Neurocomputing 408 (2020) 231–245

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

# Recent development in face recognition

Umarani Jayaraman[a], Phalguni Gupta[b,*], Sandesh Gupta[c], Geetika Arora[d], Kamlesh Tiwari[d]

[a] Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, India
[b] National Institute Technical Teachers' Training and Research, Kolkata, India
[c] University Institute of Engineering and Technology, CSJM University Kanpur, India
[d] Birla Institute of Technology and Science Pilani, Pilani, India

ARTICLE INFO

*Article history:*
Received 16 April 2019
Revised 13 July 2019
Accepted 11 August 2019
Available online 6 April 2020

Communicated by Prof. De-Shuang Huang

*Keywords:*
Biometrics
Face recognition
Digital face
Scanned face
Face features
Deep learning
Indexing schemes

ABSTRACT

Face stands out as a preferable biometric trait for automatic human authentication as it is intuitive and non-intrusive. This paper investigates various feature-based automatic face recognition approaches in detail. High degree of freedom in head movement and human emotion leads a face recognition system to face critical challenges in terms of pose, illumination and expression. Human face also undergoes irreversible changes due to aging. These factors makes the process of face recognition non trivial and hard. This paper also provides a review of the facial recognition approaches individually dealing with these issues. Applications of face recognition in the forensic domain sometimes needs identification using a scanned facial image. The scenario is quite useful to get investigative leads. Important approaches for the same are also been discussed in the manuscript. Recent developments in the low-cost image capturing devices has flooded the facial image databases with a lot of images, at the same time availability of GPU based compute power has helped develop deep learning approaches to handle the face recognition at a very accurate and massive level. The same has also been surveyed and analyzed in the manuscript.



## 1. Introduction

Face is one of the most prominent identifier of a human identity. This is why an automatic human recognition system using face as a biometric trait is more appealing than the ones using other traits such as fingerprint, iris, palmprint *etc*. Providing a facial biometric sample is non-intrusive and highly convenient as the user is not required to cooperate much for providing his face image. The biometric system uses distinguishing facial features like eyes, mouth, forehead and nose or other spatial geometry features for discriminating one user from the other. A feature based automatic human recognition systems rely on global analysis of face to construct orthonormal basis vector representation. Face recognition is widely applicable in the area of surveillance, human computer interaction, crime investigation and border control surveillance for purposes including monitoring and access control.

A typical face recognition system broadly follows three steps, (1) Detection and Extraction of face, (2) Feature Extraction and Representation, and (3) Face Recognition. Face detection starts by confirming the presence of a face in given image. Next step is to find the location of face in that image and to segment region of interest. The representation of the segmented face is done using a suitable feature vector that highlights unique characteristics of the face [1–6]. Face recognition can be done either by verification or identification. Verification involves comparison of the features related to given face with the corresponding feature template stored in database against a claimed identity. It involves only one single comparison. Identification, however, compares the given face with all the feature templates stored in the database to find the identity of a face among several possibilities that matches the most. This is why identification takes time proportional to the size of the database.

Structural appearance of different frontal faces are roughly alike and the differences between them is at very minute level. Human face cannot be treated as a unique and rigid object as the humans are very creative to make different appearances of the same. There are also various challenges to face recognition arising due to presence of uneven lightening, variable expression, pose *etc*. Fig. 1 shows the effect of these variations on the facial images present in FERET dataset [7]. Factors leading to the variation of the facial appearance can be categorized into intrinsic and extrinsic [8].

**Intrinsic Factors** are the ones that depend on the physical structure of the face.

- Expression: Humans are very efficient in communication. There is a whole palette of expression to communicate even without

* Corresponding author.
*E-mail addresses:* umarani@iiitdm.ac.in (U. Jayaraman), pg@iitk.ac.in (P. Gupta), guptasandesh@uietkanpur.org (S. Gupta), p2016406@pilani.bits-pilani.ac.in (G. Arora), kamlesh.tiwari@pilani.bits-pilani.ac.in (K. Tiwari).

https://doi.org/10.1016/j.neucom.2019.08.110
0925-2312/


**Fig. 1.** Challenges to Face Recognition System (a) Normal Face (b) Illumination Variation (c) Simultaneous Expression and Illumination (d) Pose Variation (e) Pose and Expression (f) Pose and Eye Closed (images from FERET dataset [7]).


**Fig. 2.** An example of expression with (a) neutral, (b) anger, (c) eyes closed, (d) disgust, (e) laugh, (f) fear, (g) sad and (h) surprise.

**Fig. 3.** Images in Extended Yale Database [9].

a world. Some of the prominent expressions are neutral, laugh, smile, disgust, anger, fear, surprise *etc*. Some of the facial expressions are shown in Fig. 2. Expression could be characterized by facial actions including closing eyes and/or mouth, modifying the geometry and/or texture *etc*.

- Age: Being made of a biological tissue, face is also bound to change with time. The texture and appearance of a human face varies with aging which may create an issue during the face recognition.
- Occlusion: User could cover some part of his face due to weather condition or fashion. Appearance or removal of hair/beard, wearing glasses/accessories, application of makeup *etc* may lead to occlusion of the face.
- Doctored images and image falsification: Spoofed facial image could be used to attack an automatic face recognition system. A common practice is to present an artificially constructed facial image that is a mixture of two faces to the system. A secure system should be provisioned to detect spoofs and protect its templates.

**Extrinsic Factors** are the external factors that affect the image capturing process and thus, alter the appearance of the face.

- Illumination: Lighting has a magical effect on the appearance of any object. There is a very high degree of freedom for lighting variation to appear. These variations many of the cases result in a huge change that is even more than the identity change. Fig. 3 shows some of the illumination effected face images from Extended Yale Database [9].
- Scale/Resolution: Image can be captured from different distances. This not only affects the resolution but also the attention in the image. These variations are broadly called scale variation and are dealt in a very sophisticated way.
- Pose: It is characterized by rotation of the head with respect to the image capturing plane. One can observe that it could not only hide some part of the face but also could be present all together lot different perspective to the appearance.
- Noise: It is one of the fundamental properties of the acquisition device that cannot be avoided. Therefore, some noise compensation methods could be deployed in face recognition. It can be noted that in most of the cases, noise is very minimal and sometime is difficult to perceive from any naked eye.
- Blur: It is quite common in an image and can appear due to various reasons including (1) unstable camera (2) motion of object (3) high camera exposure time *etc*.

## 2. Face databases

There are plenty of standard face databases that can be used to benchmark the performance of a new face recognition approach. This section provides a brief overview of few commonly used facial databases. Also, their performances are shown in Table 1.

1. **Labeled Faces in the Wild (LFW)** [10]: The LFW database has been widely used to study face recognition in unconstrained environment. It consists of around 13,000 labeled face images of 5749 subjects that have been collected from the web. Every

**Table 1**
Popular face databases: Sample Details with Accuracy.

| Database | Number of | | Invariant to | | | | | Best |
|---|---|---|---|---|---|---|---|---|
| | Subjects | Images | Expression | Age | Pose | Illumination | Occlusion | Accuracy |
| LFW [10] | 5749 | 13,233 | No | No | No | No | No | 99.83% [24] |
| FERET [7] | 856 | 2413 | Yes | No | No | No | No | 95.5% [25] |
| FG-NET [11] | 82 | 1002 | Yes | Yes | Yes | Yes | No | 86.5% [26] |
| AR [27] | 126 | 4000 | Yes | No | Yes | Yes | Yes | 86% [28] |
| YALE [13] | 15 | 165 | Yes | No | No | Yes | No | 99.37% [29] |
| YALE Db-B [14] | 10 | 5760 | No | No | Yes | Yes | No | 98.51% [30] |
| CMU-PIE [15] | 68 | 40,000 | Yes | No | Yes | Yes | No | 98.63% [31] |
| ORL [16] | 40 | 400 | No | No | No | Yes | No | 98% [32] |
| BROWN [17] | 4 | 152 | No | No | No | No | No | 28.1% [33] |
| MORPH [18] | 13,000 | 55,000 | No | Yes | No | No | No | 97.51% [34] |
| CFD [19] | NA | 597 | Yes | Yes | Yes | Yes | No | 73.00% [35] |
| CelebA [20] | NA | 200K | Yes | Yes | Yes | Yes | No | 83.00% [36] |
| MS-Celeb-1M [21] | NA | 100K | Yes | Yes | Yes | Yes | No | 98.8% [37] |
| Mega Face [22] | 690K | 1M | Yes | Yes | Yes | Yes | No | 72.72% [38] |
| IJB-A [23] | 500 | 5712 | Yes | Yes | Yes | Yes | No | 98.4% [39] |

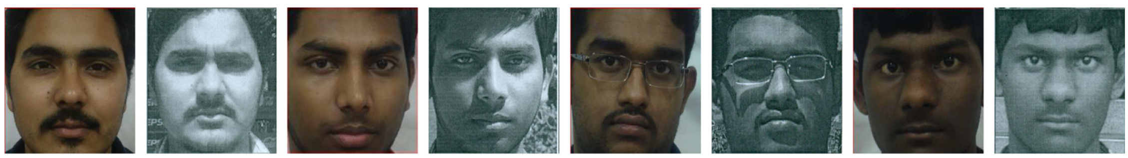
**Fig. 4.** Sample images from the FG-NET aging Database [11].


**Fig. 5.** Poses in AR Database [12].

person has a varying number of images in the database but, 1680 persons have at least two distinct images.

2. **FERET database** [7]: FERET database is a standard database for evaluating face recognition systems with varying expressions. The database consists of 2413 images of 856 subjects.
3. **FG-NET Database** [11]: This database contains images of different individuals at different age. Having with 1002 images of 82 subjects that have expression, pose and illumination variation it is the largest publicly available image database with such variations. There are roughly 12 images per subject on an average. Age of the subjects in the images lie in the range of 1 to 69 years. Some of the sample images from FG-NET database are shown in Fig. 4.
4. **AR database** [12]: It has 4000 color images pertaining to 126 subjects. The images are captured for various expressions, illumination conditions, pose and occlusion (scarf and glasses). Different poses have been shown in Fig. 5.
5. **Yale Database** [13]: The YALE database consists of 165 grayscale face images from 15 subjects. There are 11 images of each subject in different illumination conditions and expressions namely, normal, sleepy, sad, wink, surprised, with and without glasses, center-light, left-light, right-light, happy and sad.
6. **YALE database-B** [14]: This database consists of 5760 images of 10 subjects taken under single light condition but each image is taken in 9 poses and 64 illumination conditions. Also, for every subject in a particular pose, an image with ambient background was also captured. Some of the images from this database are shown in Fig. 3.
7. **CMU Database** [15]: This database consists of 40,000 face images of 68 subjects. Every subject has been imaged for four different expressions across 13 different poses and under 43 different illumination conditions. It is also called as Pose, Illumination, and Expression (PIE) database.
8. **ORL database** [16]: It consists of 400 face images taken from 40 subjects. Each subject has 10 images that have been taken at different time, under varying illumination, facial expressions and occlusion (with and without glasses)
9. **Browns Database** [17]: It is an aging database that has been generated using a book by Nixon and Galassi. The database contains images of four subjects varying over a time period of 36 years *i.e.* 36 images per subject. The images in this database are of better quality than FG-NET database and do not contain any pose variation. Fig. 6 shows few images for each individual from the database.
10. **MORPH database** [18]: It is the largest publicly available longitudinal face database consisting of 55,000 images of more than 13,000 people. The age of the subjects ranges between 16 to 77. An average of 4 images are there per subject.
11. **IITK-Student Database:** It consists of images of individuals from IIT Kanpur. In total there are 50 subjects, with 2 im-

**Fig. 6.** Sample images from Browns Database [17].



(a) IITK-Student Database

(b) IITK-Rural Database

**Fig. 7.** Sample images from IITK-Student and IITK-Rural Database.

ages for each subject, one image taken in controlled laboratory condition (with constant background) and another image taken in non-controlled condition (no constant background, taken in both laboratory as well as outdoor). The non-controlled images are printed and scanned, and used as probe images. The controlled digital images are used as gallery images. Fig. 7(a) shows sample images from the IITK-Student Database.

12. **IITK-Rural Database:** It consists of images of individuals from Villages near Kanpur. In total there are 820 subjects, with 2 images for each person, taken at a gap of 4–5 months. There are certain additional challenges in this database which include large illumination variance, slight age variation, slight pose variation in certain images and occlusion in the form of scarf covering part of the face in female subjects and beard and glasses in male subjects. Fig. 7(b) show sample images from the IITK-Rural Database.
13. **Chicago Face Database (CFD)** [19]: The version 1 of the Chicago Face Database has been released in 2014. It consists of 158 images of male and female faces of varying ethnicity and expression between the ages of 17–65 with high-resolution. Version 2 has been released in 2015, with an addition of 439 new neutral expression images with norming data. The database now includes Asian, Black, Latino, and White male and female subjects. Updated version has been released in 2016.
14. **CelebFaces Attributes Dataset (CelebA)** [20]: It is a large-scale face attributes database with more than 200K celebrity images, each with 40 attribute annotations. The images are captured in unconstrained environment. It consists of 10,177 different identities, 202,599 face images, and 5 landmark locations, 40 binary attributes annotations per image.
15. **MS-Celeb-1M** [21]: It is a public database consisting of 100K celebrity subjects and has about 10M images in total. It is invariant to illumination, age, pose and makeups.
16. **Mega Face** [22]: The MegaFace database includes one Million images that capture more than 690K different individuals. It has been collected from the Flickr set and have been made publicly available. The images vary largely in terms of capturing device, resolution, race, gender, pose and number of faces per image. Unconstrained pose, expression, lighting, and exposure.
17. **IARPA Janus Benchmark (IJB-A, B and C)** [23]: The database IJB-A has been released by the NIST in 2015. This database contains 5712 images and 2085 videos of 500 subjects captured in unconstrained environment. Each subject in the database has an average of 11 images and 4 videos per subject. The IJB-B database (2017) contains 67,000 face images, 7000 face videos. The IJB-C (2017) has 138,000 face images, 11,000 face videos.

## 3. Face detection

Face detection is an important step in face recognition that corresponds to the localization of the face in a given image. Once face is detected it is cropped for recognition. Although several methods exist for face detection, the performance was not quite satisfactory until the method proposed by Viola Jones et al. [40] appeared in the year 2000. The Viola-Jones face detector contains three main ideas that make it possible to build and run in real time: the integral image, classifier learning with AdaBoost, and the attentional cascade structure.

The integral image, also known as a summed area table, is an algorithm for quickly and efficiently computing the sum of values in a rectangle subset of a grid. It is first introduced to the computer graphics field to be used in mipmaps. Viola-Jones face detector applied the integral image for rapid computation of Haar-like features. The features are defined as the (weighted) intensity difference between two to four rectangles. Boosting is a method of finding a highly accurate hypothesis by combining many weak hypotheses, each with moderate accuracy. The attentional cascade is a critical component in the Viola-Jones detector. The key insight is that smaller, and more efficient, boosted classifiers can be built which reject most of the negative sub-windows while keeping almost all the positive examples. This method has been tested on

MIT+CMU face dataset and is found to be 15 times faster than any other previous method.

Viola and Jones is the first method which apply rectangular boxes for the face. But, it has some drawbacks as its feature size is large. In a 24 × 24 image, the total number of Haar_like features is 160,000 and also it is not handled for wild faces and frontal faces. Because of this many methods such as HOG, SIFT, SURF, and ACF have been used which are more robust features in detecting the face. A new feature called Normalized Pixel Difference (NPD) is introduced in [41] which differentiates the two pixels intensity. Another well known method is Dlib [42], in which support vector machine (SVM) is used as a classifier which increases the robustness in the face detection.

Deep Learning has some of the most explored work including automatically learning and synthesizing feature which could help detects face [43]. Multitask learning framework that integrates a ConvNet and a 3D mean face model. Further improvements have been done using CNN cascade and the region proposal network (RPN). Faster R-CNN and ResNet [44] have given significant boosts in performance on face detection benchmarks like FDDB [45]. One of the major drawbacks of deep learning is that they are not easily interpretable, because of the information being encoded in a distributed manner.

## 4. Face recognition for digital images

The face recognition for digital images can be categorized on the basis of the feature used such as global appearance based methods and local feature based methods. The other methods includes such as neural network based methods and Deep Learning based methods. The overview of the face recognition methods discussed in this paper is given in Fig. 8.

### 4.1. Global appearance based methods

These methods are based on statistical feature representation. There exists several methods to extract features from facial image. Some of them are Principal Component Analysis (PCA), kernel PCA (KPCA), 2D-Image PCA (2D-IPCA) and Fisher-Linear Discriminate Analysis (LDA), *etc*. All these methods concentrate on extracting features which is known as global features by considering the image as a whole.

#### 4.1.1. Principal component analysis

This method is based on *Karhunen-Loeve Transform (KL transform)* [46]. First, the 2D-face image matrix is transformed into a 1*D* vector which represents one column vector in the data matrix. After converting every face image into data matrix co-variance matrix is computed. A face can be represented in terms of a best coordinate system, termed as *Eigen Pictures*. The KL transform has been applied to the cropped face images including only eyes and nose. A set of cropped face images, called *face ensemble* has been represented as a vector. Here, upright face images have only been chosen. The mean face vector has been calculated and the deviation from mean has been obtained for each face image. PCA has been applied to these face ensemble images and Eigen vectors corresponding to the highest eigen values have been calculated, which are termed as the Eigen picture.

Later, the work has been extended in [47], where it has been treated as a 2D face recognition problem. A computational model has been developed which is fast, simple and accurate in constrained environment. *Eigen face* that handles faces as a whole, has emphasized on the Information Theory based approach to encode the most relevant information in a group of faces. PCA has been applied to transform face images from image space to face space. Both the methods has the drawback like lack of sufficient ensemble of face images and work only in constrained environment. Initially, it has been tested on locally available dataset in which accuracy obtained with varying orientation, illumination and size. The reported accuracy is in the range of 64% [47].

#### 4.1.2. Kernel principal component analysis

The input image is mapped to a feature space via suitable nonlinear mapping and then PCA is applied in the feature space. The approach is called kernel PCA (KPCA) [48]. Use of a polynomial kernel can permit higher correlation between input pixels in the analysis of facial images. Face recognition is performed using linear support vector machines (SVMs) on the features obtained in transformed domain. The method has been tested on the Olivetti Research Laboratory (ORL) database and has obtained better performance than a linear PCA with an accuracy of 83.3% [48].

#### 4.1.3. 2D-image principal component analysis

In *2D-IPCA*, a 2*D*-Image matrices is used rather than 1-D vector [49]. The image matrix is not transformed into a 1*D* vector as in conventional PCA, but a co-variance matrix is directly constructed using the original image matrices. The size of this matrix is small. Mathematically, an image $I$ of size m × n has to be projected onto $n$ dimensional column vector $X$ by following a linear transformation: $Y = IX$, where $Y$ is the projected vector. Now, the covariance matrix, $S_x$ of the projected feature vectors of the training samples is obtained. A nearest neighbor classifier is used once the feature vector has been obtained for an image. In [49], Euclidean distance measures has been used for classification. It has been tested on the ORL [16], AR [12] and Yale datasets [13] and achieved maximum recognition accuracy of 84% which is superior than PCA and kernel PCA [46,48,50].

#### 4.1.4. Independent component analysis

PCA exploits pairwise relationships between the pixels of the image database. In Independent Component Analysis (ICA), important information is presented with the high-order relationships among pixels. It leads to better basis images for high-order statistics. Two different architectures have been put forth in [50]. One of them considers images as random variables and their pixels as output and the second one is vice-versa. The face database has been represented as a matrix where each row is a different person. As stated earlier, for the first approach, face images are considered as random variables and their pixels as trials and then the degree of independence of two images is calculated. If two images are independent, then while moving across the pixels, it would not be possible to predict the value of a pixel on the basis of the value of same pixel in another image. While in the second approach, independence of two pixels is computed. That is, while moving across an image, it is not possible to predict the value of a pixel based on some other pixel in the same image. Performance of ICA is better when cosines is used as a similarity measure rather than euclidean distance. It has been tested on the FERET database and gave an accuracy of 86% [50] which is better as compared to PCA [46].

### 4.2. Local feature based methods

Unlike a global feature based method, the local feature based methods divide the face image into several blocks and each block is considered separately for extracting the features. The block which does not play major role in discriminating features may be ignored. So the accuracy of these methods are good comparatively global feature based methods. Finding suitable descriptor for local facial regions is challenging and is still an open issue. Any such descriptor should ideally be easy to compute and have high inter-class and low intra-class variance.

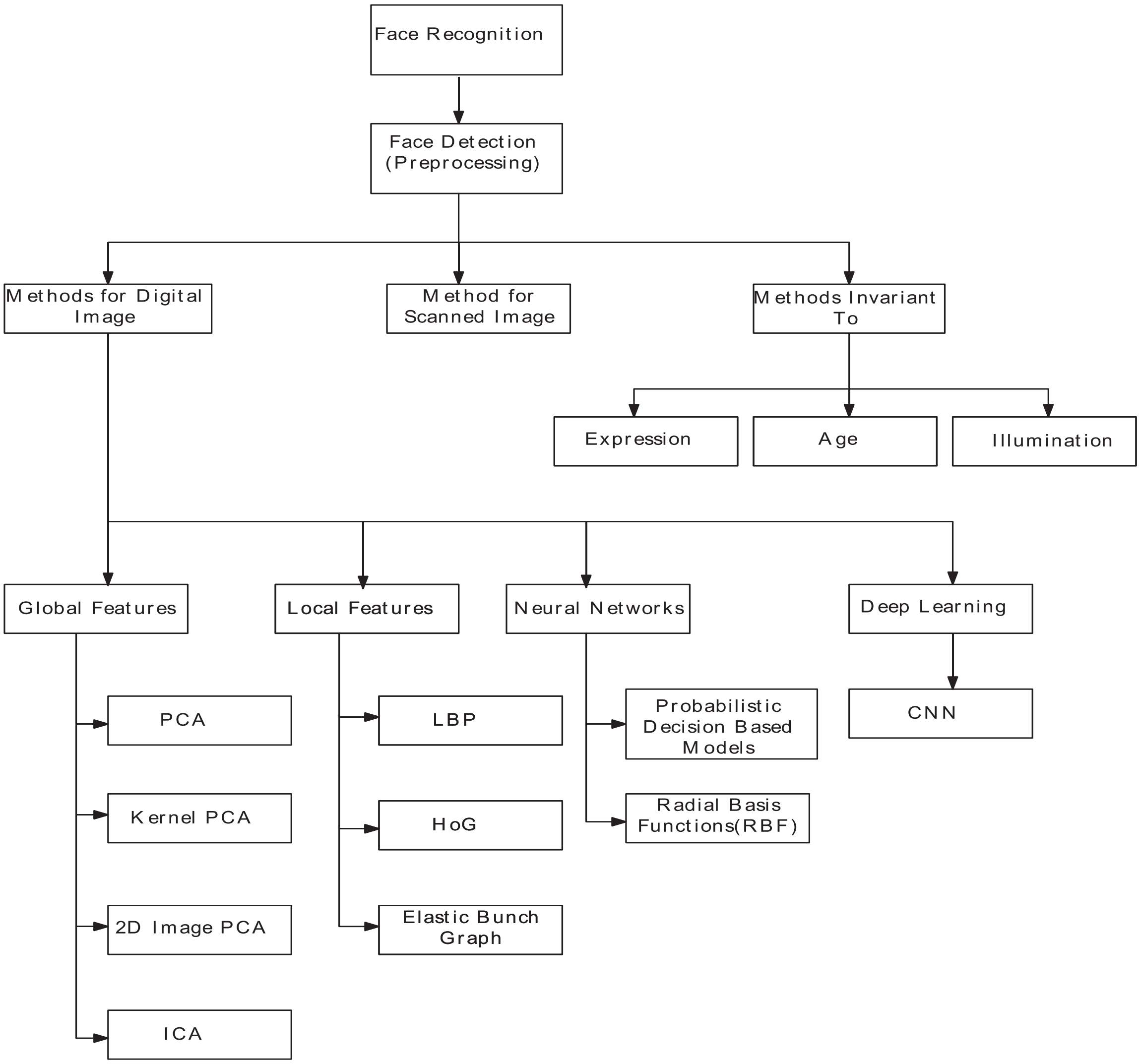



**Fig. 8.** Overview of the Face Recognition Methods.

*4.2.1. Local binary patterns (LBP)*

LBP has been primarily proposed to describe texture patterns [51]. Using it as a face descriptor is encouraged by the fact that human face is also composed of micro-patterns. LBP assigns a binary label to every pixel in the image by applying a threshold on its $3 \times 3$ neighborhood. Then, a histogram of the labels is generated that acts as a texture descriptor [52]. Some facial features (such as eyes) have been found to contribute more in human face recognition than any other features. Based on the importance of the information, weights could be assigned to the regions. Similarity for classification can be determined using Chi square distance [53]. This method has been tested on the FERET database and achieved a mean recognition rate of 95%.

*4.2.2. Histogram of oriented gradients*

HOG is an image descriptor which is invariant to 2*D* rotation and scaling [54,55]. Initially, the image is divided into small connected regions, called cells, and for each cell, a histogram of edge orientations is computed. The histogram counts are normalized to represent the final HOG descriptor [25]. The accuracy of the method on FERET [7] database has been found to be around 95.5%.

*4.2.3. Elastic bunch graph matching*

It has been proposed in [56] which is based on Dynamic Link Structures [57]. An individual face is represented by a graph by choosing a set of fiducial points on the face. A fully connected graph has been generated with each fiducial point as a node. The edge represents the distance between two corresponding joining fiducial points. A stack-like structure is obtained by combining a set of such graphs called a face bunch graph. The method on FERET [7] database is found to have accuracy of 89%, 88% and 67% for manual positioning, with phase information and without phase information respectively.

## 4.3. Neural networks based methods

As the technology in imaging and storage advances, we observe huge increase in the dimensionality and number of images in database. In these scenarios, special techniques are often needed as the conventional image processing based statistical methods [58–79] fail to attain space-time efficiency. Following are a few neural network based face recognition methods.

*4.3.1. Probabilistic decision based neural networks*

In this method [80], a facial recognition is implemented in two levels. A primary level of facial recognition consisting of eyes and nose (excluding mouth) that provides distinctive facial features yet being relatively invariant against different facial expressions, hair styles, and mouth movement. The secondary facial features such as hairline and mouth features are obtained by normalization and resolution reduction using intensity and edge values from the input face images. Feature vectors are fed into two Probabilistic Decision based Neural Networks (PDBNN). The method on two public databases, FERET [7] and ORL [16] have achieved verification accuracy around 84% and 98%.

*4.3.2. Radial basis function (RBF)*

RBF is a neural classifier that is useful in the cases when the size of training set is small and has high dimension face images. Facial features are subjected to PCA and Fisher's linear discriminant (FLD) [81] to get lower-dimensional discriminant features. Learning algorithm train a RBF neural network. PCA and FLD generates most discriminant features such that different classes of training pattern can be separated as far as possible and the same classes of patterns are kept as close as possible. In [32] simulation results have been generated on the ORL database that shows that the system achieves excellent performance in terms of both error rates of classification and learning efficiency.

## *4.4. Deep learning based methods*

The traditional methods for face recognition involve preprocessing, local descriptor extraction, and feature transformation steps. Even though these steps have been improved separately, not much increase in the accuracy for face recognition has been reported. Also, most of the methods for unconstrained face recognition tackle with only one challenge among pose, expression or illumination. Hence, traditional methods are not much capable of extracting stable features for recognition that are invariant to real world situations [82]. On the other hand, deep learning methods use a stack of layers to learn different representations at multiple levels. The features extracted from these layers show robustness towards variation in illumination, expression and pose. These methods are developed using neural network models [83–96] by developing methods for staking one useful representation over other. The features extracted from the multiple stacked layers of a deep neural network show robustness towards any kind of variation in general.

The initial deep learning architecture, called DeepFace, has been proposed in [97]. It uses nine-layered convolutional neural network (CNN) in which last two layers are fully connected. The output of the fully connected layer is passed to a $K$-way softmax function, where $K$ is the number of classes. The output of the softmax layer gives the probability distribution for various classes. This method achieved a verification accuracy of 97.35% when tested on LFW database.

Another method based on learning an euclidean embedding for every image using CNN has been proposed in [98]. The network is trained such that the $L_2$ distance of the similar face representations are less as compared to the representation of faces of different persons. It introduces online triplet mining method that utilizes triplets of roughly aligned matched and non-matched face images. This method on LFW dataset have achieved 99.63% accuracy. Liu et al. [99] claimed that combining softmax loss and euclidean margin based losses should not be motivated as they are not compatible with each other. The authors instead proposed an angular margin instead of considering the euclidean distance. The features learned by modified softmax loss are angularly distributed. This modified method achieved an accuracy of 97.88% on LFW dataset. To further maximize the inter-class discriminability, a new loss function known as, additive angular margin loss function has been introduced in [24]. This network, also called as ArcNet which have achieved 99.83% verification accuracy on LFW dataset.

A light CNN framework has been proposed in [100] that learns face representation from the large-scale data with many noisy labels. This method has proposed a Max-FeatureMap (MFM) operation, instead of using ReLU, to attain a compact representation and to separate noisy labels from the informative ones. This method has achieved an accuracy of 97.50% on LFW database. A CNN based model for face recognition consisting of one trunk network and two small branch networks has been proposed in [101]. The trunk network learns the features from different face images at different resolutions. The small branch networks learns resolution-specific coupled mappings (CM). The CM makes projects the high resolution database images and low resolution probe images in a space where their distances are minimized. Face verification accuracy using this method is 98.7% on LFW dataset when the size of probe image is 112 $\times$ 96. Zhong et al. has proposed an end-to-end deep learning framework for face recognition that does not require aligning the faces prior to giving them to a network [102]. This method has introduced alignment learning in which no prior knowledge on facial landmarks nor artificially defined geometric transformations is required. This model achieved a verification accuracy of 99.33% on the LFW dataset.

Unlike most deep-learning models that perform frontalization or normalization of face image to tackle pose variation, a model has been proposed that processes several face images using multiple pose-specific CNN models in [103]. This method firstly detects landmark points in all the images and then align them to five different reference poses. The CNN models extract features in each pose. These have been trained using CASIA webface dataset containing 500k face images. This method has been tested on IJB-A dataset and has achieved an identification rate of 86.2% and 93.1% for top-1 and top-5 matches respectively. All these methods have been summarized and tabulated in Table 2.

**Table 2**
Well-known Face Recognition Methods.

| S. No | Method | Benchmark Databases | Recognition Accuracy |
|---|---|---|---|
| 1. | LBP [51] | FERET | 95% |
| 2. | HOG [54,55] | FERET, AR, CMU Multi PIE and YALE | 95.5% |
| 3. | EBGM [56] | FERET | 89% |
| 4. | PCA [46] | LOCAL DATASET | 64–96% |
| 5. | ICA [50] | FERET | 44–86% |
| 6. | kPCA [48] | ORL | 83.3% |
| 7. | 2DPCA [49] | ORL, YALE and AR | 84% |
| 8. | PDBNN [80] | ORL, FERET, and LOCAL DATABASE | 84–98% |
| 9. | RBF [32] | ORL | 98% |
| 10. | CNN [24,82] | LFW | 97.35–99.83% |

# 5. Expression invariance

With the face recognition systems expanding their usage in less restricted environments, it has become crucial for the algorithms to be robust to expression changes. As such, most of the images and photographs are obtained in a natural setting and they tend to have various facial expressions.The present day algorithms, however, are very susceptible to changes in expression as documented in a number of evaluations. This poses a significant problem for a wide range of commercial products, especially in the scenarios where the subject is non-cooperative and it is difficult to obtain a neutral expression face image. Expressions tend to produce large distortions and variations in certain regions of the face which dramatically alters the appearance of a given face. This causes the performance of traditional methods like eigenface to drop dramatically making them less effective. Also, if a face recognition system fails to withstand changes in expression, then it would become mandatory for a test subject to be co-operative and present a poker face. This would be cumbersome and face recognition would cease to remain a passive and non-intrusive process. Therefore, it is required to develop a system for face recognition that works well even for different facial expressions. It is also a very common human nature to represent emotion and communicate using body language of which a major part happens to be facial expression.

Although a lot of work has been done in non-expressive/neutral face recognition, there is still a scope in the problems dealing with expressive face images. The presence of expression on the face tends to produce local variations. In other words, only specific regions of the face undergo a change in appearance while the others may remain more or less unaffected. The challenge arises due to the fact that the variation due to change in expressions may show up on unpredictable regions of the face and in various shapes and sizes. Darwin et al. [104] have proposed that happiness, fear, surprise, sadness, disgust and anger are basic emotions out of which, happy and sad are positive emotions while the remaining are negative emotions. Ekman [105] has proposed cues that help in understanding the distortions produced by these expressions. Later on, facial expression has received a lot of attention from psychologists, medical practitioners, actors and artists [106–108].

Expression variance can be handled by only considering variation only from a single angle. This can be done by locating one or more face sub-spaces to reduce the impact of expression variation. Face subspace using geometric features has been constructed in [109]. This method recorded a descent recognition rate and thus, template based techniques have come into dominance. Achieving an optimized representation in the subspace has been the goal of most of the traditional algorithms [110–112]. These methods apply linear or non-linear transformation on the gallery images and then prove their robustness to the global variation [113]. However these methods lack in handling local variations especially when it is more such as the case of a laughing/surprised person. Local probabilistic method has been employed in [28,114,115] to handle this problem. In these methods, gaussian representation of each subspace has been generated. Self-Organizing Maps (SOM) have been exploited in [116,117] to handle this variation.

A method has been proposed in [118] that generates a local non-negative matrix factorization or LNMF for representation. Sparse-representation classifier (SRC) has been proposed in [119] that aims to minimize the L1 norm of coefficients. This is done by capturing the underlying sparsity to its advantage. Various methods have been proposed on the basis of an assumption that the deformations are not only local but contiguous. That is, the face image can be partitioned into blocks and each block can be treated uniquely. Weighted local probabilistic subspace (WLPS) has been proposed in [114] that uses a gaussian model. This has been remodelled in [120] by adding the neural networks. These methods perform well in some situations but discrimination among the blocks become difficult when the number of available samples are less. Singh et al. [121] have proposed a method that uses Gabor features for identification. Ivanov et al. [122] has proposed a method in which facial regions are initially detected by using SVG classifiers. Later on, partial images of the faces concerned are regenerated employing the identified mouth regions and then a separate classifier is used for the purpose of final recognition.

## 6. Age invariance

Aging is a slow but invariant aspect of human life. Aging patterns are different for different individuals and are affected by various factors such as environment, lifestyle, culture, gender, disease, consumption of alcohol and anti-aging drugs *etc*. The appearance of the face of the individual alters and hence, the extracted features. The features that had been stored in the face database have to be updated to incorporate such changes and thus, improving the usability of the face recognition system.

Researchers have proposed several methods to handle age invariance in face recognition. Burt et al. [123] handled age variation in facial images by creating composite images by using their texture and shape information for multiple age groups. Furthermore, facial features affected by aging have been extracted. The authors collected 147 images from males aged between 20 and 62 and have shown effect of change in color, shape and both on the perceived age. It has been reported that the transformation in shape, color and both increments the perceived age by 5.9, 8.2 and 12.1 respectively for the person of 27 years. Wu et al. [124] have proposed a model that dynamically simulate wrinkles and skin aging on a 3-D model. Lanitis et al. [125] have proposed a method that establishes relation between age of a person and its corresponding coded representation of facial images. The relation could further be used to simulate face appearance at any given age. Also, the age of a test image can be estimated. For experimentation, 500 face images from 60 subjects have been collected that includes age-progressive images of 45 individuals. The database has been further divided into two sets, A and B, wherein A and B has 80 and 85 images respectively of 12 subjects whose age ranges from 0 to 18 years in A and 19 to 30 years in B. Two experiments have been conducted. In the first one, set A has been used for training and classification accuracy of 67% and 71% has been re-coded for Weighted Appearance-Specific and Weighted Person-Specific method respectively. While in the second method, set B has been used for training and the classification accuracy of 63% and 66% has been recorded for the respective methods. Gandhi et al. [126] have proposed a method that establishes the relation between the given image of the face and his age by using support vector regression machine. To do so, signature images have been created for various age groups lying in the range from 15 to 99 years. The face image at a target age has been simulated using these signature images and support vector regression machine. Liu et al. [127] have also proposed an age simulation model using a technique known as Image-Based Surface Detail Transfer. This technique extracts the geometric details of an image and put it onto a different surface. This has been used to produce an aged image by transferring age features like bumps and wrinkles onto the younger version of the face.

Wang et al. [128] have proposed to simulate the enrolled and probe image at a particular age for face recognition. The simulation has been done by extracting shape and texture features of the face using PCA. The method has been tested on TH aging database that has 400 images from 60 subjects of varying age. It has been shown that this method increases the face recognition rate from 80% (without age-simulation) to 91.1%. Ramanathan and Chellappa [129,130] have proposed a cranio-facial growth model to observe the shape variation that occurs due to aging. Bayesian classifier has been used for verification across ages. The method has been tested on 233 images from 109 subjects, taken collectively from FG-Net aging and an in-house database. The reported recognition accuracy is 37% and 58% for rank-5 and rank-10 respectively. It has been concluded that simulation of change due to aging is challenging and is affected by ethnicity, environment and lifestyle. Park et al. [131] have proposed to build a 3D model using the 2D image for age simulation. The simulation of texture and shape has been done separately using PCA coefficients. The method has been tested on FG-NET database and achieved an average recognition rate of 24.8% and 54.2% for $\leq 18$ and $> 18$ age group respectively. Mahalingam et al. [132] have proposed a graph based technique for age invariant recognition. The face image has been represented using a graph wherein, the vertices represent the feature points. Local Feature Analysis has been used to locate feature points on a facial image and aging pattern has been learned using the trained model graphs.

In [133], a latent factor guided convolutional neural network (LF-CNN) has been proposed to address age invariant face recognition problem. Using the designed model, the age-invariant deep face features are extracted that is invariant to the aging process over the time. The proposed has been tested on three face aging datasets, MORPH Album2 and FG-NET and achieved Rank-1 idenification rate of 97.51% and 88.1% respectively. An auto-encoder based model, called coupled auto-encoder networks (CAN) to address age invariance in face recognition has been proposed in [26]. The proposed model employing two auto-encoders is capable of learning latent representations from input images and thus, learns the complex nonlinear aging and de-aging process. This method achieved Rank-1 identification accuracy of 86.5% on FG-NET [11] and verification accuracy of 92.3% on CACD-VS. A deep learning based method to learn age invariant features has been proposed in [34]. The deep features extracted from the network have been decomposed into two orthogonal components that represents age-related and identity-related features. Further, the identity related features that are robust to age have been used for face recognition. This method achieved 98.67% Rank-1 identification rate on MORPH Album-2 database. A method that models aging variables and human identity simultaneously by utilizing Probabilistic Linear Discriminant Analysis (PLDA) has been proposed in [134]. The aging and identity subspaces are then learnt using appearance-age labels and Expectation Maximization (EM) algorithm respectively. Also, the identity features are combined using Canonical Correlation Analysis (CCA) with the aim of maximizing their correlations for the purpose of face recognition. This method has been tested on three datasets namely, FGNET [11], MORPH [18] and CACD [135] and achieved verification accuracy of 88.2%, 95.6% and 89.9% respectively.

## 7. Illumination invariance

There exists a number of methods to tackle the problem of illumination variation in face recognition. Zou et al. [136] have presented an extensive survey on existing techniques for handling illumination variation in facial images. The subspace based statistical methods have been used to tackle the problem.

A modular-PCA based method that uses gabor wavelets has been proposed in [137]. In this method, the face image is divided into sub-images, known as modules. A series of gabor wavelets have been applied on these sub-images for feature extraction at local level. PCA has been applied on all the modules separately for dimensionality reduction. It is believed that there would be some modules that would remain invariant even with variation in illumination. The method has been tested on AR and ORL databases and achieved a recognition rate of 88.8% and 87.90% respectively, considering 30 eigen vectors. A method employing independent component analysis (ICA) has been proposed in [138]. ICA is comparatively simpler, faster and has better learning capability. The method starts by pre-processing the images using PCA. ICA has been applied on the pre-processed images for extracting feature vectors; that have been used for classification. The method when tested on the Asian database achieved recognition rate of 86.5% and 100% while considering 200 and 50 components respectively. An adaptive principal component analysis (APCA) method has been proposed in [139] that handles illumination variation in face recognition. The APCA features have been used for classification. They have been extracted from facial images in three steps namely, space rotation, whitening transformation, and eigenface filtering. The method requires only one gallery image per subject. The method has been trained and tested on 251 and 284 images respectively, taken from the Asian database. It achieved a recognition rate of 83.66% on the testing images. Belhumeur et al. [140] have suggested a 3D linear subspace method for handling illumination variation in face recognition. In this method, three images are used to create a 3D basis for the subspace. It has also proposed Fisher Linear Discriminant (FisherFace) method to achieve better performance. The method has been tested on an in-house dataset containing 330 images from five subjects. Five subsets have been created out of the total images and two types of experiments- intrapolation and extrapolation has been conducted. In the intrapolation experiment, the method has been trained on subset 1 and 5 and tested on the remaining subsets. The method attained an error rate of 0.0%, 0.0% and 1.2% on subset 2, 3, and 4 respectively. While in the extrapolation experiment, training has been done on subset 1 and testing has been done on subset 1, 2 and 3. The error rate obtained on subset 1, 2 and 3 is 0.0%, 0.0% and 4.6% respectively. A method utilizing segmented linear subspace model has been proposed in [141]. The face image is segmented into sub-images according to the distribution of the surface normal. A 3D subspace model has been applied to each sub-image separately to reduce the deflection caused by the local distortions. The method has been tested on YALE-B database that has been divided into four subsets on the basis of the angle between light source and camera's axis. The subsets have 70, 120, 120 and 140 images respectively and seven images from subset 1 have been used for training. The method attains an error rate of 0.0% on all the subsets. Lee et al. has shown that there exists nine single light source directions that prove helpful for face recognition [142]. In other words, considering the images from these nine light sources is enough for handling variable lighting condition during face recognition. A nine-dimensional linear subspace has been created from these images to get a linear approximation of the illumination cone. An inhouse dataset has been created which contains images of 10 subjects under 45 different lighting conditions. The dataset has been divided into four subsets based on the angle light makes with the camera axis. The training has been done on first two subsets that cover angle from 0 to 25 degrees while testing has been done on all the 450 images. The error rate obtained by this method is 0.0%, 0.0% and 2.8% on subset 1 and 2, subset 3 and subset 4 respectively. This method holds an advantage because the subspaces have been generated from the real images directly and does not involve intermediate steps such as PCA or 3D construction.

Illumination invariant features have been explored for tackling the problem of illumination variation. Edge information has been used in [143] for feature feature representation. The authors have proposed a Line Edge Map (LEM) that extracts lines from a face edge map. The method has been tested on BERN and AR databases and achieved a recognition rate of 100% and 96.4%, respectively. A face recognition system that performs comparison of relative gradients to tackle with illumination variation has been proposed in [31]. It involves aligning the face image under different lighting conditions and then computing a similarity score for final matching. During experimentation, three templates of every subject, *i.e* left light, right light and ambient light have been used for training and the remaining 1292 mages have been used for testing. The method has been tested on CMU-PIE database and recorded an error rate of 1.47%. To handle the problem of illumination variance and rotation out of plane simultaneously, a method using symmetric-shape-from-shading algorithm to extract shape information has been proposed in [144]. This method has used two face databases, Yale and Weizmann for experimentation. Two phases of experiments have been conducted for each face database. In the first phase, the training set contains only one image, while in the second phase, the training is done using two images of each subject. The illumination-invariant measure for Yale database has been recorded as 83.3% and 90.0% in phase 1 and 2 respectively. For the Weizmann database, the recorded illumination-invariant measure in phase 1 and 2 is 81.3% and 97.9% respectively. A concept called as Self Quotient Image (SQI) based on retinex algorithm [145] has been proposed in [146]. SQI applies anisotropic smoothing instead of isotropic smoothing and outputs a light-invariant version of the 3D objects. The method has been tested on YALE-B database and the images have been divided into four subsets on the basis of the increasing illumination angle. The recorded recognition rate is 100%, 96%, 98% and 97% on subset 1, 2, 3 and 4 respectively. Qing et al. has exploited phase information of the gabor filters as they found it more discriminative than phase of FFT [147]. Later, the difference between the phases has been computed using bayesian analysis. CMU-PIE and YALE datasets have been used for the experimentation. The datasets have been decomposed into training and testing partition in which the train set has images from 34 persons (21 images per subject) from the illumination set of the CMU-PIE. The test set contains the images from the remaining 34 persons from the CMU-PIE and all the images from the YALE-B dataset. The YALE-B dataset has been divided into five subsets based on the illumination angle with the camera axis. The method attained an error rate of 0.0% on light and illumination set of CMU-PIE dataset and 0.0%, 0.0%, 0.0%, 2.8% and 4.7% respectively on the five subsets of the YALE-B dataset. An image normalization method based on gamma intensity correction (GIC) and histogram equalization has been proposed in [148]. A quotient illumination relighting (QIR) has been proposed to create an image in normal-lighting condition corresponding to the given image in not normal lighting conditions. The method has been tested on YALE-B dataset which has been divided into five subsets as mention before. Subset 1 images have been used as gallery images and the mean recognition rate achieved on the remaining subsets has been recorded as 91.8%. Another method that pre-processes the illuminated image by applying normalization techniques has been proposed in [149]. The new pre-processed image has been claimed to have same pixel values as that of the normal non-illuminated image. This method has been trained on 15, 121, 10 and 68 images of YALE, AR, YALE-B and PIE datasets respectively. The method, when tested on 30, 363, 640 and 1564 images of the four mentioned datasets, recorded a recognition rate of 93.3%, 86.2%, 99.7% and 100%, respectively. Discrete Cosine Transformation (DCT) in logarithmic domain has been used by Chen et al. [150] to compensate for illumination variation. Two databases, CMU-PIE and YALE-B has been used for experimentation. Training has been done on frontally illuminated images of CMU-PIE and subset 1 of the five subsets from the YALE-B dataset. The testing has been done on remaining 20 images of the CMU-PIE and four subsets from YALE-B dataset. The method attained an error rate of 0.36% on CMU-PIE and 0.0%, 0.0%, 0.18% and 1.71% on subset 2, 3, 4 and 5 of YALE-B dataset. Kao et al. has proposed a method that enhances contrast locally to compensate for uneven lighting [30]. The method then selects the important features that could be used for illumination invariant face recognition and finally, classification is performed using SVM. The method has been tested on Yale and Yale-B datasets and achieved a recognition rate of 96.67% and 98.51%, respectively. Liu et al. proposed a new method that uses ratio image [151]. Ratio image is quotient between an illuminated image which has to be normalized and the reference image. Blurriness has been introduced in both the images using gaussian filter. An iterative image restoration method has also been applied on the reference image. The aim of this method is to restore an arbitrarily illuminated image to a frontally illuminated one. The method when tested on YALE-B dataset gives a recognition rate of 81.7% with the restored images while the recognition rate dropped to 55% when raw images were used. The advantage of this method was that it did not require feature extraction. Choi et al. has utilized the fact that human faces have similar composition *i.e*. they have nose, mouth and two eyes and each of these, results in shadowing based on the direction of the light [152]. The shadowing characteristics have been used to generate a corresponding frontally illuminated image of the face, also known as compensated image. In addition to the compensated images, local features have also been exploited to see the effect on the recognition performance. The method has been tested on 45 images of 10 subjects taken from YALE-B database. These images have been further divided into four subsets and this method shows a recognition rate of 97% with he shadow compensated features. However, using local features with the compensated image increases the recognition rate to 99.6%. A double-density dual-tree complex transform (DD - DTCWT) based method has been proposed in [153]. The (DD - DTCWT) produces more of high frequency components. Since the illumination information lies in low-frequency bands, the high frequency components are suppressed to keep only the low-frequency ones. For restoration, the low frequency components have been then subtracted from the illuminated image. This method has been trained on subset 1 of the YALE-B dataset and tested on the remaining four subsets. The recorded recognition rate is 100% on all the subsets (subset 2, 3, 4 and 5). Li et al. have extracted illumination-invariant local features in the logarithmic domain by computing differential of every pixel with its neighbors [154]. The authors have used subset 1 of the YALE-B dataset for training while the remaining four subsets have been used for testing. The method attained a recognition rate of 99.56%, 92.31%, 88.40% and 75.91% on subset 2, 3, 4 and 5 respectively. A local image normalization method has been used in [155] wherein the normalization has been applied point by point on the whole image. The method has been tested on three databases, YALE, AR and YALE-B and has achieved an average recognition rate of 92.1%. A method based on wavelet analysis fusion and local binary patterns has been proposed in [156]. Firstly, the illumination component is removed followed by enhancement of the reflectance component. Both the processed images are then combined using wavelet fusion method. The method has been tested on YALE-B and CMU-PIE databases. The CMU-PIE database contains images of 67 subjects and three images of each subject have been randomly chosen for training. The remaining eight have been used for testing. The equal error rate reported by this method on CMU-PIE is 13.11%. The YALE-B database contains images of 31 subjects and the experiment has been conducted in two parts. The first experiment has used four randomly chosen images as training while the second has used seven images from each of the subject. the testing in both the parts has been done on the remaining 12 images of each subject. The EER in both the experiments have been reported as 8.75% and 11.69% respectively. Teoh et al. [156] have used wavelet fusion and local binary patterns to provide illumination invariance. The illumination component has been removed through logarithmic wavelet approximation and reflectance component has been enhanced using LBP histogram. The resulting images are combined using wavelet image fusion.

To address the problem of illumination variation, [157] has introduced the concept chromaticity space to suppress shadow in 2D facial images. It has proposed a novel method for lighting normalization to handle illumination variation. The method has achieved a verification rate equal of 92.5% at FAR = 0.1% on VGG-Face dataset. Cheng *et.al* [158] have proposed two methods namely, non-directional local reflectance normalization (NDLRN) and the fused multi-directional local reflectance normalization (fMDLRN) that depend on the ratio-relationship of two adjacent pixels in a face image. This method preserves the facial details while decreasing the illumination difference. The method has achieved 81.36 $\pm$ 4.11% recognition rate (mean $\pm$ standard deviation) when tested on YALE dataset [13] having one image as a training image. A gradient based descriptor to address illumination variations has been proposed in [29]. The method handles magnitude, lighting direction and the spectral wavelength together. Two illumination invariant components have been extracted and integrated into a histogram based feature representation. The accuracy achieved by this method is 99.37% on YALE dataset [13].

**Table 3**
Important Face Recognition Methods.

| SN | Authors | Scheme/ Methods | Database | Evaluation Parameters | Result |
|---|---|---|---|---|---|
| 1 | Ahonen et al. [52] | LBP | FERET | Mean Recognition Rate | 95% |
| 2 | Dniz et al. [25] | Histogram Of Oriented Gradients | FERET, AR, CMU Multi-PIE and Yale | Accuracy | 95.5% |
| 3. | Sirovich et al. [46] | PCA | In-house dataset | Accuracy | 64–96% |
| 4. | Bartlett et al. [50] | ICA | FERET | Accuracy | 44–86% |
| 5. | Kim et al. [48] | kPCA | ORL | Accuracy | 83.3% |
| 6. | Yang et al. [49] | 2dPCA | ORL, AR and Yale | Accuracy | 84% |
| 7. | Er et al. [32] | RBF | ORL | Accuracy | 98% |
| 8. | Wang et al. [128] | Age invariant | TH aging face database (400 images from 60 subjects at different age) | Rank-1 identification rate | 63.3% |
| 9. | Park et al. [131] | 3D age modeling | FG-NET, MORPH Album-1 and BROWNS | Rank-1 identification rate | 37.4%, 66.4% and 28.1% respectively |
| 10. | Mahalingam et al. [132] | graph based feature representation + Gaussian Mixture Models | FG-NET | Equal Error Rate | 25.4% for age range [18,69] |
| 11. | Felix et al. [165] | Walsh-Hadamard transform encoded local binary patterns (WLBP) | FG-NET | Rank-1 identification rate | 100% |
| 12. | Zhou et al. [134] | Probabilistic Linear Discriminant Analysis (PLDA) | FGNET, MORPH and CACD | Verification Accuracy | 88.2%, 95.6% and 89.9% respectively. |
| 13. | Wen et al. [133] | latent factor guided convolutional neural network (LF-CNN) | MORPH Album-2 and FG-NET | Rank-1 identification rate | 97.51% and 88.1% respectively |
| 14. | Wang et al. [34] | Deep learning based orthogonal components | MORPH Album-2 | Rank-1 identification rate | 98.67% |
| 15. | Xu et al. [26] | coupled auto-encoder networks (CAN) | FG-NET | Rank-1 identification rate | 86.5% |
| 16. | Karande et al. [138] | PCA and ICA for illumination invariant | Asian face dataset | Recognition rate | 100% for both respectively (first 50 principle and independent components) |
| 17. | Batur et al. [166] | Segmented linear subspaces | YALE face dataset-B | Error rate | 0% |
| 18. | Lee et al. [142] | Nine Points of Light (9PL) | YALE extended face | Average error rate | 0.93% |
| 19. | Wei et al. [31] | K-NN Face Localization | CMU-PIE | Error rate | 1.47% |
| 20. | Qing et al. [147] | Probabilistic model of gabor phase | CMU-PIE and YALE-B | Average error rate | 0% and 1.5% respectively |
| 21. | Shan et al. [148] | Quotient Illumination Relighting (QIR) | YALE-B | Average recognition rate | 91.8% |
| 22. | Kao et al. [30] | Local contrast enhancement | YALE-B | Average recognition rate | 98.51% |
| 23. | Zhang et al. [157] | chromaticity space | VGG-Face dataset | Verification rate | 92.5% at FAR = 0.1%% |
| 24. | Cheng et al. [158] | fused multi-directional local reflectance normalization (fMDLRN) | YALE | Recognition rate (mean ± std. deviation) | 81.36 ± 4.11% |
| 25. | Zhu et al. [29] | histogram based feature representation | YALE | Accuracy | 99.37% |
| 26. | Martınez et al. [28] | motion estimation | AR database | Average recognition rate | 85% (for happy, angry and scream) |
| 27. | Tan et al. [117] | Self-organizing map (SOM) and soft k-NN | FERET database | Top-1 recognition rate | 91.5% (for k value set between 10% - 20%) |
| 28. | Valery et al. [159] | Deriche algorithm | In-house database of 100 passport and camera images. | EER | 9% |
| 30. | Ahonen et al. [161] | Translation Invariant Discrete Wavelet Transform, LDA and k-NN | In-house database having images from 25 subjects | EER | 9.89% |
| 31. | Wu et al. [100] | Light- CNN framework | LFW | Verification Accuracy | 97.50% |
| 32. | Deng et al. [24] | CNN | LFW | Verification Accuracy | 99.83% |
| 33. | Lu et al. [101] | CNN | LFW | Verification Accuracy | 98.7% |
| 34. | Zhong et al. [102] | CNN | LFW | Verification Accuracy | 99.83% |

## 8. Face recognition for printed and scanned images

Biometric systems can be used by law enforcement agencies to identify a person and to get an investigating lead. In many such scenarios, face recognition system is presented with non-digital images in the form of newspaper clippings, driver license photos or passport photos. However, the images of criminals/missing people are available in the form of digital images in the database. The non-digital images are usually scanned in order to obtain digital equivalent. These scanned images are of low quality and contain a lot of noise. Traditional face recognition methods usually fail at matching these images with the images in the database thus, a specialized method is required for the same.

The problem of face recognition on scanned face images has been sparsely studied so far. Valery et al. [159] have suggested a method for matching scanned passport images with digital images. A ranking transform has been applied to the gray scale edge map of the face image obtained using Deriche algorithm [160]. Then, the rank correlation coefficient is used to obtain dissimilarity score between two images. An EER of 9% has been obtained on a database of 100 passport and camera images. Another method has been described in [161] where digital face images are matched with scanned passport images, as well as passport images photographed using a camera. The emphasis of the method is on handling the problem of watermark traces present in face images using de-noising. The method starts with de-noising of image using

Translation Invariant Discrete Wavelet Transform, followed by feature extraction using LDA, and classification using k-nearest neighbor method has been used. This method has obtained an EER of 5.49% on a database of 25 subjects while matching digital controlled images with passport images taken using a camera. An EER of 9.89% has been obtained when matching scanned passport images with images taken in the controlled environment. In both the cases, two controlled images have been used in the gallery, and one scanned passport image in probe.

Image denoising attempts to recover a digital image which has been contaminated by noise. Noise in scanned image is due to paper defects caused while printing and scanning of the images. Special artifacts and noise can be removed by using wavelet based methods [162]. However, traditional method of denoising using DWT leads to visual artifacts, because of lack of translation invariance of the wavelet basis. Translation Invariant Discrete Wavelet Transform using cyclic spinning method used in [163] averages out translation dependence. This is based on shifitng the image, denoising the image, and unshifting the image again. The shifting is done for a range of values, and then the final denoised image is obtained by averaging out the images. This leads to less visual artifacts than those obtained in the earlier method of denoising used in [164].

All major methods for face recognition discussed in this paper have been summarized in Table 3. The table discusses the methods along with the dataset that it has been tested on and the obtained results.

## 9. Conclusions

In this paper, an extensive survey of face recognition, together with a brief discussion of some of the face recognition challenges such as expression, age and illumination are addressed. The following is a concise summary with conclusions and possible future direction of face recognition system.

- Face recognition is currently a very active research area in the field of biometrics. The last couple of years have shown great advances in algorithms dealing with complex environments such as low quality gray-scale images, group photographs (recognition of multiple persons) and cluttered backgrounds, etc. Some of the best algorithms are still too computationally expensive to be applicable for real-time processing, but this is likely to change with coming improvements in computer hardware.
- Feature based methods are applicable for real-time systems where color and motion is available. Since an exhaustive multiresolution window scanning is not always preferable, feature-based methods can provide visual cues to focus attention.
- Image based methods are the most robust techniques for processing gray-scale static images. These methods are based on multiresolution window scanning to detect and recognize faces at all scales, making them computationally expensive. Multiresolution window scanning can be avoided by combining the image based method with a feature based method as a preprocessor with the purpose of guiding the search based on visual clues such as skin color.
- At last, the most important challenges in face recognition systems such as expression, age and illumination have been discussed. Face recognition methods have reached a point where the recognition of a single face in an image with fair resolution is close to being a solved problem. However, accurate recognition of multiple faces in group photographs and to deal the challenges such as pose, expression, illumination and ageing is still a hard problem to solve in face recognition.

### *9.1. Limitations of visible spectrum in FR*

Dealing with Visual Spectrum (390–750 nm) images for face recognition face challenges such as illumination, pose, facial expression changes, and facial disguises. Even though many methods have been developed by different researchers to handle these challenges it is still remains an open problem. There is no single method which could address all the above challenges at once. For example, the pose problem is minimized by method named as 2D warp. The problem of illumination is minimized by using the two techniques called as Image Processing Filters and Statistical Facial Models. To handle all these challenges at once the infrared spectrum (IRS) may be used in face recognition because it has its own special characteristics.

### *9.2. Characteristics of infrared spectrum (IRS) over visible spectrum (VS)*

The images captured IRS cameras have more characteristics than Visual Spectrum (VS). The IR images can be captured in different type of lighting environment; even in the complete dark night which show the strength of IR images to represents the facial variances. Another characteristic of IR images is the effect of absorption and scattering on the energy, less as compared to the VS. Including exterior information of the face like expression, pose variation, illumination the anatomical information such as vascular network of a face can be drawn strongly using IRS images. The thermal vision of IR images provides help to find the disguises in the face which are not supported in the VS. Therefore a lot of scope of research and applications can anticipate in IRS domain. Simultaneously, the present study emphasizes the use of three dimensional cubic dataset (i.e.) Multi/ Hyper-spectral Images in face recognition. The IR based Multi/ Hyper-spectral Imaging System can minimize the several limitations arise in the existing and classical face recognition system because the skin spectra derived with cubic dataset depicts the unique features for an individual.

### *9.3. Future of face recognition in IR*

Multi-spectral imaging is the process of concurrent acquisition of a set of images. The every image in the multi-spectral imaging corresponds to a different band in the electromagnetic spectrum. The simple example of multispectral image is colour image in the visual spectrum having RGB sensations and can also be observed by human eyes. The hyper-spectral images include more levels in a particular sub-band as compared to the multispectral images. Hyper-spectral Imaging System (HIS) provides helpful discriminants for individual face recognition that cannot be achieved by any existing imaging system. Recent research has proven that Multi-spectral/Hyper-spectral Imaging System would be the future of human FR.

## References

[1] B. Li, C. Wang, D.S. Huang, Supervised feature extraction based on orthogonal discriminant projection, Neurocomputing 73 (1–3) (2009) 191–196.
[2] C.Y. Lu, D.S. Huang, Optimized projections for sparse representation based classification, Neurocomputing 113 (2013) 213–219.
[3] L. Shang, D.S. Huang, J.X. Du, C.H. Zheng, Palmprint recognition using fastica algorithm and radial basis probabilistic neural network, Neurocomputing 69 (13–15) (2006) 1782–1786.
[4] Y. Zhao, D.S. Huang, W. Jia, Completed local binary count for rotation invariant texture classification, IEEE Trans. Image Process. 21 (10) (2012) 4492–4497.
[5] J.-X. Mi, D.-S. Huang, B. Wang, X. Zhu, The nearest-farthest subspace classification for face recognition, Neurocomputing 113 (2013) 241–250.
[6] B. Li, C.-H. Zheng, D.-S. Huang, Locally linear discriminant embedding: an efficient method for face recognition, Pattern Recognit. 41 (12) (2008) 3813–3821.

[7] P.J. Phillips, H. Wechsler, J. Huang, Jeffery, P.J. Rauss, The FERET database and evaluation procedure for face-recognition algorithms, Image and Vision Computing 16 (5) (1998) 295–306.
[8] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Comput. Surv. 35 (4) (2003) 399–458.
[9] A. Georghiades, P. Belhumeur, D. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 643–660.
[10] G.B. Huang, M. Mattar, T. Berg, E. Learned-Miller, Labeled faces in the wild: A database forstudying face recognition in unconstrained environments, in: Proceedings of the Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, 2008.
[11] The FG-NET aging database, Available at http://sting.cycollege.ac.cy/~alanitis/fgnetaging/index.htm.
[12] A. Martinez, R. Benavente, in: The Ar Face Database Cvc Tech Report, 1998.
[13] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.
[14] T. Zhang, Y.Y. Tang, B. Fang, Z. Shang, X. Liu, Face recognition under varying illumination using gradientfaces, IEEE Trans. Image Process. 18 (11) (2009) 2599–2606.
[15] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression (pie) database, in: Proceedings of the IEEE International Conference on Automatic Face Gesture Recognition, IEEE, 2002, pp. 53–58.
[16] Y. Xu, B. Zhang, Z. Zhong, Multiple representations and sparse representation for image classification, Pattern Recognit. Lett. 68 (2015) 9–14.
[17] N. Nixon, P. Galassi, The brown sisters: Thirty-Three years, Museum of Modern Art, 2007.
[18] K. Ricanek, T. Tesafaye, Morph: A longitudinal image database of normal adult age-progression, in: Proceedings of the International Conference on Automatic Face and Gesture Recognition, IEEE, 2006, pp. 341–345.
[19] chicagofacest, https://chicagofaces.org/default/.
[20] Celeba, http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html.
[21] Megaface, http://megaface.cs.washington.edu/publications.html.
[22] Ms-celeb-1mt, https://arxiv.org/pdf/1607.08221.pdfe.
[23] Janus benchmark (ijb-a, ijb-b and ijb-c) t., https://www.nist.gov/programs-projects/face-challenges.
[24] J. Deng, J. Guo, N. Xue, S. Zafeiriou, Arcface: additive angular margin loss for deep face recognition, arXiv:1801.07698(2018).
[25] O. Dniz, G. Bueno, J. Salido, F.D. la Torre, Face recognition using histograms of oriented gradients, Pattern Recognit Lett 32 (12) (2011) 1598–1603.
[26] C. Xu, Q. Liu, M. Ye, Age invariant face recognition and retrieval by coupled auto-encoder networks, Neurocomputing 222 (2017) 62–71.
[27] A. Martinez, R. Benavente, in: The AR Face Database, Technical Report 24, Computer Vision Center, 24, 1998.
[28] A.M. Martınez, Matching expression variant faces, Vis. Res. 43 (9) (2003) 1047–1060.
[29] J.-Y. Zhu, W.-S. Zheng, F. Lu, J.-H. Lai, Illumination invariant single face image recognition under heterogeneous lighting condition, Pattern Recognit. 66 (2017) 313–327.
[30] W.-C. Kao, M.-C. Hsu, Y.-Y. Yang, Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition, Pattern Recognit. 43 (5) (2010) 1736–1747.
[31] S.-D. Wei, S.-H. Lai, Robust face recognition under lighting variations, in: Proceedings of the International Conference on Pattern Recognition, 1, IEEE, 2004, pp. 354–357.
[32] M.J. Er, S. Wu, J. Lu, H.L. Toh, Face recognition with radial basis function (rbf) neural networks, IEEE Trans. Neural Netw. 13 (3) (2002) 697–710.
[33] U. Park, Y. Tong, A.K. Jain, Age-invariant face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (5) (2010) 947–954.
[34] Y. Wang, D. Gong, Z. Zhou, X. Ji, H. Wang, Z. Li, W. Liu, T. Zhang, Orthogonal deep features decomposition for age-invariant face recognition, in: Proceedings of the European Conference on Computer Vision, 2018, pp. 738–753.
[35] F. Fuentes-Hurtado, J.A. Diego-Mas, V. Naranjo, M. Alcaiz, Automatic classification of human facial features based on their appearance, PLoS ONE 14 (1) (2019) 1–20.
[36] Z. Liu, P. Luo, X. Wang, X. Tang, Deep learning face attributes in the wild, CoRRabs/1411.7766 (2014).
[37] X. Wu, R. He, Z. Sun, T. Tan, A light cnn for deep face representation with noisy labels, IEEE Trans. Inf. Forensics Secur. 13 (11) (2018) 2884–2896.
[38] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, L. Song, Sphereface: deep hypersphere embedding for face recognition, CoRRabs/1704.08063 (2017).
[39] R. Ranjan, A. Bansal, J. Zheng, H. Xu, J. Gleason, B. Lu, A. Nanduri, J. Chen, C.D. Castillo, R. Chellappa, A fast and accurate system for face detection, identification, and verification, IEEE Transactions on Biometrics, Behavior, and Identity Science 1 (2) (2019) 82–96.
[40] P. Viola, M.J. Jones, Robust real-time face detection, Int. J. Comput. Vis. 57 (2) (2004) 137–154.
[41] S. Liao, A.K. Jain, S.Z. Li, A fast and accurate unconstrained face detector, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2) (2016) 211–223.
[42] D.E. King, Dlib-ml: a machine learning toolkit, J. Mach. Learn. Res. 10 (2009) 1755–1758.
[43] S. Zhan, Q.Q. Tao, H.H. Li, Face detection using representation learning, Neurocomputing 187 (2016) 19–26.
[44] X. Sun, P. Wu, S.C. Hoi, Face detection using deep learning: an improved faster RCNN approach, Neurocomputing 299 (2018) 42–50.
[45] V. Jain, E. Learned-Miller, FDDB: A Benchmark for Face Detection in Unconstrained Settings, Technical Report UM-CS-2010-009, University of Massachusetts, Amherst, 2010.
[46] L. Sirovich, M. Kirby, Low-dimensional procedure for the characterization of human faces, Josa a 4 (3) (1987) 519–524.
[47] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognit. Neurosci. 3 (1) (1991) 71–86.
[48] K.I. Kim, K. Jung, H.J. Kim, Face recognition using kernel principal component analysis, IEEE Signal Process Lett. 9 (2) (2002) 40–42.
[49] J. Yang, D.D. Zhang, A.F. Frangi, J.-y. Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2004) 131–137.
[50] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face recognition by independent component analysis, IEEE Trans. Neural Netw. 13 (6) (2002) 1450–1464.
[51] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. (7) (2002) 971–987.
[52] T. Ahonen, A. Hadid, M. Pietikäinen, Face recognition with local binary patterns, in: Proceedings of the European Conference on Computer Vision, Springer, 2004, pp. 469–481.
[53] A.H. Toderici, M. Stamp, Chi-squared distance and metamorphic virus detection, J. Comput. Virol. Hack. Tech. 9 (1) (2013) 1–14.
[54] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Proceedings of the International Conference on Computer Vision & Pattern Recognition, 1, IEEE Computer Society, 2005, pp. 886–893.
[55] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[56] L. Wiskott, J.-M. Fellous, N. Kruger, C. von der Malsburg, Face recognition by elastic bunch graph matching, Intell. Biometr. Tech. Fingerprint Face Recognit. 11 (5) (1999) 355–396.
[57] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. Von Der Malsburg, R.P. Wurtz, W. Konen, Distortion invariant object recognition in the dynamic link architecture, IEEE Trans. Comput. (3) (1993) 300–311.
[58] D. Huang, The Study of Data Mining Methods for Gene Expression Profiles, Science Press of China, Beijing, China, 2009.
[59] D.S. Huang, W. Jiang, A general CPL-ADS methodology for fixing dynamic parameters in dual environments, IEEE Trans. Syst. Man Cybern. Part B (Cybern.) 42 (5) (2012) 1489–1500.
[60] D.S. Huang, J.-X. Mi, A new constrained independent component analysis method, IEEE Trans. Neural Netw. 18 (5) (2007) 1532–1535.
[61] D.S. Huang, X.M. Zhao, G.B. Huang, Y.M. Cheung, Classifying protein sequences using hydropathy blocks, Pattern Recognit. 39 (12) (2006) 2293–2300.
[62] J. Zhang, D.-S. Huang, T.-M. Lok, M.R. Lyu, A novel adaptive sequential niche technique for multimodal function optimization, Neurocomputing 69 (16–18) (2006) 2396–2401.
[63] K.H. Liu, D.S. Huang, Cancer classification using rotation forest, Comput. Biol. Med. 38 (5) (2008) 601–610.
[64] L. Shang, D.S. Huang, C.H. Zheng, Z.L. Sun, Noise removal using a novel non-negative sparse coding shrinkage technique, Neurocomputing 69 (7–9) (2006) 874–877.
[65] X.F. Wang, D.S. Huang, H. Xu, An efficient local Chan–Vese model for image segmentation, Pattern Recognit. 43 (3) (2010) 603–618.
[66] X.F. Wang, D.S. Huang, A novel density-based clustering framework by using level set method, IEEE Trans. Knowl. Data Eng. 21 (11) (2009) 1515–1531.
[67] X.F. Wang, D.S. Huang, J.X. Du, H. Xu, L. Heutte, Classification of plant leaf images with complicated background, Appl. Math. Comput. 205 (2) (2008) 916–926.
[68] D.-S. Huang, Systematic Theory of Neural Networks for Pattern Recognition, 201, Publishing House of Electronic Industry of China, Beijing, 1996.
[69] B. Li, D.-S. Huang, C. Wang, K.-H. Liu, Feature extraction using constrained maximum variance mapping, Pattern Recognit. 41 (11) (2008) 3287–3294.
[70] C.-H. Zheng, D.-S. Huang, L. Shang, Feature selection in independent component subspace for microarray data classification, Neurocomputing 69 (16–18) (2006) 2407–2410.
[71] C.-H. Zheng, D.-S. Huang, K. Li, G. Irwin, Z.-L. Sun, Misep method for postnonlinear blind source separation, Neural Comput. 19 (9) (2007) 2557–2578.
[72] C.-H. Zheng, D.-S. Huang, Z.-L. Sun, M.R. Lyu, T.-M. Lok, Nonnegative independent component analysis based on minimizing mutual information technique, Neurocomputing 69 (7–9) (2006) 878–883.
[73] F. Han, D.-S. Huang, Improved extreme learning machine for function approximation by encoding a priori information, Neurocomputing 69 (16–18) (2006) 2369–2373.
[74] J.-X. Du, D.-S. Huang, X.-F. Wang, X. Gu, Computer-aided plant species identification (CAPSI) based on leaf shape matching technique, Trans. Inst. Meas. Control 28 (3) (2006) 275–285.
[75] W. Jiang, D.-S. Huang, S. Li, Random walk-based solution to triple level stochastic point location problem, IEEE Trans. Cybern. 46 (6) (2016) 1438–1451.
[76] X.F. Wang, D.S. Huang, A novel multi-layer level set method for image segmentation, J. Univers. Comput. Sci 14 (14) (2008) 2428–2452.
[77] Z.-L. Sun, D.-S. Huang, C.-H. Zheng, L. Shang, Optimal selection of time lags for TDSEP based on genetic algorithm, Neurocomputing 69 (7–9) (2006) 884–887.

[78] Z.-L. Sun, D.-S. Huang, Y.-M. Cheung, J. Liu, G.-B. Huang, Using FCMC, FVS, and PCA techniques for feature extraction of multispectral images, IEEE Geosci. Remote Sens. Lett. 2 (2) (2005a) 108–112.
[79] Z.-L. Sun, D.-S. Huang, Y.-M. Cheun, Extracting nonlinear features for multispectral images by FCMC and KPCA, Digit Signal Process. 15 (4) (2005b) 331–346.
[80] S.H. Lin, S.Y. Kung, L. J. Lin, Face recognition/detection by probabilistic decision-based neural network, IEEE Trans. Neural Netw. 8 (1) (1997) 114–132.
[81] C. Xiang, X.A. Fan, T.H. Lee, Face recognition using recursive fisher linear discriminant, IEEE Trans. Image Process. 15 (8) (2006) 2097–2105, doi:10.1109/TIP.2006.875225.
[82] I. Masi, Y. Wu, T. Hassner, P. Natarajan, Deep face recognition: a survey, in: Proceedings of the Conference on Graphics, Patterns and Images, IEEE, 2018, pp. 471–478.
[83] F. Han, Q.-H. Ling, D.-S. Huang, An improved approximation approach incorporating particle swarm optimization and a priori information into neural networks, Neural Comput. Appl. 19 (2) (2010) 255–261.
[84] D.S. Huang, H.H.S. Ip, K.C.K. Law, Z. Chi, Zeroing polynomials using modified constrained neural network approach, IEEE Trans. Neural Netw. 16 (3) (2005) 721–732.
[85] J.X. Du, D.S. Huang, G.J. Zhang, Z.F. Wang, A novel full structure optimization algorithm for radial basis probabilistic neural networks, Neurocomputing 70 (1–3) (2006) 592–596.
[86] Z.Q. Zhao, D.S. Huang, W. Jia, Palmprint recognition with 2DPCA+ PCA based on modular neural networks, Neurocomputing 71 (1–3) (2007) 448–454.
[87] D.-S. Huang, Radial basis probabilistic neural networks: model and application, Int. J. Pattern Recognit. Artif. Intell. 13 (7) (1999) 1083–1101.
[88] W.-B. Zhao, D.-S. Huang, J.-Y. Du, L.-M. Wang, Genetic optimization of radial basis probabilistic neural networks, Int. J. Pattern Recognit Artif. Intell. 18 (08) (2004) 1473–1499.
[89] J.-X. Du, D.-S. Huang, X.-F. Wang, X. Gu, Shape recognition based on neural networks trained by differential evolution algorithm, Neurocomputing 70 (4–6) (2007) 896–903.
[90] F. Han, Q.-H. Ling, D.-S. Huang, Modified constrained learning algorithms incorporating additional functional constraints into neural networks, Inf. Sci. (Ny) 178 (3) (2008) 907–919.
[91] F. Han, D.-S. Huang, A new constrained learning algorithm for function approximation by encoding a priori information into feedforward neural networks, Neural Comput. Appl. 17 (5–6) (2008) 433–439.
[92] D.-S. Huang, W.-B. Zhao, Determining the centers of radial basis probabilistic neural networks by recursive orthogonal least square algorithms, Appl. Math. Comput. 162 (1) (2005) 461–473.
[93] D.-S. Huang, S.-D. Ma, Linear and nonlinear feedforward neural network classifiers: a comprehensive understanding, J. Intell. Syst. 9 (1) (1999) 1–38.
[94] D.-S. Huang, J.-X. Du, A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks, IEEE Trans. Neural Netw. 19 (12) (2008) 2099–2115.
[95] Z.Q. Zhao, D.S. Huang, A mended hybrid learning algorithm for radial basis function neural networks to improve generalization capability, Appl. Math. Model. 31 (7) (2007) 1271–1281.
[96] D.S. Huang, A constructive approach for finding arbitrary roots of polynomials by neural networks, IEEE Trans. Neural Netw. 15 (2) (2004) 477–491.
[97] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, Deepface: Closing the gap to human-level performance in face verification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708.
[98] F. Schroff, D. Kalenichenko, J. Philbin, Facenet: a unified embedding for face recognition and clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815–823.
[99] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, L. Song, Sphereface: deep hypersphere embedding for face recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 212–220.
[100] X. Wu, R. He, Z. Sun, T. Tan, A light CNN for deep face representation with noisy labels, IEEE Trans. Inf. Forensics Secur. 13 (11) (2018) 2884–2896.
[101] Z. Lu, X. Jiang, A. Kot, Deep coupled resnet for low-resolution face recognition, IEEE Signal Process Lett 25 (4) (2018) 526–530.
[102] Y. Zhong, J. Chen, B. Huang, Toward end-to-end face recognition through alignment learning, IEEE Signal Process Lett 24 (8) (2017) 1213–1217.
[103] I. Masi, F.-J. Chang, J. Choi, S. Harel, J. Kim, K. Kim, J. Leksut, S. Rawls, Y. Wu, T. Hassner, et al., Learning pose-aware models for pose-invariant face recognition in the wild, IEEE Trans. Pattern Anal. Mach. Intell. 41 (2) (2018) 379–393.
[104] C. Darwin, The expression of the Emotions In Man and Animals, Oxford University Press, USA, 1998.
[105] P. Ekman, W.V. Friesen, Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues, Ishk, 2003.
[106] B. Fasel, J. Luettin, Automatic facial expression analysis: a survey, Pattern Recognit. 36 (1) (2003) 259–275.
[107] P. Dulguerov, F. Marchal, D. Wang, C. Gysin, Review of objective topographic facial nerve evaluation methods, Otol. Neurotol. 20 (5) (1999) 672–678.
[108] V. Bettadapura, Face expression recognition and analysis: the state of the art, arXiv:1203.6722(2012).
[109] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042–1052.
[110] L. Sirovich, M. Kirby, Low-dimensional procedure for the characterization of human faces, J. Opt. Soc. Am. 4 (3) (1987) 519–524.
[111] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 1994, pp. 84–91.
[112] D. Valentin, H. Abdi, A.J. O'Toole, G.W. Cottrell, Connectionist models of face processing: a survey, Pattern Recognit. 27 (9) (1994) 1209–1230.
[113] C. Liu, Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance, IEEE Trans. Pattern Anal. Mach. Intell. 28 (5) (2006) 725–737.
[114] A.M. Martinez, Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class, IEEE Trans. Pattern Anal. Mach. Intell. 24 (6) (2002) 748–763.
[115] A.M. Martinez, Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1, IEEE, 2000, pp. 712–717.
[116] X. Tan, S. Chen, Z.-H. Zhou, J. Liu, Face recognition under occlusions and variant expressions with partial similarity, IEEE Trans. Inf. Forensics Secur. 4 (2) (2009) 217–230.
[117] X. Tan, S. Chen, Z.-H. Zhou, F. Zhang, Recognizing partially occluded, expression variant faces from single training image per person with som and soft K-NN ensemble, IEEE Trans. Neural Netw. 16 (4) (2005) 875–886.
[118] S.Z. Li, X.W. Hou, H.J. Zhang, Q.S. Cheng, Learning spatially localized, parts-based representation, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, IEEE, 2001, pp. I–207.
[119] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2) (2009) 210–227.
[120] A.M. Martinez, The ar face database, CVC Technical Report 24 (1998).
[121] R. Singh, M. Vatsa, A. Noore, Face recognition with disguise and single gallery images, Image Vis Comput 27 (3) (2009) 245–257.
[122] Y. Ivanov, B. Heisele, T. Serre, Using component features for face recognition, in: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, IEEE, 2004, pp. 421–426.
[123] M.D. Burt, D.I. Perrett, Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information, R. Soc. Lond. B 259 (1995) 137–143.
[124] Y. Wu, N.M. Thalmann, D. Thalmann, A dynamic wrinkle model in facial animation and skin aging, J. Vis. Comput. Animat. 6 (4) (1995) 195–202.
[125] A. Lanitis, C. Taylor, T. Cootes, Toward automatic simulation of aging effects on face images, IEEE Trans. Pattern Anal. Mach. Intell. 24 (4) (2002) 442–455.
[126] M. Gandhi, A Method for Automatic Synthesis of Aged Human Facial Images, McGill University, 2004.
[127] Z. Liu, Z. Zhang, Y. Shan, Image-based surface detail transfer, IEEE Comput. Graph. Appl. 24 (3) (2004) 30–35.
[128] J. Wang, Y. Shang, G. Su, X. Lin, Age simulation for face recognition, in: Proceedings of the International Conference on Pattern Recognition, 3, IEEE Computer Society, 2006, pp. 913–916.
[129] N. Ramanathan, R. Chellappa, Face verification across age progression, IEEE Trans. Image Process. 15 (11) (2006a) 3349–3361.
[130] N. Ramanathan, R. Chellappa, Modeling age progression in young faces, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, IEEE Computer Society, Washington, DC, USA, 2006b, pp. 387–394.
[131] U. Park, Y. Tong, A.K. Jain, Age-invariant face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (5) (2010) 947–954.
[132] G. Mahalingam, C. Kambhamettu, Age invariant face recognition using graph matching, in: Proceedings of the IEEE International Conference on Biometrics: Theory Applications and Systems, IEEE, 2010, pp. 1–7.
[133] Y. Wen, Z. Li, Y. Qiao, Latent factor guided convolutional neural networks for age-invariant face recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4893–4901.
[134] H. Zhou, K.-M. Lam, Age-invariant face recognition based on identity inference from appearance age, Pattern Recognit. 76 (2018) 191–202.
[135] X. Hou, S. Ding, L. Ma, Robust feature encoding for age-invariant face recognition, in: Proceedings of the IEEE International Conference on Multimedia and Expo, IEEE, 2016, pp. 1–6.
[136] X. Zou, J. Kittler, K. Messer, Illumination invariant face recognition: a survey, in: Proceedings of the IEEE International Conference on Biometrics: Theory, Applications, and Systems, IEEE, 2007, pp. 1–8.
[137] N. Gudur, V. Asari, Gabor wavelet based modular PCA approach for expression and illumination invariant face recognition, in: Proceedings of the Applied Imagery and Pattern Recognition Workshop, 2006, p. 13.
[138] K. Karande, S. Talbar, Face recognition under variation of pose and illumination using independent component analysis, Int. J. Graph. Vis. Image Process. 8 (4) (2008) 1–6.
[139] B.C. Lovell, S. Chen, T. Shan, Robust adapted principal component analysis for face recognition, Int. J. Pattern Recognit. Artif. Intell. 23 (3) (2009) 491–520.
[140] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[141] A.U. Batur, M. Hayes, Linear subspaces for illumination robust face recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2001, pp. 296–301.
[142] K.-C. Lee, J. Ho, D. Kriegman, Nine points of light: Acquiring subspaces for face recognition under variable lighting, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, Citeseer, 2001, pp. I–519.

[143] Y. Gao, M.K.H. Leung, Face recognition using line edge map, IEEE Trans. Pattern Anal. Mach. Intell. 24 (6) (2002) 764–779.
[144] W.Y. Zhao, R. Chellappa, Illumination-insensitive face recognition using symmetric shape-from-shading, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1, IEEE, 2000, pp. 286–293.
[145] D. Jobson, Z.-u. Rahman, G. Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, IEEE Trans. Image Process. 6 (7) (1997) 965–976.
[146] H. Wang, S.Z. Li, Y. Wang, Face recognition under varying lighting conditions using self quotient image, in: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, IEEE, 2004, pp. 819–824.
[147] L. Qing, S. Shan, X. Chen, W. Gao, Face recognition under varying lighting based on the probabilistic model of gabor phase, in: Proceedings of the International Conference on Pattern Recognition, 3, IEEE, 2006, pp. 1139–1142.
[148] S. Shan, W. Gao, B. Cao, D. Zhao, Illumination normalization for robust face recognition against varying lighting conditions, in: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, IEEE, 2003, pp. 157–164.
[149] X. Xie, K.-M. Lam, An efficient illumination normalization method for face recognition, Pattern Recognit. Lett. 27 (6) (2006) 609–617.
[150] W. Chen, M.-J. Er, S. Wu, Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 36 (2) (2006) 458–466.
[151] D.-H. Liu, K.-M. Lam, L.-S. Shen, Illumination invariant face recognition, Pattern Recognit. 38 (10) (2005) 1705–1716.
[152] S.-I. Choi, C. Kim, C.-H. Choi, Shadow compensation in 2d images for face recognition, Pattern Recognit 40 (7) (2007) 2118–2125.
[153] A. Baradarani, Q. Wu, M. Ahmadi, Illumination suppression for illumination invariant face recognition, in: Proceedings of the International Conference on Pattern Recognition, IEEE, 2012, pp. 1590–1593.
[154] S. Li, F. Long, X. Cheng, J. Song, Illumination invariant face recognition by local image descriptor in logarithm domain, in: Proceedings of the International Symposium on Computational Intelligence and Design, 1, 2012, pp. 201–204.
[155] X. Xie, K.M. Lam, An efficient method for face recognition under varying, Illumination 4 (2005) 3841–3844.
[156] A.B.J. Teoh, Y.Z. Goh, M. Goh, Illuminated face normalization technique by using wavelet fusion and local binary patterns, in: Proceedings of the International Conference on Control, Automation, Robotics and Vision, IEEE, 2008, pp. 422–427.
[157] W. Zhang, Z. Xi, J.-M. Morvan, L. Chen, Improving shadow suppression for illumination robust face recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2018) 611–624.
[158] Y. Cheng, L. Jiao, X. Cao, Z. Li, Illumination-insensitive features for face recognition, Vis. Comput. 33 (11) (2017) 1483–1493.
[159] V. Starovoitov, D. Samal, B. Sankur, Matching of faces in camera images and document photographs, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 6, IEEE, 2000, pp. 2349–2352.
[160] O. Monga, R. Deriche, G. Malandain, J.P. Cocquerez, Recursive filtering and edge tracking: two primary tools for 3d edge detection, Image Vis. Comput. 9 (4) (1991) 203–214.
[161] T. Bourlai, A. Ross, A. Jain, On matching digital face images against scanned passport photos, in: Proceedings of the IEEE International Conference on Biometrics, Identity and Security, 2009.
[162] J. Joy, S. Peter, N. John, Denoising using soft thresholding, Int. J. Adv. Res. Electr. Electron Instrum. Eng. 2 (3) (2013) 1027–1032.
[163] R.R. Coifman, D.L. Donoho, Translation-invariant De-noising, in: Wavelets and Statistics, Springer, 1995, pp. 125–150.
[164] D.L. Donoho, De-noising by soft-thresholding, IEEE Trans. Inf. Theory 41 (3) (1995) 613–627.
[165] F. Juefei-Xu, K. Luu, M. Savvides, T.D. Bui, C.Y. Suen, Investigating age invariant face recognition based on periocular biometrics, in: Proceedings of the International Joint Conference on Biometrics, IEEE, 2011, pp. 1–7.
[166] A.U. Batur, M. Hayes, Linear subspaces for illumination robust face recognition, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2, 2001, pp. 296–301.

**Umarani Jayaraman** did her Ph.D. from IIT Kanpur and started her career in 2014 by joining the institute “Indian Institute of Information Technology, Design & manufacturing, kancheepuram”, Chennai, India as a faculty in the Department of Computer Science and Engineering. She published five journal papers and more than ten international conference papers in the field of biometrics and pattern recognition. She also been associated with the project sponsored by the Department of Information Technology, Government of India, Delhi, India. She is the reviewer for the top most journals and conferences happening all over the world. Her research interest includes Digital Image Processing, Biometrics and Pattern Recognition.

**Dr. Phalguni Gupta** did his Ph.D. from IIT Kharagpur and started his career in 1983 by joining in Space Applications Centre (ISRO) Ahmedabad, India, as a Scientist. In 1987, he joined the Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, India. Currently he is director at National Institute of Technical Teachers' Training & Research Kolkata and also a Professor in the department of CSE IIT Kanpur. He works in the field of Data Structures, Sequential Algorithms, Parallel algorithms, Online Algorithms, Image Analysis, and Biometrics. He has published more than 250 papers in International Journals and Conferences. He has dealt with several sponsored and consultancy projects which are funded bythe Government of India. Some of these projects are in the area of Biometrics, System Solver, Grid Computing, Image Processing, Mobile Computing, and Network Flow. During this period he has proved himself a well known researchers in theoretical computer science; especially in the field of Biometrics.

**Dr Sandesh Gupta** is an Assistant Professor at University Iinstitute of Engineering and Technology Kanpur. He works on the field of Digital Image Processing, Biometric Security and Medical Imaging. He holds patent for the algorithms deloped towards Cancer detection from Lung X-ray radiographs. He has organized many International and National conference/workshops/symposia. He is In-Charge of operations for ACM International Collegiate Programming Contest (ICPC) as Associate Regional Contest Director.

**Ms. Geetika Arora** is currently a Ph.D. candidate at Birla Institute of Technology and Science Pilani, India. She received M. Tech. degree in Computer Science and Engineering from the Department of Computer Science and Engineering, at Banasthali Vidyapith Rajasthan, India, in 2016. Her research interest includes Machine Learning, Biometrics, System Securities, and Image Processing.

**Dr. Kamlesh Tiwari** is working as an assistant professor at department of Computer Science and Information System at Birla Institute of Technology and Science Pilani. He have earned his Ph.D. degree from the department of Computer Science and Engineering of Indian Institute of Technology Kanpur in 2016. His Ph.D. work was in the area of biometrics. He is a member of IEEE and Signal Processing Society (SPS). His research interests include Machine Learning, Computer Vision, Multimodal Biometric (Fingerprint, Face, Palmprint, knuckleprint) and Security. He is co-incharge of Advanced Data Analytics & Parallel Technologies Lab at BITS Pilani. He is also an active member of Multimedia & HCI Laboratory.