

# A Comprehensive Study on Artificial Intelligence-Based Face Recognition Technologies



Sachin Kolekar, Pratiksha Patil, Pratiksha Barge, and Tarkeshwari Kosare

**Abstract** Face recognition technology is a technique that recognizes and authenticates individuals based on their unique facial features. It has various applications, including security, access control, identity verification, and social media tagging. These systems use algorithms to analyze facial traits and create a facial template for comparison with a database of known faces. Advances in machine learning and computer vision have improved the accuracy of face recognition technology. However, concerns about the privacy and security implications of biometric data collection and storage have arisen. This paper provides an overview of the history, techniques, algorithms, applications, performance evaluation metrics, challenges, and future directions of face recognition technology.

**Keywords** Convolutional neural network · Deep convolutional neural network · Principal component analysis · Linear discriminant analysis · Deep learning

## 1 Introduction

Due to its use in security, surveillance, access control, and personalized marketing, face recognition is a rapidly expanding topic. It involves extracting and recognizing human faces from photos or videos using computer algorithms. There has been an increase in research targeted at creating more precise, effective, and reliable face recognition systems as the need for face recognition technology keeps growing. From basic machine learning approaches to more complex deep learning architectures, these systems rely on a range of methodologies. This survey paper's goal is to give readers a broad overview of facial recognition technology as it stands right now. The

---

S. Kolekar · P. Patil (✉) · P. Barge · T. Kosare  
JSPMs Rajarshi Shahu College of Engineering, Pune, Maharashtra 411033, India  
e-mail: [pratikshap2404@gmail.com](mailto:pratikshap2404@gmail.com)

S. Kolekar  
e-mail: [smkolekar@jspmrscoe.edu.in](mailto:smkolekar@jspmrscoe.edu.in)

most widely used algorithms, approaches, and datasets will be covered, along with the most recent research and breakthroughs in the area.

Face recognition technology enables a computer system to identify or confirm a person's identity by examining and comparing their facial features to a database of recognized faces. The system employs algorithms that examine and contrast many facial characteristics, including the separation between a person's eyes, the nose's shape, the breadth of the mouth, and the curves of the face. Face recognition has several applications, including social networking, biometric authentication, security systems, and law enforcement. It frequently works in conjunction with other biometric technologies, such as iris and fingerprint scanning, to provide more reliable and accurate identification systems.

Face recognition is a complex process that involves multiple steps. The system uses object detection techniques in the initial step of face detection to find and locate human faces in an image or video. The system aligns the discovered faces to a standardized position and orientation in the subsequent stage of face alignment. The algorithm then extracts the important facial characteristics from each face, such as the separation between the eyes, the contours of the nose, and the face's contours, which are often represented as mathematical vectors. The system then does face matching, where it checks the feature vectors of a detected face against a database of recognized faces to find a match. Deep neural networks and support vector machines are two common machine learning algorithms used for this. The system finally decides based on the degree of similarity between the feature vectors of the discovered face and the database of recognized faces. The system will recognize the face as a match if the score rises above a particular cutoff. If not, it will not accept the face. A face recognition system's performance and accuracy are influenced by a number of parameters, including the quality of the input photos, the complexity of the facial features, and the quantity and caliber of the database of recognized faces. However, there are a number of difficulties that need to be overcome while face matching, including changes in lighting, position, expression, and occlusion. Researchers are creating more sophisticated algorithms that can manage these difficulties and use extra features, such as 3D face models, to improve the durability and accuracy of face matching.

The face recognition process typically involves the following steps:

1. Face detection
2. Feature extraction
3. Face matching
4. Verification or identification.

## 2 Related Work

Survey of various studies are shown in Table 1.

**Table 1** Literature survey

S. No.	Year Dataset used	Technique/methods used	Performance metrics	Challenges/future scope
1	2013 Dataset MOBIO database	(1) Principal component analysis [1]		Future work will focus on biometrics, a system that integrates the results of face and speaker verification
2	2015	(1) Principal component analysis (2) LDA (3) ICA	93.7% [2]	PCA improvements are not considered in this work. Basic face recognition by integration with specific filter, such as contrast enhancement techniques and nerve training. In the future, we plan to integrate some of the proposed methods. Known noise reduction and image enhancement methods
3	2016 Dataset: Image database	(1) Viola Jones algorithm (2) PCA	90% [3]	Face detection and identification will become more precise in the future. The Viola Jones Algorithm is a fast face recognition algorithm. One of the drawbacks of this algorithm is that it has some false positive values for photographs with occluded faces. Thus future research will focus on lowering false positives
4	2016 Dataset MS1MV2	(1) Deep CNN (2) SGD algorithm (3) Multi face training mechanism	96.08% [4]	
5	2016 Dataset YouTube faces database	(1) Deep learning (2) Convolutional neural networks [5]		The next research direction should be the mobile face unlimited recognition system
6	2017 Dataset ImageNet 2012	(1) Deep reinforcement learning (DRL) (2) Convolutional neural networks (CNN) [6]		The facial image recognition accuracy of existing CNN schemes can be improved using the model proposed in this paper

(continued)

**Table 1** (continued)

S. No.	Year Dataset used	Technique/methods used	Performance metrics	Challenges/future scope
7	2018 Dataset Celebrity face dataset	(1) VGGNet arch [7]	95%	Future work in this area will include improving the experience of facial age, classification of human facial expressions to support facial recognition, detection of facial disorders, images, social media images, and more
8	2019 Dataset of Single Person	(1) CNN (2) Deep learning (3) Fisher faces (4) Eigen faces (5) LPBH	97.95%. [8]	Further research can be done to suggest a more robust and accurate face detection algorithms in a variety of uncontrolled and difficult scenarios
9	2019	(1) Haar cascade (2) Eigen faces (3) LBPH (4) Fisher faces [9]	2) 78% 3)82% 4)86%	A smart camera that recognizes face traits in videos and stores them in a database for security purposes. For schools and universities, an automatic attendance system is available. In a controlled environment, face recognition is possible. Neural networks will be able to separate people with great resemblance in Faceswins by increasing the amount of levels and information per class
10	2020 Dataset ORL	(1) PCA (2) LDA	94% [10]	If you consider other databases like YALE and GTF datasets, you will find other facial recognition issues such as reorientation, lighting, poses, and facial expression changes. The results obtained through this study can be improved by applying and testing different other facial recognition techniques

(continued)

**Table 1** (continued)

S. No.	Year Dataset used	Technique/methods used	Performance metrics	Challenges/future scope
11	2020 Dataset ahs	(1) FaceNet	90% [11]	In future this work can prove to be helpful to police in order to recognize the person within the fraction of time. It can also be used to recognize an individual using video capture that will help to identify an individual from CCTV cameras. It can also be used in different security systems like: Home security system • Visitors analysis system In future face detection as well as time detection while faces appear in video can be implemented
12,	2021 Dataset (1) WIDER FACE (2) LFW (3) FDDB (4) FERET	(1) R-CNN (2) R-FCN (3) SSD (4) Deep learning	91.8% [11]	
13	2021 Dataset Kaggle's medical mask dataset	(1) Bottleneck (2) Convolutional neural network [12]	(1) 0.9264 score (2) 0.93 <i>F1</i> score	
14	2021 Dataset Real-world masked face dataset (a) RMFRD (b) SMFRD	(1) VGG-16 (2) Alex Net (3) ResNet-50	(a) 91.3% (b) 88.9% [13]	In future work, we will consider adding a pre-trained model and applying a deep ensemble model to improve accuracy

(continued)

**Table 1** (continued)

S. No.	Year Dataset used	Technique/methods used	Performance metrics	Challenges/future scope
15	2021	(1) Deep convolutional neural network [14]		To overcome these issues, we offer MultiFace, a simple yet effective training technique in which the original high-dimensional characteristics are approximated by an ensemble of low-dimensional features
16	2022 Dataset MS-Celeb-1M VGGface2 Megaface WebFace260M	(1) Face pre-processing (2) Deep feature extraction (3) Training loss for face matching	(a) 97% [15]	

### 3 Techniques Used

#### 3.1 Deep Convolutional Neural Networks

In order to analyze and interpret pictures or other forms of structured data, deep convolutional neural networks (CNNs) were created. Multiple layers of convolutional and pooling processes are the foundation of a CNN's design, which is followed by one or more fully linked layers. The local patterns and characteristics in the input data, such as edges, corners, and textures, are recognized by the convolutional layers. These layers take the input data and run it through a series of learnt filters, creating a number of feature maps that represent various facets of the original image. The feature maps are subsequently downsampled by the pooling layers, which lowers their spatial resolution while maintaining the most crucial characteristics. The network's fully connected layers employ the high-level properties discovered by the convolutional layers to predict outcomes based on the input data. For instance, the fully connected layers may provide a probability distribution across a set of potential classes in an image classification problem. Deep CNNs have demonstrated outstanding performance across a range of computer vision applications, including segmentation, object identification, and picture classification. They've also been used with other kinds of structured data, such audio signals and text written in natural language structure of CNN network is shown in Fig. 1.

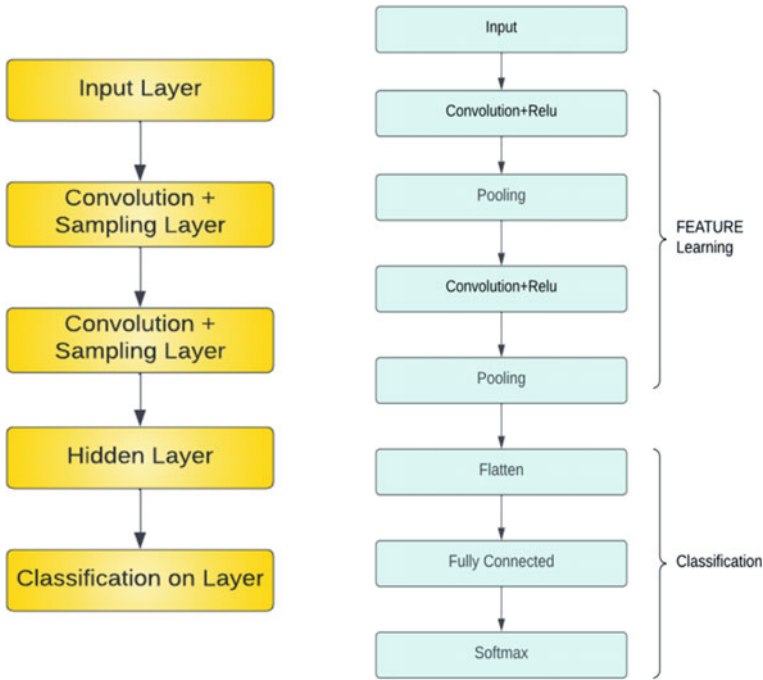


Fig. 1 Structure of CNN Network

### 3.2 Deep Face

DeepFace begins by recognizing faces in images with a pre-trained face detector. It then normalizes for changes in position and illumination by aligning the faces to a canonical stance. The aligned faces are then sent into the deep neural network, which learns to extract discriminative high-level properties for diverse faces. These characteristics are then utilized to calculate a similarity score between faces, which is subsequently used for face recognition. DeepFace was trained on a huge dataset of over 4 million labeled faces, and it performed well on multiple benchmark face recognition datasets, including Labeled Faces in the Wild (LFW) and YouTube Faces (YTF). On the LFW dataset, it achieved an accuracy of 97.35%, exceeding other classical and deep learning-based face recognition systems.

The DeepFace algorithm in face recognition involves the following steps:

- Face detection
- Face alignment
- Feature extraction
- Similarity scoring
- Face recognition
- Training.

### 3.3 *VGG-Face*

VGG-facial is a deep convolutional neural network architecture optimized for facial recognition applications. It is built on the VGG-16 network architecture and was trained on the VGG-Face dataset, a large-scale collection of faces. VGG-Face is made up of 13 convolutional layers and 3 fully linked layers. A face picture is sent into the network, which is subsequently processed via the layers to extract characteristics for face recognition. The purpose of face recognition is to compare a face image to a database of known faces and see whether there is a match. The VGG-Face network is capable of extracting features from both the probe (input) and gallery (database) faces. To see if there is a match, the characteristics can be compared using a similarity measure such as cosine similarity. Overall, VGG-Face has shown to be a highly successful face recognition architecture, delivering cutting-edge outcomes on a range of benchmark datasets.

### 3.4 *Capsule Networks*

Capsule Networks (CapsNets) are a form of neural network architecture developed as an alternative to classic convolutional neural networks (CNNs) for image recognition applications. Capsule Networks include a unique building component known as a capsule, which is a collection of neurons that represent the instantiation characteristics of a single item or feature in an image. Capsule Networks may be used in face recognition to identify and recognize facial features such as the eyes, nose, mouth, and other facial traits. Capsules in the network can learn to represent these features as unique entities with connections to one another, rather than merely as individual pixels or picture attributes. Face recognition tasks have showed potential for capsule networks, particularly in situations where there may be significant fluctuations in lighting, position, and facial expression. The network's ability to accurately capture the spatial connections between the various facial characteristics by modeling them as capsules helps increase identification rates. Although Capsule Networks have demonstrated some potential, further study is required to fully grasp their advantages and disadvantages in face recognition and other applications.

### 3.5 *3D Face Recognition*

A sort of face recognition technology called 3D face recognition utilizes 3D reconstructions of people's faces to recognize and identify them. 3D face recognition employs a 3D model of the face to capture the form, texture, and other physical aspects of the face, in contrast to standard 2D face recognition, which uses 2D photographs of the face. Lighting circumstances, position fluctuations, and facial



emotions are some of the 2D face recognition's drawbacks that can be overcome using 3D face recognition. Since 3D models record the form of the face, they may be used to identify faces in a variety of lighting and viewing angles. Several techniques, such as structured light scanning, stereo photogrammetry, and laser scanning, can be used to create 3D representations of the face. These techniques record the facial form from several angles, which may be merged to create a 3D representation of the face. A 3D model of the face may be used to extract features for facial recognition once it has been created. These characteristics might include the how the face is shaped, how the nose and forehead curve, and how deep the eyes and lips are. In general, 3D face recognition has showed promise as a more reliable and accurate method of face recognition, especially in difficult situations where 2D face recognition may struggle. It can be more difficult to implement in some applications since it also calls for more complicated hardware and computation.

### ***3.6 Principal Component Analysis***

Using principal component analysis (PCA), face recognition software may decrease the dimensionality of face photographs while still preserving key facial traits. The method involves projecting the data onto a lower-dimensional space based on the key elements or traits that make up the majority of the variance in the dataset. In order to extract a set of eigenfaces from a huge number of face photos, PCA is frequently employed in face recognition. These eigenfaces, which may be used to categorize and identify novel faces, represent the most crucial characteristics or patterns in the faces. Finding the most similar eigenface from a fresh face picture projected onto the eigenface space is the first step in the recognition process.

Algorithm for PCA:

- Step 1: Standardization.
- Step 2: Computation of covariance matrix.
- Step 3: Calculate the eigenvectors and eigenvalues of the covariance matrix to identify principal components.
- Step 4: Feature vector.
- Step 5: Recast the data along the principal component axes.

Face recognition used PCA is shown in Fig. 2.

### ***3.7 Linear Discriminant Analysis***

Another popular face recognition method, linear discriminant analysis (LDA), is similar to PCA but has a different objective. LDA is used for feature extraction and classification whereas PCA is used to reduce dimensionality. In order to distinguish

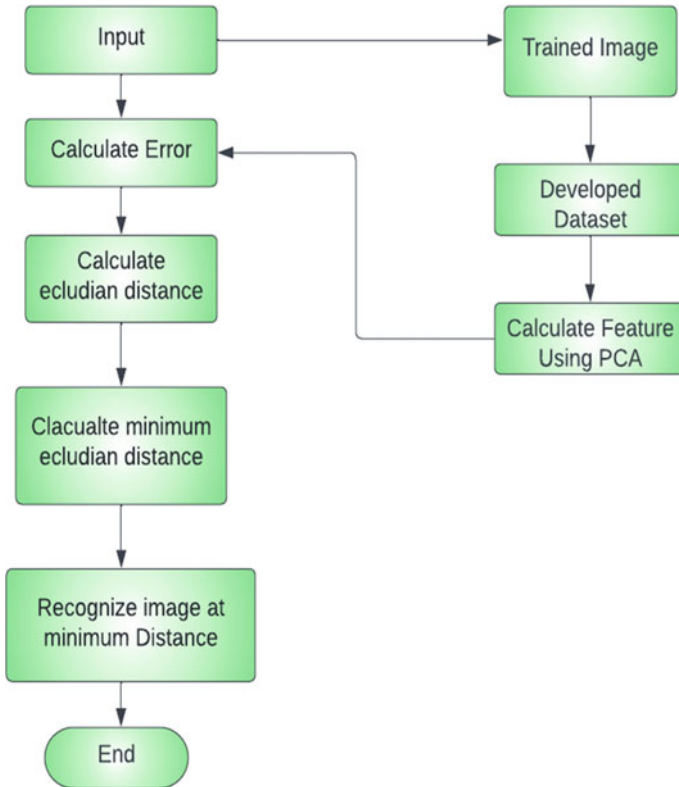


Fig. 2 Face recognition used PCA

between distinct classes of faces, LDA in face recognition seeks out a linear combination of characteristics. In order to do this, the distribution of the face photos is modeled in a high-dimensional space, and the data is then projected onto a lower-dimensional space that maximizes the separation between the classes. The resulting traits, known as “fisherfaces,” can be used to classify and recognize new faces. To determine the closest match, a new face image is projected onto the fisherface space and compared to the existing fisherfaces. LDA has been found to outperform PCA in face recognition tasks, particularly when working with datasets with a large number of classes. However, labeled training data is required for LDA to perform properly, which can be a constraint in some applications. Face recognition LDA Subspace is shown in Fig. 3. Principle and PCA of LDA approach is shown in Figs. 4 and 5.

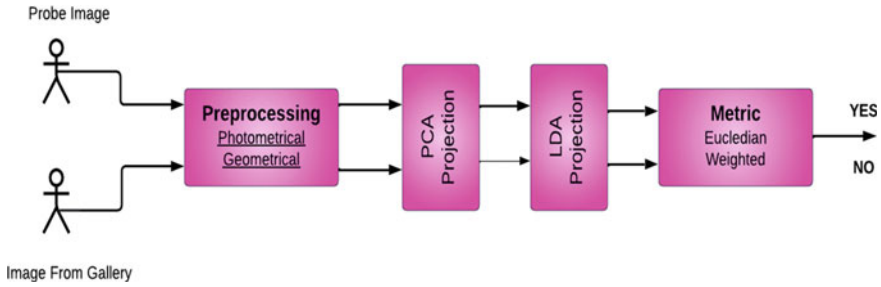
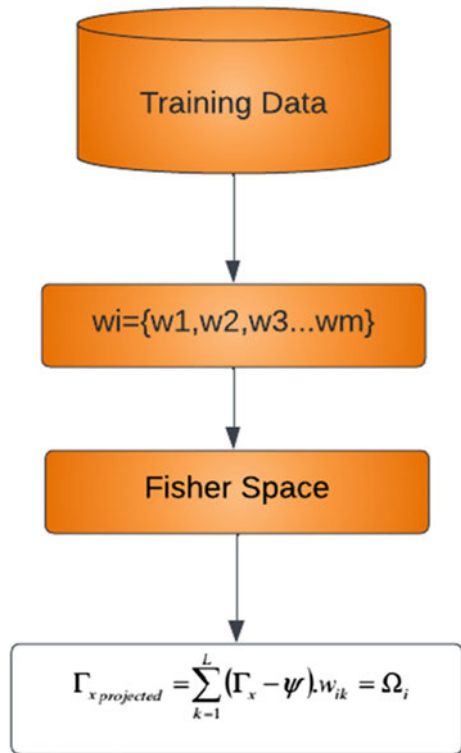


Fig. 3 Face recognition LDA subspace

Fig. 4 Principle of LDA approach



### 3.8 FaceNet

FaceNet is a deep learning-based facial recognition system created in 2015 by Google researchers. It learns the characteristics and patterns of faces using a deep neural network and encodes each face as a high-dimensional vector in a feature space.

FaceNet’s facial recognition algorithm consists of the following steps:

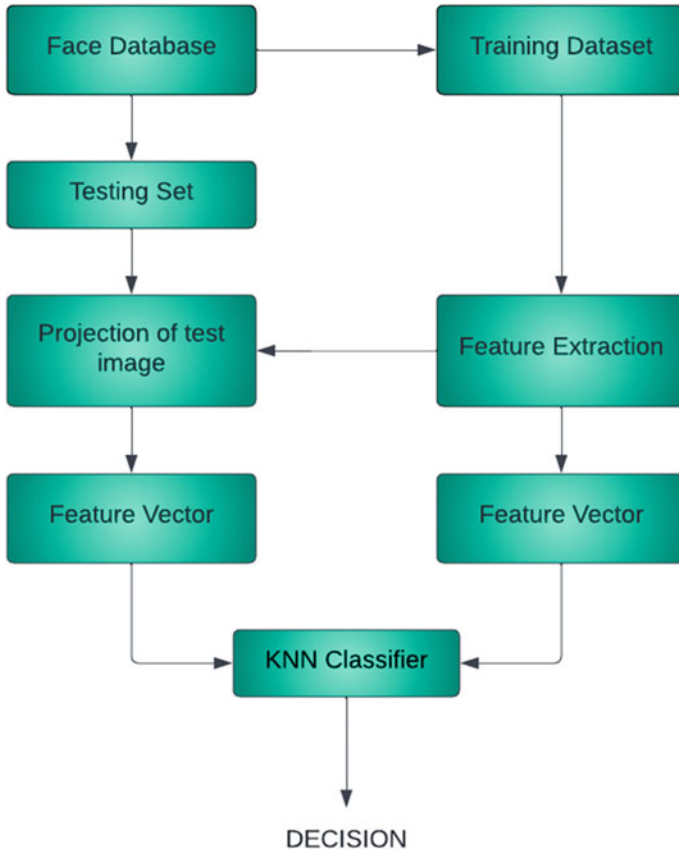


Fig. 5 PCA LDA approach

1. Face detection and alignment.
2. Triplet loss function.
3. Training.
4. Face representation.
5. Face recognition.

### 3.8.1 Euclidean Distance

A series of facial traits from each face, such as the eye location, nose shape, and mouth size, must first be retrieved in order to determine the Euclidean distance between two faces. In a high-dimensional space, these features are often represented as a collection of coordinates.

The Euclidean distance between the two faces can be determined using the formula below once the feature vectors for the two faces have been obtained:

$$\text{distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + \dots + (n_2 - n_1)^2}$$

Where  $x_1, y_1, \dots, n_1$  represent the coordinates of the first face's feature vector, and  $x_2, y_2, \dots, n_2$  represent the coordinates of the second face's feature vector. The square root of the sum of the squared differences is the Euclidean distance.

## 4 Proposed Model

The proposed model for face recognition authentication in a library is a computer vision system that utilizes deep learning techniques to identify and verify individuals based on their facial features. Only authorized users will be able to access library resources and services thanks to the model's ability to recognize faces properly.

The system also includes a database of every book in the collection, together with information about where each book is located on the shelves and its borrowing history. For the convenience of library patrons and librarians, the system stores and manages this data.

A camera for taking pictures of library patrons' faces, a deep neural network trained on a large dataset of facial images to recognize patterns and features of human faces, and an authentication mechanism that compares the input image with the images of authorized users stored in the database make up the system.

An image of the user's face must be taken in order to begin the face recognition procedure. The user's face is captured by the camera, which then sends the image to the deep neural network for analysis. Convolutional neural networks (CNNs) are used by the neural network to extract important elements from the image, such as the separation between the eyes, the profile of the nose, and the curve of the lips.

The system compares these properties to those in the database once the neural network has extracted them. A collection of pre-registered faces of library patrons who have permission to utilize the library's resources and services can be found in the database. The system authenticates the user and gives access to the library services if the features of the input image match the features of any of the registered faces.

Users of the system can browse and search the library's book collection using an intuitive interface on a PC or mobile device. Additionally, they may view each book's availability and shelf location.

Each book's borrowing history is recorded by the system, which makes it possible to determine the most read titles in the collection and inform future book purchase choices. By keeping track of borrowing and return records, it also aids in preventing theft or loss of library books.

The proposed facial recognition authentication model for library services is, in general, a trustworthy and secure system that can offer seamless and practical access to library resources while guaranteeing the security and safety of the library's patrons and their data.

### **4.1 Future Scope in Proposed Model**

To improve user experience and offer personalized services, the suggested model can potentially be expanded to include further elements like age and gender detection. It can also be improved by adding extra security features like liveness detection to thwart spoofing and hacking attempts.

## **5 Application**

### **1. Real-Time Face Recognition**

Real-time facial recognition systems are projected to become increasingly widespread and widely employed in applications such as security, surveillance, and access control as high-performance computing and advanced algorithms become more commonly available.

### **2. Mobile and Wearable Devices**

Face recognition will very certainly be integrated into mobile and wearable devices, enabling for more easy and secure device authentication and unlocking.

### **3. Biometric Passports and Travel Documents**

Face recognition is currently being used in biometric passports and travel papers, and it is expected to become more common in the future, offering travelers with more security and convenience.

### **4. Personalized Advertising and Marketing**

Face recognition software can analyze facial expressions and emotions, enabling for more personalized advertising and marketing efforts that are tailored to individual tastes and requirements.

### **5. Healthcare**

Face recognition technology may be utilized in medical diagnosis, monitoring, and therapy, such as recognizing illness symptoms and remotely monitoring patient status.

## **6 Conclusion**

This survey paper has focused on various face recognition techniques. The evolution of face recognition from 2013 to 2021 is described using various algorithms. The challenges in various fields are also described such as Masked face detection,

Gender detection, Mobile face detection, Attendance system, Security systems. The solution is also provided for the challenges and the accuracy is obtained using various algorithms and methods.

## References

1. Günther M, Costa-Pazo A, Ding C, Boutellaa E, Chiachia G, Zhang H, de Assis Angeloni M et al (2013) The 2013 face recognition evaluation in mobile environment. In: 2013 international conference on biometrics (ICB). IEEE, pp 1–7
2. Kaur R, Himanshi E (2015) Face recognition using principal component analysis. In: 2015 IEEE international advance computing conference (IACC). IEEE, pp 585–589
3. Meena D, Sharan R (2016) An approach to face detection and recognition. In: 2016 international conference on recent advances and innovations in engineering (ICRAIE). IEEE, pp 1–6
4. Stoimenov S, Tsenov GT, Mladenov VM (2016) Face recognition system in Android using neural networks. In: 2016 13th symposium on neural networks and applications (NEUREL). IEEE, pp 1–4
5. Li H, Zhu X (2016) Face recognition technology research and implementation based on mobile phone system. In: 2016 12th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD). IEEE, pp 972–976
6. Wang P, Lin W-H, Chao K-M, Lo C-C (2017) A face-recognition approach using deep reinforcement learning approach for user authentication. In: 2017 IEEE 14th international conference on e-business engineering (ICEBE). IEEE, pp 183–188
7. Dhomne A, Kumar R, Bhan V (2018) Gender recognition through face using deep learning. *Procedia Comput Sci* 132:2–10
8. Khan S, Ahmed E, Hammad Javed M, Shah SAA, Ali SU (2019) Transfer learning of a neural network using deep learning to perform face recognition. In: 2019 international conference on electrical, communication, and computer engineering (ICECCE). IEEE, pp 1–5
9. Onyema EM, Shukla PK, Dalal S, Mathur MN, Zakariah M, Tiwari B (2021) Enhancement of patient facial recognition through deep learning algorithm: ConvNet. *J Healthcare Eng* 2021
10. Sharma S, Bhatt M, Sharma P (2020) Face recognition system using machine learning algorithm. In: 2020 5th international conference on communication and electronics systems (ICCES). IEEE, pp 1162–1168
11. Manna S, Ghildiyal S, Bhimani K (2020) Face recognition from video using deep learning. In: 2020 5th international conference on communication and electronics systems (ICCES). IEEE, pp 1101–1106
12. Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J (2021) SSDMNV2: a real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustain Cities Soc* 66:102692
13. Hariri W (2022) Efficient masked face recognition method during the covid-19 pandemic. *SIViP* 16(3):605–612
14. Xu J, Guo T, Xu Y, Xu Z, Bai K (2021) MultiFace: a generic training mechanism for boosting face recognition performance. *Neurocomputing* 448:40–47
15. Dalvi J, Bafna S, Bagaria D, Virmodkar S. A survey on face recognition systems