Tech Science Press

Criminal Persons Recognition Using Improved Feature Extraction Based Local Phase Quantization

P. Karuppanan[1,*] and K. Dhanalakshmi[2]

[1]Department of CSE, Pandian Saraswathi Yadav Engineering College, Sivaganga, 630561, Tamilnadu, India
[2]Department of CSE, PSNA College of Engineering and Technology, Dindigul, Tamilnadu, India
*Corresponding Author: P. Karuppanan. Email: prabukaruppanan@gmail.com

Abstract: Facial recognition is a trending technology that can identify or verify an individual from a video frame or digital image from any source. A major concern of facial recognition is achieving the accuracy on classification, precision, recall and F1-Score. Traditionally, numerous techniques involved in the working principle of facial recognition, as like Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Subspace Decomposition Method, Eigen Feature extraction Method and all are characterized as instable, poor generalization which leads to poor classification. But the simplified method is feature extraction by comparing the particular facial features of the images from the collected dataset namely Labeled faces in the wild (LFW) and Olivetti Research Laboratory (ORL) dataset. In this paper, the feature extraction is based on local phase quantization with directional graph features for an effective optimal path and the geometric features. Further, Person Identification based deep neural network (PI-DNN) has proposed are expected to provide a high recognition rate. Various performance metrics, such as recognition rate, classification accuracy, accuracy, precision, recall, F1-score is evaluated. The proposed method achieves high-performance values when it is compared with other existing methods. The novelty of this paper explains in understanding the various features of different types of classifiers used. It is mainly developed to recognize the human faces in the crowd, and it is also deployed for criminal identification.

Keywords: Face recognition; face detection; criminal identification; local phase quantization (LPQ); directional graph; deep neural network

1 Introduction

Face recognition can also be defined as biometric artificial intelligence relied on the application that is exclusively designated to recognize a person by exploring sequences based on the individual's facial shape and texture. Even though facial recognition has a wide variety of applications, it is typically essential to use it in the detection of criminal and forensic investigations and inquiries. Applying the digital images for the identification of victims can be challenging at specific occasions. Especially when the individual is wearing a mask or tattoo, or if the environment has background noises, camera distortion, insufficient

storage, inadequate computing techniques, occlusion, detection of identical images and low-resolution images [1]. Nowadays, deep learning explored the hidden sights of advanced computing applications and executed it in the right place to get an efficient expected output. It is trained similar to the human brain to analyze multiple human faces and store it in a database for future use. Generally, human utilizes all his sensory organs to recollect and analyze the input data, a similar computer process, and the input captured images into various features by a different algorithm. To get the exact desired result and confirm it by comparing it with the original picture by an integral section of biometrics and matching of essential traits. The system involved dual methods, such as face detection and face recognition [2]. Face detection is used to search and find one face by image processing, and face recognition is performed by comparing and analyzing the processed image with similar images from the database.

The method of face recognition in the proposed research comprises feature extraction and feature classification methods. In feature extraction, Phase quantization, directional graph-based process, and Geometric based feature extraction are performed. The feature classification is achieved by using Convolutional neural network (CNN) for criminal identification. The mandatory methods for effective facial recognition should calibrate the facial expression from the given images. The facial expression can be derived from eyes, mouth, eyebrows, cheeks, and it's represented as salient regions. But few emotions like surprise and sadness acquired with only one salient feature, but anger and fear cannot be obtained from a single salient feature. In feature extraction, Phase quantization is applied to extract the features from blurred images by short Fourier transform to analyze the structural properties of the image with promising values of accuracy and efficacy [3]. Directional graph-based and geometric based methods in feature extraction are applied to extract the facial feature in Region of interest (ROI) along with error filtering techniques to provide prominent face components. It is used to locate the edges and relative size and position of mandatory expression components such as mouth, eyes, nose, and eyebrows. Then it is differentiated into the unimportant and meaningful part and also the conversion of grayscale distribution into feature vector based on the value of the pixel. The feature extraction is based on image pixel distribution, orientation selectivity, and spatial localization. The extracted images are then fed into CNN for further classification by deep learning. It is composed of multiple layers to detect necessary edges, intricate shapes, and by further processing, the final layer can catch the entire face and confirm it by verifying with the high dimensional dataset. Hence it is mostly preferred in real-time applications.

Face Recognition is mainly deployed in genetic engineering and biometric devices for identification, authorization, verification, and authentication. Many companies have adopted face recognition for security purposes by implementing CCTV to manages and control the organization. Apart from security system, it have other exciting applications such as unlocking the phones, more refreshing advertising, finding missing people, protecting the law enforcement, identifying person on social media, diagnosing various types of diseases, spotting VIP at events, protecting school and government institutions from threat, and controls the access to sensitive regions. In recent days Deep learning techniques are highly utilized in recognition of human face from the large datasets [4]. By considering this the major contribution of the study involves,

- The applied machine learning algorithm, such as Local Phase quantization with directorial graph-based features and Geometric based feature extraction have been applied in selected LFW and ORL datasets to perform effective human face recognition.
- To find the optimal path to achieve the best features using directional graph-based features.
- To implement the Person identification based Deep Neural network considered an effective classifier and applied in the real-time dataset for criminal identification.

The organization of the research is as follows. Section 2 explains the related work of face recognition on different feature extraction and classification methods. Section 3 is based on the detailed description of the

proposed work. Section 4 represents the analysis of the performance of the proposed method, and Section 5 concludes the results of the research work and explains its future extension of the work.

2 Literature Review

In modern ages, face recognition has achieved extensive attention from research and market, but still, it has some challenging task to implement in real-time applications. In past decades, a wide variety of face recognition machine algorithm is developed. But the proposed techniques in feature extraction and classification achieves high accuracy in face recognition. The paper is based on CNN in an image recognition process for the input of the Gait energy picture. It is composed of dual pattern trios' images of pooling, convolutional, and normalization cover and double connected layers with high efficacy. Classification is the last stage of the face expression recognition where the classifier divides the expression like sad, smile, surprise, fear, anger, neutral and disgust. This method involves silhouette extraction, extraction of features, and matching of images [5,6]. The silhouette extraction is made to avoid being affected by clothing, texture, and color. Then the Gait energy image (GEI) picture is fed into CNN of 88 * 128 pixels [7]. The face recognition method involves face area acquisition, localization, and segmentation to define the facial expression, face shape, and position of components, and the quality of the image [8]. Then feature detection and extraction of the face are obtained by geometric approach by designing feature vectors to locate the corners of the various face feature components. Then it is fed as the input image for the backpropagation neural network model for facial expression recognition. But this method fails to work if there is any alteration in eye areas [9].

The link prediction in face recognition is divided into two types, such as feature extraction and feature learning methods. Similarity-based methods, probabilistic methods, and likelihood-based methods fall under feature extraction techniques. In similarity-based models, it follows global, local, and Quasi-local approaches by calculating the similarity between nodes. Depending on the structure of the network, a statistical model is designed to calculate the possibility of unnoticed connections to happen by using maximum likelihood models such as Hierarchical structural design and stochastic block model. These techniques are time-consuming and derive an accurate result. In feature learning methods, CNN extract features from the graph to check the connection between linked nodes and extract the data from the topology of the system [10]. The awareness of this learning model is to establish the structure of the localized neighbor by disintegrated feature data as an alternative to creating the entire graph by multiple source link prediction of input images [11]. The detection and localization of blurred face using local phase quantization applying Fourier transform phase in resident neighborhoods [12]. Under certain conditions, the period can be represented as a blur invariant attribute. So in the analysis, a histogram of local phase quantization labels calculated in resident areas is executed as face descriptor, related to broadly applied local binary pattern for a description of face images. But to limit the impact of lighting alteration in face pictures, procedure on illumination normalization is applied that comprises of contrast equalization, gamma correction, and Gaussian filtering that increases the result of recognition accuracy [13]. Some investigations made on a pair of parent and child to verify the effectiveness of the local phase quantization. The standardized and cropped images of parent and child were fed as input and converted to grayscale. Then local phase quantization is applied to retrieve the local features of every mode [14]. Encrypted images are differentiated into k non-overlap rectangle scratches, and histogram values define each stretch. Each histogram has allocated to the high dimensional feature vector. Then cosine similarity is analyzed between the images by projecting into transformed subspace, and the resultant score is compared with threshold value from Receiver Operator Characteristic (ROC) by performance measures. With the obtained result, it is decided whether the person belongs to the same family or not [15]. The face is an essential attribute considered in a security system. Even though it is applied vastly, there is some limitation such as illumination, pose, and condition of the pictures [16]. To overcome this issue,

face recognition is performed under non-uniform illumination by CNN that can learn exclusively local patterns from information is utilized for identification of a face. The symmetry of face data is processed to eliminate and enhance the performance of the system by accepting the horizontal reflection of the face images. The experiment was made using the Yale dataset under various illumination settings and showed moderate enhancement in the performance of CNN. The production remains undisturbed by the horizontal reflection of images. But it is failed when it is applied with fisher discriminant analysis and fisher linear discriminant analysis that is based on Gabor wavelet transform. Human gait is soft biometric that support to identify an individual by their walking gestures [17]. This research focused on deep neural network for the effective face recognition system in which the new face alignment technique has been implemented for face key point localization. The deep features dimensionality has reduced by the principal component analysis for removing the contaminated and redundant visual features. The feature vectors similarity has been evaluated by Bayesian context and the better classification accuracy attained. Different face recognition attacks have been handled based on different frameworks. [18] Personal face images have been differentiated by Deep Face. 4000 unique identities have been verified in this study. Deep convolutional neural networks have been used for learning the deep identification DeepID. This study utilizing the LFW dataset. [19] AlexNet network has been utilized to evaluate the ORL dataset for realizing the face recognition. The proposed convolutional kernel network with recognition rate depicts higher rate compared with AlexNet network in face recognition. The gender-based classification is also made possible using a supervised machine learning approach. The diverse classifier used in this approach is the support vector machine, convolutional neural network, and adaptive boost algorithm. The input images are fed into pre-processing for removing noisy information, and feature extraction is based on a geometric approach to extract meaningful data from the processed images. It is fed into the combination of a classifier to get the resultant output. It is used to recognize the gender of the picture.

3 Proposed Methodology

In the proposed work, the features from the image are extracted using Local phase quantization based on directional graph features and geometric based features and classified using Person identification based Deep Neural Network (PI-DNN) whereas it is not found in previous texture and feature extraction algorithms like edge based method, geometric feature-based method, texture feature-based method, patch-based method, global and local feature-based method. Gabor filter, Vertical Time Backward, Gaussian Filter, Weber Local Descriptor. Supervised Descent Method Discrete Contourlet Transform, Graphics-processing unit based Active Shape Model, Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Local Binary Pattern (LBP) are the text descriptors. In this study further, the criminal faces are identified from the selected LFW and ORL datasets using the proposed model shown in Fig. 1.

3.1 Local Phase Quantization for Feature Extraction

In general, Local Phase quantization is insensitive to centrally symmetric texture images. So this research considered both spatial and temporal coordinates. In statistical analysis, to overcome the loss of generality, the covariance matrix and cross-correlation are evaluated for the transformed coefficients. This method is based on the property of blur invariance of the Fourier phase spectrum. It utilizes local phase data that is derived using a two-dimensional discrete Fourier transform or Short-Term Fourier Transform (STFT) calculated on a rectangular matrix $M \times M$ at each pixel position x of image f(x). It can be formulated as below Eq. (1),

$$F(q, a) = \sum_{b \in N_a} f(a - b)e^{-j2\pi q^T b} = w_q^T f_a \tag{1}$$

Here, w_q represents the basis vector of frequency q. f_a Vector represents all image samples from neighborhood N_a.
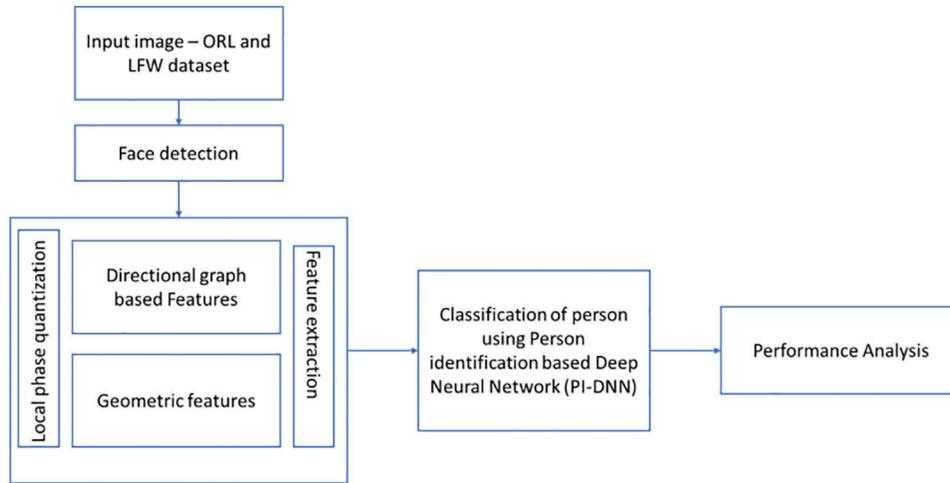
Figure 1: Architecture of proposed system

Local Phase quantization is defined as follows,

The original image is compared with the input image, and it is formulated as follows,

$$g(a, b) = f(a, b) \otimes h(a, b) + n(a, b) \tag{2}$$

From Eq. (2), where $h(a, b)$ is represented as a Point Spread Function (PSF), $n(a, b)$ is described as an additive noise, then the Fourier transformation is performed to convert the image from the range of high intensity to low intensity. It is formulated as below,

$$G(m, n) = F(m, n).H(m, n) \tag{3}$$

From Eq. (3), where $G(m, n)$, $F(m, n)$, $H(m, n)$ Fourier are transforms of $g(a, b)$, $f(a, b)$, and $h(a, b)$ respectively. In order to predict the intensity range of an image, phase information and magnitude can be extracted as follows,

$$\angle G(a, b) = \angle F(a, b) + \angle H(a, b) \tag{4}$$

From Eq. (4), where $\angle G(a, b)$, $\angle F(a, b)$, $\angle H(a, b)$ are phases of $g(a, b)$, $f(a, b)$, and $h(a, b)$ respectively.

To analyze the impulsive response of an image, Point Spread Function is calibrated as below,

$$\angle H(a, b) = \begin{cases} 0, & if \ H(a, b) \geq 0 \\ \pi, & otherwise \end{cases}, \tag{5}$$

To calculate the relative translation between two pixels, phase correlation is estimated as below,

$$\angle G(a, b) = \angle F(a, b), \ for \ all \quad H(a, b) \geq 0. \tag{6}$$

Then, Short Term Fourier Transform is executed at first to extract phase information for every pixel. It can be calculated as below,

$$G(a,\ b) = \sum_{a \in N_b} \sum_{b \in N_b} g(a,\ b) e^{\frac{-j2\pi(ma+nb)}{M}} \tag{7}$$

The final results are organized as below,

$$V = [G(m_1),\ G(m_2),\ G(m_3),\ G(m_4)] \tag{8}$$

and

$$W = [Re\{V\},\ Im\{V\}] \tag{9}$$

From Eq. (9), where $Re\{V\}$, $Im\{V\}$ represent real and imaginary parts correspondingly. Lastly, the encoded value (x) distribution histogram within an image is built to attain the local phase quantization in two hundred and fifty six dimensions. It is given by the below Eq. (10).

$$x = \sum_{i=1}^{8} q_i 2^{i-1} \tag{10}$$

Here q_i denotes the quantization corresponding to the ith element in W and is given by the below Eqs. (11) and (12).

$$q_i = \begin{cases} 1, & if\ W_i \geq 0 \\ 0, & otherwise \end{cases} \tag{11}$$

$$G_a = V^T F_a \sum_{i=0}^{N-1} f(i) g(a,\ b) \tag{12}$$

Though, the local quantization measures the short-term Fourier Transform coefficients, this strategy tends to confirm the blur insensitivity with a slight minimum discriminative power. On the other hand, it is just an indication of STFT coefficients that impact the texture characterization.

3.1.1 Directional Graph-based Features

In this process, the input image is represented as a graph where each pixel in the picture is mapped as times series of respective vertex [20]. In this context, the image retrieval concept rely on the Graph Theory (GT). Moreover, the edge detection is implemented as an alternative to the map based representation. The main aim of the edge concentration is to obtain the rapid alterations in the brightness of the image. This representation permits a direct accessibility to the picture's edge node without segmentation and edge point search. Another advance is less data consumption; only data for nodes and their connections are needed, which is important in large database applications for good scalability. Typically, a time series $\{x_i\}_{(i=1,\ldots,n)}$ is mapped into a graph $G(V,\ E)$ where a time, Point x_i is mapped into a node $v_i \in V$. The relation between any two points (x_i, x_j) is represented as an edge $e_{ij} \in E$, and the value is defined as below,

$$e_{ij} = \begin{cases} 1, & x_k < x_i \wedge \quad x_k < x_j \quad \forall k \in (i,\ j) \wedge i < j \\ 0, & otherwise \end{cases} \tag{13}$$

From Eq. (13), Where e_{ij} implies the presence of edge. If $e_{ij} = 0$, implies that the edge does not exist. There are several steps in assessing directional graphs, which are explained further.

Directional Graph Generation*:*

Step 1: Link Density for Directed Graph

The increase in the number of edges increases the graph density. It can be formulated as below,

$Edges = size(E_{ij})$ // Number of formed edges,

$vertices = size(V_{ij})$ // Number of formed vertices,

$$Dense_{graph} = \frac{2 * Edges}{vertices\ (vertices\ - 1)} \tag{14}$$

Step 2: Average of closeness centrality

The graph can effectively detect the nodes centrality that have capability to spread the information by calculating the closeness of centrality. It can be formulated as below,

$$Closeness_{central}(a) = \frac{size(E_{ij})}{D(a)} \tag{15}$$

where, $D\ (a) = \lim\limits_{a \to 1:size\ (V_{ij},1)} \sum\limits_{b=1}^{size\ (V_{ij},1)} Dist(a,\ b);$

$Dist\ (a,\ b) = \lim\limits_{b \to 1:size\ (V_{ij},1)} \sqrt{(X(V_{ij}(a)) - X(V_{ij}(b)))^2}$

Step 3: Graph Entropy

To characterize the texture of input image, graph entropy is calculated. It can be formulated as below,

$$Entropy = -\sum_{i=1}^{n} P(i) * \log(P(i)) \tag{16}$$

where,

P-Degree Distribution of graph, $P = \frac{n_k}{n}$

n_k-number of k connected nodes to the particular n^{th} node.

n-Number of vertices

Step 4: Average Distribution Weight of the Graph

To detect the similarities or dissimilarities between the vertices connected by the edges, weight is calculated using the below Eq. (17),

$$Weight_{avg} = mean(W) \tag{17}$$

From Eq. (17), where W can be represented as the weight of all edges in a graph.

Step 5: Graph Hilbert Transform (GHT):

To convert the image and to extract the global features, the Hilbert transform method is applied in this work. The Hilbert transform of the graph of a function f(x) can be defined as below,

$$F(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(x)}{t - x} dx \tag{18}$$

When Hilbert transform is computed in more than a single dimension and if any of the dimensions remain unchanged (constant), then the transform will be a numerical value near to zero or zero. Thus, the Hilbert transform corresponding to the constant is 0.

3.1.2 Geometric Features

In a geometric-based approach, the local features that are based on local statistics and locations like bodily features are extracted. It can be implemented in the following ways.

- Convex Area Of Image
- Eccentricity
- Filled Area
- Axis Length
- Orientation

Convex Area of Image

The binary image covering the convex area gives the outline of the curved image. The convex area can be denoted as several pixels in a convex image. The convex area can be computed as shown in Fig. 2.
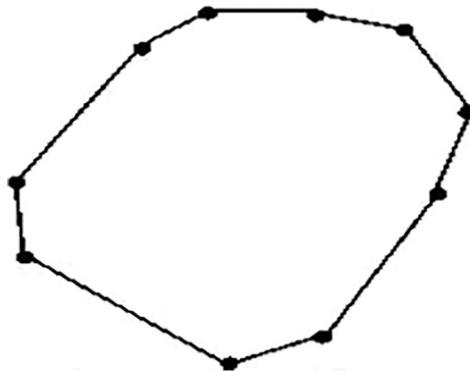
Figure 2: Convex area of image

The Eccentricity of an Image

Eccentricity is the ratio of length off short or minor axis to the length of the major or long axis of an image. For a linked region of a digital image, it is defined by neighbor graph and calibrated metrics. The number that measures and look of the ellipse by illustrating its flatness is also called as eccentricity. The eccentricity can be calculated by using the given equation,

$$\text{eccentricity} = \frac{\text{lenght of the short axis}}{\text{length of the long axis}} \tag{19}$$

Filled Area

Full Image can be denoted as a binary image of the same size similar to the bounding box of the region. Filled Area is defined as the number of pixels in that particular filled image.

Length of the Major Axis

The long line's endpoints denotes the major axis. It is constructed through an object or an image. These endpoints are retrieved by computing the distance of the pixel amongst individual boundary pixel grouping in the boundary of an object. It also determines the pair with an improved length. The length of the major axis affords the object length or image value.

Length of the Minor Axis

The minor axis endpoint is the longest line that can be constructed through the image or object, whereas remaining perpendicular along with the major axis. Minor Axis Length gives the value of the image or object width.

Orientation

The angle of the major axis is computed from an angle amongst the major axis and X-axis of the corresponding image. It can differ from 0° to 360°. Orientation refers to the overall direction of the object shape in an image.

3.2 Deep Neural Network for Person Identification

The proposed architecture is based on deep neural network for the person identification (PI-DNN) shown in Fig. 3. From the geometric based feature vectors and local phase quantization, the input is accepted from the PI-DNN. Fully connected dense nodes of 512 comprised in dense layer which are included after the input layer. The dense layer output has expressed as,

$$H_1 - U.V \tag{20}$$

From the above Eq. (20), V-dimensional weight matrix of 25×512. At every epoch, A batch has been selected and the size of A is T. The H_1 dimensional is then calculated as $T \times 512$.
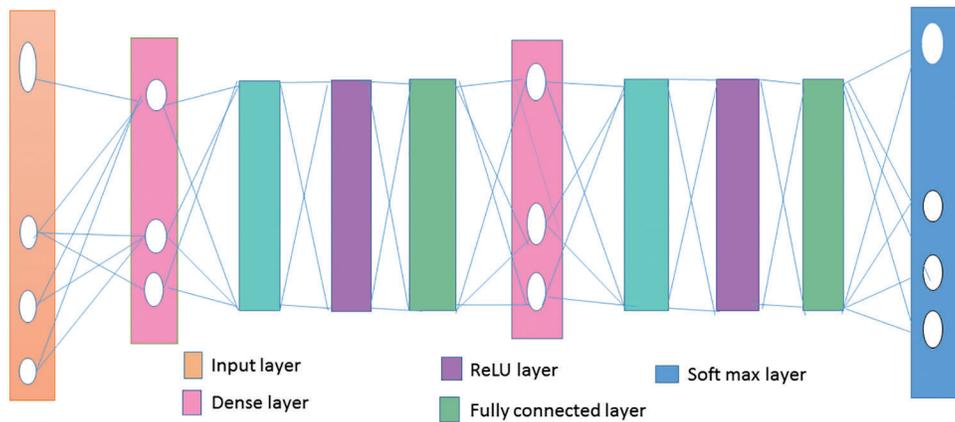
Figure 3: PI-DNN architecture

The over-fitting is prevented and the performance, stability and speed have been improved by the batch normalization layer inclusion after dense layer. The A mean measured at batch normalization layer is measured as,

$$\sigma_A - \frac{1}{T} \sum_{T-1}^{T} H_1(t) \tag{21}$$

From Eq. (21), $H_1(t)$ is defined as Tth row vector of H_1 and 1×512 resulted as the σ_A dimension. The A variance with σ_A measured as,

$$-\gamma_A^2 - \frac{1}{T} \sum_{T-1}^{T} \{H_1(t) - \sigma_A\}^2 \tag{22}$$

From Eq. (22), 1×512 is the γ_A^2 dimension.

Let consider $H_1(t, u)$ as the $H_1(t)$'s uth element is the σ_A and γ_A^2 (u)'s uth element and it is the γ_A^2's uth element. Further, normalized $H_1(T, u)$ resulted in,

$$\widetilde{H_1}(t, u) - \frac{H_1(t, u) - \sigma_A(u)}{\sqrt{\gamma_A^2(u) + \epsilon}} \qquad (23)$$

From Eq. (23), $t \epsilon \{ 1, 2, \ldots, T-1, T\}$, $u \in \{1, 2, \ldots, 511, 512\}$ and small constant value is termed as ε included for numerical stability. $\epsilon = 0.001$ value is set.

In the batch normalization layer output, $\widetilde{H_1}(t, u)$ is termed as,

$$H_2(t, u) - \lambda(u). \widetilde{H_1}(t, u) + \mu(u), \qquad (24)$$

From Eq. (24), the parameters $\mu(u)$ and $\lambda(u)$ have been learned during the training phase based optimization process.

In the proposed PI-DNN model, after the batch normalization layer the activation layer is included. The rectified linear unit activation function-Rectified Linear Unit (ReLU) is utilized at an activation layer which are established for adding non-linearity to proposed PI-DNN model. The activation layer output is expressed as,

$$H_3(t, u) - \max(H_2(t, u), 0), \qquad (25)$$

From Eq. (25), $\max(H_2(t, u), 0)$ is considered as function which returns a for the $H_2(t, u) > 0$ and 0 otherwise.

Dropout layer is included after the activation layer. At the drop out layer, during the training phase, at specified rates the input units have been dropped randomly. The dropout rate is 0.1 in PI-DNN proposed model. As explained above, batch normalization layer, second dense layer, dropout layer and activation layer have been added successfully. Dense nodes in first dense layer is similar as such in the second dense layer. ReLU activation has utilized at second activation layer. In second drop out layer, dropout rate is 0.1. For every person, the classification of labels is included in soft max layer at PI-DNN end of the model. As an activation function, the softmax function has been performed in this layer. Hence, from softmax layer, activation values have been normalized through

$$\widetilde{Z}_\theta - \frac{\exp(Z_\theta)}{\sum_{\theta-1}^{\theta} \exp(z_\theta)}, \ \forall \theta \ \epsilon\{1, 2, \ldots \theta - 1, \theta\} \qquad (26)$$

From Eq. (26), where exp(.)-exponential function, label-θ for θ th person. The normalized activation value is $\widetilde{Z}_\theta$ and the un-normalized activation values are Z_θ. The total number of labels are θ, if the dataset of person identification comprised with total subjects. $1 \times \theta$ dimensional vector refer as softmax layer output in which the vector θ th value is $\widetilde{Z}_\theta$. Sparse categorical cross entropy loss property is applied to an objective function for PI-DNN training.

4 Performance Analysis

The experimental analysis is performed by calculating various performance metrics such as accuracy, precision, recall, F1-score, classification accuracy, and recognition rate. It is verified with a different existing method such as FaceNet, Clinical Document Architecture (CDA), Transform and Two Dimensional Logarithmic (TDL) algorithm k class feature transfer KCFT, Gabor-Based Feature Extraction and Feature Space Transformation (G-FST), and Support Vector Machine (SVM) classifier trained with Histogram of Oriented Gradients (HOG) features and image segmentation (HOG + SVM) () by using LFW and ORL dataset. The criminal images are also analyzed to identify the criminals in the images using the proposed model.

4.1 Dataset Description

- Labeled Faces in the Wild (LFW) dataset contains 13,000 images to study the problem of unconfined face recognition, and it is detected by Viola-Jones face detector. The LFW dataset is obtained from http://vis-www.cs.umass.edu/lfw/.
- ORL dataset contains the pictures captured from 1992 to 1994 of different subjects under various lighting conditions with several facial expression such as open or closed eyes, smiling and not smiling. The ORL dataset is obtained from https://paperswithcode.com/dataset/orl.

4.2 Recognition Rate

The recognition rate is divined as the summation of correctly identified acquisition image by total number acquisition images by various experimental analysis. Tab. 1 shows the exploratory analysis of the recognition rate for the proposed and different existing methods is shown.

Table 1: Multi-directional local gradient descriptor: A new feature descriptor for face recognition using different existing methods and the proposed method

Methods	Recognition rate %
Xing [21]	74.64 to 80.82
DML-eig	82.28 to 87.94
SILD	80.07 to 86.04
ITML	79.98 to 85.94
LDML	80.65 to 86.64
KISSME	83.37 to 88.92
DLML	85.35 to 91.15
PFFT features	97.07
MLGD [22]	97.25
Proposed	99.12

4.3 Classification Accuracy

Classification accuracy is defined as the ratio of total quantity of accurate prediction images to the total quantity of input images by using various experimental analysis. Tab. 2 shows the result of classification accuracy by using existing and proposed methods. It is proved that the proposed method gives give the value of accuracy.

Table 2: Ensemble of texture and shape descriptors by existing and proposed method

References	Classification accuracy (%)
Mowbray et al. [23]	87.55
Ensemble texture descriptors [24]	91.8
DCSD [24]	92.5
Proposed	99.12

4.4 Accuracy

Accuracy is computed to validate the proposed system's performance metrics. It is shown in Tab. 3 that the proposed method gives high efficiency when compared with another existing method. From the Figs. 4 and 5, it is shown that the value of accuracy is high in the proposed method using a realtime dataset from the image Fig. 8 and compared to other existing methods such as FaceNet, CDA, TDL, KCFT, G-FST, and HOG + SVM by using LFW and ORL dataset.

Table 3: Enhanced human face recognition using Local Binary Pattern Histogram (LBPH), descriptor, Multi k-Nearest Neighbor algorithm (KNN), and back-propagation, neural network and proposed method

Reference	Method	% Accuracy
[25]	PCA+BPNN	88
[26]	LDA	89.5
[27]	Gabor+NMF	95
[28]	DCT	91.5
[29]	FKNN	87
[30]	WT+PCA	95
[31]	CASNN, FFNN	86.5
[31]	PCA+DCT+Corr+PIFS	86.8
[31]	LBPH, Multi-KNN & BPN	98
Proposed	LPQ, DG, CNN	98.023
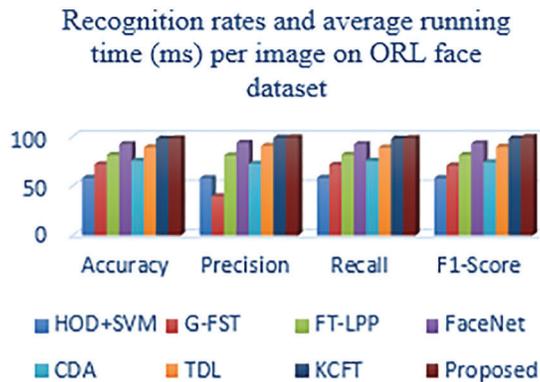
Figure 4: Comparison of performance metrics using the different existing method and proposed method in ORL dataset
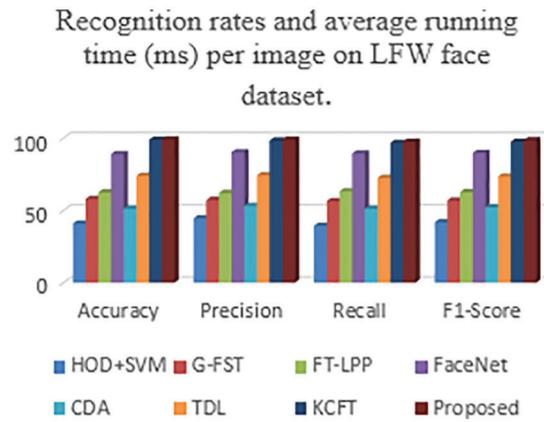
Figure 5: Comparison of performance metrics using the different existing method and proposed method in LFW dataset
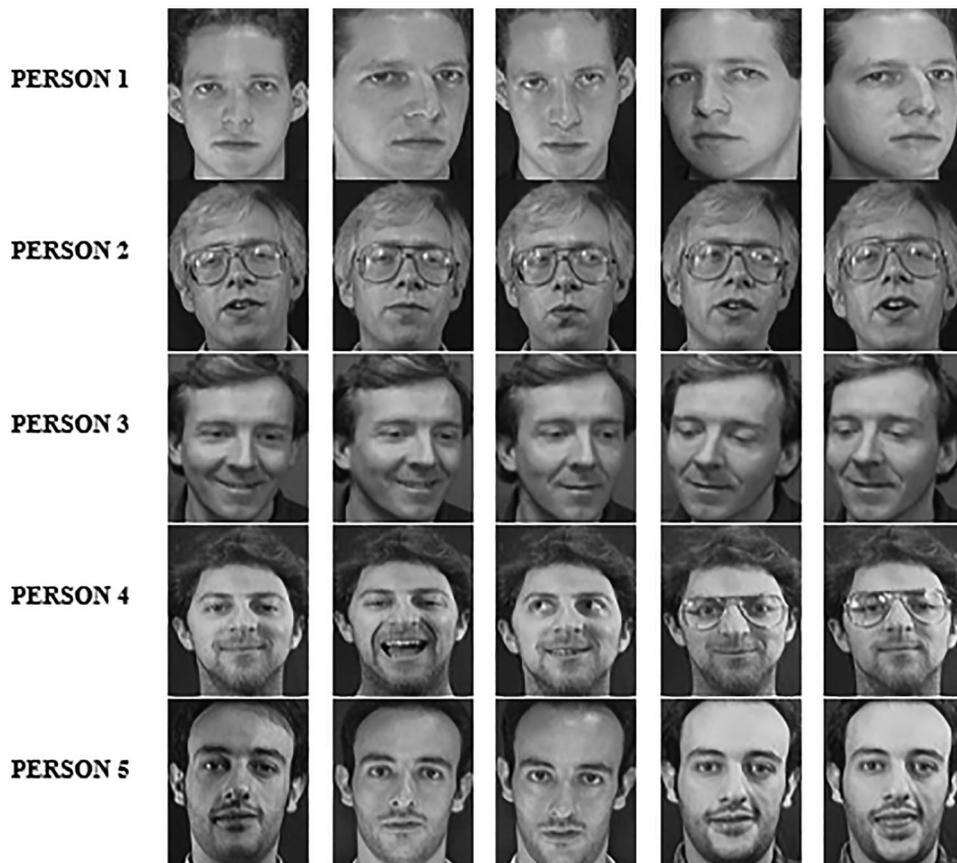
Figure 6: ORL dataset of various person

4.5 Precision

Precision is a significant performance metric calculated to find the positive predicted values. It is the fraction of related instances from the extracted cases. From the Figs. 6 and 7, it is shown that the value of precision is high in the proposed method using a real-time dataset from the image Fig. 8 and compared to other existing methods such as FaceNet, CDA, TDL, KCFT, G-FST, and HOG + SVM by using LFW and ORL dataset.

Figure 7: LFW dataset of various person

4.6 Recall

The recall is termed as sensitivity that is the ratio of total quantity of related instances that are truly retrieved. From the Figs. 9 and 10, it is shown that the value of recall is high in the proposed method using real-time dataset from the image and it is compared to other existing methods such as FaceNet, CDA, TDL, KCFT, G-FST and HOG + SVM by using images of ORL and LFW dataset

4.7 F1-Score

It is one of the critical performance metrics to test the accuracy of the proposed method. It is shortly defined as the weighted harmonic mean of tested values of recall and precision. It is observed that the value of recall is high in the proposed method by using input real-time dataset images from the Fig. 8. It is compared to other existing methods such as FaceNet, CDA, TDL, KCFT, G-FST, and HOG + SVM by using the pictures of ORL and LFW dataset.
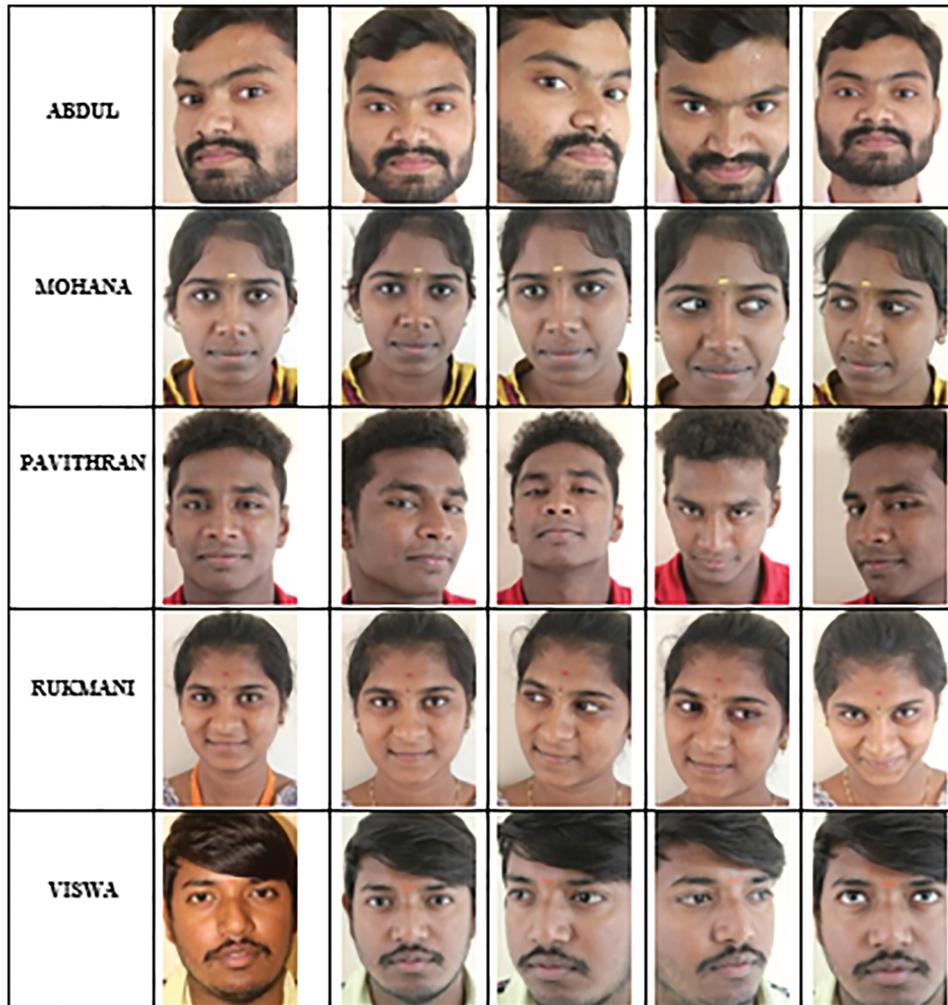
Figure 8: Real-time datasets of various person

Figure 9: LFW criminal dataset

4.8 Comparative Analysis Results Based on Criminals Identification on LFW and ORL Datasets

From the below Tabs. 4 and 5 it shows the comparative analysis of the proposed Multi-directional local gradient descriptor with DNN and existing methods [32] based on LFW dataset. The standard models are resolution face (resface) (0.5) and resface64. The existing models from earlier state of art researches are analyzed and compared with proposed model. The basic model is FaceNet [33] and in LFW, it shows

accuracy as 99.63%. Further, Sphere Face [34] and Center Face [35] shows lesser results because they have been trained on minor dataset namely CASIA Webface [36]. ArcFace [37] techniques outcomes have also been depicted for the comparisons and it attained better performance while utilizing the probabilistic embedding techniques. The ProbFace technique results are shows higher than PFE model. The low quality images are difficult in handling by the smaller model.

Figure 10: ORL criminal dataset

Table 4: Comparative analysis on LFW dataset

Methods	LFW
FaceNet	99.63
CenterFace	99.28
SphereFace	99.42
CosFace	99.73
MobileFaceNet, ArcFace	99.50
LResNet50E IR, ArcFace	99.80
LResNet100E IR, ArcFace	99.77
IR 50, Pytorch, ArcFace	99.78
IR 152, Pytorch, ArcFace	99.82
SphereFace PFE	99.82
ResFace64	99.80
ResFace64 PFE	99.82
ResFace64 ProbFace	99.85
ResFace64(0.5)	99.72
ResFace64(0.5) PFE	99.82
ResFace64(0.5) ProbFace	99.82

Table 5: Comparative analysis on ORL dataset

Algorithm	Features	Rate of recognition
(With minimum samples) [38]	Two-Phase Test Sample Sparse Representation (TPTSSR)	93.33%
Wang et al. [39]	Deep CNN	91.0%
[40]	one-Level DWT/or two-dimensional Locality Preserving Projection (LPP)	93.57%
Liu et al. [41]	CSDL-SRC-Power Norm	95.5%
Gupta et al. [42]	SIFT	95.5%
Proposed	PI-DNN	99.12%

The proposed model has been compared with the several existing approaches on ORL dataset as shown in above table. Higher recognition rate resulted from proposed PI-DNN classifier compared with several techniques on ORL dataset. Due to inefficiency in handling the images occlusion, the existing studies [42] shows not better performances. Assessing the performance of accuracy, precision, recall and F1-score in Deep learning CNN method by testing them on other standard databases and improving the result is considered as the future work.

5 Conclusion

In this research, the novelty in deep neural network for person identification (PI-DNN) has been proposed that gives a high accuracy rate of face recognition by using Local Phase Quantization (LPQ), Geometric-based features and direction graph-based feature extraction methods, and Convolutional Neural Network. This proposed model is used for effective person recognition that is implemented for criminal identification and to evaluate the human faces in crowd taken from LFW and ORL dataset. Various experimental and performance analyses such as classification accuracy, recognition rate, recall, precision, and F1-score performance value of the proposed technique outperforms other existing methods by high efficiency and recognition value ensuring its supremacy over existing methods. This verifies that the proposed approach is more efficient in the feature extraction and more stable, less over fitting the training data or better generalization. The proposed framework performed well with respect to seven considered metrics by showing an outstanding performance than other methods which is confirmed through results. This efficient performance has proved as an effective system for face detection and recognition.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Dutta, A. K. Gupta and N. Narayan, "Identity crime detection using data mining," in *2017 3rd Int. Conf. on Computational Intelligence and Networks (CINE)*, Bhabaneshwar, India, pp. 1–5, 2017.

[2] B. Amos, B. Ludwiczuk and M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," *CMU School of Computer Science*, vol. 6, pp. 1–22, 2016.

[3] Y. Xiao, Z. Cao, L. Wang and T. Li, "Local phase quantization plus: A principled method for embedding local phase quantization into fisher vector for blurred image recognition," *Information Sciences*, vol. 420, pp. 77–95, 2017.

[4] F. Zhao, J. Li, L. Zhang, Z. Li and S. G. Na, "Multi-view face recognition using deep neural networks," *Future Generation Computer Systems*, vol. 111, pp. 375–380, 2020.

[5] P. Kamencay, M. Benčo, T. Mizdos and R. Radil, "A new method for face recognition using convolutional neural network," *Advances in Electrical and Electronic Engineering*, vol. 1, pp. 1–15, 2017.

[6] P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran and R. Jidhesh, "Intelligent face recognition and navigation system using neural learning for smart security in internet of things," *Cluster Computing*, vol. 22, pp. 7733–7744, 2019.

[7] S. Ouellet and F. Michaud, "Enhanced automated body feature extraction from a 2D image using anthropomorphic measures for silhouette analysis," *Expert Systems with Applications*, vol. 91, pp. 270–276, 2018.

[8] O. Laiadi, A. Ouamane, E. Boutellaa, A. Benakcha and A. T. Ahmed, "Kinship verification from face images in discriminative subspaces of color components," *Multimedia Tools and Applications*, vol. 18, pp. 1–23, 2018.

[9] M. A. Abuzneid and A. Mahmood, "Enhanced human face recognition using LBPH descriptor, multi-KNN, and back-propagation neural network," *IEEE Access*, vol. 6, pp. 20641–20651, 2018.

[10] H. El Khiyari and H. Wechsler, "Age invariant face recognition using convolutional neural networks and set distances," *Journal of Information Security*, vol. 8, pp. 174, 2017.

[11] Y. X. Yang, C. Wen, K. Xie, F. Q. Wen and G. Q. Sheng, "Face recognition using the SR-CNN model," *Sensors*, vol. 18, pp. 4237, 2018.

[12] H. Wechsler and H. E. Khiyari, "Age invariant face recognition using convolutional neural networks and set distances," *Journal of Information Security*, vol. 8, no. 3, pp. 174–185, 2017.

[13] Y. Said, M. Barr and H. E. Ahmed, "Design of a face recognition system based on convolutional neural network (CNN)," *Engineering, Technology & Applied Science Research*, vol. 10, pp. 5608–5612, 2020.

[14] S. Garg and R. Kaur, "Face recognition using improved local directional pattern approach for low resolution images," *International Journal of Advanced Research in Computer Science*, vol. 7, pp. 1–12, 2016.

[15] K. Kapoor, S. Rani, M. Kumar, V. Chopra and G. S. Brar, "Hybrid local phase quantization and grey wolf optimization based SVM for finger vein recognition," *Multimedia Tools and Applications*, vol. First Online, pp. 1–39, 2021.

[16] D. Ghimire, J. Lee, Z. -N. Li and S. Jeong, "Recognition of facial expressions based on salient geometric features and support vector machines," *Multimedia Tools and Applications*, vol. 76, pp. 7921–7946, 2017.

[17] C. Li, X. Min, S. Sun, W. Lin and Z. Tang, "Deepgait: A learning deep convolutional representation for view-invariant gait recognition using joint Bayesian," *Applied Sciences*, vol. 7, pp. 210–232, 2017.

[18] S. S. Mohamed, W. A. Mohamed, A. Khalil and A. Mohra, "Deep learning face detection and recognition," *International Journal of Advanced Science and Technology*, vol. 29, pp. 1–6, 2020.

[19] J. X. Tong, H. Li and S. L. Yin, "Research on face recognition method based on deep neural network," *International Journal of Electronics and Information Engineering*, vol. 12, pp. 182–188, 2020.

[20] E. C. Mutlu and T. A. Oghaz, "Review on graph feature learning and feature extraction techniques for link prediction," *Computer Science*, vol. 22, no. 2, pp. 45–56, 2020.

[21] C. Ding, J. Choi, D. Tao and L. S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 518–531, 2015.

[22] V. C. Kagawade and S. A. Angadi, "Multi-directional local gradient descriptor: A new feature descriptor for face recognition," *Image and Vision Computing*, vol. 83, pp. 39–50, 2019.

[23] S. D. Mowbray and M. S. Nixon, "Automatic gait recognition via Fourier descriptors of deformable objects," in *Proc. Int. Conf. on Audio-and Video-Based Biometric Person Authentication*, Guildford, UK, pp. 566–573, 2003.

[24] P. VenkateswarLal, G. R. Nitta and A. Prasad, "Ensemble of texture and shape descriptors using support vector machine classification for face recognition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 32, pp. 1–8, 2019.

[25] H. S. Dadi, G. K. M. Pillutla and M. L. Makkena, "Face recognition and human tracking using GMM, HOG and SVM in surveillance videos," *Annals of Data Science*, vol. 5, pp. 157–179, 2018.

[26] L. Li, H. Ge, Y. Tong and Y. Zhang, "Face recognition using gabor-based feature extraction and feature space transformation fusion method for single image per person problem," *Neural Processing Letters*, vol. 47, pp. 1197–1217, 2018.

[27] J. Pan, X. S. Wang, and Y. H. Cheng, "Single-sample face recognition based on LPP feature transfer," *IEEE Access*, vol. 4, pp. 2873–2884, 2016.

[28] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Transactions on Multimedia*, vol. 17, pp. 2049–2058, 2015.

[29] Y. Zhang and H. Peng, "Sample reconstruction with deep autoencoder for one sample per person face recognition," *IET Computer Vision*, vol. 11, pp. 471–478, 2017.

[30] J. Zeng, X. Zhao, J. Gan, C. Mai, Y. Zhai et al., "Deep convolutional neural network used in single sample per person face recognition," *Computational Intelligence and Neuroscience*, vol. 2018, pp. 1–17, 2018.

[31] R. Min, S. Xu and Z. Cui, "Single-sample face recognition based on feature expansion," *IEEE Access*, vol. 7, pp. 45219–45229, 2019.

[32] K. Chen, Q. Lv, T. Yi and Z. Yi, "Reliable probabilistic face embeddings in the wild," *Transfer Learning for Human &AI*, vol. 1, pp. 1–10, 2021.

[33] F. Schroff, D. Kalenichenko and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Boston, USA, pp. 815–823, 2015.

[34] W. Liu, Y. Wen, Z. Yu, M. Li and B. Raj, "Sphereface: deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Honolulu, USA, pp. 212–220, 2017.

[35] Y. Wen, K. Zhang, Z. Li and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Proc. European Conf. on Computer Vision*, Amsterdam, The Netherlands, pp. 499–515, 2016.

[36] Y. Sun, X. Wang and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1891–1898, 2014.

[37] J. Deng, J. Guo, N. Xue and S. Zafeiriou, "Arcface: additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, CA, USA, pp. 4690–4699, 2019.

[38] J. Ke, Y. Peng, S. Liu, J. Li and Z. Pei, "Face recognition based on symmetrical virtual image and original training image," *Journal of Modern Optics*, vol. 65, pp. 367–380, 2018.

[39] W. Wang, J. Yang, J. Xiao, S. Li and D. Zhou, "Face recognition based on deep learning," in *Proc. Int. Conf. on Human Centered Computing*, Toronto, Canada, pp. 812–820, 2014.

[40] G. Li, B. Zhou and Y. N. Su, "Face recognition algorithm using two dimensional locality preserving projection in discrete wavelet domain," *The Open Automation and Control Systems Journal*, vol. 7, pp. 1–16, 2015.

[41] B. D. Liu, B. Shen, L. Gui, Y. X. Wang and X. Li, "Face recognition using class specific dictionary learning for sparse representation and collaborative representation," *Neurocomputing*, vol. 204, pp. 198–210, 2016.

[42] S. Gupta, K. Thakur and M. Kumar, "2D-Human face recognition using SIFT and SURF descriptors of face's feature regions," *The Visual Computer*, vol. 20, pp. 1–10, 2020.