




## Research Article

# A Real-Time Framework for Human Face Detection and Recognition in CCTV Images

**Rehmat Ullah** <sup>1</sup>, **Hassan Hayat**,<sup>2</sup> **Afsah Abid Siddiqui**,<sup>2</sup> **Uzma Abid Siddiqui**,<sup>2</sup> **Jebran Khan** <sup>3</sup>, **Farman Ullah**,<sup>2</sup> **Shoaib Hassan**,<sup>2</sup> **Laiq Hasan**,<sup>1</sup> **Waleed Albattah** <sup>4</sup>, **Muhammad Islam**,<sup>5</sup> and **Ghulam Mohammad Karami** <sup>6</sup>

<sup>1</sup>Department of Computer Systems Engineering, University of Engineering and Technology Peshawar, Peshawar, Pakistan

<sup>2</sup>Department of Electrical and Computer Engineering, COMSATS University Islamabad, Attock Campus, Attock, Pakistan

<sup>3</sup>Department of Artificial Intelligence, AJOU University, Suwon, Republic of Korea

<sup>4</sup>Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

<sup>5</sup>Department of Electrical Engineering, College of Engineering and Information Technology, Onaizah Colleges, Al-Qassim, Saudi Arabia

<sup>6</sup>SMEC International Pvt. Limited, Kabul 1007, Afghanistan

Correspondence should be addressed to Ghulam Mohammad Karami; [ghulam.karami@smec.com](mailto:ghulam.karami@smec.com)

Received 2 October 2021; Accepted 15 December 2021; Published 3 March 2022

Academic Editor: Nouman Ali

Copyright © 2022 Rehmat Ullah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to develop a machine learning and deep learning-based real-time framework for detecting and recognizing human faces in closed-circuit television (CCTV) images. The traditional CCTV system needs a human for 24/7 monitoring, which is costly and insufficient. The automatic recognition system of faces in CCTV images with minimum human intervention and reduced cost can help many organizations, such as law enforcement, identifying the suspects, missing people, and people entering a restricted territory. However, image-based recognition has many issues, such as scaling, rotation, cluttered backgrounds, and variation in light intensity. This paper aims to develop a CCTV image-based human face recognition system using different techniques for feature extraction and face recognition. The proposed system includes image acquisition from CCTV, image preprocessing, face detection, localization, extraction from the acquired images, and recognition. We use two feature extraction algorithms, principal component analysis (PCA) and convolutional neural network (CNN). We use and compare the performance of the algorithms K-nearest neighbor (KNN), decision tree, random forest, and CNN. The recognition is done by applying these techniques to the dataset with more than 40K acquired real-time images at different settings such as light level, rotation, and scaling for simulation and performance evaluation. Finally, we recognized faces with a minimum computing time and an accuracy of more than 90%.

## 1. Introduction

Today's organizations face significant security challenges; they need several specially trained personnel to achieve the required security. However, humans make mistakes that affect safety. Closed-circuit television (CCTV) is currently used for various purposes in everyday life. The development of video surveillance has transformed simple passive monitoring into an integrated intelligent control system.

Face detection and its new applications for secure access control, financial transactions, etc. Biometric systems (faces, palms, and fingerprints) have recently gained new importance. With advances in microelectronics and vision systems, biometrics has become economically viable. Facial recognition is an essential part of biometrics. In biometrics, human fundamentals are mapped to current data. The facial features are hauled out and implemented using an efficient algorithm, and some variations are made to improve the

existing algorithm model. Face recognition from the computer can be applied to a variety of applied applications, including crime ID, security systems, and authentication. A facial recognition system typically involves steps of face detection where the face of the input image is detected, and then the image process cleans the face image for easy recognition.

In this modern age, face recognition has become a necessity as the individual's identification increases daily with globalization. Since the last two decades, face recognition has received much attention because of its various applications, invaluable image analysis, and understanding domains. Face recognition is also becoming important in other fields like image processing, animation [1], security [2], human-computer interface [3], and medicine [4]. Face recognition is natural, noninvasive, and easy to use. The face recognition system has a wide choice of applications in public safety, entertainment, attendance management, and financial payment. While today's facial recognition systems work well in relatively controlled environments, they suffer from significant problems when used in existing surveillance systems due to image resolution, background clutter, lighting variations, and face and expression posture.

Face recognition systems consist of three steps, such as preprocessing of the image, feature extraction, and classification technique for recognition [5]. Features extracted from the face, such as the mouth, nose, eyebrows, etc., are geometric features. The detected and processed face is compared to a database of known faces to determine who the person is. The surveillance system needs people to monitor it. Human monitoring involves reliability issues, scalability issues, and the inability to identify everyone.

Facial occlusions, such as beards and accessories (glasses, hats, and masks), involve evaluating facial recognition systems, making the subject diverse and challenging to function in a nonsimulated environment. Another essential factor to consider is the different terminologies of the same distinct: macro and microterminologies find their place on someone's face because of changes in an emotional state, and because of the many expressions of this type, effective recognition becomes difficult. A perfect face recognition system should be able to tolerate changes in lighting, expressions, poses, and occlusions and can scale for many users who need to capture the fewest images simultaneously.

The overall contributions of the research paper can be summarized as follows:

- (i) A machine learning-based framework for detecting and recognizing faces in CCTV images with various clutter backgrounds and occlusion
- (ii) A dataset of 40K images with different environmental conditions, clutter backgrounds, and occlusion
- (iii) Performance comparison of classical machine learning and deep learning algorithms for faces recognition in CCTV images

The rest of the paper is organized as follows: Section 2 briefly introduces the related works. Section 3 explains the

methodology, and the results are discussed in Section 4. Finally, we conclude the paper in Section 5.

## 2. Related Work

In this section, we briefly introduce the related works about face detection and recognition using classical approaches and deep learning.

### 2.1. Face Detection Algorithms

*2.1.1. Geometric Methods for Face Detection.* In the early stages of computer vision, researchers explored many algorithms that extracted the image characteristics and utilized geometric requirements to comprehend the provisions of all features. This was partly due to very limited computational resources. The reduction of information from the extraction of functionality has made computer vision possible in the first computers [6, 7].

*2.1.2. Template-Based Face Detection [8].* Most of the face detection algorithms are model-based, they encode facial images directly on the basis of pixel intensity. Probabilistic models are mostly used for the characterization of these images of facial images also by neural networks or by some other mechanisms. The parameters of these models are automatically adjusted by sample images or manually.

*2.1.3. Simple Templates.* If you are using a skin-based method and another skin color is found in the image (like arms and hands), these algorithms show false results. Many researchers tried to overcome this by using simple models to integrate results from the color matching of skin. These models have varied from some ovals related to the image of the edge of the entrance to the correlation models for the regions of skin color and skin color (like lips, hands, or eyes). However, these techniques can enhance the robustness of detectors by color, but with also the enhancement of speed.

*2.2. Face Recognition algorithms.* Face recognition is a technique that has now attained consideration in machine learning and artificial intelligence. It plays an essential role in many social security applications. There are many studies and practices now under research that can solve the problem of face recognition. Vivek and Guddeti [9] proposed combining cat swarm optimization (CSO), particle swarm optimization (PSO), and genetic algorithm (GA). This hybrid technique has inspired many others to work similarly. Ali et al. combined SVM, higher-order spectral (HOS), and random transformation (RT) [10].

*2.2.1. Iterative Closest Point-Based Alignment.* The objective of the alignment approach [11, 12] is based on the closest iterative point to determine the translation and the rotation parameters in an iterative way to convert the point cloud. Clouds' mean square error becomes minimal while both point clouds are aligned. So, distance among point clouds is

reduced to a minimum by translating and rotating one of the point clouds with respect to others, also determine by identifying the distance with every point in the initial point clouds every second, also calculating the average of all distances. An important disadvantage of the alignment approach based on the closest iterative point is that it needs an initial alignment of the convergence course. This approach is computationally very expensive, so that's another disadvantage.

*2.2.2. Simulated Annealing-Based Alignment.* It is an algorithm based on a stochastic process used for local research [13]. The difference between hill-climbing and simulated annealing is that it can compute an even worse solution than the current one in the iteration process. As simulated annealing is not bounded by local minima, it is more likely that you will find a solution. Six parameters are required for simulated annealing (in which three for every translation also the rotation referencing to a 3D coordinate system) which is used to define transformation matrix which is used for an alignment between two 3D faces. This approach aligns images of the face in three phases: (1) alignment in initial level, (2) alignment in an approximate level, and (3) alignment in the last level [14]. Initially, the center of the two-sided mass is being aligned. By using this approach, it serves to minimize an approximation measure which uses the consensus of multiple estimators  $M$  (MSAC) together with the mean square error corresponding point of two faces that will compare. Then, an accurate alignment is obtained with the mean of a search algorithm that is based upon simulated annealing, which uses the measurement of the interpenetration of surfaces (SIM) as an estimation criterion. The disadvantage of alignment based on simulated annealing is its more calculation time which is comparable to the alignment based on the nearest iterative point.

*2.2.3. Average-Based Face Model.* This alignment is based on the medium-based face model [15]. First of all, the reference points are on the face automatically or manually. Subsequently, the average of pivotal coordinates calculated, followed by procrustes examination and transformed milestone [16], are again mediated to obtain a face model. While in this method, the image of the probe face aligns with the average model using an alignment on the nearest iterative point. A notable weakness of the alignment based on the medium face model is the low precision index [17] and part of the spatial material lost during the creation of the medium face model.

The first step in face recognition is preprocessing. Images taken from a camera or in real-time video surveillance setups may suffer from various degradations during the process of capture, transformation, conversion, or compression [18]. For instance, blurry, noisy, and low-resolution images affect the face recognition process. Such issues may lead to significant challenges in the face recognition scheme and decrease its performance. Therefore, pre-processing is an essential step in any face recognition system. Many color normalization, statistical, and convolutional methods are

used as preprocessing tools [19]. Another big problem in face recognition through surveillance cameras is that too many images of a person are collected and applying a face recognition algorithm to each of them proves costly in terms of processing and energy consumption. Vignesh et al. [20] presented a technique for image quality assessment (IQA) using CNN to take the person's best image. Tudavekar et al. [21] proposed video inpainting to fill the missing regions in a video by dual-tree complex wavelet transformation.

PCA is the most widely used technique in signal and image processing. They are also known as eigenfaces, the orthogonal vectors that help in face recognition. Drume and Jalal proposed a two-level classification technique that uses principal component analysis (PCA) in level one and boosts its results by support vector machine (SVM) at level two [22]. Kanade employed image processing techniques to extract 16 facial parameters with the ratio of distance, angle, and area and used the method of Euclidean distance to achieve a performance of 75% [23]. On this basis, a method called eigenface for face recognition was proposed for the first time [24]. This method leads to the formation of an algorithm called principal component analysis (PCA). From then on, PCA gathered a lot of attention and became the most effective approach for face recognition. Many improvements have been made in the PCA algorithm to get its best results [25–30].

Rala used PCA and Kernel-PCA for feature extraction and face recognition, respectively. They explore the non-linear kernel function for the improvement of PCA [31]. Abdullah et al. optimized the PCA time complexity without affecting the performance of the algorithm [32]. Another approach includes hexagonal feature detection, which works on the principle of edge detection [33]. A part-based method in [34] utilizes PCA, NMF, ICA, LDA, etc., under partial occlusion. Another effective algorithm called AFMC shows the results to be more accurate with the reduced computational cost and proposes eliminating the SSS problem [35]. Viola–Jones algorithm was also presented with the smoothed invalid regions and excluded near-ear regions [36].

Deep hidden ID entity feature (DeepID), a face representation based on CNN, is suggested in [37]. Unlike DeepFace, which learns features from a single large CNN, DeepID learns features from an ensemble of tiny CNNs that are utilized for network fusion. Similarly, a face recognition pipeline, WebFace, is proposed in [38], which uses CNN to learn the face representation. The convolutional neural network (CNN) [39] has been one of the most prominent approaches in computer vision over the last decade, with applications including image classification [40], object identification [41], and face recognition [38]. Different methods, such as PCA-based eigenfaces [42] and LDA-based Fisherfaces [43] employ the nearest neighbor (NN) classifier and its variants [44]. In a face recognition system, supervised classifiers such as support vector machines (SVM) [45] and neural networks [46] are also proposed. Huang et al. [47, 48] developed a novel learning technique for single hidden layer feedforward networks (SLFNs) called the extreme learning machine (ELM), that can be utilized in regression and

classification applications [42, 49–51]. Yang et al. [52] proposed a re-enforcement-based deep learning algorithm for multirobot path planning. Table 1 depicts the summary of the literature review.

### 3. Proposed Framework for Face Detection and Recognition in CCTV Images

The proposed method consists of four significant steps: (i) image acquisition, (ii) image enhancement, (iii) face detection, and (iv) face recognition, as shown in Figure 1. We performed different machine learning techniques for recognition purposes that include random forest, decision tree, K-nearest neighbor (KNN), and convolutional neural network (CNN).

**3.1. Image Acquisition.** In this phase, we acquire an image. Images need to be restored from the source (usually a hardware source) camera, making it the first step in the workflow sequence because processing is not possible. Our CCTV constantly reads images, which is our preprocessed input.

**3.1.1. Camera Interfacing.** An Internet protocol (IP) camera, Hikvision DS-2CD2T85FWD-15/18, is used for image acquisition. It is an 8-megapixel camera and captures 15 frames per second video with a resolution of 1248 \* 720. Firstly, the camera will capture the image, which will be saved and accessed using some software tool, such as MATLAB. Table 2 shows the CCTV camera specification used for image acquisition.

The face database includes the faces of those whom it will recognize. Because facial recognition involves classification algorithms, each image in the dataset is labeled. Images of each person's faces have their own unique labels. We have more than 41,320 images of 90 people. Thus, the label of these classes (persons) is from 1 to 90. It means that each label has multiple images. Given below is the dataset description.

So, label 1 has 775 images approximately and same as others displayed in the figure (classes on the  $x$ -axis and number of images on the  $y$ -axis). Figure 2 shows the sample images in the dataset.

**3.2. Preprocessing.** After the image acquisition, preprocessing of the image prepares it for further handling. Preprocessing includes two main steps: gray scale conversion and edge detection techniques.

**3.2.1. Grayscale Conversion.** From the camera, we acquire the RGB image (R for red, G for green, and B for blue). An RGB pixel has 1 pixel of red combined with pixels of blue and green. The RGB image made computation expansive as 1 pixel is of 8 bits, so in RGB, it would become 24 bits. In a grayscale image, each pixel is a scalar, so it will be an 8-bit image. So, the equation that converts RGB to grayscale is

$$\text{Grayscale} = 0.3 * R + 0.59 * G + 0.11 * B. \quad (1)$$

Here R, G, and B represent red, green, and blue pixels, respectively.

**3.2.2. Canny Edge Detection.** The Canny filter detects edges in pictures by detecting abrupt changes in color in photos. We are using this to enhance the edges of the images. The more the advantages are improved, the more accuracy we can achieve in recognizing facial expressions. The filter consists of Gaussian and Sobel filters. Firstly, a Gaussian filter with a predefined value of  $\sigma$  is applied to grayscale images to smooth edge finding.

$$G = \frac{1}{(2\pi\sigma^2)} e^{-(x^2+y^2)/2\sigma^2}. \quad (2)$$

In the second step, the Sobel filter is applied for finding the edges in the images. The filter used for finding the horizontal edges is

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}. \quad (3)$$

For horizontal edges, the filter is

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (4)$$

The horizontal and vertical edges are calculated in order to find all the edges in the filter.

$$A = x = \sqrt{G_x^2 + G_y^2}. \quad (5)$$

The third and last step of the canny edge detector, the hysteresis threshold, is applied to images containing the edges. The threshold is expressed as

$$H = \frac{1}{1 + e^{-x}}. \quad (6)$$

The maximum and minimum thresholds are selected initially. If the pixel's value is greater than the specified threshold, then one is assigned to the pixel, and if the value of the pixel is less than the threshold, then it is set to 0. Another case is when the value is the same as the threshold; it remains the same. Lastly, the edges are added to the original image to get the final enhanced image. Thus, detection and extraction of facial features become easy and increase the efficiency of the overall system.

**3.3. Face Detection.** The next step after getting the image from the camera is to detect the face from the images by the Viola–Jones algorithm that distinguishes the face and nonface regions. Then, for further processing, the face region is extracted.

TABLE 1: Literature review.

Ref. no.	Algorithm	Accuracy	Dataset
[53]	Principal component analysis, local binary patterns histograms, K-nearest neighbor, and convolutional neural network	85.6%, 88.9% 81.4%, and 98.3%	400 images for 40 persons
[42]	Local binary pattern	93.3% and 90.8%	30 images over 10 people, 5040 images over 120 people
[43]	Convolutional neural network and support vector machine	97.5%	1400 images for 200 persons
[54]	Virtual geometry group (VGG) face model	92.1%	2.6M images over 2.6K people
[55]	Nearest neighbor	87.3%	14,000 images of over 1000 people
[56]	Recurrent regression neural network	95.6%	4207 images for 337 persons
[57]	Binary quality assessment	95.56%	494 414 images for 10 575 persons
[58]	Eigenfaces, Fisherfaces, and Laplacian faces	79.4%, 94.3%, and 95.4%	41 368 images of 68 persons
[59]	SRC, NN, NS, and SVM	98.4%, 72.7%, 94.4%, and 95.4%	4000 images for 126 persons
[60]	Fisher vector space and deep face	93.1% and 97.3%	2.6M images of 2622 persons

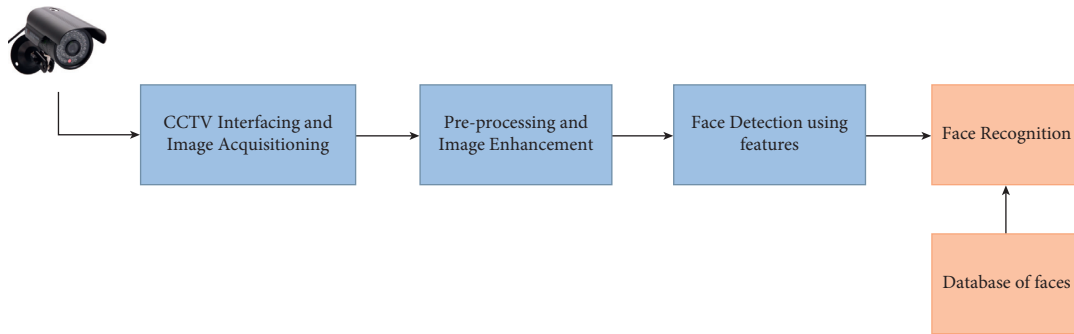


FIGURE 1: Process flow of the proposed system.

TABLE 2: Camera properties.

DS-2CD2T85FWD-15/18
Up to 8 megapixel high resolution
Digital noise reduction
Day and night vision
Max. resolution 3840 × 2160

3.3.1. *Face Detection Using Viola–Jones Algorithm.* Viola–Jones algorithm is the first algorithm that provides competitive object detection rates in real-time. It provides robustness with high detection rates, easy for real-time applications as it can process two frames per second. After applying this, different classification techniques are used to recognize the image. The main steps include the following:

- (1) Haar feature
- (2) Integral image
- (3) Ada boost training
- (4) Cascading classifiers

3.3.2. *ROI Extraction and Resizing.* The face detected by the Viola–Jones technique is extracted and resized as a 40 × 40 image, then used by various feature extraction techniques to find the features.

3.4. *Features Extraction from Detected Face Images.* We have used the principal component analysis (PCA) technique to extract features of the face in order to detect the face in later steps.

3.4.1. *PCA-Based Facial Feature Extraction.* PCA is a technique used to reduce the dimensions of the images in our dataset. It finds the characteristics of images, the difference and variance in pixels in one column from the other [58]. PCA has the following steps as shown in Figure 3:

- (1) Mean of each column

In this step, we have calculated the mean value of each column. The sum of the means of the columns are expressed as

$$\gamma_i = \sum_{i=1}^n \frac{a_{1i} + a_{2i} + a_{3i} + \dots + a_{mi}}{m}. \quad (7)$$

Here,  $\gamma_i$  is the mean of  $i$ -th column.

- (2) Covariance matrix

The second step is calculating the covariance of the matrix. The variance of the pixels is calculated as



FIGURE 2: Sample of face images used for recognition.



FIGURE 3: PCA steps for feature extraction.

$$\text{cov}(X_i, X_j) = \frac{1}{n} \sum_{k=1}^m (X_i^k - \gamma_i)(X_j^k - \gamma_j). \quad (8)$$

In the above equation,  $i$  is the number of columns in the original image matrix,  $j$  is the second column in the image, and  $k$  is the number of rows. The following equation shows the result.

$$\begin{bmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \dots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \dots & \text{cov}(X_n, X_n) \end{bmatrix}. \quad (9)$$

### (3) Eigenvalues

After the covariance matrix is calculated, the eigenvalues of the covariance matrix can be calculated by.

$$|\text{covariance} - \gamma I_n| = 0. \quad (10)$$

### (4) Eigenvectors

Using the eigenvalues calculated in the previous step, we can find the eigenvectors from the following equation:

$$|\text{covariance} - \gamma_i I_i| * X_i = 0. \quad (11)$$

Eigenvalues are the features of an extracted face. These values will be used for recognition.

## 3.5. Face Recognition Using Machine Learning Algorithms

**3.5.1. Random Forest.** This is a machine learning approach for solving classification and regression problems. It makes use of ensemble learning, a technique used for solving difficult problems by combining many classifiers. Many

decision trees make up a random forest algorithm. The random forest algorithm's produced "forest" is trained via bagging or bootstrap aggregation. Bagging is a meta-algorithm that enhances accuracy by grouping them together.

**3.5.2. Decision Tree.** For classification and regression, the decision tree is a nonparametric supervised learning approach. The objective is to learn basic decision rules from data characteristics to construct a model that predicts the value of a target variable. It is a flowchartlike tree structure in which each internal node represents an attribute test, each branch indicates the outcome, and each leaf node (terminal node) carries a class label.

**3.5.3. K-Nearest Neighbor.** We have used 5, 10, and 15 eigenvectors as our features. The dataset is created with these vectors, and the new face image will pass through all the steps of PCA. Then, we will calculate its distance with the features of other images in the dataset, and the nearest one will be our prediction. We have used the Manhattan distance formula to calculate distance as it is more accurate. The Manhattan distance formula is

$$D(Z, B) = \sum_{x=1}^n |z_x - b_x|. \quad (12)$$

Here,  $z$  is for the dataset, and  $b$  is for the test image. Then we will check which instance in the dataset has the minimum distance with the test image, which will be our prediction.

**3.6. Face Recognition Using Convolutional Neural Network.** Convolutional neural networks consist of convolutional layers, pooling layers, and, at the end, a fully connected layer. A CNN has a much different architecture than a simple neural network. It has an input layer, a convolutional layer, a

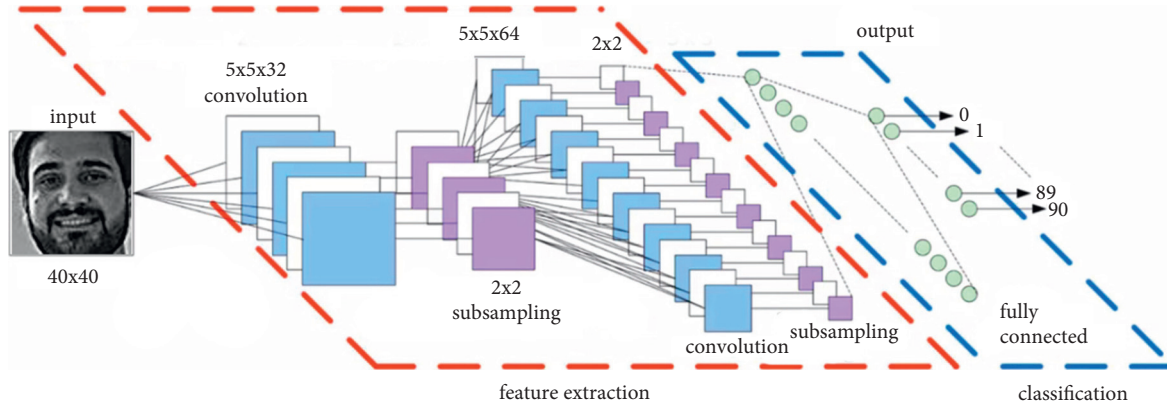


FIGURE 4: The architecture of CNN.

max-pooling layer, and at the end, a fully connected neural network as shown in Figure 4.

We have used Adam optimizer for training in optimizing weights.

### 3.6.1. Adam Optimizer

$$\begin{aligned}
 v_t &= \beta_1 * v(t-1) - (1 - \beta_1) * g_t, \\
 s_t &= \beta_2 * s(t-1) - (1 - \beta_2) * g_t^2, \\
 \Delta\omega_t &= -\eta \frac{v_t}{\sqrt{s_t + \epsilon}} * g_t, \\
 \omega_{t+1} &= \omega_t + \Delta\omega_t,
 \end{aligned}
 \tag{13}$$

where  $\eta$ : learning rate (0.001),  $g_t$ : gradient at time  $t$ ,  $v_t$ : exponential average of the gradient,  $s_t$ : exponential average of the square of Gradient, and  $\beta_{1,2}$ : hyperparameters.

## 4. Results and Discussion

When we apply PCA, we get eigenvectors; these eigenvectors are our features. We have used different features, such as we have used 5, 10, and 15 eigenvectors.

**4.1. K-Nearest Neighbour (KNN) Algorithm Results.** Results obtained by simulating different values of  $k$  are as shown in Table 3.

Figure 5 with 5 eigenvectors shows the results obtained having a maximum accuracy of 94.7%. When we increase the value of  $K$  the accuracy decreased. For  $K = 1$ , with Manhattan distance, we get approximately 95% accuracy, and with Euclidean distance, we get 89% accuracy.

In Figure 6, PCA features with 10 coefficients are shown. With 10 eigenvectors, we obtained a maximum of 93.7% accuracy with Manhattan distance and with Euclidean distance, we obtained 87.6%. Then the accuracies decreased as the value of  $K$  increased. Here we have also noted that Manhattan distance performs better than Euclidean distance. And if the eigenvectors increase, the accuracy also

decreases because the starting eigenvectors show maximum feature importance.

In Figure 7, PCA features with 15 coefficients are shown. Same case here, as the features increase, accuracy decreases. And the same with the value of  $k$ .

**4.2. Decision Tree Result.** For the decision tree, the results obtained for different features are given below, both in tabular form Table 4 and graphical form Figure 8.

**4.3. Random Forest Results.** The random forest shows the highest accuracy of 93.20% with 5 eigenvectors in Table 5 and Figure 9.

**4.4. CNN Results.** As in CNN, we must train our dataset. We have trained our data in 5000 steps and obtained 95.7% accuracy with only 30 images for testing and 30 for training.

**4.4.1. With 50% Training and Testing Data.** We have obtained a maximum of 95.67% accuracy with 50% data of training and testing. We trained it in 4000 steps. In some steps, the training steps, the accuracy increased, and at some points, it decreased, but at the end, we have obtained a maximum accuracy of 95.67%, accuracy as shown in Figure 10.

**4.4.2. With 90% Training and 10% Testing Data.** Now we have obtained 95% accuracy in this section, maybe because testing data is much less than training. And we have obtained this accuracy in 300 steps, as shown in graph Figure 11.

**4.4.3. With 80% Training and 20% Testing Data.** Now we have obtained 97.5% accuracy in this section, which may be because testing data is much less than training data. And we have trained data in 5000 steps, as shown in the graph Figure 12.

TABLE 3: Results for KNN.

No. of features	Training data	Numerical methods	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
5	90	Euclidean	89.0115%	89.7889%	79.0841%	76.5861%	75.278%
		Manhattan	94.7623%	90.0457%	88.6113%	86.9326%	86.0086%
	80	Euclidean	87.8664%	80.2338%	77.7842%	75.2137%	73.5403%
		Manhattan	93.7989%	89.0839%	87.6401%	85.9456%	84.8975%
10	90	Euclidean	88.3589%	79.7717%	77.8163%	76.1214%	75.0927%
		Manhattan	93.7989%	89.4494%	88.4072%	87.1582%	77.0927%
	80	Euclidean	86.8185%	77.905%	75.9567%	74.377%	73.4407%
		Manhattan	93.6811%	88.3288%	87.335%	86.0392%	85.3475%
15	90	Euclidean	86.383%	76.6452%	74.7327%	73.293%	72.5651%
		Manhattan	93.9484%	88.2299%	87.3221%	86.1644%	85.618%
	80	Euclidean	84.5773%	74.3589%	72.5164%	71.1934%	70.4926%
		Manhattan	92.8172%	86.6614%	85.9425%	84.7646%	84.1484%

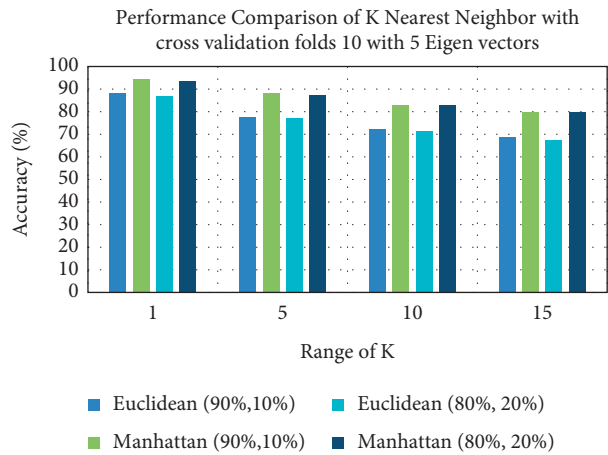


FIGURE 5: Comparison of KNN results for 5 eigenvalues.

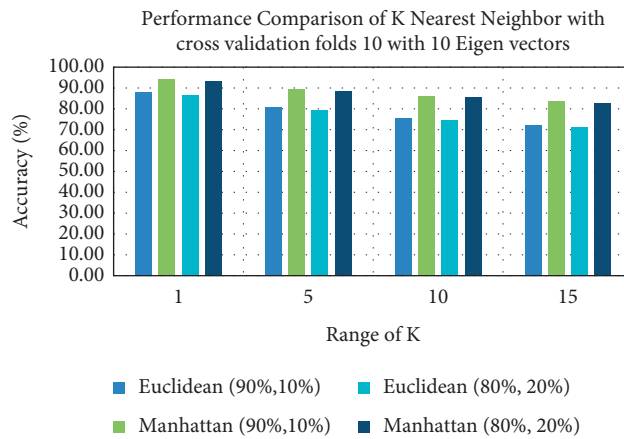


FIGURE 6: Comparison of KNN results for 10 eigenvalues.



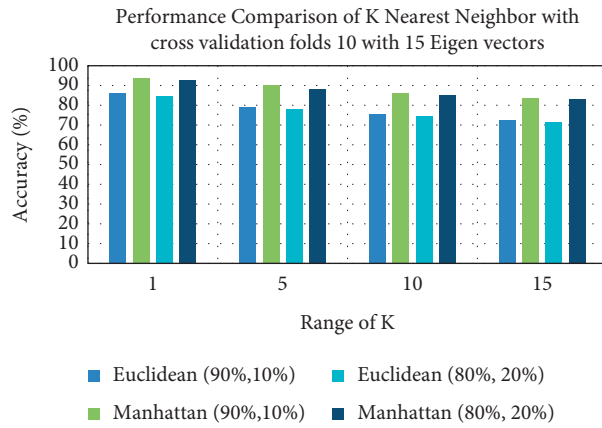


FIGURE 7: Comparison of KNN results for 15 eigenvalues.

TABLE 4: Results for decision tree.

No. of features	Training data	Testing data	Accuracy
5	90%	10%	70.34%
	80%	20%	68.75%
10	90%	10%	68.88%
	80%	20%	68.39%
15	90%	10%	68.64%
	80%	20%	68.28%

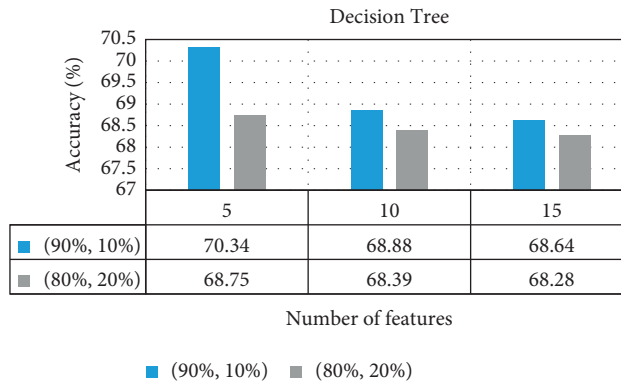


FIGURE 8: Comparison of decision tree results.

TABLE 5: Results for random forest.

No. of features	Training data	Testing data	Accuracy
5	90%	10%	93.20%
	80%	20%	92.65%
10	90%	10%	91.38%
	80%	20%	90.71%
5	90%	10%	89.95%
	80%	20%	88.60%

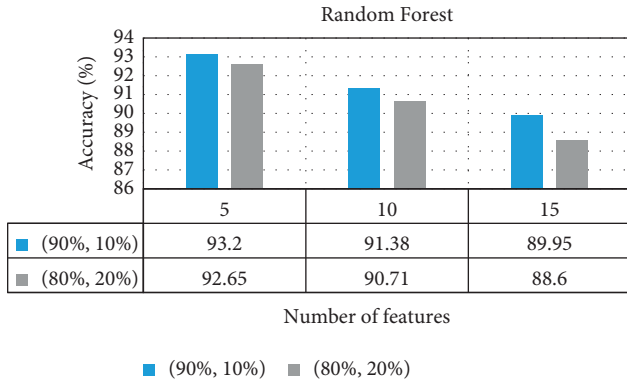


FIGURE 9: Comparison of random forest results.

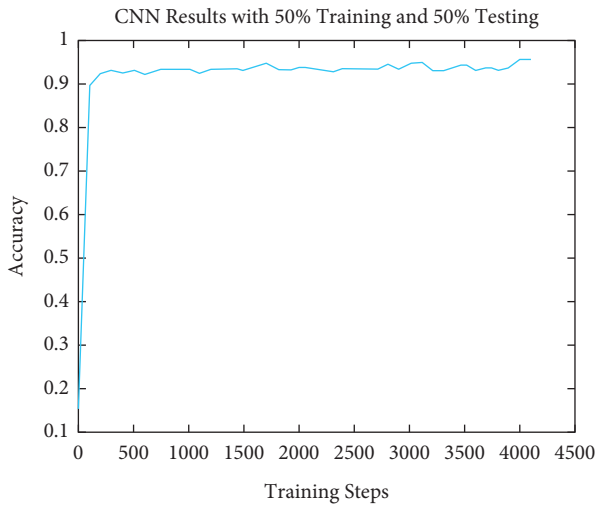


FIGURE 10: Results of 50% training and 50% testing data using CNN.

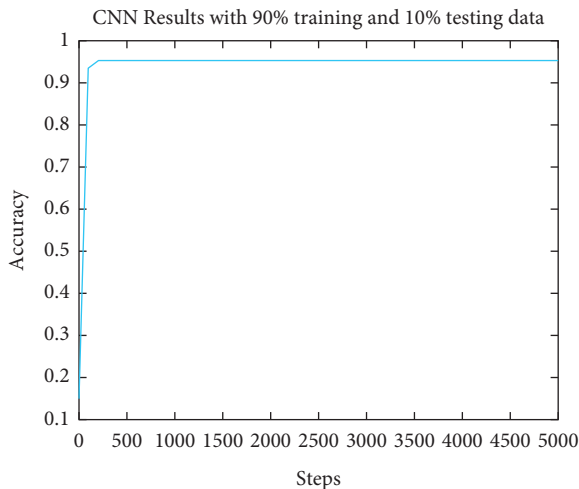


FIGURE 11: Results of 90% training and 10% testing data using CNN.

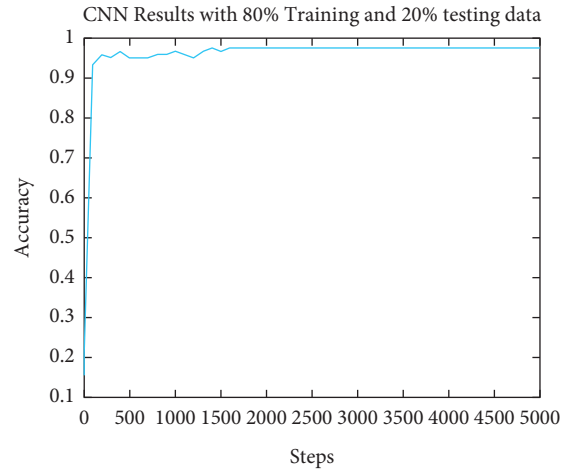


FIGURE 12: Results of 80% training and 20% testing data using CNN.

### 5. Conclusion

We have developed a framework for automatic face recognition based on CCTV images using different machine learning algorithms in this work. One of the objectives of this work is to collect more than 40,000 face images and compare the performance of algorithms to obtain the highest recognition accuracy. We have implemented different algorithms and have obtained high accuracy for CNN. CNN is much more reliable than PCA with DT, RF, and KNN. KNN is a lazy algorithm, and it checks all the instances in the dataset for prediction while CNN recognizes in very little time from its model. The other reason is that we have used 41,320 images for 90 classes for PCA, and for CNN, we have used ten classes and 30 images per class, and we obtained good accuracy compared to PCA. We collected more than 41,320 images. We will enhance this system by making it a complete security system. We recognize a single face from the image; our next step is to recognize multiple faces in a live-streaming video.

### Data Availability

The data are available with the first author and will be provided on request for research purposes.

### Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

### References

- [1] A. Deepali, A. Colburn, G. Faigin, L. Shapiro, and B. Mones, "Modeling stylized character expressions via deep learning," in *Proceedings of the Asian Conference on Computer Vision*, pp. 136–153, Springer, Taipei, Taiwan, November 2016.
- [2] S. T. Saste and S. M. Jagdale, "Emotion recognition from speech using MFCC and DWT for security system," vol. 1, pp. 701–704, in *Proceedings of the International conference of*

- Electronics, Communication and Aerospace Technology (ICECA)*, vol. 1, IEEE, Coimbatore, India, April 2017.
- [3] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis et al., "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, 2001.
  - [4] J. Edwards, H. J. Jackson, and P. E. Pattison, "Emotion recognition via facial expression and affective prosody in schizophrenia," *Clinical Psychology Review*, vol. 22, no. 6, pp. 789–832, 2002.
  - [5] S. Umer, B. Chandra Dhara, and B. Chanda, "Face recognition using fusion of feature learning techniques," *Measurement*, vol. 146, 2019.
  - [6] C. Lin and K.-C. Fan, "Human face detection using geometric triangle relationship," vol. 2, pp. 941–944, in *Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 2, pp. 941–944, IEEE, Barcelona, Spain, September 2000.
  - [7] K. T. Talele and S. Kadam, "Face detection and geometric face normalization," in *Proceedings of the TENCON 2009-2009 IEEE Region 10 Conference*, pp. 1–6, IEEE, Singapore, January 2009.
  - [8] J. Miao, W. Gao, Y. Chen, and J. Lu, "Gravity-center template based human face feature detection," in *Proceedings of the Advances in Multimodal Interfaces - ICMI 2000*, pp. 207–214, Springer, Beijing, China, October 2000.
  - [9] T. V. Vivek and G. R. M. Reddy, "A hybrid bioinspired algorithm for facial emotion recognition using CSO-GA-PSO-SVM," in *Proceedings of the 5th Int. Conf. Commun. Syst. Netw. Technol.*, pp. 472–477, Gwalior, India, April 2015.
  - [10] H. Ali, M. Hariharan, S. Yaacob, and A. H. Adom, "Facial emotion recognition based on higher-order spectra using support vector machines," *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 6, pp. 1272–1277, 2015.
  - [11] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
  - [12] X. Wang, Q. Ruan, Y. Jin, and G. An, "Three-dimensional face recognition under expression variation," *EURASIP Journal on Image and Video Processing*, vol. 6, pp. 281–289, 2014.
  - [13] C. C. Queirolo, L. Silva, O. R. P. Bellon, and M. P. Segundo, "3d face recognition using simulated annealing and the surface interpenetration measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 2, pp. 206–219, 2010.
  - [14] C. C. Queirolo, L. Silva, O. R. P. Bellon, and M. P. Segundo, "3d face recognition using the surface interpenetration measure: a comparative evaluation on the frgc database," in *Proceedings of the International Conference on Pattern Recognition*, pp. 1–5, Tampa, FL, USA, December 2008.
  - [15] N. Alyüz, B. Gökberk, and L. Akarun, "Regional registration for expression resistant 3-d face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 425–440, 2010.
  - [16] C. Goodall, "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society: Series B*, vol. 53, no. 2, pp. 285–321, 1991.
  - [17] B. Gokberk, H. Dutagaci, A. Ulas, L. Akarun, and B. Sankur, "Representation plurality and fusion for 3-d face recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 1, pp. 155–173, 2008.
  - [18] C. Ding and D. Tao, "A comprehensive survey on pose-invariant face recognition," *ACM Transactions on intelligent systems and technology (TIST)*, vol. 7, no. 3, 2016.
  - [19] T. Heseltine, N. Pears, and J. Austin, "Evaluation of image preprocessing techniques for eigenface-based face recognition," in *Proceedings of the 2nd International Conference on Image and Graphics*, vol. 4875, July 2002.
  - [20] S. Vignesh, K. V. S. N. L. Manasa Priya, and S. S. Channappayya, "Face image quality assessment for face selection in surveillance video using convolutional neural networks," in *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, IEEE, Orlando, FL, USA, December 2015.
  - [21] G. Tudavekar, S. R. Patil, and S. S. Saraf, "Dual-tree complex wavelet transform and super-resolution based video inpainting application to object removal and error concealment," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 4, pp. 314–319, 2020.
  - [22] Y. L. Tian, T. Kanade, and J. F. Cohn, *Handbook of Face Recognition*, Springer, New York, NY, USA, pp. 247–275, 2005.
  - [23] T. Kanade, *Picture Processing System by Computer Complex and Recognition of Human Faces*, PhD. Thesis, Kyoto University, Japan, 1973.
  - [24] X. Xiang, J. Yang, and Q. Chen, "Color face recognition by PCA-like approach," *Neurocomputing*, vol. 152, pp. 231–235, 2015.
  - [25] R. Gottumukkal and V. K. Asari, "An improved face recognition technique based on modular PCA approach," *Pattern Recognition Letters*, vol. 25, no. 4, pp. 429–436, 2004.
  - [26] K. Sushel, V. B. Semwal, and R. C. Tripathi, "Real time face recognition using adaboost improved fast PCA algorithm," 2011, <https://arxiv.org/abs/1108.1353>.
  - [27] C. Li, J. Liu, A. Wang, and K. Li, "Matrix reduction based on generalized PCA method in face recognition," in *Proceedings of the 5th International Conference on Digital Home*, IEEE, Guangzhou, China, November 2014.
  - [28] C. Liu, T. Zhang, D. Ding, and C. Lv, "Design and application of Compound Kernel-PCA algorithm in face recognition," in *Proceedings of the 35th Chinese Control Conference (CCC)*, July 2016.
  - [29] M. Peter, J.-L. Minoi, and H. M. H. Irwandi, *3D Face Recognition Using Kernel-Based PCA Approach*, Springer, Singapore, 2019.
  - [30] L. H. Tran and L. H. Tran, "Tensor sparse PCA and face recognition: a novel approach," 2019, <https://arxiv.org/abs/1904.08496>.
  - [31] M. E. Rala, "Feature extraction using PCA and kernel-PCA for face recognition," in *Proceedings of the 8th International Conference on INFOrmatics and Systems Computational Intelligence and Multimedia Computing Track*, Giza, Egypt, May 2012.
  - [32] M. Abdullah, M. Wazzan, and S. Bo-saeed, "Optimizing face recognition using PCA," 2012, <https://arxiv.org/abs/1206.1515>.
  - [33] M. Sharif, A. Khalid, M. Raza, and S. Mohsin, "Face detection and recognition through hexagonal image processing," *Sind University Research Journal*, vol. 44, no. 2, pp. 541–548, 2012.
  - [34] A. Azeem, M. Sharif, M. Raza, and M. Murtaza, "A survey: face recognition techniques under partial occlusion," *The International Arab Journal of Information Technology*, vol. 11, no. 1, pp. 1–10, 2014.
  - [35] M. Murtaza, M. Sharif, M. Raza, and J. Shah, "Face recognition using adaptive margin Fisher's criterion and linear discriminant analysis," *The International Arab Journal of Information Technology*, vol. 11, no. 2, pp. 1–11, 2014.

- [36] R. Agada and J. Yan, "Edge based mean LBP for valence facial expression detection," in *Proceedings of the 2015 IEEE Int Conf Electr Comput Commun Technol ICECCT*, Coimbatore India, March 2015.
- [37] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1891–1898, IEEE, Columbus, OH, USA, June 2014.
- [38] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," 2014, <https://arxiv.org/abs/1411.7923>.
- [39] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [41] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 779–788, Las Vegas, NV, USA, June 2016.
- [42] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [43] W. Zhao, A. Krishnaswamy, and R. Chellappa, "Discriminant analysis of principal components for face recognition," in *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 336–341, Nara, Japan, April 1998.
- [44] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.
- [45] P. J. Phillips, "Support vector machines applied to face recognition," in *Proceedings of the Advances in Neural Information Processing Systems II*, pp. 803–809, Denver, Colorado, USA, 1999.
- [46] S.-H. Lin, S.-Y. Kung, and L.-J. Lin, "Face recognition/detection by probabilistic decision-based neural network," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 114–132, 1997.
- [47] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2004)*, vol. 2, pp. 985–990, Budapest, Hungary, July 2004.
- [48] H.-J. Rong, G.-B. Huang, and Y.-S. Ong, "Extreme learning machine for multi-categories classification applications," in *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*, pp. 1709–1713, IEEE World Congress on Computational Intelligence, Hong Kong, China, June 2008.
- [49] Y. Lan, Y. C. Soh, and G.-B. Huang, "Extreme learning machine-based bacterial protein subcellular localization prediction," in *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*, pp. 1859–1863, IEEE World Congress on Computational Intelligence, Hong Kong, China, June 2008.
- [50] T. Helmy and Z. Rasheed, "Multi-category bioinformatics dataset classification using extreme learning machine," in *Proceedings of the Eleventh Conference on Congress on Evolutionary Computation (CEC 09)*, pp. 3234–3240, Trondheim, Norway, May 2009.
- [51] C.-W. T. Yeu, M.-H. Lim, G.-B. Huang, A. Agarwal, and Y.-S. Ong, "A new machine learning paradigm for terrain reconstruction," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 382–386, 2006.
- [52] Y. Yang, J. Li, and L. Peng, "Multi-robot path planning based on a deep reinforcement learning DQN algorithm," *CAA Transactions on Intelligence Technology*, vol. 5, no. 3, pp. 177–183, 2020.
- [53] A. S. Mian, "Representations and matching techniques for 3d free-form object and face recognition," Ph. D. Thesis, , Ph. D. Thesis 2006.
- [54] P. Kamency, M. Benco, T. Mizdos, and R. Radil, "A new method for face recognition using convolutional neural network," *Digital Image Processing And Computer Graphics*, vol. 15, pp. 663–672, 2017.
- [55] T. Bao, C. Ding, M. Zhu, and Y. Wang, *Face Recognition in Real-World Surveillance Videos with Deep Learning Method*, Department of Information and Technology University of Science and Technology of China Hefei, China, 2017.
- [56] A. Gorman, *Cctv Facial Recognition Analysis*, Santa Clara University COEN 150 Project, California, CL, USA, 2011.
- [57] Y. Li, W. Zheng, Z. Cui, and T. Zhang, *Face Recognition Based on Recurrent Regression Neural Network* Research Center for Learning Science, Southeast University, Nanjing, Jiangsu, China, 2017.
- [58] J. Yu, K. Sun, F. Gao, and S. Zhu, "Face biometric quality assessment via light cnn," pp. 25–32, 2018, *Pattern Recognition Letters* 107.
- [59] X. Yan, S. Hu, P. Y. Niyogi, and H. J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 328–340, 2005.
- [60] M. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13, pp. 1450–1464, 2002.