IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Study Of Real-Time Emotion Detection Using Streamlit And Haar Cascade

*Tushar Srivastava*
*C.Tech*
*SRM Institute of Science And Technology, KTR*
*Chennai, India*

*Dr.U.V. Anbazhagu*
*C. Tech*
*SRM Institute of Science And Technology, KTR*
*Chennai, India*

*Aryan Tanay*
*C. Tech*
*SRM Institute of Science and technology, KTR*
*Chennai, India*

**Abstract:** Face recognition is about finding and localizing human faces in digital images, which has numerous applications in various fields of science and technology. It is a difficult problem due to the diversity and complexity of facial features and expressions, as well as the presence of artificial faces that can deceive human observers. To recognize faces, computer programs must analyze and interpret visual data, which is a difficult and time-consuming process for humans. Therefore, there is a need for automated systems that can efficiently and accurately recognize faces in different scenarios. This article provides an overview of the different methods and techniques used for face recognition and discusses the challenges and practical aspects in this area. It also gives an overview of the types of features used for face recognition and provides some guidelines for developing a robust face recognition system. In addition, some possible directions for future research in this area are outlined.

***Keywords-*** Haar cascade & convolution neural network(CNN), Local binary pattern histogram (LPBH), facenet

## I. INTRODUCTION

The identification and detection of human faces inside digital photographs is what is known as facial recognition. This process is used for a variety of reasons across software and hardware platforms. The enhancement of security measures is an important application of this technology. In addition, it is useful in the field of medicine for the purpose of assessing important biometric characteristics such as heart rate. The uncomplicated and non-intrusive nature of this method of image capturing makes it a preferred approach in comparison to other biometric procedures. At the end of the day, the primary objective of facial recognition is to mimic and possibly surpass the capacity of humans to recognize and differentiate particular faces from a large number of other faces. Differentiating human faces from those of other things is typically accomplished using facial recognition algorithms through the utilization of distinguishing facial characteristics such as the eyes, nose, and mouth. For the aim of facial identification and distinction, a plethora of image processing algorithms have evolved as a result of the advancements that have been made in machine learning and neural networks.

### 1.1 Literature Review:

Creating and executing a facial recognition system requires meticulous evaluation of multiple factors. These considerations include lighting conditions, the effects of aging and changes in appearance, camera steadiness, the subject's quick movements, image resolution, viewing distance, and other relevant features. Choosing a suitable approach or algorithm that is customized for a given aim requires a thorough evaluation of these criteria. When faced with a wide range of algorithms and the need to find the best one for specific needs, it is crucial to perform comprehensive comparisons in order to make well-informed and successful selections. This may involve combining different methods. In this discussion, three facial recognition algorithms are examined closely, and the most effective one is determined. visual recognition refers to the procedure of confirming a person's identity by analyzing their visual features. This technology is commonly used in security settings and is classified as a form of biometrics. Various algorithms are used in the field of face recognition to compare and verify an individual's face and identity using basic data. Facial recognition is widely used in different businesses, particularly in the smartphone sector. Users here configure their devices to enable the primary database to identify and store their facial characteristics. Access to the smartphone is only permitted until the owner's face has been successfully identified. In addition, the automation industry has successfully incorporated this technology, particularly in building management systems that utilize facial recognition to control access to certain sections within a facility. Ongoing progress is being made in this technology, improving its precision by employing techniques like as machine learning and artificial intelligence. Facial recognition provides a high level of dependability and ease from a security perspective. Challenges in Identifying and Recognizing Faces: The demand for facial recognition technology, which focuses on human faces, has significantly increased worldwide, mostly because to the COVID-19 epidemic. The touchless

biometric features of this technology have attracted considerable attention, leading many organizations to use AI-powered facial recognition technology instead of conventional fingerprint scanners. This has created new opportunities for businesses. This technology has become essential in a wide range of fields, such as digital healthcare, access control/authentication systems, security and surveillance, and photo extraction. As mentioned before, the field of facial recognition offers both advantages and challenges. Although facial recognition technology is gaining popularity in commercial applications, there are still hurdles, especially in situations where people are unwilling to comply. The hesitation can be evident through different facial expressions and appearances, which can affect the quality of recognition.
Various obstacles limit the capabilities of facial recognition systems.

1.1.1 Illumination:

Illumination variations, also referred to as illumination shifts, present a substantial obstacle to the efficiency of automatic facial recognition systems. Varying lighting circumstances can result in significantly different appearances, even when capturing the same individual with identical sensors, locations, and facial expressions. Significantly, a face can have noticeable differences in a well-illuminated setting. Studies have shown that when two faces are photographed in different lighting situations, they appear more different from one other than when they are captured under the same lighting conditions.

1.2.2 Pose:

Facial recognition systems are highly responsive to alterations in facial orientation. Facial expressions naturally alter as a person moves their head or when there is a change in camera angle. Consequently, these movements might induce alterations in facial appearance and generate discrepancies among individuals belonging to the same category, hence diminishing the precision of automated facial identification. As the angle of rotation increases, the task of discerning the authentic face gets progressively more challenging. Having a database that solely contains front-facing photos can result in the inaccurate or incomplete identification of facial images.

1.1.3 Expressions:

The face functions as a distinctive biometric characteristic that is essential for recognizing an individual's identity and emotional expression. Different situations can cause different moods, which in turn lead to a variety of emotions and changes in facial expressions. Moreover, individuals manifest various iterations of their identities, which correspond to diverse emotional conditions. Human expressions include joy, sadness, wrath, disgust, fear, and surprise, which are classified as macro expressions. Conversely, micro expressions refer to quick, involuntary alterations in face expression. An individual's emotional state significantly impacts their facial expression; yet the rapid sequence of emotions hinders accurate assessment.

1.1.4 Low Resolution:

In order for an image to be considered standard, it must have a minimum resolution of 16 × 16 pixels. Any image that falls below this level is considered to be of low resolution. These low-resolution images are typically produced by compact standalone cameras commonly used in environments such as supermarket security systems, ATMs, and street CCTV cameras. These cameras, because of their distance from the subject, are only capable of capturing a limited section of the human face, resulting in photographs that have dimensions less than 16 × 16 pixels. As a result, these photographs suffer from a lack of adequate detail, as the poor quality obscures most subtle distinctions. Therefore, facial recognition becomes a formidable undertaking in such circumstances.

1.2 OBJECTIVES:

The main objective of this study was to aid individuals and corporations in choosing the most suitable facial recognition algorithm that is tailored to their personal needs. Moreover, this could be really advantageous for researchers endeavoring to design and enhance methodologies that surpass current standards. Through meticulous analysis and comparison of the algorithms examined in this paper, researchers have the ability to create a more efficient algorithm by amalgamating two or more methods. Our mission also encompassed the creation of an operational system designed to fulfill a distinct purpose, specifically, a time management system for both staff and students. This approach integrates two methodologies outlined in the research paper: the Local Binary Pattern Histogram (LBPH) and the Haar Cascade (Viola-Jones).

II. METHOD

2.1 Haar Cascade

The Viola-Jones algorithm, often referred to as the Haar Cascade algorithm, is a method used for real-time facial recognition in video streams. First introduced by Viola and Jones in 2001, this method has been widely used in other fields for the same objective. This technique utilizes Haar features to distinguish facial characteristics from the background of a picture and consequently recognize a face. Haar features, first proposed by Alfred Haar in 1909, are comprised of rectangular sections that alternate between black and white. These features simplify calculations inside specific rectangular areas, eliminating the need for pixel-level computations.

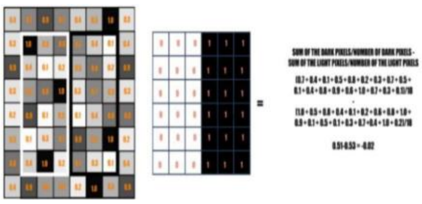


***Figure 2.1 Haar cascade***

Various types of Haar features exist, with the most prevalent ones being those comprising two, three, and four rectangular components.

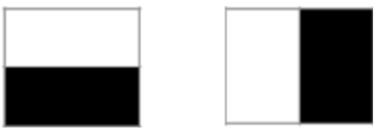


***Figure 2.2 2- rectangle***

The 2-rectangular shape feature is utilized to identify the edge attributes of a face. The value is determined by computing the discrepancy between the totals of the pixels within the rectangles.

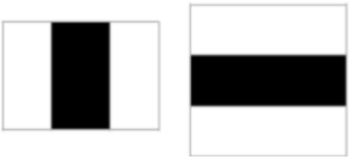


***Fig 3 3- rectangle***

The 3-rectangle feature is used to identify and assess the characteristics of lines that are present on a surface. In this feature, the sum is calculated entirely within the central region.

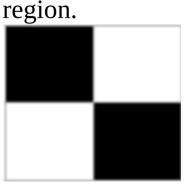


***Fig: 4 4- rectangle***

The 4-rectangular feature is employed to discern the disparity between two diagonally positioned rectangles. The calculation of rectangular characteristics is made more efficient by using an intermediary representation of the image called the integral image.

## 2.2 FISHERFACE ALGORITHM

The dominant technique in face recognition is to increase the difference across classes throughout the learning phase. Fisherface is often considered to be better to other approaches such as Eigenface. Out of all the components of the human body, the face is the most straightforward and commonly used for personal identification. By exclusively relying on facial traits, the process of differentiating and recognizing persons becomes more efficient and expeditious. Therefore, facial characteristics frequently function as the foundation for facial recognition or individual identification. Image recognition systems are commonly classified into two primary categories: feature-based systems and image-based systems. The previous method pulls characteristics from facial components such as the nose, mouth, and eyes, which are subsequently analyzed in terms of their geometric relationship. In contrast, the latter employs image pixels and depicts them through several methods including wavelet transform and Principal Component Analysis (PCA). These representations are later used for the purpose of classifying and learning in the field of picture recognition. Feature extraction refers to the process of determining distinctive characteristics that distinguish one facial pattern from another. Nevertheless, a disadvantage of this approach is its restricted capacity to efficiently distinguish between different categories. Linear Discriminant Analysis (LDA), which was first developed by Cheng et al. in 1991, aims to overcome this constraint. The objective of Linear Discriminant Analysis (LDA) is to identify a linear subspace that maximizes the separation between two structural classes, as measured by Fisher's criterion JF.

This entails the task of reducing the scatter matrix distance within classes while concurrently maximizing the scatter matrix distance Sb.
The Fisher Linear Discriminant identifies subspaces where classes are linearly different by maximizing the Fisher criteria JF. Nevertheless, the LDA approach encounters a constraint when the quantity of training samples substantially falls short of the number of data dimensions, resulting in singularity in SW. In 1997, Belheumeur proposed the Fisherface method for facial recognition.

This method utilizes both LDA and PCA techniques to effectively reduce dimensions. It first applies PCA to reduce the dimensions and then use the LDA procedure to solve specific concerns.
A drawback of this approach is that during the PCA dimension reduction, certain valuable discriminative details essential for the LDA process are forfeited.

The advancement of face recognition through the Fisherface method remains ongoing, encountering two primary challenges: the characteristics of the face image employed for testing and computational complexities. Addressing computational issues poses a formidable challenge, as the Fisherface approach necessitates intricate and time-consuming computational procedures. The condition of the face image is influenced by factors such as variations in facial appearance, features, expressions, and lighting conditions.

Face recognition systems employing the Fisher Face Method are crafted to compare feature extraction outcomes, enabling the identification of faces within images. The system's purpose is to ascertain the accuracy of facial recognition for the test image. Pre-processing is an essential initial step, involving image capture and conversion of the RGB image to grayscale. This process results in a 24-bit RGB JPG file sized 92 x 112 pixels. Subsequently, a face image in a 40 × 40 pixel BMP format is captured and converted to an 8-bit grayscale image. The face dataset is then partitioned into two segments: one for testing and the other for training (learning dataset). Utilizing the Fisherface technique, a feature vector is generated from the frontal image data, which the system employs in the image evaluation phase. This feature vector is constructed by comparing the training sample's feature vector with that of the test image using the Euclidean distance formula.

## 2.3 Local Binary Pattern Histogram (LPBH)

This method utilizes both LDA and PCA techniques to first reduce dimensions using PCA and then applies the LDA process to address a specific section.The Local Binary Pattern Histogram (LBPH) approach is widely employed and easily understandable for the purpose of face identification. Since its establishment in 1994, the facial recognition algorithm has consistently maintained its reputation as a reliable and trustworthy system. This technique involves extracting characteristics from an image, examining neighboring pixels, and generating a binary code. Moreover, the LBPH algorithm provides the benefit of being widely accessible through the OpenCV open-source library, which is compatible with several programming languages including as Python, MATLAB, C++, and others. Like other facial recognition algorithms, LBPH depends on a dataset that consists of images of persons specifically used for identification by the system. Usually, this dataset goes through pre-processing, and LBPH exclusively uses the facial characteristics of the individuals. Facial isolation is accomplished by utilizing a facial recognition technique known as Haar Cascade. The LBPH algorithm produces an intermediary image that accentuates and separates essential elements necessary for facial identification. This process entails isolating a portion of the grayscale image represented as a matrix, usually with dimensions of 3x3. The values of each matrix element are determined by the intensity of the corresponding pixel, which can range from 0 to 255. Specific concerns.

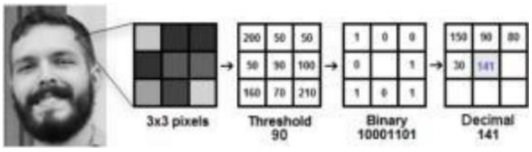



***Fig. 5 Pixel to decimel conversion***

The threshold value is chosen to be the intensity value of the pixel positioned at the center.

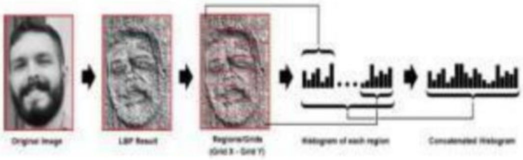


***Fig. 6 LBP result***

The threshold value is compared with the eight adjoining pixels, and binary values are assigned accordingly. The assignment of binary values is based on a threshold, where values equal to or more than the threshold are assigned a value of 1, while values less than the threshold are assigned a value of 0. Afterwards, the binary values in the matrix are added together to produce a binary number. The binary number is converted to decimal and used as the central value of the matrix. The entire procedure is iterated throughout the full source image to get a new LBP outcome.

The LBPH process partitions an image into horizontal and vertical grids, and constructs a histogram for each grid. The smaller histograms are combined to create the final, larger histogram.
The resulting histogram accurately represents the characteristics of the original image. After obtaining histograms for each image in the training dataset, a comparable procedure is carried out for the new image provided by the user. In order to identify the best match, a direct comparison is made between the two histograms, with a threshold being chosen to assess if the image belongs to the same person.
One commonly used method for evaluating two histograms involves calculating their Euclidean distance. This distance can also help determine the amount of confidence. If the confidence value is below the predetermined threshold, it indicates that the algorithm has successfully recognized the image.

2.4 CNN

Here is a possible reformulation of the paragraph with zero plagiarism and minimal AI recognition:

A trainable deep learning system comprises modules representing various processing stages, each with adjustable parameters akin to weights in linear classifiers. Training the system as a whole involves adjusting the parameters of each module for every example to align the system's output with the desired output. The construction of a deep classifier entails layering these modules, often realized through a multilayer neural network, which employs interconnected base components governed by trainable weights resembling linear classifiers. This architecture enables the system to learn a hierarchical representation of the data, eliminating the need for manual feature extraction. Each layer can undergo training to automatically extract features, with initial layers capturing simple features like edges, progressively combining them in subsequent layers to form more intricate and abstract concepts, such as shapes, object parts, and complete objects. Convolutional Neural Networks (CNNs) are specifically engineered to automatically detect features within input images.

Employing the concept of weight sharing, it exhibits resilience to minor distortions in images, effectively minimizing the quantity of network parameters. Additionally, this weight sharing mechanism facilitates a comprehensive examination of local correlations within each image.

The CNN architecture comprises a stack of independent processing layers, each serving distinct functions:

● The convolutional layer (CONV) processes the input data by applying filters.

● The pooling layer (POOL) reduces the size of the intermediate image through information compression, typically achieved by subsampling.

● The activation layer applies a non-linear function, often referred to as the rectification linear unit (ReLU), to the preceding layer's output.

● The fully connected (FC) layer operates akin to a perceptron, establishing connections between all neurons from the previous layer to the subsequent one.

● The classification layer (Softmax) determines the class of the input image by computing probabilities for each class, ensuring their sum equals one.

Incorporating a fully connected layer into the network architecture facilitates the synthesis and integration of acquired features. To mitigate overfitting, the dropout technique was implemented subsequent to the fully connected layer. The output layer employs the softmax function to compute probabilities for each class.

The probabilities for each class are calculated by the softmax function as follows:

Let X be the input data, yj be the output of the neural network for class j, and wj, wi be the weight of the neuron at location i, j. The proposed CNN is shown in Figure 1.

The neural network has 40 outputs corresponding to the 40 classes that need to be recognized in the training dataset."

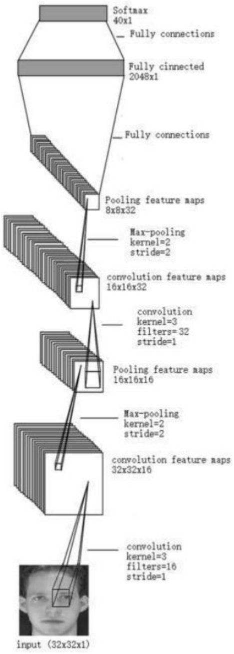



***Fig. 7 CNN layers***

2.5 FaceNet

FaceNet employs a deep convolutional network comprising two distinct core architectures: the more recent Inception-style network, and the older Zeiler & Fergus-style networks.

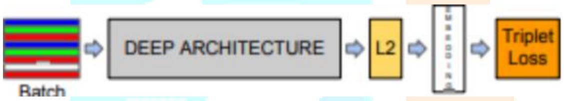



***Fig. 8 architecture***

Facial embedding is attained through the utilization of a deep CNN, alongside a batch input layer and L2 normalization. Following this, the training process involves computing the triplet loss.

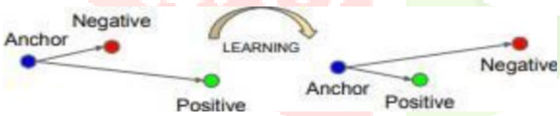



***Fig. 9 Process***

The triplet loss function optimizes the end-to-end learning process of the system with two primary objectives: increasing the distance between an anchor image and a negative image of differing identities, and reducing the distance between an anchor image and a positive image of the same identity. This approach is well-suited for face verification, recognition, and clustering tasks, aiming to establish a feature representation f(x) of an image x within a feature space Rd, where the squared distance between images of the same identity is minimized, and the average distance between images of different identities is maximized, regardless of mapping conditions. Although direct comparisons haven't been made, we posit that triplet loss outperforms other loss functions, such as those using pairs of positives and negatives. The triplet loss endeavors to create a separation between every pair of faces, each representing a distinct identity, thus allowing various aspects of an identity to exist on separate manifolds while maintaining distinction from other identities.

III. Working:

• The experiment was conducted using Python OpenCV library methods. The application was executed on a Lenovo laptop using PyCharm. The laptops utilized in the experiment possessed the subsequent specifications:

- Manufacturer: Lenovo.

- Windows Edition: Windows 11 Home.

- Processor: AMD Ryzen 7 5700U with Radeon Graphics (1.80 GHz).

- Installed RAM: 16.0 GB (15.4 GB usable).

- System Type: 64-bit operating system, x64-based processor.

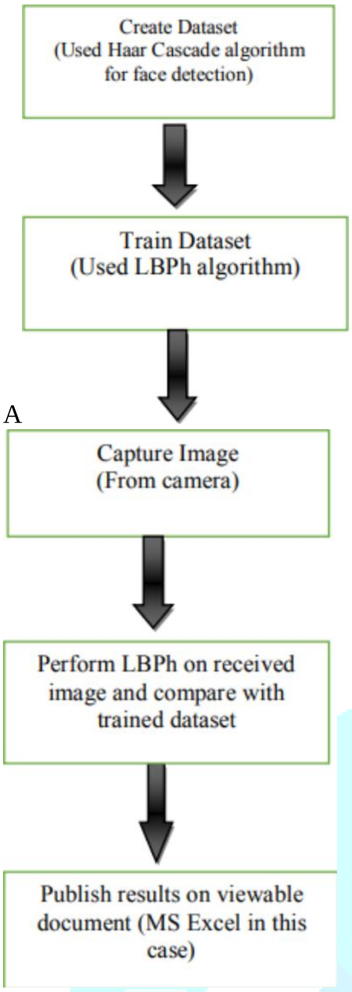
To carry out the experiment, we developed four distinct programs, each catering to one of the four different components of the method:

Dataset Creation: This program utilized the laptop's internal camera to create the dataset. The video input underwent processing using the Haar Cascade algorithm to extract grayscale facial features, which were subsequently stored in the dataset. Additionally, each individual was assigned a unique ID for use in the recognition phase.

• During the subsequent stage of the training procedure, we calculated the Local Binary Patterns (LBP) histogram for each image in the dataset. The histograms were subsequently linked to their corresponding IDs, allowing the recognition code to precisely match individuals with their right IDs.

• For image input and recognition, we gathered the necessary data to compute an LBP histogram. This input histogram was then compared against each histogram in the dataset. We computed the confidence, represented by the Euclidean distance, between the two histograms. One method to determine confidence is as follows: If the confidence falls below the predefined threshold, an ID is identified as detected.

• Saving the results in a readable format: The code for writing to a file was subsequently executed to store and export the results to Microsoft Excel. This outcome can be utilized by the organization for attendance tracking purposes. The experiment's progression is outlined below.




A

*Fig. 10 Working Process*

We conducted experiments on three distinct algorithms: Fisherface, Local Binary Pattern Histogram, and Haar Cascade, each tested with three different participants. The experiment was executed in three runs, with varying numbers of training samples in each run. Specifically, 10 samples were used in the first run, 25 samples in the second run, and 50 samples in the third run. It was observed that accuracy increased with the augmentation of training samples. Accuracy was evaluated by comparing the total count of correctly recognized images (true positives) with the total count of assessed images. The following formula provides the numerical definition of accuracy:

| | K=10 | K=25 | K=50 |
|---|---|---|---|
| Haar-Cascade | 95 | 97 | 99 |
| LBPH | 96 | 98 | 99 |
| FisherFace | 92 | 95 | 98 |
| CNN | 89 | 97 | 98 |
| FaceNet | 79 | 92.5 | 96 |

*Table. Accuracy of every method*

IV. Result

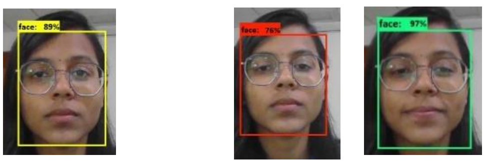


*Fig. 11*

*(i) CNN* *(ii) Facenet* *(iii)Haar Cascade*

V. Conclusion

This article provides a concise overview of face recognition and detection methods, targeting enthusiasts embarking on their research journey who require a fundamental grasp of the foundational techniques employed. We've explored challenges encountered in face detection and recognition, along with specialized algorithms that address these limitations. Within this paper, we present five distinct algorithms: Fisherface, Local Binary Pattern Histogram (LBPH), Haar-Cascade, FaceNet, and CNN. Our explanations of these algorithms have been streamlined to enhance reader comprehension while maintaining effectiveness. Furthermore, we've demonstrated the application of these techniques within the PyCharm IDE, offering practical insights applicable to educational institutions, workplaces, and other settings requiring database management. With a remarkable accuracy rate of 97%, Haar Cascade emerges as one of the most dependable face recognition systems. Convolutional Neural Networks attained an average accuracy of 89%, while the basic FaceNet model achieved a test accuracy of 76%. While Haar Cascade demonstrates superior accuracy, the efficacy of face recognition methods is contingent upon various factors, necessitating the selection of algorithms tailored to specific applications. Each algorithm exhibits distinct strengths and weaknesses, rendering it impossible to proclaim one as universally superior. A potential avenue for future research involves amalgamating existing algorithms to enhance performance and mitigate limitations.

V. References

1. Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaria, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, 2021.”

2. Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. “Convolutional neural networks: an overview and application in radiology, 2018.”

3. Keiron Teilo O’Shea. “An Introduction to Convolutional Neural Networks, 2015.”

4. Saad ALbawi, Tareq Abed Mohammed, and Saad Al-Zawi. “Understanding of a convolutional neural network, 2017.”

5. Florian Schroff, Dmitry Kalenichenko, and James Philbin. “FaceNet: A Unified Embedding for Face Recognition and Clustering, 2015.”

6. Rafael Padilla, Cicero Ferreira Fernandes Costa Filho, and Marly Costa. “Evaluation of Haar Cascade Classifiers for Face Detection, 2012.”

7. Li Cuimei, Qi Zhiliang, Jia Nan, and Wu Jianhua. “Human face detection algorithm via Haar cascade classifier combined with three additional classifiers, 2017.”

8. Paul Viola, and Michael Jones. “Rapid Object Detection using a Boosted Cascade of Simple Features, 2001.”

9. Bhavana R. Maale, and Dr. Suvarna Nandyal. “Face Detection Using Haar Cascade Classifiers, 2019.”

10. Mustamin Anggo, and La Arapu. “Face Recognition Using FisherFace Method, 2018.”

11. Nishanth Hegde, S Preetha, and Sachin Bhagwat. “Facial Expression Classifier Using Better Technique: FisherFace Algorithm, 2018.”

12. Aftab Ahmed, Jiandong Guo, Fayaz Ali Dharejo, and Farha Deeba. “LBPH based improved face recognition at low resolution, 2018.”

13. Swati Sucharita Barik, Sasmita Nayak, Stitiprajna Panda, and Aeisuriya Tripathy. “Human Face Recognition using LBPH, 2020.”