Comparison of Real-Time Face Detection and Recognition Algorithms

Yasar Borkar,, Reeve Mascarenhas, Shubham Tambadkar, and Jayanand P. Gawande*

Department of Instrumentation, Ramrao Adik Institute of Technology, D.Y Patil University, Nerul, Navi Mumbai, Maharashtra, 400706.

Abstract. With the phenomenal growth of video and image databases, there is a tremendous need for intelligent systems to automatically understand and examine information, as doing so manually is becoming increasingly difficult. The face is important in social interactions because it conveys information. Detecting a person's identity and feelings Humans do not have a great deal of ability to identify. Machines have different faces. As a result, an automatic face detection system is essential.in face recognition, facial expression recognition, head-pose estimation, and human–computer interaction, and so on Face detection is a computer technology that determines the location and size of a person's face. It also creates a digital image of a human face. Face detection has been a standout topic in the science field This paper provides an in-depth examination of the various techniques investigated for face detection in digital images. Various face challenges and applications. This paper also discusses detection. Detection features are also provided. In addition, we hold special discussions on the practical aspects of developing a robust face detection system, and finally. This paper concludes with several promising research directions for the future.

1 INTRODUCTION

Face detection and recognition is an integral part of computing systems since its inception in the early 1960s. One of the major areas of application of Face recognition is for security purposes [10]. It may also be used in the medical field to measure biometrics such as heartrate etc. This method is particularly popular for public use since images are the easiest to collect and are the one of the least intrusive biometric methods. The main objective of face detection and recognition is to emulate, and eventually exceed a human being's ability to identify and distinguish a particular face from multiple others. Face detection algorithms usually use universal features like eyes, nose and mouth to distinguish human faces from other objects. Nowadays, with advancement in machine learning and neural networks multiple image processing algorithms are available for face detection and recognition.

1.1 Literature Review

While designing and implementing a face recognition algorithm, many factors, like illumination, ageing and changing physical appearance, unstable placement of the camera, rapid movement of the subject, image resolution, viewing distance, etc. are considered [8]. All these factors are taken into account while choosing a suitable method/algorithm for a specific purpose. Because of the availability of so many algorithms and the many factors that go into selecting a particular method for a specific purpose, it is quite important to analyse these algorithms in reference to each other to efficiently choose a one, or a combination to suit our application.

In this paper, we will be discussing three face detection algorithms. We will conclude on which of the three the most efficient is.

Facial detection and recognition are methods of verifying the identity of an individual through their facial features. It is used for security purposes and falls under the category of biometrics.

The detection part commences with initially collecting either a dataset or a primary reference data through a camera which in turn helps in recognition or identification. Once we have the primary data, an already set program verifies the face through various algorithms and the identity of a person is confirmed.

Face recognition is used in various parts of the industry. The most common example is the smartphone industry. The device initially has to be set up by the owner. Their face is detected and recorded as the primary database. The smartphone unlocks only after recognizing the owner's face. The automation industry has used this technology in an innovative way. Building management systems use facial recognition devices to control access to the restricted areas in a building.

This technology uses various artificial intelligence and machine learning techniques and is rapidly evolving, getting even more accurate by the day. As for security concerns, facial recognition is of great reliability and convenience.

*Corresponding author: reevemascarenhas@gmail.com

1.2 Challenges faced in face detection and recognition

Because it deals with human faces, face recognition has always been one of the most fascinating and intriguing technologies. The Covid-19 outbreak has accelerated the world's transition to touchless facial recognition technology. Because of its contactless biometric features, it is gaining a lot of traction all over the world. Companies are abandoning traditional fingerprint scanners in favour of AI-based facial recognition technology, which is opening up new business opportunities. Security and surveillance, authentication/access control systems, digital healthcare, photo retrieval, and so on are some of the applications where its use has become critical. [6]

As previously stated, opportunities and challenges coexist. Growing commercial interest in face recognition is encouraging, but it also proves to be a difficult endeavour when it comes to problems associated with it, which have continuously hampered its quality of delivery. These difficulties arise when situations are uncooperative, resulting in the various facial appearances/expressions.

The challenges that limit the potential of a Facial Recognition System are listed below.

1.2.1 Illumination

Illumination means light variations. A minor change in lighting conditions poses a significant challenge to automated face recognition and can have a significant impact on its results. If the lighting varies and the same person is captured with the same sensor and almost identical facial expression and pose, the results may appear quite different.

Illumination drastically alters the appearance of the face. It has been discovered that the difference between two identical faces taken under different illuminations is greater than the difference between two different faces taken under the same illumination.

1.2.2 Pose

Facial Recognition Systems are extremely sensitive to variations in pose. When a person's head movement and viewing angle change, the pose of his or her face changes. Head movements or varying camera POVs can invariably cause changes in face appearance and generate intraclass variations, causing automated face recognition rates to plummet dramatically. When the rotation angle increases, identifying the true face becomes more difficult. If the database only has a frontal view of the face, it may result in incorrect or no recognition.

1.2.3 Occlusion

Occlusion is defined as a blockage that occurs when one or more parts of the face are blocked and the entire face is not available as an input image. One of the most difficult challenges in face recognition systems is occlusion.

It occurs as a result of a beard, moustache, or accessories (goggles, cap, mask, etc.), and it is common in real-world scenarios. The presence of such components diversifies the subject, making automated face recognition a difficult nut to crack.

1.2.4 Expressions

Face is one of the most important biometrics because its distinctive features play an important role in determining human identity and emotions. Different situations induce different moods, which result in a variety of emotions and, eventually, a change in facial expressions.

Another important factor to consider is the different manifestations of the same individual. Human expressions, in particular macro-expressions such as happiness, sadness, anger, disgust, fear, and surprise, are examples of human expressions. Micro-expressions are those that exhibit rapid facial patterns and occur involuntarily.

Macro and micro expressions appear on someone's face as a result of changes in one's emotional state, and in the aftermath of such emotions, which are numerous, efficient recognition becomes difficult.

1.2.5 Low Resolution

Any standard image should have a minimum resolution of 16*16 pixels. A picture with a resolution less than 16*16 is referred to as a low-resolution image. These low-resolution images can be obtained from small-scale standalone cameras such as CCTV cameras in streets, ATM cameras, and supermarket security cameras. These cameras can only capture a small portion of the human face area, and because the camera is not very close to the face, they can only capture a face region of less than 16*16 pixels. Because most of the details are lost in such a low-resolution image, it does not provide much information. The process of recognising faces can be difficult.

1.2.6 Ageing

Face appearance and textures changes over time and reflects as ageing, which is another challenge for facial recognition systems. Human face features, shapes, lines, and other aspects change as we get older. It is done after a long period of time for visual observation and image retrieval.

The dataset for a different age group of people over a period of time is calculated for accuracy checking. In this case, the recognition process is based on feature extraction, which includes basic features such as wrinkles, marks, brows, hairstyles, and so on.

The face is the most important part of the human body, and its distinctive features make it even more important in human identification. Various algorithms and technologies are used around the world to improve the accuracy and reliability of face recognition. The applications of this ever-expanding technology are also expanding in healthcare, security, defence, forensics, and transportation, all of which require greater precision. However, some challenges are universal in the development of face recognition technology, such as pose, occlusion, expressions, ageing, and so on, as discussed above in the paper.

1.3 Objectives

One of the main motivations with which we undertook this review is assist to individuals and organizations looking to use face detection algorithms in a specific application to select the method most suitable for their use. This could also be beneficial for researchers intending to design better and more efficient methods than already existing ones. The review and comparison provided in this paper might help researchers combine two or more algorithms mentioned here and develop an algorithm that is more efficient than when those methods are used individually. Another objective we had in mind was to implement a fully functional system to cater a specific application – in this case, a student/employee attendance system. Out of the four methods discussed in the paper we have used two – Haar Cascade (Viola-Jones) and Local Binary Patter histogram (LBPH) in above mentioned implementation.

1.4 Significance of this paper

The information derived from this paper can be used for educational purposes and will help fellow research members. This paper discusses three different algorithms of face detection and recognition. It gives a basic understanding and overview of the algorithms which can be easily understood by not only the industry experts but from a layman point of view as well. Apart from theoretical knowledge, this paper contains an experimental setup in which we have used Haar Casscade and Local Binary Pattern Histogram algorithms to detect and recognize faces. This experiment has helped us with better understanding of the two mentioned algorithms and the provided code can be implemented in the PyCharm IDE to run and check it for yourself to learn more about the selected techniques.

2 METHODOLOGIES

2.1 Haar-Cascade

Haar-cascade algorithm, popularly known as the viola-jones algorithm is used to recognize faces in a real time video [12]. It was initially proposed by Viola and Jones in 2001. The haar cascade algorithm is extensively used in several applications for the same purpose. It uses Haar features to differentiate facial features from the rest of the image to detect a face. Haar features were proposed by Alfred Haar in 1909. They are rectangular black and white regions which help in calculation in a specific rectangular region rather than calculations at pixel level.

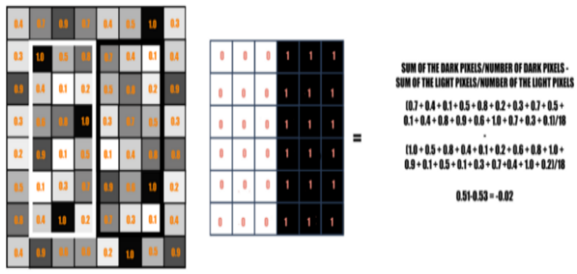


Figure 2

They come in several types, out of which the most common are 2 rectangular features, 3 rectangular features and 4 rectangular features.

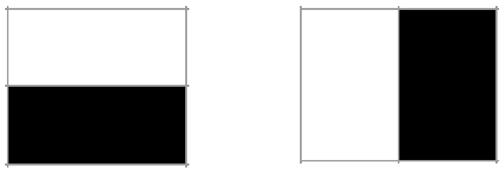


Figure 2: 2 Rectangle Feature

The 2-rectangle feature is used to detect the edge features in a face. The difference between the sum of the pixels is considered to be the value of the 2-rectangular feature.

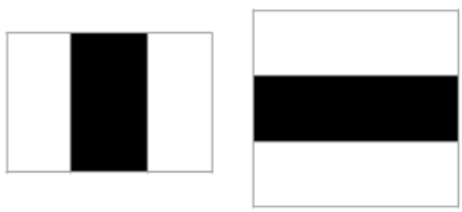


Figure 3: 3 Rectangle Feature

The 3-rectangle feature is used to detect the line features in a face. The sum is computed in the centre region in a 3-rectangle feature.

Figure 4: 4 Rectangle Feature

This is used to detect the 4-Rectangle features and computes the difference between diagonal pair of rectangles.

Rectangle features can be computed very rapidly using an intermediate representation for the image which we call the integral image. The integral image at location x,y contains the sum of the pixels above and to the left of x,y, inclusive:

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

2.2 Fisherface Algorithm

Fisherface is a popular face recognition algorithm that is widely thought to be superior to other techniques such as eigenface due to the effort made during the training process to maximise the separation between classes.

The face, as a part of the human body, is the simplest and most frequently used to determine an individual's identity. Humans can be distinguished and recognised more quickly and easily based on their faces.As a result, the face is used as a means of identifying a person or for face recognition.

In general, image recognition systems are classified into two types: feature-based systems and image-based systems. The first system extracts feature from the eye image components, nose, mouth, and so on, which are then geometrically modelled to determine the relationship between these features. In the second system, image pixels are used to represent them in various methods such as Principal Component Analysis, wavelet transformation, and so on, which are then used for training and image identification classification.

The process of obtaining characteristics that distinguish one face sample from another is known as feature extraction. As the principal component analysis method (PCA) used for face recognition was introduced by Turk and Pentland in 1991, reliable feature extraction techniques are critical in solving pattern recognition problems.

The PCA method seeks to project data in the direction with the greatest variation (indicated by the eigenvector), which corresponds to the covariance matrix's largest eigenvalues. The method's weakness is that it is less effective in class separation.

Cheng et al. introduced the Linear Discriminant Analysis (LDA) method for face recognition in 1991.According to the Fisher Criterion JF, this method attempts to find a linear subspace that maximises the separation of two pattern classes. This is obtained by simultaneously minimising the distance of the within class distribution matrix and maximising the split matrix spacing between the Sb classes, resulting in a maximum Fisher Criterion JF. By maximising the Fisher Criterion JF, the Fisher Linear Discriminant will find subspaces where classes are linearly separated. Sw will be singular if the data dimension is much larger than the number of training samples. This is a shortcoming of the LDA method.

Belheumeur introduced the fisherface method for face recognition in 1997. This method is a hybrid of the PCA and LDA methods. Before performing the LDA process, the PCA method is used to solve singular problems by reducing the dimensions. The disadvantage of this method is that the PCA dimension reduction process results in some loss of discriminant information useful in the LDA process.

The between-class scatter matrix is defined as

$$S_B = \sum_{i=1}^{c} N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T$$

The within-class scatter matrix be defined as

$$S_W = \sum_{i=1}^{c} \sum_{\mathbf{x}_k \in X_i} (\mathbf{x}_k - \boldsymbol{\mu}_i)(\mathbf{x}_k - \boldsymbol{\mu}_i)^T$$

However, in its early stages of development, face recognition with the fisherface method still has some issues, such as computation issues and the condition of the face image into input that will be used for image testing.

The computation problem in face recognition using the fisherface method becomes a problem due to the extremely complicated and complex computation process. While the diversity of the light of the face image, the attributes of the face image, the expression of the face image, and the variation of the position of the image of the face itself are the problems that affect the condition of the face image.

Face recognition systems based on the fisherface method are intended to recognise face images by matching the results of feature extraction. The system is expected to determine whether or not the image to be tested is correctly recognised.

The face image that will be used must first have to go through pre - processing stage. This stage consists of image acquisition as well as RBG image reconfigured into grayscale image. Using a camera, capture images of people's faces. This acquisition's image is a 24-bit RGB JPG file with a resolution of 92 x 112 pixels. Conversion of acquisition face image from RGB to 8-bit grayscale, BMP format, 40 x 40 pixels.

Furthermore, the face dataset must be split up into two (two) parts, with one part of the image serving as the training image (training dataset) and the other serving as the test image (testing dataset).

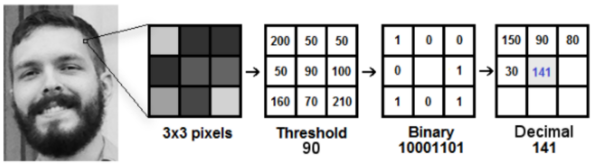
During this image evaluation phase, the Fisherface technique will be used to create a feature vector of frontal image data that can be used by a system, and then to match the vector of traits of the training sample with the vector of characteristics of the test image through using Euclidean distance formula.

2.3 Local Binary Pattern Histogram (LPBH)

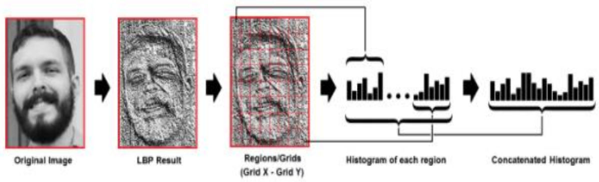
Local Binary Pattern histogram (LBPH) is one of the most common and easily available face recognition algorithms. It was first developed in 1994 and has been since used as an efficient facial recognition algorithm. It uses features extracted from an image and thresholding its neighbouring pixels and getting the result as a binary number. Another advantage about the LBPH algorithm is that it is easily available in the open-source library (OpenCV) and can be used by many programming languages like Python, MATLAB, C++, etc.

The LBPH, like most facial recognition algorithms uses a dataset with the images of the subjects we need to recognize using the algorithm. Usually, the dataset is already pre-processed and only the facial features of the subject are used in the LBPH algorithm. This facial isolation is done using a face detection algorithm – for ex. Haar Cascade.

The LBPH algorithm creates an intermediate image that details and isolates the necessary features for facial recognition. This is done by getting a part of the grayscale image in a matrix format (usually 3x3) where the value of the elements of the matrix are the intensities of each pixel (from 0 to 255). The central pixel is selected and its intensity value is chosen as the threshold value. Now, its 8 neighbouring pixels are compared with the threshold and each are given binary values. The binary values are selected by giving 1 for values greater than or equal to the threshold and 0 for values lower than threshold. Next, the binary values from the matrix are concatenated and a binary number is formed. This binary number is then converted to decimal and set as the central value of the matrix. This procedure is repeated, covering the whole original image and a new LBP result is produced. [13]



Now, the image generated from LBP procedure is divided into vertical and horizontal grids. The histograms of each grid are created and all the resulting histograms are concatenated to create the final bigger histogram. This final histogram represents the characteristics of the original image.



Having obtained the histograms for every image in the training dataset, we repeat the same procedures for the new image obtained from the user input. To find the correct match, simply the two histograms are compared and a threshold value is set passing which the image is concluded to be the same person. One of the most common approaches for comparing the histograms is calculating the Euclidean distance between them. This can be done by the formula:

$$D = \sqrt{\sum_{i=1}^{n}(hist1_i - hist2_i)^2}$$

This distance can also be used as confidence measurement. When a confidence value is lower than the given threshold, it can be assumed that the algorithm has successfully recognized the image.

3 EXPERIMENTAL RESULTS

To conduct the experiment, we used the functions available in the OpenCV library available for Python. We run the program on a laptop computer through PyCharm. The specifications of the computer are:

Company: HP
Windows Edition: Windows 11 Home Single Language
Processor: AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10 GHz
Installed RAM: 8GB
System Type: 64-bit operating system, x64-based processor

https://doi.org/10.1051/itmconf/20224403046

Four separate programs were written for the four different aspects of the method. One program each was used for:

- Creation of Dataset:

The dataset creation was done using the onboard camera of the computer, running the video input received by the camera through the Haar Cascade algorithm and isolating the grayscaled facial features and storing in the dataset. For every person, an ID is provided so that it can be used in the recognition stage.

- Training of Dataset:

Next, the training is done by collecting the LBP histograms of each mage in the dataset. These histograms are stored against their respective IDs so that the recognition code can match the appropriate IDs to its relevant person.

- Input and recognition of image:

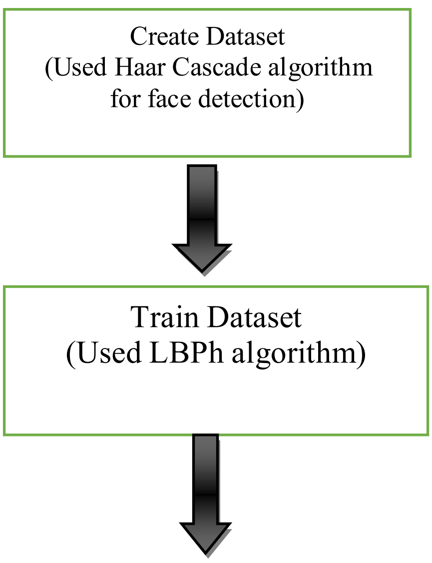
In the recognition stage, the camera is once again activated and an input video stream is collected whose LBP histogram is created. This input histogram is compared with all histograms from the dataset. The confidence (Euclidean distance) between the two histograms is calculated. The confidence can be calculated by:

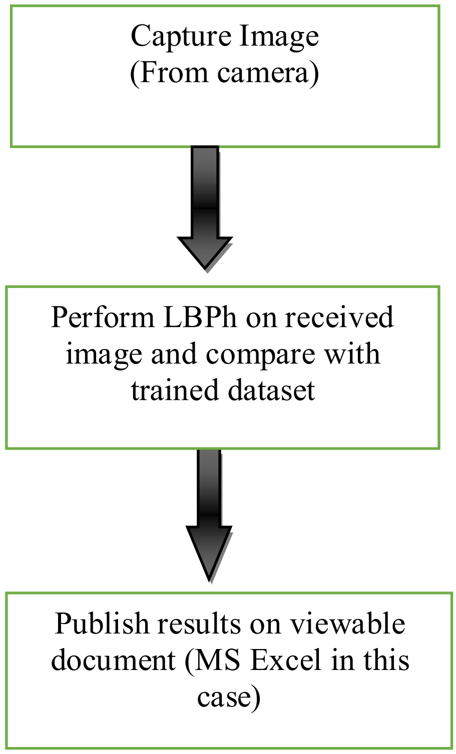
If the confidence is lower than the provided threshold value, that ID is considered to be recognized.

- Writing the results onto a readable file:

Finally, the write to file code is run to save and publish the results onto MS Excel. This result can be monitored by the organization for attendance purposes.

The flow of the experiment is given below for better understanding.





Process Flow Diagram

We tested the Haar Cascade, Local Binary Pattern Histogram and Fisherface algorithms with 3 different subjects for each algorithm. [18]

The experiment was performed three times with varying number of training samples in each case. The first case was conducted with 10 samples, second with 25 and the third with 50 samples. It was observed that the accuracy improves with increasing number of training samples. Accuracy was calculated by comparing the total number of pictures recognized correctly (true positives) with total number of pictures tested. It can be mathematically defined by the following expression:

$$\text{Accuracy (\%)} = \frac{\text{Total no. Of T.P}}{\text{Total no. of T.S}}$$

Where,

T.P stands for True Positive

T.S stands for Training Samples

The results obtained are shown below in tabular form:

	3 Subjects		
	Accuracy in %		
	N = 10	N = 25	N = 50
Haar-Cascade	95	97	99
LBPH	96	98	99
Fisherface	92	95	98

Where N = no. of training samples used.

4 CONCLUSION

This paper gives us a general idea about facial recognition and detection algorithms. It has been written keeping in mind enthusiasts who wish to begin with their research and need a basic understanding of the initially used algorithms for the same. We have discussed some challenges faced in the field of face detection and recognition and how certain algorithms can be used to overcome specific limitations. Three different algorithms have been reviewed in this paper, namely, Haar-Cascade, Local Binary Pattern Histogram (LPBH) and Fisherface. These algorithms have been explained in a simple yet efficient manner so that the reader will understand the concepts with ease.

Apart from the theoretical explanation, we have implemented the above-mentioned algorithms in the PyCharm IDE for an application-based point of view. By real-time implementation, it was observed that out of the three reviewed algorithms, Local Binary Pattern Histogram (LBPH) was the most accurate with an average accuracy of 97.66% followed by Fisherface with an average accuracy of 95.21%. Haar Cascade has proved to be one of the most reliable facial detection algorithms with an accuracy of 97.69%. This program can be implemented by schools, offices or any other organization in the industry that requires a database to keep record of.

Even though LBPH was found to be the most accurate algorithm, various factors are taken in account for various reasons in the face detection and recognition techniques. Every algorithm has its respective benefits and limitations. Hence, it cannot be empirically justified that any one of the existing algorithms is best suited for all kinds of applications. A future avenue for research could be combining these existing algorithms for greater efficiency and minimizing the effects of certain limitations.

5 REFERENCES

[1] NURULHUDA ISMAIL, MAS IDAYU MD. SABRI Department of System and Computer Technology (Multimedia), Faculty of Computer Science and Information Technology, University of Malaya, 50603 LembahPantai, Kuala Lumpur, MALAYSIA, Page 1-11

[2] UMAMAHESWARI, K., SUMATHI, S., SIVANANDAM, S.N., and PONSON, T. (2007). Artificial Neural Network Based Image Data Mining with Hybrid Recognizers. *International Journal of Soft Computing Applications*. Issue 1, Page. 27-
49.

[3] PRASEEDA, L. and SASIKUMAR, M. (2008). RBF Based Face Recognition and Expression Analysis. In Proceedings of World Academy of Science, Engineering and Technology. Vol. 35.

[4] MOHAMED, D. (2008). A Simple Face Recognition Algorithm using Eigeneyes and a Class-Dependent PCA Implementation. *International Journal of Soft Computing*. Vol. 3, No. 6, Page. 438-449.

[5] MANJUNATH, B.S. and CHELLAPPA, R. (1993). A Unified Approach to Boundary Perception: Edges, Textures and Illusory Contours. In Proceedings of *IEEE Transactions Neural Networks*. Vol. 4, No.1, Page. 96–107.

[6] Preeti Chatterjee, Subhadeep Jana & Souradeep Ghosh, Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 21 Issue 2 Version 1.0 Year 2021

[7] UMAMAHESWARI, K., SUMATHI, S., SIVANANDAM, S.N., and PONSON, T.(2007). Artificial Neural Network Based Image Data Mining with Hybrid Recognizers. *International Journal of Soft Computing Applications*. Issue 1, Page. 27-49.

[8] K. Delac and M. Grgic, "A survey of biometric recognition methods," in 46th International Symposium Electronics in Marine, vol. 46, pp. 16– 18, 2004.

[9] SINGH,S.K.,CHAUHAN, D.S., VATSA, and SINGH, R. (2003). A Robust Skin Colour Based Face Detection Algorithm. *Tamkang Journal of Science and Engineering*.Vol. 6, No. 4, Page. 227-234.

[10] Sasha Callaway Jeffrey Cheng Alexander Contratti, David Fu Harshika Gelivi Jacob Wachulec, Sonia Purohit, Comparative Analysis of Image Processing Algorithms for Airport Security, New Jersey's Governor's School of Engineering and Technology, July 18, 2020.

[11] TURK, M., and PENTLAND, A. (1991). Face Recognition Using Eigenfaces. In Proceedings of IEEE Conference on *Computer Vision and Pattern Recognition*.Page. 586-591.

[12] Michael Jones, MA 02139 Cambridge, MA 02142, Mitsubishi Electric Research Labs Compaq CRL 201 Broadway, Paul Viola, 8th FL One Cambridge Center Cambridge, Rapid Object Detection using a Boosted Cascade of Simple Features.

[13] Kelvin Salton Do Prado, Facial Recognition: Understanding LBPH Algorithm, November 11, 2017.

[14] J.Brownlee. “How to Control Neural Network Model Capacity with Nodes and Layers.” Machine Learning Mastery, 10 Jan. 2020.

[15] ROWLEY, H., BALUJA, S., and KANADE, T. (1998). Neural Network- Based Face Detection. In Proceedings of IEEE Transactions on *Pattern Analysis and Machine Intelligence*. Vol. 1, No. 20, Page.23–28.

[16] W. Koehrsen. “Beyond Accuracy: Precision and Recall.” Medium, Towards Data Science, 10 Mar. 2018.

[17] Peter N. Belhumeur, Joao~ P. Hespanha, and David J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, JULY 1997.

[18] Amir Dirin, Nicolas Delbiaggio, Janne Kauttonen Haaga-Helia University of Applied Science, Helsinki, Finland, Comparisons of Facial Recognition Algorithms Through a Case Study Application.