectangular blocks.

OBJECTIVES

i. To verify faces of the photos or videos taken with the existing photos in database.
ii. To enable 3D Face security.
iii. To ensure safety.
iv. To extract the Face properly.
v. To extract feature of Face.
vi. To unlock a secure system with proper recognition of face.

Background

Previous Work

Current methods in the field of image recognition work either by using LBP as a primary algorithm or by using algorithms that include the use of colour segmentation or matching models for facial recognition and Eigen-Fischer for facial recognition [IV], [VI]. These algorithms have been compiled using MATLAB and then implemented on a droid phone.

Skin-based facial detection is very popular and a useful technique. In this process, each pixel is defined as either a skin or a non-skin based on its colour composition. Although this approach works well, it has some flaws. In order to minimise the time of computation, the image is always scaled down by a

few factors. This could have an effect on overall results. This approach still uses an RGB classifier that is very sluggish compared to its new generation counterparts.

Related Work

Our work is inspired by the Viola-Jones Object Detection System [X], which analyses the picture in order to differentiate a face from a non-face. In this context, the object detection system uses the AdaBoost algorithm to pick the best features and train the classifiers to be used. The algorithm uses a linear combination of "weakly-approved classifiers" to create a classifier which is a good acceptable classifier.

$$h(\mathbf{x}) = \text{sgn}\left(\sum_{j=1}^{M} \alpha_j h_j(\mathbf{x})\right)$$

Each weak classifier is a threshold function based on the feature fj.

$$h_j(\mathbf{x}) = \begin{cases} -s_j & \text{if } f_j < \theta_j \\ s_j & \text{otherwise} \end{cases}$$

Based on the algorithm (Viola-Jones)[X] described above, additional references include Js Zhang and X. Xiao and their face Recognition algorithm based on multi-layer weighted LBP [VI]. In addition, references used in our work give priority to the Haar Cascade as opposed to the LBP [IX][X] method, which lacks the accuracy needed for the protection framework. The home safety feature of our project is influenced by N. A. Othman and I are. Aydin supports improved security work in smart homes and cities using IoT (Internet of Things)[VII]. Here, multiple modules such as raspberry pi, etc. are used for image capture and analysis. Later, when a match is found, a notification is sent to the respective user via a telegram application.

Another piece of work suggested by D. Hey, Pertsau and A. Uvarov is a more sophisticated variant of the Haar[II] algorithm. Their framework modifies the GPU OpenCV algorithm. This (GPU programming) can be done using the CUDA (Compute Unified Device Architecture). The modified algorithm makes the device not only to process HD streams,

In real time, quicker, but also lowers the total Processor load. Other similar works are N. It's Erdogmus and S. Marcel retrieves paper on how a face recognition device can be spoofed by using high-grade 3D face masks.

The related work referred to above uses the Hair Cascade Classifier for identification and recognition of the face, but lacks any module that supports the ability to recognise the face[II],[III],[V]. It is also mandatory for the protection system to function, both in terms of identification and recognition, which is available in our proposed system.

Literature Survey

Some essential literature reviews are as follows:

V Latham et al. suggested the latest successful face search findings of Multiple CovNets or Deep CovNets. Recent findings, as shown by Yi Sun et.al, have shown that the methods that usually exist address the FR issue in two phases: extraction of the properties (for a better picture we must build or learn from each face of an image picture) and identification (calculates identical characteristics between the two face with a rep). Although CovNets shows positive FR performance, a well-designed architecture in CovNet remains unclear, since a specific classification project lacks theoretical guidance.

Cox et al. reported that Brunelli and Poggio have determined geometrical characteristics such as mouth width and position, noise condition face size and chin form. They reported a recognition rate of 90% in a 47-person survey. However, we find that 100 per cent identification of the identical dataset is created by a rather simple method matching framework. The interrogatory database of 95 pictures of 30 manually extracted functions describing each face observed the mixing distance technique introduced by Mr Cox.

Pentland et al. defines database (95 percent of 200 of the 3000 recognition) which incorporates best performance. Breakdown results are complicated since several images of the people appeared to be the same. The closest stored graph using elastic graph matching, applied a dynamic link architecture to define the distortive, invariable entity. With a sample of 87 individuals and 150 years old research images of different expressions and faces, we have seen strong findings. A 23-transuter parallel method uses a computer-cost methodology that takes around 25 years to compare 87 saved items. Eigen faces are thus an algorithm that is rapid, easy and functional. The pixel intensity for training and test images may therefore be decreased, because a high correlation is needed for the optimal performance. Another means of identifying the face is by visuals.

Wikott et al. used better methods for the Facial Recognition Technology (FERET) database and contrasted 300 of the same men's sides against three hundred separate sides. You reported a consciousness score of 97.3 percent. In restricted conditions for example local binary models and local step quantization, handmade findings have been respectable in the FR. Furthermore the efficiency is significantly reduced when used in pictures taken under uncontrolled conditions such as multiple face postures, voice and lighting.

Turkaand and Pentland said that there is also a 2-dimention method for image processing. They suggested a way to classify the face portraits by projecting the principal elements of the initial trainers' photos. The resulting Eigen faces are ranked, as opposed to recognized individuals. High-level identification is usually based on multiple processing layers, for example the Marr processing framework for fitted objects from images to surfaces to 3D versions.

V Nagpal et al. suggested such approaches used by deep neural network analysis only for the features derived by the FR. Their paper suggests that pixel values be moved to CovNet by utilizing a haircascade (frontier face) as an input to the detection of the neonatal network.

Thet face identification is better achieved by Tom Holland et al. who has a range of measurements for any section of face. hair cascades feature. All features in their report were also applied and their accuracy rates were 96 percent.

Danial et al. reported that Brunelli and Poggio were evaluating physical characteristics such as mouth width and position, noise location facial size and chin form. They have a detection rate of 96 percent in a 69-person survey. However, we find

that 100% identification for the like data set is focused on a very simple matching framework. The interrogatory database of 95 pictures of 30 manually extracted functions describing each face observed the mixing distance technique introduced by Cox et al.

Christopher et al. said face detection improves the protection of a project tremendously as the face detection functionality is quite difficult to compromise. Haar Cascade Detection provides new facial detection protection that is not broken yet.

Title	Author	Proposed Methods	
A overview of Face Detection Strategies	**V Latham *et al.***	1.	Haar Cascade
		2.	Frontal Face
		3.	LPBH
Ultimate Face Detections And Recognitions	**Cox *et al.***	1.	Eigen Face
		2.	Fisher Faces
		3.	Haar Cascade
A overview of Face Detections Techniques	**Pentland *et al.***	1.	Eigen Face
		2.	Haar Cascade
Face Detection advance methods	**Wikott *et al***	1.	LPBH
		2.	Fisher Faces
		3.	Haar Cascade
Haar Cascading Face Detections	**V Nagpal *et al.***	1.	Haar Cascade
		2.	Frontal Eyes
Frontals Face Detection Algorithms	**Tom Holland *et al.***	1.	Texture Method
		2.	Haar CAscade

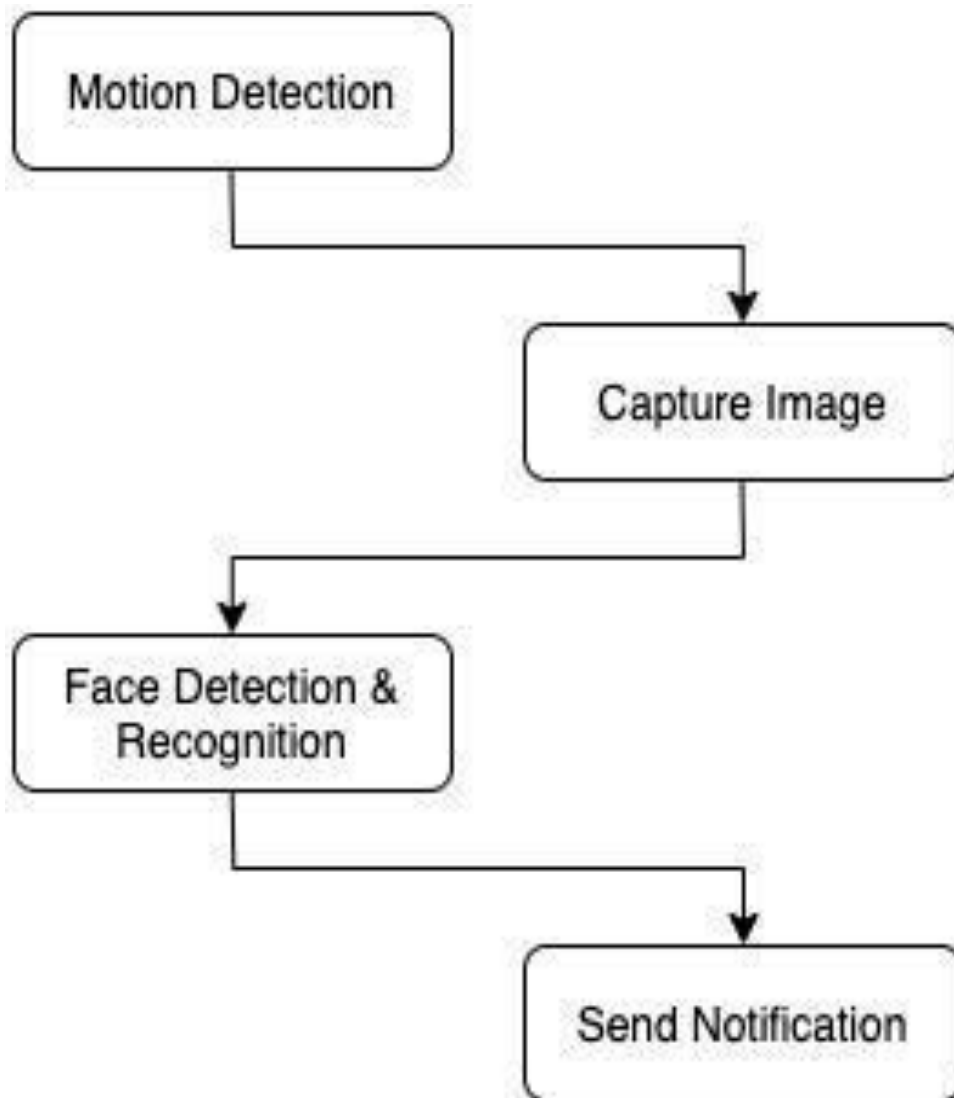| Haar Cascading Face Recognition process and algorithms	**Danial** *et al.*	1. Haar Cascade 2. LPBH 3. Fisher Faces
A overview of Face Detections Algorithms	**Christopher** *et al.*	1. Image Segmentation 2. Haar Cascade
Facials Features analysis and studies	**Brendon** *et al.*	1. Edge Detection 2. Haar Cascade

Figure 1. Module Flow Diagram

There are a total of four modules in which this method can be divided. Each of them is equally relevant to the other. Motion detection module is correlated with any motion in the vicinity (by human or any other object). The entire device remains idle when no motion is observed in the vicinity. As soon as the PIR sensor senses motion, the camera begins operating with other modules. The Image Capture Module deals with the acquisition of photographs using a camera module. The picture

227

is then submitted for processing by additional modules. First, the picture will be examined for the faces. The module will not proceed any further if there are no faces present in the picture. The recognizer module comes into play after the face has been detected successfully. The recognition module uses the KNN (K-nearest-neighbor) algorithm to carry out the recognition[I]. After a good match, the notification module will send a message to the administrator indicating that an entry has been made. This is done to counter any suspicious entry. The notification module can also request a PIN authentication if the recognizer is not confident that an entry request is being made. Fig.1 displays the flow diagram of the modules. The diagram shows the hierarchical flow of the modules proposed by our method.

Detecting the Motion

To detect the presence of any nearby object, we are trying to use the PIR sensor (passive infrared sensor). This sensor is capable of detecting any human activity between 5m and 12m. This is required to determine whether a person has entered or left the field. This module triggers the machine to start the process of image capture when some motion is detected
Capturing the Image

The camera module present in the device captures the image of the object identified by the PIR sensor. Depending on the distance of the object from the camera, the lighting conditions and the quality of the picture obtained from previous attempts, the camera can take more than one photo.
Detecting and Recognizing Face

The machine learning module is now called into action to detect and identify the face   present in   the image from the camera module. The Hair algorithm is used to detect a face. The Hair Cascade Classifier extracts unique features

from the image, taking into account positive (images of faces) and negative images (images without faces). Each function (Edge, Line and Four-Rectangle) is a single value obtained by subtracting the sum of the pixels. This measures all of the important features in the picture and discards the irrelevant features.

The KNN algorithm (K-nearest-neighbor) is used for face recognition. The KNN does not make its own decisions, the way it works depends entirely on the training of the machine. It is a memory-based algorithm that learns by storing the representation of training examples.
Sending Notification

This last module analyses the outcome and sends a notification to the administrator. Either access has been given or refused, it shall be reported to the administrator. This means that every unauthorised person enters the premises. If the system partially recognises an individual, it will automatically send a message asking for a PIN authentication.

Approach

Once a face is identified, the next step will be to decide whether or not the human ID (image) at the door is registered. This step is the Face Recognition Step, where the data already uploaded by the approved users is used to enable their entry. The product of the Face Detection Module is an xml file containing a list of human features at the entrance. The newly generated xml file is compared to the current xml file of the approved users. The two files are compared to decide whether or not the newly generated file data is already present in the system. The KNN (K-Nearest Neighbor) algorithm is used to compare the features.
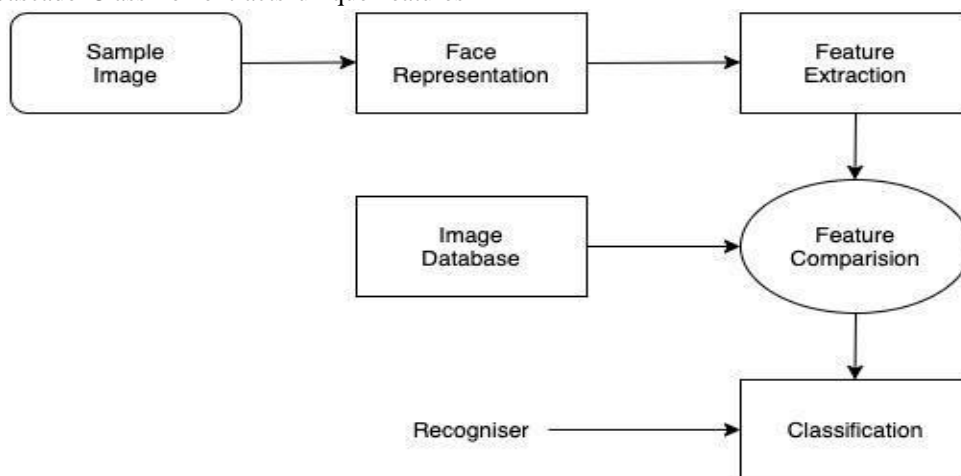

Figure 2. Image Identification and Recognition

KNN is used to detect related facial features in this situation. A vector representation of the features is generated and then

compared using a distance metric (Euclidean Distance). If similar features are found with the utmost expertise, it indicates that the person at the door is an approved person who

has access to the door. Else, the human being is classified as an approved individual who does not have access to the door and is identified as a visitor.

If the person at the door is an approved staff member, the door will be opened for him and the admin will be informed of his entry into the building. Furthermore, his ID will be disabled once he is inside the building. This means that his ID (face features) will be locked down before he stays inside the building.

If the person at the door is identified as a visitor by the system, the admission of the person depends on the action taken by the administrator when prompted by the system. If the admin allows its entry, the guest framework will create a temporary ID. The temporary ID will remain locked until it is inside the building. The temporary ID will be removed from the database as soon as the visitor leaves the house. The admin can also have a time bound, which restricts the device from removing a temporary ID and preserves it until the time is up.

An intellectual algorithm focused on a neural network designed to identify patterns in different datasets. For conversion of neural networking machine vectors, real life details like images, text sounds, pictures etc. are used. A multi-layer neural network consists of several nodes in each stratum. Depending on the model type, the neural network attempts to assess the weight of all data in a node. The weight defines the importance of the input results. The weighted value is determined and the output of the node is measured according to such threshold biases.

Certain active functions for triggering the output and input are used. Every f-function can be rendered complicated by a neural network. A simple function classifying,

$$z = f(b) \quad ....(1)$$

is to classify the inputedadatum x in class y, while the neuralsanets classify a default parametera, aβ and thus
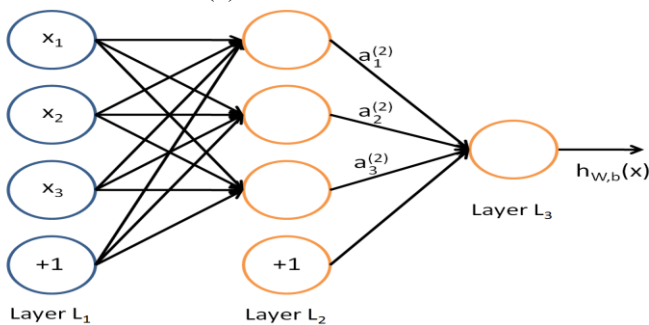
$$z = f(a\beta) \quad ....(2)$$

Figure 3. Small Neural Networks

This can be expressed by the network of these functions:

$$f = f(x)f2(x)f6(x)f3(y)f7(z) \quad .....(3)$$

The first layer in the row is called f1 and the second layer is called f2 etc. The neural network depth is defined by length by chain. The finishing layer is classified as the output layer. Fig2 displays the picture of the neural network. During the exercise, the output of the target layer is not apparent such that the main layer is called the secret layer. The DNN is the artificial neural network's multi-hidden (ANN) layer and a more abstract feeding method.

Algorithm

Based on the research, we proposed an algorithm for Face Recognition.

Before applying the algorithm, we need to follow these steps.

STEP 1: 1024 outputs with reluaactive function and dropoutawith 0.6 are provided by the 1st layer that is densed layer.

STEP 2: 1024 outputs with reluaactive function and dropouta with 0.6 are provded by the 2nd layer that is densed layer.

STEP 3: 600 outputs with reluaactive function and dropouta with 0.6 are provided by the 3rd layer that is densed layer.

STEP 4: 18 outputs with soften maxaactive function and dropout with 0.6 are provided by 4th layer that is output layer is densed layer.

Pixel dimensions are uploaded in database

We have to upload all the measurements of dimensions in pixels to out Facial Database.

Using HAAR CASCADE, recognize facial dimensions of all image.

Recognization will be done based upon Haar Cascade functions which will indicate all the parts of face such as Nose, Eyes, Mouth.

Croping the facial part from the image.

It is very much important to crop only thr Face portion from the entire image before detection.

Break the data in the ratio of 9.6 :1 and cross authenticate it.

It will be helpful if we segment the facial portion in the proportion of 9.6:1 which will further gives more segmentation.

Train the Neural Networks with the measurement of e-poch= 60.

Neural Nets will work properly only when te developers will train it with all the basic and advanced features.

Using test and train outputs, plotting the graphs.

After training Neural Nets, we need to perform all the testing purposes and based on that we need to put all the graphs of a 3-Dimensional axis.

Calculate the modded average(mean).
We have to calculate the mod of all the values and then take the average which will give the most appropriate result.

Write down the final result.
Finally, we have to note down the result which will be almost accurate and plot it in histogram.

Results

To date, the images used to validate the proposed system have produced 2% more hit rates than the current LBP systems. The overall result could be changed progressively depending on the number of images trained. The proposed system needs high-end PC rigs to produce a better outcome as it will gradually increase the computational speed resulting in faster image training. This is perhaps the most prolific downside in this method as a typical home-based PC lacks the processing power needed to train images for the required data collection. A line graph comparing the two (LBP and Haar) algorithms was shown. The graph clearly shows how well Haar performs over the current LBP. As the number of objects is increased for identification, the LBP will not work. Whereas, the consistency of the hair is retained.

We have performed the experiment and got that it can detect and recognise face very properly.
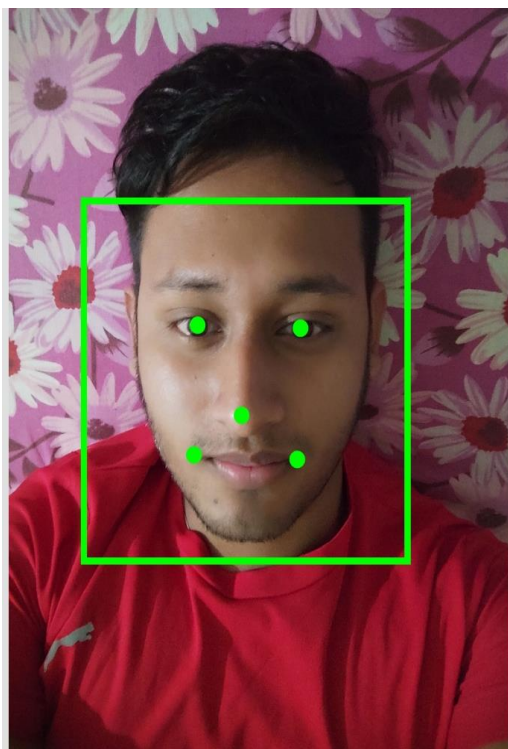

Figure 4. Result Image

A table, pointing out major differences between proposed (Haar) and existing algorithm (LBP), is shown in Table 1.

Table-1: Key Differences between Haar & LBP

Algorithm	Advantages	Disadvantages
Haar	1. High Detection accuracy (about 1-2% better at detecting faces). 2. Low false positive and false negative rate.	1. Computationally complex. 2. Limitations in difficult lightening conditions.
LBP	1. Computationally simple and 3x faster than Haar cascade. 2. Provides slightly better results in poor lightening.	1. Less accurate. 2. High False positive and false negative rate.

Conclusion

Briefly put, the process of detection and recognition using our methodology is effective in the accurate identification of human faces. The method is able to work marginally better than the LBP algorithm and has a good scope for potential implementations. The whole system is also doing its hardest to exclude any unnecessary costs. The motion sensor used keeps the machine idle, resulting in lower overall power consumption. The proposed device would be able to detect high-level 3D spoofing if it is fitted with a depth sensor and is trained accordingly. The primary objective of this project is to strengthen and remove any discrepancy in existing security systems and to achieve results as precisely as possible.

Acknowledgements

References

[1]. C. Panjaitan, A. Silaban, M. Napitupulu, J.W. Simatupang, "Comparison K-Nearest Neighbors (K-NN) and Artificial Neural Network (ANN) in Real Time Entrants Recognition". 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)

[2]. D. Pertsau, A. Uvarov, "Face detection algorithm using haar-like feature for GPU architecture". 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS). Vol. 2, 2013.

[3]. D. T. P. Hapsari, C. G. Berliana, P. Winda, M. A. Soeleman, "Face Detection Using Haar Cascade in Difference Illumination". 2018 International

Seminar on Application for Technology of Information and Communication.

[4]. N. Jiang, Y. Lu, S. Tang, S. Goto, "Rapid face detection using a multi-mode cascade and Separate Haar Feature". 2010 International Symposium on Intelligent Signal Processing and Communication Systems.

[5]. Wikipedia.org/wiki/viola-jones_object_detection_framework.