3D Modeling and Facial Regions based Unconstrained Face Recognition

Asif Raza Butt, Sajjad Manzoor*, Muhammed Sajid,

Department of Electrical Engineering, MUST, Mirpur AJK, Pakistan
*Corresponding author: sajjad.ee@ must.edu.pk

Abstract

Unconstrained face image recognition is one of the most challenging research issues in the field of computer vision and intelligent machines (IM) due to different problems such as variation in pose, expression, illumination, and resolution. Although several face recognition algorithms have been suggested by researchers, recognition of unconstrained images presents low recognition rates. This study presents an approach of 3D modeling of 2D unconstrained face images, using intelligent machine tools for pose correction and to increase the number of training images. Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (FLDA) have been employed to extract features from various face regions to develop weighted PCA and weighted FLDA. Furthermore, the matching scores level fusion of these two techniques is used to enhance recognition accuracy. The experimental results on two standard face image databases, Surveillance Cameras face (SCface) and Labeled Faces in the Wild (LFW), show the highest rank-1 identification accuracy of 39.23% and 42.22%, respectively, which is better than state-of-the-art approaches.

Keywords—Unconstrained face images, 3D modeling, pose variation, facial regions, weighted PCA, weighted FLDA

1 Introduction

Face recognition is a significant area of research, with extensive studies documented in the literature [1-5]. While face recognition in controlled environments has reached a mature stage, achieving relatively high recognition accuracy [6], recognizing faces in unconstrained settings remains a developing and challenging field. This difficulty arises due to real-world factors like varying poses, expressions, lighting conditions, and image quality [7-9]. Among these, pose variation is regarded as the most critical and unresolved challenge [10]. Another significant issue in unconstrained face recognition is the limited availability of only one image per individual, commonly known as Single Sample per Person (SSPP) [11, 12].

Although various methods have been proposed to recognize unconstrained face images [13], some issues require attention of the research community, such as (i) recognition of face images in scenarios where only an SSPP is available for training, and (ii) pose variation in test images. Motivated by these facts, we use a 3D face modeling technique to handle SSPP and pose variations. We propose an algorithm incorporating the significance of different facial regions to recognize unconstrained face images. In addition, we propose the fusion of two holistic feature representation methods for better accuracy. The contribution of this research is summarized as follows:

1) Firstly, by using 3D modeling, the pose for a frontal unconstrained face image is corrected. In applications, such as CCTV footage-based surveillance, a full-frontal face image is not available. Such work may help in the improvement of pose for face recognition.
2) A large dataset is synthesized from a single sample per person by using 3D modeling, which can give better results for face recognition.
3) Incorporating the recognition accuracy of individual facial regions (weighted PCA and weighted FLDA) to improve the overall face recognition accuracy.
4) Finally, the fusion of weighted PCA and weighted FLDA is used to further boost the recognition accuracy of unconstrained face images compared to state-of-the-art approaches.

The rest of the paper is organized as follows: related work is given in Section 2. A description of

ISSN: 2523-0379 (Online), ISSN: 1605-8607 (Print)
DOI: https://doi.org/10.52584/QRJ.2302.05

databases used and 3D modeling, using an intelligent machine tool and pre-processing, has been provided in Section 3. The algorithm proposed for face recognition is also elaborated in this section. Section 4 is dedicated to the results and discussions, while the last section (Section 5) concludes the paper.

2 Related Work

A biometric recognition is termed as a human's identity established on behavioral or physical characteristics [14]. The most commonly used physical traits are fingerprints [15], iris [16], and face [17]. The nonintrusive biometric human face is extensively used for recognition in surveillance and social applications, as compared to the iris and fingers. Face recognition in an unconstrained environment is an emerging research area in the field of computer vision [18]. One of the challenges of face recognition for an unconstrained environment is SSPP [11].

The methods used to address the SSPP problem are generally categorized into three groups [19]: (i) virtual image creation, (ii) generic learning, and (iii) image partitioning-based techniques. Among these, approaches based on virtual sample generation have shown promising results [20]. In [21], a new NMF-MSB virtual sample generation method is proposed by combining the non-negative matrix factorization (NMF) reconstruction strategy with mirror transform, sliding window, and bit plane (MSB) sample extension methods. In [22], the authors employed the rotated face model (RFM) to generate additional virtual samples by rotating the image and applying bilinear interpolation. While synthesized images are generated from the original and its corresponding mirror image, with the original being replaced by the synthesized version for recognition, in [23].

Approaches based on 3D modeling of 2D images have seen considerable development in virtual image generation [24]. In [25] a texture mapped 3D face model is developed by projecting a 2D image onto a 3D face shape. Moreover, the virtual samples are obtained by rotating a 3D face model. A 3D Generic Elastic Model (GEM) is used to develop a face model to create more virtual images. Furthermore, motivated by the periocular region containing the most stable and discriminant features, these regions of different synthesized and test images are extracted for matching [26]. In [27], commercial software FaceGen Modellar is used for 3D modeling of a single 2D image to generate different poses of varied synthesized images. A 3D modeling is also considered an effective face alignment method for mark improvement in unconstrained face recognition [28]. A commercial off-the-shelf (COTS) 3D face modeling SDK is used to construct a 3D face model of 2D face images for pose correction [29]. An algorithm based on an intrinsic coordinate system for pose correction of 3D faces is proposed in [30]. To address the issue of pose variation, an ample analysis of state-of-the-art approaches is presented in [31]. In the light of existing literature, the current study is first of its kind to (i) address the problem of SSPP by synthesizing training images with varying poses and pose correction in testing stage and (ii) investigating the significance of different face regions to recognize unconstrained face images on two benchmark facial databases, the SCface [32] and LFW [33].

3 Material And Methods

3.1 Unconstrained Face Images Databases

In this research, two databases of unconstrained face images are used for experimental purposes. The first database used is the SCface [32] database, which is developed to investigate unconstrained face image problems. It is the most challenging facial database, containing 4160 face images of 130 individual subjects. All the unconstrained face images have been acquired under uncontrolled conditions, such as varying poses, illuminations, different qualities, and resolutions, to mimic the real-world conditions. These images have been taken from 5 individual surveillance cameras, including the following three distances: D1 = 1.0 meters, D2 = 2.60 meters, and D3 = 4.2 meters. Thus, the probe set consists of 1950 unconstrained face images. The database also contains a single high quality frontal image of each subject taken from a standard camera under controlled conditions, used as a training set.

The second database used is LFW [33], another database of static images of human faces taken in an unconstrained environment. It contains 13000 face images of 5749 individual subjects. We have selected a subset of 45 randomly selected subjects, as reported in [27]. A single frontal or near frontal image is used for the training set, while the set S contains 174 unconstrained face images of these subjects, which are used for testing. Some examples of face images from these databases are shown in Fig. 1.

3.2 F

ig. 2 depicts the complete framework of our proposed model, which consists of 3D modeling for the creation of different poses to handle SSPP and correction of pose variation of a 2D unconstrained face image, preprocessing, extraction of features from different facial regions, matching, and score level fusion.

Fig. 1: Example unconstrained face images with different pose variations, (a-c) SCface database of subject 1 at distances D1, D2, and D3, respectively, (d) 5 different subjects from the LFW database

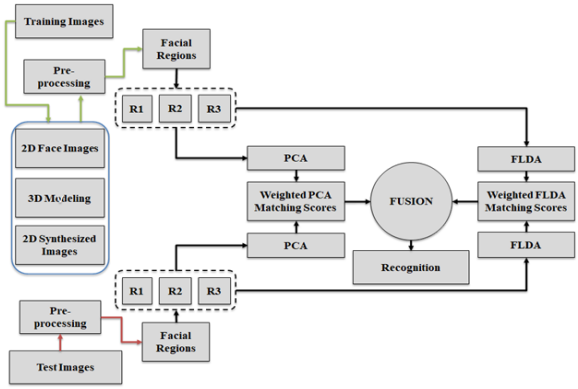

Fig. 2: Block diagram of proposed 3D modeling and region based unconstrained face recognition algorithm.

3.2.1 Modeling

In unconstrained face recognition jargon, pose variation of underlying facial images is a major challenge because varying poses result in reduced face recognition accuracy. In order to cope with pose variations, 3D modeling is one of the precise methods for pose correction, as this method can create a pose-corrected frontal face image from a given pose varied single face image acquired in an unconstrained environment.

A 3D face shape can be reconstructed from a 2D image by establishing a mapping between 3D and 2D geometry. Since a 2D image has two coordinates while a 3D face consists of three coordinates, the coordinate mapping can be accomplished if the third coordinate of depth for a 2D image is estimated. We have used 11 selected facial landmarks for depth estimation. Once depth is estimated, the texture is mapped onto the 3D face shape according to the selected facial locations of the 2D face image. Then, with the control of these points, face texture is projected onto the surface of the selected 3D. Fig. 3 shows the implementation of 3D based modeling on a given 2D face image. The 3D face model of pose varied unconstrained face images from the SCface database and the LFW database are generated. Fig. 4 shows the 2D pose corrected synthesized images of examples shown in Fig. 1. To handle the SSPP problem, the 3D face model is generated from a single input frontal or near frontal image. In the case of the SCface database, the surveillance cameras are installed at a height of 2.25 meters, slightly above the head of the subject; the most variation in poses of probe images is of downward head movement. The only available frontal image of each subject in training is used to create a 3D face model. This 3D face is rotated in 3D space to produce two down looking 2D synthesized images with considerable variation. The LFW database consists of high-resolution face images with unconstrained variations in poses. A single frontal or near frontal image is selected for the training set. A 3D face model is generated from this single image to create 8 virtual face images of different poses with down head movement, left-down, right-down, left, left, left, right, upward, left-upward, and right-upward as shown in Fig. 5. In this paper, we have used a commercially available state-of-the-art intelligent machine tool, FaceGen modeler, to generate a 3D face models available at [34]. It is a Facial Action Coding System (FACS) to anatomically generate a 3D face model using various facial expressions.

3.2.2 Pre-processing

After the generation of 3D faces, the pre-processing of pose-corrected 2D face images is performed as follows.

1) All images are rotated in the x-y plane such that they are vertically upright,
2) Alignment is performed, such that inter-pupillary distance is the same for all face images,

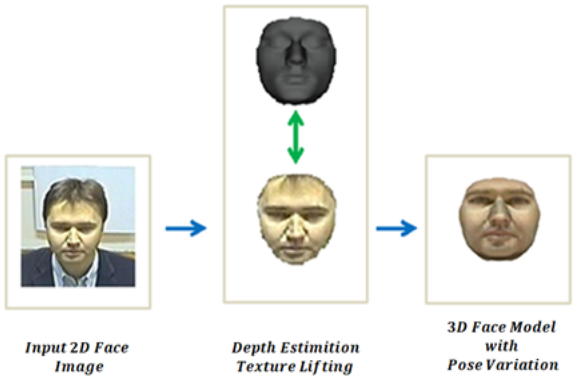

Fig. 3: A 3D based modeling of a 2D unconstrained face image selected from the SCface database.

3) Histogram equalization is performed to remove unwanted lighting variations,
4) Finally, face images are cropped to a size of 120×120 pixels.

3.2.3 Proposed Face-Recognition Algorithm

This section outlines the feature extraction methods, emphasizes the importance of different face regions, and presents the proposed face recognition algorithm.

3.2.3.1 Facial Feature Representation: Two subspace algorithms, namely Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (FLDA) [35] have been used to extract facial features as illustrated in the following subsection.

Principal Component Analysis (PCA): PCA is a subspace-based method [36]. It is widely used for feature extraction in face recognition. A PCA based set of eigenvectors, representing eigenfaces, can be derived such that for $X_1, X_2, X_3 \ldots X_M$ are the face images in the training set, having m×n as image size, the mean of face images is shown in equation (1);

$$\mu = \frac{1}{M}\sum_{i=1}^{M} X_i \tag{1}$$

Each face image differs from the mean face image, which is given in equation (2).

$$Y_i = \mu - X_i \tag{2}$$

where μ is an average face image, and $i = 1, 2, \ldots\ldots\ldots\ldots, M$, the covariance matrix C is defined as shown in equation (3):

$$C = \frac{1}{M}\sum_{i=1}^{M} Y_i Y_i^T = ZZ^T \tag{3}$$

where Z is a matrix that consists of face images subtracted from the mean. The matrix Z can be decomposed into eigenvectors as shown in equation (4).

$$Ce_i = ZZ^T e_i = \lambda_i e_i \tag{4}$$

where e_i and λ_i represent corresponding eigenvectors and eigenvalues, respectively. However, in Equation (3), the size of the covariance matrix is very large. For a manageable size of the covariance matrix, we calculate $Z^T Z$ with eigenvectors decomposition shown in equation (5).

$$Z^T Z f_i = \lambda_i f_i \tag{5}$$

By multiplying Equation (5) by Z on both sides, we get Equation (6):

$$ZZ^T f_i = \lambda_i Z f_i \tag{6}$$

From equation (6), it is obvious that if f_i is an eigenvector of $Z^T Z$, then $e_i = Z f_i$ corresponds to the eigenvector of the covariance matrix with a much more adaptable size. After calculating the eigenvectors, the normalized basis vectors B having size $m\ddot{O}n$ are computed as mentioned in Equation (7):

$$B = Z \times eigenvectors\ of\ C \tag{7}$$

Out of the total eigenvectors, we can select the largest eigenvector L, corresponding to the largest eigenvalues, resulting in a projection matrix. P of size $mn \times L$. Finally, the N number of templates of size $L \times S$ each is calculated as shown in equation (8).

$$N = P^T \times I \tag{8}$$

where I denotes the S number of image vectors projected into P. The above PCA based subspace method, in turn, reduces the dimensions of a given feature vector from d to d' where $d' < d$.

Fisher Linear Discriminant Analysis (FLDA): FLDA is another subspace based method [35] which is extensively used in face recognition. It finds a projection matrix ϕ, in order to maximize the inter-class separation and minimize the intra-class (S_w) separation while recognizing face images. The inter-class separation is the measure of the between-class scatter (S_B), as shown in equation (9):

$$S_B = \sum_{j=1}^{c} n_j(\mu_j - \mu)(\mu_j - \mu)^T \tag{9}$$

Now, intra-class separation is the amount of within-class scatter(S_w) as described in equation (10):

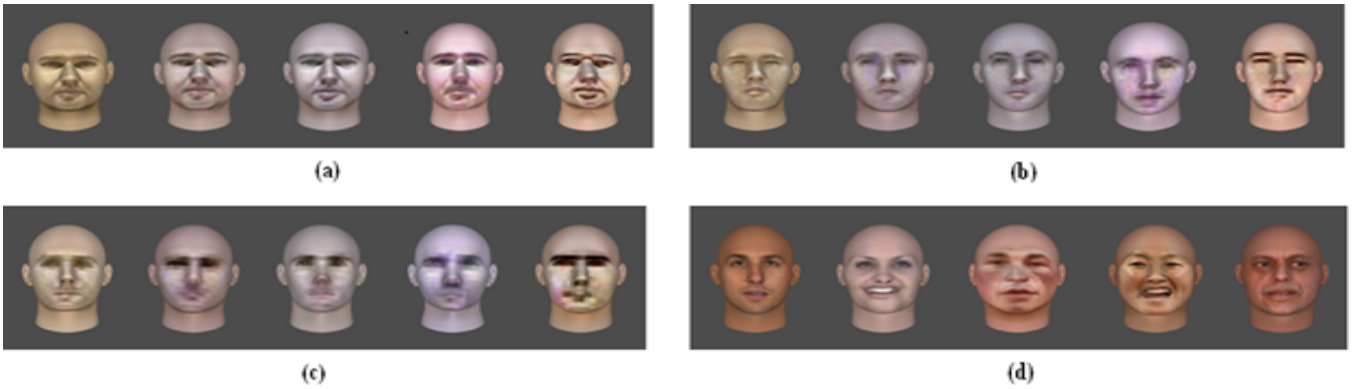

Fig. 4: Face images after pose correction by use of 3D models (a-c) SCface database (d) LFW database

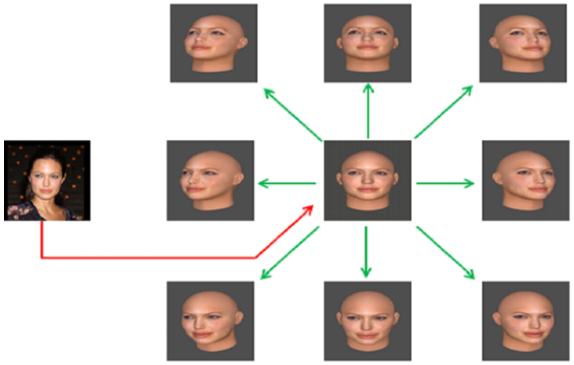

Fig. 5: Creation of poses based on 3D modeling of a single near frontal image selected from the LFW database

$$S_w = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \tag{10}$$

A subspace-based projection can be learnt by selecting a training set of n face images with a minimum of two images of each subject. For x_i^j, the i^{th} sample in training, we can calculate μ_j as shown in equation (11).

$$\mu_j = \frac{1}{n_j} \sum_{j=1}^{n_j} x_i^j \tag{11}$$

where, x_i^j is the i^{th} sample of the j^{th} class, μ_j is the mean of the class, n_j number of samples in class j. The mean of all classes μ is calculated in equation (12):

$$\mu_j = \frac{1}{n_j} \sum_{j=1}^{c} n_j \mu_j \tag{12}$$

where c is the number of classes, and n is the number of all samples. Finally, the projection matrix ϕ is selected to maximize the ratio of the determinant of S_B to the determinant of S_w using Fisher's criterion as shown in equation (13).

$$\Phi = \arg\max_{\Phi} \frac{|\Phi^T S_B \Phi|}{|\Phi^T S_W \Phi|} \tag{13}$$

In order to calculate ϕ, which is a set of generalized eigenvectors of S_w and S_B corresponding to the largest eigenvalues (λ), is shown in equation (14).

$$S_B \phi = \lambda S_W \phi \tag{14}$$

FLDA reduces the feature vector dimension from d to d'corresponding to c, where $d' < c$ by preserving the class specific information.

3.2.3.2 The Significance of Different Face Regions to Recognize Unconstrained Face Images: Before recognizing unconstrained face images, we aim to determine the significance of different face regions in recognizing unconstrained face images. Motivated by varying recognition accuracies of different facial regions (for example, the eyes, nose bridge, lips, etc.) [37]. We divide each face image into three regions, R1 (forehead and eyes), R2 (nose bridge and cheeks), and R3 (lips and chin).

An individual region to recognize unconstrained face images in subsequent face recognition experiments performed in this study. For this purpose, each 120×120 face image is divided into three equal regions of size 40×120 each, as shown in Fig. 6. For each facial region, we extract PCA and FLDA feature vectors used to calculate the corresponding matching scores. We use each region separately to calculate recognition accuracies of face images. Individual recognition accuracies are calculated for each region, both for Scface and LFW databases, as recorded in Table 1. From Table 1, it is obvious that R1 (forehead and eyes) gives the highest recognition accuracy, while R3 (lips and chin) gives the least recognition accuracy in unconstrained

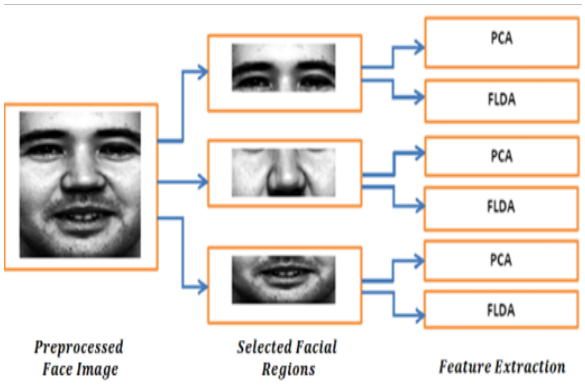

Fig. 6: Feature extraction from different facial regions of a pre-processed face image.

face images. The knowledge learned from individual recognition accuracies will be used in recognizing unconstrained face images in recognition experiments.

3.2.3.3 Proposed Face Recognition Algorithm: After investigating the contribution of different regions to recognizing unconstrained face images, the proposed face recognition algorithm, shown in Fig. 2, is explained below:

1) All the face images are preprocessed as described in section 3.2.
2) For each 120x120 face image, we extract PCA and FLDA feature vectors.
3) Each face image is divided into three regions as explained in section 3.3.2, with PCA and FLDA feature vectors calculated for each 40x120 facial region.
4) Once feature vectors have been extracted, PCA and FLDA based similarity scores are calculated. In the case of three facial regions, a region-wise matching score calculation is performed, and weights are assigned to each region based on individual recognition accuracies calculated in Table 1.

Weights W_R of each region are calculated using equation 15.

$$W_R = \frac{A_R}{A_T} \tag{15}$$

Where A_R is the individual accuracy of a specific region, and A_T is the total accuracy of all regions. A combined similarity score $M_{matching}$ is calculated representing all three regions using the weighted sum rule as shown in equation 16.

$$Mmatchimg = \sum_{i=1}^{3} w_i m_i \tag{16}$$

where m_i represent similarity scores, and w_i represent corresponding weights.

The weighted similarity scores learnt above are called weighted PCA and weighted FLDA. For several images S in training and T in testing, each matching algorithm results in a similarity matrix M_i of size $S \times T$, where i denotes the individual algorithms. An element of a matrix M_i denotes the similarity score between the probe number and the gallery number of face images. Combined similarity scores $M_{combined}$ are calculated by fusing weighted PCA and weighted LDA similarity scores, using the simple sum rule as shown in equation 17.

$$M_{combined} = sum_{j=1}^{3} M'_{combined} \tag{17}$$

Where $M'_{combined}$ represents the normalized combined similarity scores of individual algorithms obtained using the min-max rule [38]. These normalized similarity scores are used to calculate rank-1 recognition accuracies in the following face recognition experiments.

4 Experiments and Results

In this section, we evaluate our proposed face recognition method by conducting experiments on SCface and LFW databases. The performance of the experimental setup is evaluated using rank-1 identification accuracy, and the details of the experiments conducted on both databases are provided in the following subsections.

4.1 Experiments on SCface Database

For the experiments on the SCface database, we have used the sets $S1$ (1 original frontal and 2 corresponding pose varied synthesized face images by using 3D modeling) as training and $S2$ (1 original unconstrained face image and 1 corresponding pose corrected synthesized face image by using 3D modeling) for probe. We have applied our proposed algorithms to the above-mentioned experimental setup to compute Rank-1 identification accuracy. In the case of our proposed weighted PCA and weighted FLDA, the weights relating to each region are assigned using the corresponding recognition accuracies of regions $R1$, $R2$, and $R3$ as recorded in Table 1. For our proposed combined approach, we have computed the combined similarity scores $M_{combined}$ by fusing the proposed weighted PCA and weighted FLDA as explained in Subsection 3.3.

The performance of PCA and weighted PCA, LDA and weighted LDA, and the proposed approach combined with other state-of-the-art approaches is given in Table 2, Table 3, and Table 4 corresponding to unconstrained face images acquired at distances 1.0 meters, 2.60 meters, and 4.20 meters, respectively.

TABLE 1: Individual regional accuracies for SCface and LFW databases

Database	Probe set		Rank-1 identification accuracy (%)					
			PCA			FLDA		
			R1	R2	R3	R1	R2	R3
SCface	D1	Cam1	15.38	12.31	10	19.23	16.15	11.54
		Cam2	14.62	10.77	8.46	15.38	13.85	9.23
		Cam3	8.46	7.69	6.92	13.08	11.54	8.46
		Cam4	13.08	10	8.46	14.62	12.31	9.23
		Cam5	14.62	11.54	9.23	16.92	17.69	7.69
	D2	Cam1	20.77	17.69	13.85	29.23	18.46	14.62
		Cam2	18.46	16.92	12.31	19.23	17.69	11.54
		Cam3	13.85	14.62	10.77	20.77	19.23	12.31
		Cam4	15.38	12.31	9.23	21.54	19.23	13.08
		Cam5	16.92	12.31	10	19.23	16.92	11.54
	D3	Cam1	13.08	11.54	7.69	14.62	12.31	8.46
		Cam2	10.77	9.23	6.92	12.31	10.77	7.69
		Cam3	9.23	10	6.15	13.08	11.54	8.46
		Cam4	10	10.77	5.38	13.85	12.31	9.23
		Cam5	8.46	9.23	4.62	13.08	10.77	6.92
LFW	S		19.38	13.13	9	20.11	17.78	11.11

TABLE 2: Rank-1 identification accuracies for the SCface database at a distance of 1.0 meters

Database	Training	Probe	Algorithms	Rank-1 Identification accuracy (%)					
				Cam1	Cam2	Cam3	Cam4	Cam5	Mean
SCface	S1'	S2'	**PCA[32]**	6.18	3.9	7.7	8.5	5.4	6.34
	S1	S2'	**PCA[27]**	7.69	6.92	6.92	6.15	9.23	7.38
	S1	S2	**PCA**	16.92	14.62	8.46	13.08	12.31	13.08
	S1	S2	**Weighted PCA**	19.23	16.15	10	16.92	14.62	15.38
	S1	S2'	**FLDA[27]**	12.31	15.38	10.77	12.31	13.85	12.92
	S1	S2	**FLDA**	21.54	16.15	14.62	19.23	18.46	18
	S1	S2	**Weighted FLDA**	23.85	18.46	16.15	23.85	18.46	20.15
	S1	S2	**Proposed Combined**	24.62	19.23	18.46	24.62	23.85	22.16

TABLE 3: Rank-1 identification accuracies for the SCface database at a distance of 2.60 meters

Database	Training	Probe	Algorithms	Rank-1 Identification accuracy (%)					
				Cam1	Cam2	Cam3	Cam4	Cam5	Mean
SCface	S1'	S2'	**PCA [32]**	7.7	7.7	3.9	3.9	7.7	6.18
	S1	S2'	**PCA [27]**	14.62	7.69	11.54	12.31	11.54	11.54
	S1	S2	**PCA**	22.31	19.23	15.38	16.92	20	18.77
	S1	S2	**Weighted PCA**	25.38	24.62	18.46	22.31	23.85	22.93
	S1	S2'	**FLDA[27]**	23.85	16.92	20.77	23.85	18.46	20.77
	S1	S2	**FLDA**	34.62	26.92	23.85	24.62	27.69	27.54
	S1	S2	**Weighted FLDA**	37.69	30.77	27.69	28.46	31.54	31.23
	S1	S2	**Proposed Combined**	39.23	31.54	29.23	30.77	33.08	32.77

TABLE 4: Table 4 Rank-1 identification accuracies for the SCface database at a distance of 4.20 meters

Database	Training	Probe	Algorithms	Rank-1 Identification accuracy (%)					
				Cam1	Cam2	Cam3	Cam4	Cam5	Mean
SCface	S1'	S2'	**PCA [32]**	2.3	3.1	1.5	0.7	1.5	1.82
	S1	S2'	**PCA [27]**	5.38	6.15	7.69	7.69	6.23	6.63
	S1	S2	**PCA**	13.85	11.54	10.77	11.54	9.23	11.39
	S1	S2	**Weighted PCA**	16.15	13.85	11.54	13.85	11.54	13.39
	S1	S2'	**FLDA [27]**	13.85	10.77	11.53	16.15	10	12.46
	S1	S2	**FLDA**	15.38	13.85	14.62	15.38	12.31	14.31
	S1	S2	**Weighted FLDA**	18.46	14.62	15.38	16.92	13.85	15.85
	S1	S2	**Proposed Combined**	21.69	19.23	18.46	19.23	17.69	19.26

4.2 Experiments on LFW Database

For the experiments on the LFW database, the sets S1 (1 original frontal or near frontal face image and 8 pose varied synthesized face images using 3D modeling) and S2 (1 original unconstrained face image and 1 pose corrected synthesized face image by using 3D modeling) are used as training and probe sets, respectively. Similar to the experiments on SCface, PCA and weighted PCA, LDA and weighted FLDA, and the proposed combined approach have been applied to compute rank-1 identification accuracies. The performance of our proposed approaches with other state-of-the-art approaches is given in Table 5.

4.3 Comparison with Existing Algorithm

For performance comparison with other state-of-the-art approaches, the related experimental setup of the SCface database is arranged in the following sets: $S1'$ (1 original frontal face image), $S1$ (1 original frontal and 2 pose varied synthesized face images by using 3D modeling) of training data and set $S2'$ (1 original unconstrained face image) for probe. In [32] PCA is implemented on the experimental setup $S1'$ as training and $S2'$ as a probe. PCA and FLDA are implemented on the training set S1 and probe set $S2'$ as proposed in [27]. The rank-1 identification accuracies for PCA, weighted PCA, FLDA, weighted FLDA, and the proposed approach, combined with these state-of-the-art approaches, are presented in Table 2-4. The mean of Rank-1 recognition accuracies of all cameras at three distances is also calculated for complete comparison of the proposed and above mention approaches is shown in Fig. 7-9.

Similarly, for performance evaluation with other state-of-the-art approaches, the related experimental setup of LFW database is organized in following sets: sets $S1'$ (1 original frontal or near frontal face mage), $S1''$ (1 original frontal or near frontal face image and 8 pose varied synthesized face images using Rotated Face Model), $S1$ (1 original frontal or near frontal face image and 8 pose varied synthesized face images using 3D modeling) for training and S2' (1 original unconstrained face image) for probe. The rank-1 identification accuracy PCA ($S1'$ $S2'$) is computed by implementing PCA on the $S1'$ and $S2'$ as training and probe sets, respectively. The PCA and FLDA are used to calculate rank-1 accuracy on the training set $S1''$ and probe $S2'$ set as suggested in [22]. In [27], the training set $S1$ and probe set $S2'$ are used for the implementation of PCA and FLDA. We have computed the rank-1 identification accuracies for experiments on the LFW database and compared them with the state-of-the-art approaches as given in Table 5. The mean rank-1 recognition accuracies of the proposed methods are compared with those of the other methods in Fig. 10.

4.4 Results and Discussion

The proposed combined, weighted FLDA and weighted PCA approaches outperformed the PCA and FLDA algorithms. The highest rank-1 accuracy is 39.23%, 37.69%, and 25.38% on SCface, and 42.22%, 40%, and 37.78% on the LFW database is achieved. FLDA [27] and PCA [27] and PCA[32] have achieved the accuracy of 23.85%, 14.62%, and 7.7% on the SCface database, respectively. Similarly, FLDA[27], FLDA[22], PCA[27], PCA[22], and PCA($S1'$ $S2'$) have attained the accuracy 23.56%, 17.24%, 20.11%, 16.67%, and 15.52% on the LFW database, respectively.

The mean accuracies of all cameras at three distances are also calculated to evaluate the significance of 3D modeling for pose correction of unconstrained face images, the different facial regions, and the fusion of proposed methods. At a distance of 1.0 meter for experiments on the SCface database, 9.04%, 8%, and

TABLE 5: Rank-1 identification accuracy for LFW database

Database	Training	Probe	Algorithms	Rank-1 Identification accuracy (%)
LFW	S1'	S2'	**PCA**	15.52
	S1"	S2'	**PCA [22]**	16.67
	S1	S2'	**PCA [27]**	20.11
	S1	S2	**PCA**	28.89
	S1	S2	**Weighted PCA**	37.78
	S1"	S2'	**FLDA [22]**	17.24
	S1	S2'	**FLDA [27]**	23.56
	S1	S2	**FLDA**	33.33
	S1	S2	**Weighted FLDA**	40
	S1	S2	**Proposed Combined**	42.22

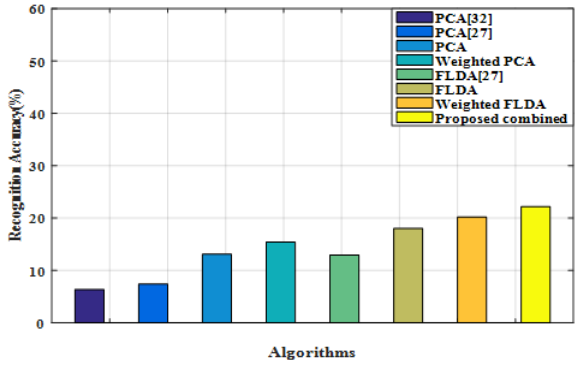

Fig. 7: Mean rank-1 recognition accuracies of different algorithms of all surveillance cameras at a distance of 1.0 meters

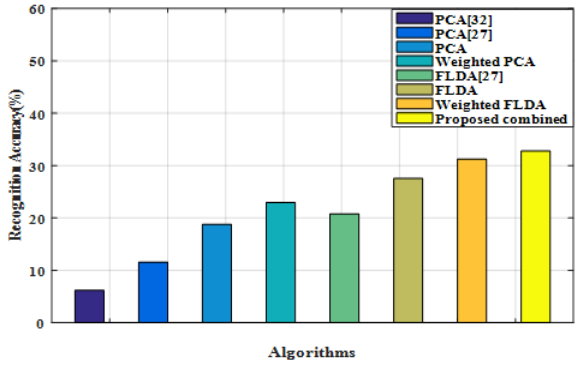

Fig. 8: Mean rank-1 recognition accuracies of different algorithms of all surveillance cameras at a distance of 2.60 meters

6.74%, and 5.7% improvement in mean accuracy by using weighted PCA and PCA as compared to PCA [32] and PCA[27]. Similarly, by implementing weighted FLDA and FLDA, an increase in mean accuracy of 7.23% and 5.08% is achieved, respectively, as compared to FLDA [27]. In addition, by using the proposed combined method, the mean accuracy is increased by 15.82%, 14.78%, and 9.24% as compared to PCA[32], PCA[27], and FLDA[27], respectively, as shown in Fig. 7.

The mean rank-1 identification accuracy at a distance of 2.60 meters is improved by using weighted PCA, and PCA is 16.75%, 11.39%, 12.59%, 7.23% as compared to PCA[32] and PCA[27], respectively. The weighted FLDA and FLDA mean accuracy is improved by 10.46% and 6.77%, respectively, as compared to FLDA [27]. Similarly, the mean accuracy increased by using the proposed combined method is 26.59%, 21.23%, and 12% as compared to PCA[32], PCA[27], and FLDA [27], respectively, as shown in Fig. 8.

Similarly, at a distance of 4.20 meters, by using weighted PCA and PCA, the mean accuracy is improved 11.57%, 6.76%, and 9.57%, 4.76% as compared to PCA[32] and PCA [27], respectively. The mean accuracy is increased by using weighted FLDA, and FLDA is 3.38% and 1.85%, as compared to FLDA[27]. The mean accuracy 17.44%, 12.63%, and 6.8% is improved by using the proposed combined method compared with PCA[32], PCA [27],27nd FLDA[27], respectively, as shown in Fig. 9. From Table 2-4 and Fig. 7-9, it is obvious that the highest rank-1 recognition accuracies were achieved at a distance of 2.60 meters. The least recognition achieved on images was at a far distance, 4.20 meters, which is because of the very low quality of the images.

For experiments on the LFW database, the mean accuracy 22.26%, 21.11%, 17.67%, and 13.37%, 12.22%, 8.78% is improved by using weighted PCA and PCA compared with PCA ($S1'$ $S21$), PCA [22], and PCA [27], respectively. By using weighted FLDA

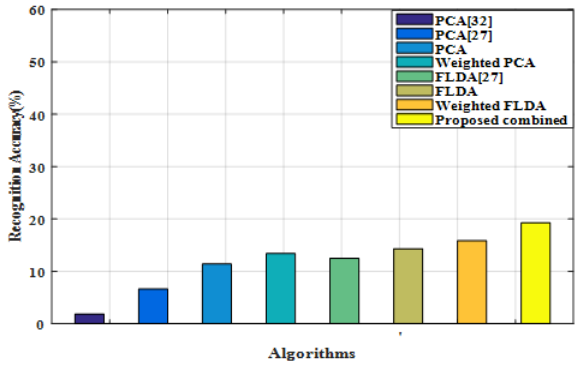

Fig. 9: Mean rank-1 recognition accuracies of different algorithms and surveillance cameras at a distance of 4.20 meters

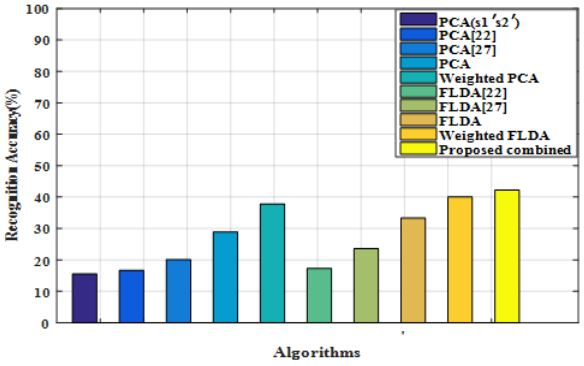

Fig. 10: Mean rank-1 recognition accuracies of different algorithms on the LFW dataset

and FLDA, the mean accuracy is increased 22.76%, 16.44%, 16.09%, and 9.77% as compared to FLDA [22] and FLDA [27], respectively. Similarly, the mean accuracy by the proposed combined is improved 26.77%, 25.55%, 22.11%, 24.98%, and 18.66% as compared to PCA ($S1'$ $S21$), PCA [22], PCA [27], FLD [22] and FLDA [27], respectively, as shown in Fig. 10. From the above experiments on Scface and LFW, the comparison clearly shows the improvement when recognition is performed on original train and probe sets, by increasing images in the training and original probe sets, by increasing training and correction of poses in probe sets, and the contribution of different regions.

5 Conclusion

In this study, we have investigated the challenging problem of recognizing face images acquired under unconstrained scenarios. The issues, such as pose variation and availability of only a single sample per person, are taken into consideration. A new 3D modeling-based approach is used to improve the pose of the image where a frontal face image is not available. The number of samples for training purposes has also increased using the proposed 3D modeling scheme. Using the proposed algorithm for an SSPP unconstrained image, the effectiveness is demonstrated for the recognition accuracy of different facial regions in recognizing unconstrained face images. It has been observed that incorporating the knowledge learnt from different facial regions improves the recognition accuracy, rather than using a whole face image. It is also evaluated that the recognition accuracy on two challenging facial databases, the SCface and the LFW database. In Table 6, a comparison of proposed algorithms and existing work is given. It can be observed that the proposed fusion algorithm gives better accuracy, and even in terms of PCA and FLDA (weighted PCA and FLDA give even better results). The proposed future work involves the use of more effective classification measures to boost the recognition performance of unconstrained face images.

TABLE 6: Comparison between existing research and the proposed algorithm

Sr#	Reference	Algorithms		
		PCA	FLDA	Fusion Algorithm
1	Proposed	28.89%	34.67%	42.22%
2	[22]	16.67%	17.24%	-
3	[27]	14.62%	23.85%	-
4	[32]	7.7%	-	-

6 Acknowledgment

This work was supported by the Technology Development Fund (Grant No. TDF02-203), Higher Education Commission, Pakistan. Moreover, the authors would also like to thank the University of Zagreb, Faculty of Electrical Engineering and Computing, for providing the database for facial images. In addition, some portions of the research in this paper use the SCface database of facial images.

References

[1] N. El Fadel, "Facial recognition algorithms: A systematic literature review," *Journal of Imaging*, vol. 11, no. 2, p. 58, 2025.

[2] X. Wang, Y. C. Wu, M. Zhou, and H. Fu, "Beyond surveillance: Privacy, ethics, and regulations in face recognition technology," *Frontiers in Big Data*, vol. 7, p. 1337465, 2024.

[3] R. A. Waelen, "The struggle for recognition in the age of facial recognition technology," *AI and Ethics*, vol. 3, no. 1, pp. 215–222, 2023.

[4] K. Juneja and C. Rana, "An extensive study on traditional-to-recent transformation on face recognition system," *Wireless Personal Communications*, vol. 118, no. 4, pp. 3075–3128, 2021.

[5] V. Tomar, N. Kumar, and A. R. Srivastava, "Single sample face recognition using deep learning: A survey," *Artificial Intelligence Review*, vol. 56, Suppl. 1, pp. 1063–1111, 2023.

[6] A. Verma, A. Goyal, N. Kumar, and H. Tekchandani, "Face recognition: A review and analysis," in *Computational Intelligence in Data Mining*. Singapore: Springer, 2022, pp. 195–210.

[7] H. Qiu *et al.*, "SynFace: Face recognition with synthetic data," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 10880–10890.

[8] A. J. Shepley, "Face recognition in unconstrained conditions: A systematic review," *arXiv preprint arXiv:XXXX.XXXX*, 2019.

[9] O. Elharrouss, N. Almaadeed, and S. Al-Maadeed, "A review of video surveillance systems," *J. Vis. Commun. Image Represent.*, vol. 77, p. 103116, 2021.

[10] S. B. Ahmed *et al.*, "On the frontiers of pose invariant face recognition: A review," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2571–2634, 2020.

[11] A. R. Butt *et al.*, "On-the-move heterogeneous face recognition in frequency and spatial domain using sparse representation," *PLOS ONE*, vol. 19, no. 10, p. e0308566, 2024.

[12] M. T. Siddique *et al.*, "An automatic photometric augmentation technique to recognize faces with single sample per person," in *Proc. 8th Brunei Int. Conf. Eng. Technol.*, 2023.

[13] M. Günther, L. El Shafey, and S. Marcel, "Face recognition in challenging environments: An experimental and reproducible research survey," in *Face Recognition Across the Imaging Spectrum*. Cham: Springer, 2016, pp. 247–280.

[14] A. K. Jain, D. Deb, and J. J. Engelsma, "Biometrics: Trust, but verify," *IEEE Trans. Biometrics, Behavior, and Identity Science*, vol. 4, no. 3, pp. 303–323, 2021.

[15] A. Peng and R. Huang, "Research progress on the application of deep learning in fingerprint recognition," *Pattern Recognition*, p. 112216, 2025.

[16] K. Nguyen, H. Proença, and F. Alonso-Fernandez, "Deep learning for iris recognition: A survey," *ACM Computing Surveys*, vol. 56, no. 9, pp. 1–35, 2024.

[17] H. Imaoka *et al.*, "The future of biometrics technology: From face recognition to related applications," *APSIPA Trans. Signal Inf. Process.*, vol. 10, p. e9, 2021.

[18] A. R. Butt *et al.*, "Unconstrained face recognition using infrared images," *Int. J. Image Graph.*, vol. 25, no. 6, p. 2550056, 2025.

[19] R. X. Ding *et al.*, "Variational feature representation-based classification for face recognition with single sample per person," *J. Vis. Commun. Image Represent.*, vol. 30, pp. 35–45, 2015.

[20] L. Li *et al.*, "A survey of virtual sample generation technology for face recognition," *Artificial Intelligence Review*, pp. 1–20, 2017.

[21] F. Li *et al.*, "Face recognition in single sample per person using multi-scale features and virtual sample generation," *Frontiers in Applied Mathematics and Statistics*, vol. 8, p. 869830, 2022.

[22] X. Hu, W. Yu, and J. Yao, "Multi-oriented 2DPCA for face recognition with one training face image per person," *J. Comput. Inf. Syst.*, vol. 6, pp. 1563–1570, 2010.

[23] C. Shao *et al.*, "Converted-face identification using synthesized images," *Multimedia Tools Appl.*, vol. 76, pp. 6641–6661, 2017.

[24] X. Ning *et al.*, "Multi-view frontal face image generation: A survey," *Concurrency Computation Practice Experience*, vol. 35, no. 18, p. e6147, 2023.

[25] D. Jiang *et al.*, "Efficient 3D reconstruction for face recognition," *Pattern Recognition*, vol. 38, pp. 787–798, 2005.

[26] F. Juefei-Xu *et al.*, "Spartans: Single-sample periocular-based recognition," *IEEE Trans. Image Process.*, vol. 24, pp. 4780–4795, 2015.

[27] X. Hu *et al.*, "Surveillance video face recognition with single sample per person based on 3D modeling," *Neurocomputing*, 2017.

[28] Y. Taigman *et al.*, "DeepFace: Closing the gap to human-level performance," in *Proc. CVPR*, 2014, pp. 1701–1708.

[29] L. Best-Rowden *et al.*, "Unconstrained face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 9, pp. 2144–2157, 2014.

[30] N. I. Ratyal *et al.*, "3D face recognition based on pose invariant alignment," *Computers & Electrical Engineering*, vol. 46, pp. 241–255, 2015.

[31] C. Ding and D. Tao, "A comprehensive survey on pose-invariant face recognition," *ACM Trans. Intell. Syst. Technol.*, vol. 7, p. 37, 2016.

[32] M. Grgic, K. Delac, and S. Grgic, "SCface—Surveillance cameras face database," *Multimedia Tools Appl.*, vol. 51, pp. 863–879, 2011.

[33] G. B. Huang *et al.*, "Labeled faces in the wild," Univ. Massachusetts, Tech. Rep. 07-49, 2007.

[34] S. I. Inc., "FaceGen," Dec. 2025. [Online]. Available: https://facegen.com/contact.htm

[35] M. Taneti, "Secure face recognition system using Eigenfaces and Fisherfaces," *Int. J. Multidisciplinary Educational Research*, vol. 11, 2022.

[36] H. R. Yazdani and A. R. Shojaeifard, "Facial recognition using eigenfaces and PCA," *Mathematics and Computational Sciences*, vol. 4, no. 1, pp. 29–35, 2023.

[37] P. Tome *et al.*, "Identification using face regions," *Forensic Science International*, vol. 233, pp. 75–83, 2013.

[38] M. Singh, R. Singh, and A. Ross, "A comprehensive overview of biometric fusion," *Information Fusion*, vol. 52, pp. 187–205, 2019.

[39] B. Li *et al.*, "Facial expression recognition via ResNet-18," in *Proc. ICMTEL*, 2021, pp. 290–303.