International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT228558

A Survey on Face Recognition Techniques in Machine Learning

Shantanu Jain*1, Vrushaket Chaudhari2, Rushikesh Chuadhari3, Tanvesh Chavan4 , Prof. Priyanka Shahane5

[*1,2,3,4]Student, Department of Computer Engineering, SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India

[5]Assistant Professor, Department of Computer Engineering, SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India

Article Info

Publication Issue :
Volume 8, Issue 6
November-December-2022

Page Number : 50-66

Article History
Accepted: 01 Nov 2022
Published: 05 Nov 2022

ABSTRACT

In the digital world, biometrics is used for authentication or recognition to examine and confirm a person's distinguishing physical or behavioral attributes. There are many authentication systems available today that use iris, fingerprint, and face features for identification and verification. Face recognition-based systems are the most popular since they don't always need the user's assistance, are more automated, and are simple to use.

Face recognition paves the way for an innovative way to perceive a human face. Face recognition and identification have been used in access control systems, which have become widely used in security frameworks during the past few years. With the help of biometrics, a facial recognition system can extract facial details from a picture or video. The data is compared to a database of recognized faces to identify a match. Personal identity can be confirmed through facial recognition. This review paper offers a comparison of various facial recognition methods.

Keywords: Image recognition, Face Recognition, Computer Vision, Artificial Intelligence, Convolution Neural Networks, Deep Learning

I. INTRODUCTION

Ensuring security in sectors such as commercial establishments, airports and banks is a difficult task, especially when human identification is required. According to [1], traditional human identification approaches such as PINs, tokens, and passwords used in magnetic credit cards are now considered obsolete because they can be easily bypassed. For example, a password or PIN can be forgotten or discovered by malicious individuals and used for illegal activities. Tokens that are already magnetic credit cards can be lost, stolen, or duplicated. Several techniques aimed at human identification are being researched in academia accurately and reliably to minimize problems with these approaches. Among these technologies, biometrics stands out, a field that uses multiple technologies to identify humans by measuring various physical and behavioral aspects of humans. Physical features include the face, iris, fingerprints, palms, palm veins, hand geometry, and DNA. Behavioral features are associated with signing and typing behaviors known as biometric modalities. Biometrics offers a solution that overcomes traditional identification techniques, as biological characteristics generally cannot be faked. Among the modalities mentioned above, the face is the most commonly used modality and is well-accepted in recognition systems. Since the advent and integration of photography, government agencies and private organizations have maintained databases of personal photographs for personal identification documents such as passports and ID cards.

With the advent of deep learning, many problems that were once solved by traditional machine learning models have performed better with this approach, but in facial recognition environments, deep learning outperforms traditional models. Feature extractor when applied to various face data databases. Therefore, the purpose of this paper is to conduct a comparative study of various traditional machine learning models and some deep learning techniques that can be used for face recognition.

This work combines machine learning techniques such as Support Vector Machines (SVMs), extreme learning machines, Artificial Neural Networks (ANNs), and several different feature extractors such as Local Binary Patterns (LBP) with deep learning. It is intended to conduct comparative studies - the architecture is used like a convolutional Neural Network model, all applied to face databases.

II. STEPS INVOLVED IN THE PROCESS OF FACE RECOGNITION

Facial recognition can be viewed as a method of authentication and verification. In this sense, new unknown faces are matched against various other faces present in the database. All of these faces have known entities. After this comparison, a result is returned indicating whether the face was recognized. Face identities are confirmed or denied by results obtained in comparison with face data available in databases. The facial recognition process consists of two main components to carry out the whole process: facial detection and facial recognition.

1. Face Detection

Detect faces or all faces in specific images or videos using a variety of detection techniques. Robustness to pose, lighting, and background removal improve face recognition. Viola-Jones [2] is the most commonly used face detector based on hair-like features and shows better results for frontal faces in real-time implementations. Several methods based on deep learning, such as the sliding window idea [3], Regions with Convolutional Neural Networks (R-CNN) [4], and single-shot detector (SSD) [5], are also used and have been used for face recognition. and gives good results.

2. Feature Extraction

By detecting and locating feature points of facial organs, the location of facial feature points (human facial features)can be determined. On the other hand, facial shape and organ-specific descriptive information can be determined by an algorithm, and an algorithm can obtain facial feature descriptions based on this information.

3. Face Recognition

The facial feature description information obtained in the previous step is compared with the facial information in the face database for facial recognition. One-to-many face matching and one-to-one authentication processing are mainly performed.[6]

III. ALGORITHMS

1. Traditional Face Recognition Technology

1.1 Face Recognition based on Principal Component Analysis (PCA)

Feature extraction methods are now an important part of face recognition. Principal Component Analysis (PCA) is one of the most widely used feature extraction methods. As an important statistical technique, PCA is widely used in fields such as image processing[7] and signal processing. The principle of principal component analysis is to isolate the main elements of the original spatial data and ignore the redundant data. This allows the extracted data to be processed in reduced-dimensional feature space to solve practical application problems.

Eigenface is commonly used in face detection algorithms based on principal component analysis. In this application, the image area containing the face can be viewed as a random vector and the facial features can be obtained after transformation. During recognition, the face map is mapped into the vector subspace formed by the feature vectors, and the positions of the recognized face map are compared in the feature space.

The specific steps of this algorithm are: to get a training set of face graphs and get the characteristic faces. Then we get the vector subspace formed by the characteristic faces and wait for the test. Import a new image under test, map it onto the vector subspace formed by the feature faces, and obtain facial feature data for the face image under test. Determine if a face is present in the image-under-test by determining the distance from the image-under-test to the vector subspace formed by the feature planes (if so, the behavior of objects in the database). Whether there is or not is judged based on the weight, and a concrete judgment is made. This step does not affect the algorithm and can be ignored.)

PCA in combination with other algorithms has been widely used for facial recognition applications over the past decades. kings. [8] For facial recognition, LBP and PCA in combination with the ABAS algorithm were used. Using a combination of LBP and PCA as feature extraction techniques, and the ABAS algorithm to optimize the Neural Network, here using the softmax function, we have constructed multiple algorithms to perform the face recognition process. Reduced face classification time.

Here, they used the ORL [9] dataset to test the proposed model and demonstrate its ability to handle multifaceted classification.

Wang et al. [10] used the F-2D-QPCA technique for face recognition. They used the F-norm to maximize the image variance and used a greedy iterative algorithm to improve the convergence and robustness of the method. Experimental results of this model on several facial image color databases showed its effectiveness and accuracy compared with other existing models. Because their method uses the image as a quaternion matrix that uses the color and spatial information of the image.

Kong et al. [11] proposed a CSGF (2D) 2PCANet algorithm for face recognition. He used CSGF for the proposed model to overcome the computational time and data redundancy problems of existing models. It consists of one stage for nonlinear output using linear SVM and two stages for feature extraction with good locality using 2DPCA. The proposed model showed higher detection rates when tested with AR [12], ORL databases, etc. It also had stable robustness against image fluctuations, which improved the accuracy of the model

Study & Pub. Year	**Method/Algorithm Dataset**	**Accuracy**
2020, [8]	Novel Multi-face ORL, ExtYaleB [14] and Recognition, ABASNet FERET [15]	ORL: 99.35%, ExtYaleB[14]: 99.54% FERET[15]: 99.18%
2020, [10]	A Quaternion PCA Method,	F-2D-QPCA- 72.29%,

	GT [16], GT-noise, F-2DQPCA GToutlier, FT and FT-outlier	QRR72.29%, QSR-71.86% (for 30 features)
2018, [11]	Deep Learning, XM2VTS [17], ORL, CSGF(2D)2PCANet Extend YaleB, LFW[18] and AR	XM2VTS [17]: 99.58%, ORL: 97.50%, Extend YaleB : 100%, LFW: 98.58%, AR: +97.50%
2017, [13]	Stacking-based CNNs, FERET, LFW and YTF PCANet+ [19]	94.23%
2016, [20]	LBP	MIT face database Without noise: 99.3%, With noise:98.28%

Table 1. PCA-based methods for face recognition.

1.2 Linear Discriminant Analysis (LDA) based Face Recognition

This is a texture operator that labels image pixels by evaluating each pixel's neighborhood with a threshold, and the resulting binary number is used to represent the local features of the image. It is believed to be useful for texture classification and improves detection performance when used in conjunction with Histogram of Oriented Gradients (HOG) histograms. Robust to monotonous grayscale conversions give excellent results in controlled environments, and he is one of the simplest face recognition algorithms.

Dalali et al. [20] used the discrete wavelet transform as a preprocessing method to extract important features. The Daubechies wavelet helps extract approximate coefficients with one level of decomposition, thus removing information for face detection. The main focus is to reduce the information to less significant factors, resulting in less memory usage. In this paper, the dataset considered is the MIT face dataset. Two types of performance results were obtained. That is images with noise and images without noise. An accuracy of 99.3% was achieved for the noise-free image, and an accuracy of 98.28% was achieved for the noisy image.

Tang et al. [21] After extracting features from facial textures using the LBP operator, 10 CNNs with 5 different neural structures were used to extract further features for training purposes or to improve network parameters. and getting classification results using the softmax function after the layers have been created. Fully connected. Finally, a majority vote was used to generate the final face recognition results using parallel ensemble learning. The face recognition rate on ORL improved to 100% using this method and 97.51% on the Yale-B dataset.

To mitigate the impact of facial image variation on feature extraction performance, Muqeet et al. [22] proposed a method using directional wavelet transform (DIWT) and LBP to overcome the effect. LBP histogram features were extracted from the selected top-level DIWT subbands as a locally descriptive feature set. The proposed method was tested with the ORL, FEI [23], and GT databases. Results showed that the proposed method was more efficient than his Local Gabor binary patterns (LGBP), and Local Steerable Pyramid Binary Pattern (LSPBPS) methods.

Zhang et al. [24] proposed a face spoofing prevention strategy using LBP, Discrete Wavelet Transform (DWT), and Discrete Cosine Transform (DCT) together with an SVM classifier. In this article, we have created a DWT-LBP function that contains

information about the spatial details of a block of video files, and finally, an SVM classifier with an RBF kernel is trained to protect against spoofing.

To improve the speed and accuracy of 3D face recognition, Shi et al. [25] presented a combination of SVM and LBP. LBP was used to extract feature details from 3D face-depth images. Afterward, the information was classified using SVM. The databases used for the experiment were the Texas 3DFace Recognition 3D face depth database and a custom depth database. As a result, it was found that the proposed method is less time-consuming and has a fast recognition speed.

Study & Year of Publish	Method/Algorithm	Dataset and Accuracy
2020, [25]	LBP and SVM	Texas 3DFRD3-D face 96.83% depth and self-made depth databases.
2020, [24]	DWT-LBP-DCT	REPLAY-ATTACK [26] REPLAY-ATTACK: 7.361%, and CASIA-FASD [27] CASIAFASD: 93.84%
2020, [21]	CNN and LBP	ORL and Yale-B ORL: 100%, Yale-B: 97.5%
2017, [22]	LBP	ORL , GT and FEI ORL: 97%, GT: 82.25%, FEI: 91.14%
2016, [20]	LBP	MIT face database Without noise: 99.3%, With noise:98.28%

Table 2. LBP-based methods for face recognition.

1.3 Histogram of Oriented Gradients (HOG) based Face Recognition

A simple feature descriptor is used in image processing to extract features from images and detect objects. Function descriptors simplify the image by extracting only the necessary information and discarding the rest. The HOG function is useful in the first stage of object detection. Gradient-based representations are derived from pixel-based representations and are used in conjunction with linear classification methods and multi-scale pyramids for object detection.

Zemgulys et al. [28] proposed an image segmentation method using HOG and SVM algorithms for classification. Two approaches for detecting the referee's hand gestures have been discussed. In other words, it achieved an accuracy of 97.5% and an F1 score of 94.95% with a wearable sensor and computer vision that recognize the referee's signal in a basketball game.

Rameswari et al. [29] implemented an access control system in which face detection and recognition are the main parameters for access control, using HOG for feature extraction and the Facenet algorithm for face recognition. In addition, the system security was strengthened by applying RFID face recognition technology. In this system, the FaceNet algorithm achieved a high accuracy of 97% compared to other face detection algorithms such as LBPH, FisherFace, etc.

Lakshmi et al. [31] used LBP with a modified HOG function for facial expression recognition and a multi-class SVM algorithm for classification and recognition. Two data sets were used: the JAFFE [32] and CK+ [33] data sets. The accuracy of the CK+ data set was 97.66%.

Yan et al. [34] used a combination of HOG, Adaboost, and SVM for real-time vehicle detection. We used HOG to extract features and then trained an AdaBoost classifier on a combination of HOG features and the dataset used to train the classifier on a group

of processed images. The accuracy of the HOG and AdaBoost combination was 97.24%, reaching 96.89% when using the HOG function with the SVM classifier.

Study & P. Year	**Method/Algorithm Dataset**	**Accuracy**
2021, [31]	Modified HOG along with JAFFE and CK+ JAFFE LBP and SVM	90.83
2020, [29]	HOG and FaceNet Private	97
2019, [30]	HOG and SVM Private	80
2018, [28]	HOG and SVM Private	97.50%
2016, [34]	HOG with AdaBoost and GTI vehicle [35] AdaBoost SVM classifier SVM	96.89

Table 3. HoG-based methods for face recognition.

1.4 Support Vector Machine (SVM) based Face Recognition

SVM brings a new dimension to the task of pattern recognition. It can solve face recognition problems in both linear and non-linear SVM training models. It is commonly used in ML classification problems because it requires less processing power.

Zhang et al. [36] extracted multi-scale features from images of 20 subjects, each with different poses and seven facial expressions, using bio-orthogonal wavelet entropy to extract multi-scale features. They also used a rigorous validation model using stratified cross-validation. Using a fuzzy multi-class support vector machine as a classifier, they achieved results that outperformed the three state-of-the-art methods with 96.77% accuracy.

The key aspect according to Pham et al. [37] is to overcome the problem that arises in CNNs when we have unbalanced training data points for classes by increasing the number of training samples for a minority class. They used the Action Units (AU) feature set to create images with similar facial expressions. To improve the overall performance of the model, we combine the AU function with the CNN function to train the SVM for classification.

Omara et al. [38] developed a multimodal biometric recognition system using a hybrid learning distance measurement model and a directed acyclic graph SVM model. The model was tested on the AR face dataset and achieved an accuracy of 99%.

Study & P. Year	**Method/Algorithm Dataset**	**Accuracy**
2021, [38]	Distance Metric and DAG AR face dataset SVM	99.85%
2020, [39]	SVM and SFFS 8000 face images (Private)	Above 90%
2019, [37]	SVM fused with CNN RAF, Fer2013, ExpW (DenseNet) and AU features	RAF: 91.37%, Fer2013: 71.01%, ExpW: 72.84%

2016, [36]	Fuzzy SVM and Stratified 20 subjects X 7 different Cross Validation expressions (Private)	96.77+-0.10%

Table 4. SVM-based methods for face recognition

2. Artificial Intelligence Algorithms

2.1. Neural Network

Since the 1980s, with the advancement of science and technology, artificial intelligence (AI) has gradually entered the general perception, making a tremendous impact on society and artificial Neural Networks (ANNs, also known as artificial neural

networks or Neural Networks), the field of artificial intelligence is also progressing with the rise of artificial intelligence. As the name suggests, artificial Neural Networks are abstract models of how Neural Networks in the human brain work. Input an activation function (also called an excitation function) and map the nodes to the output ends. The "memory" of an artificial Neural Network consists of weighted values of connecting signals between nodes. The output of an algorithm is often determined by its signal weight values, activation functions, and connection modes of "neurons".

With the increasing maturity of Artificial Neural Networks, they are widely used in many fields such as pattern recognition, biology, robotics, and medicine, solving many practical application problems that ordinary computers cannot solve, Artificial Neural Networks also play an integral role in today's face recognition field, and many of their advanced algorithms (such as convolutional Neural Networks) play an important role in advancing and contributing to the field of face recognition. In this work, artificial Neural Network-based face recognition algorithms such as DeepFace and DeepID are presented, their strengths and weaknesses are compared with other face recognition methods, and their future development directions are then predicted.

2.2 Convolutional Neural Network (CNN) based Face Recognition

A convolutional Neural Network consists of several layers, consisting of an input layer, convolutional layers, a pooling layer, a fully connected layer, and an output layer [52]. Convolutional and pooling layers.

In the field of face recognition, the conventional face recognition algorithms based on the facial feature vector introduced earlier need to extract facial features, select features, and select classifiers during operation, which is cumbersome. Manipulation is required, and no algorithm can quite achieve that well-defined goal. However, convolutional Neural Networks can make up for such shortcomings and extract high-level features of human faces to improve the expressiveness of features, effectively improving the accuracy and efficiency of face recognition.

CNN's have shown tremendous growth in recent years and are therefore recommended for solving computer vision problems. Convolutional and pooling layers of

CNN can be used for face recognition to extract the maximum amount of facial features compared to standard algorithms. As the amount of training data increases in this digital world, we need deep learning models that can drastically reduce the time it takes to train a model.

Shafiza et al. [40] challenged the main factors (illumination variance, pose, facial expression, and occlusion) that affect the performance of face recognition algorithms by proposing a robust four-layer CNN architecture. The system achieved an accuracy of 99.5% on the AR database and 85.13% on the FERET database (35 subjects). The most important feature of this system is that the facial recognition process is completed in less than 0.01 seconds.

Zangene et al. [41] used a joint mapping method architecture for high- and low-resolution face images. This has two branches of deep convolutional Neural Networks, one for each type of resolution that is transformed into the common space. The branch associated with high-resolution to common space conversion consists of a 14-layer network, and the branch associated with low-resolution face image conversion consists of an additional 5-layer network associated with the 14-layer network. Tested on FERET, LFW, and MBGC datasets [42], the proposed architecture showed 5% better accuracy. An image with a resolution of 6*6 pixels.

Im et al. [43] proposed an authentication system that protects the privacy of smartphone users from malicious clients by encrypting and storing facial feature vectors. A Euclidean distance-based matching score is calculated each time someone attempts to access a private vector on a remote server. Validation of CFP[44] and ORL records takes only 1 second, while real-time secure facial recognition takes 1.3 seconds. To further improve the calculated value, we

used the Catalano-Fiore transformation, which converts the linear homomorphic
encryption scheme to a quadratic scheme.

Goel et al. [45] used a high-level feature extraction method based on the DCNN-Optimized Kernel Extreme Learning Machine algorithm. The Particle Swarm Optimization (PSO) algorithm is used for parameter optimization along with the polynomial function kernel ELM classification algorithm. The results obtained without normalization on the AT&T, CMU-PIE [46], Yale [47], and UMIST [48] datasets were error rates of 0.5, 8.89, 0 & 21. This method has the shortest training time compared to other DLNs.

Zhao et al. [49] addressed various face-rendering attacks by proposing Deep-Architecture to increase the accuracy of multi-view human face recognition. Here the author proposed his CNN that extracts facial features and uses facial registration algorithms to further identify key points of the face. PCA was used for feature dimensionality reduction and the Joint Bayesian Framework
(JBF) was proposed to evaluate the similarity of feature vectors. On the CAS-PEAL [50] dataset, he achieved an accuracy of 98.52%.

To address the challenges of automatic age estimation in real-time applications, Al-Shannaq et al. [51] proposed a model for estimating human age using a fine-tuned CNN model. Two types of datasets were used to evaluate ideas. The FG NET dataset (limited) had an MAE of 3.446 and the UTKFace dataset (unlimited) had an MAE of 4.867. Using the Adience dataset, we fine-tuned the model for the age stratification task and the model achieved an overall accuracy of 61.4%.

Study& P. year	**Method/Algorithm Dataset used**	**Accuracy**
2020, [51]	CNN FG NET, UTKFace, Adience	MAE: 3.446, MAE: 4.867, 61.4%
2020, [49]	CNN + PCA,JBF CAS-PEAL	98.52%
2020, [45]	OKELM algorithm, PSO, AT&T ,CMUPIE, polynomial function KELM YALE,UMIST	EER: 0,0,6.67,10.9
2020, [43]	Euclidean distance-based, CFP,ORL CatalanoFiore transformation	EER: 1.17 ,0.37
2019, [41]	Coupled mapping method, FERET, LFW and MBGC DCNNs	FERET: 99.2%, LFW: 76.3% MBGC: 68.64%
2014, [40]	4-layer CNN architecture AR,	AR: 99.5% FERET: 85.13%

	FERET	

Table 5. CNN-based methods for face recognition.

2.3. DeepFace Algorithm

The deep face was proposed by Facebook in 2014. It is an optimized face recognition network structure based on the Deep learning algorithm, and its steps can be roughly divided into Recognition - Alignment - Display - Classification. In the detection step, Deep face used a fiducial point detector based on the detection points, first finding her six fiducial points (eye center, nose, mouth) on the face. Reference points were obtained by associated learning (such as using her SVR via the LBP function).

The registration step used the 3D model to process the 2D face into a 3D face and triangulated the face model. The triangulated faces were then 3D processed and deflected to finally obtain the faces [53].

For face representation, an image was sent to a Convolutional Neural Network (CNN) for processing, and the feature vector of the face image was obtained and finally classified. In practice, this algorithm can generate relatively compact and sparse descriptors, greatly improving recognition accuracy. In the Face Recognition Test Database (LFW) test, the accuracy of this algorithm reached 97.35%, reducing the error of existing techniques by 27% [53]. However, the drawback is that the model needs to be adjusted to extract specific features. Without alignment, recognition accuracy can only reach 87.9% in our tests. Additionally, in the process of training the model, a loss function was constructed with the results of the classifier. In other words, using a model meant that its classifiers and properties were bound, and the classes split by the classifier were simultaneously bound to the size of the input data classes, requiring different training. Input different data.

2.4. DeepID Algorithm

DeepID is one of the mainstream face recognition algorithms based on convolutional Neural Networks proposed by Professor Tang Xiaoou of the Chinese University of Hong Kong in 2014. Using 3 layers of convolution and 3 layers of max pooling, the last fully connected layer is connected to the 3rd and her 4th layers of the convolution [54].

In face recognition, it is obvious that factors such as brightness, angle, age, and expression of the face in the recognized image affect the recognition accuracy, and the same image shows even higher recognition accuracy. Many conventional face recognition algorithms cannot accurately recognize face models when dealing with the above practical problems due to their algorithms or other reasons. For such problems, the peculiarities of the DeepID structure connect the fully connected layers to her 3rd and her 4th convolutional layers. This particular property has the advantage that the model network can detect more scale features,

continuous sampling avoids missing information in the final convolutional layers, and improves information integration. Ultimately, DeepID outputs a dense, internally validated 160-degree vector that can be used directly for face recognition [53]. In a real combat environment, using the CeleFace dataset as a training model, DeepID achieved 97.45% accuracy in LFW records under weak alignment conditions, which is 0.05% lower than the human eye (97.50%). [53].

2.5. DeepID2 algorithm and DeepID2+ Algorithm

After the DeepID mentioned above improved the accuracy of facial recognition, DeepID2 also came out in the same year that DeepID came out. Compared to DeepID, DeepID2 made some improvements by investigating the effects of recognition and authentication signals on Neural Networks when adding facial recognition and facial recognition signals to the network structure [53]. Adding these two signals together can reduce the difference between the same face and different image sources.

The presence of these two signals can not only reduce the differences between different face image samples but also improve the diversity of different samples. In practical use, DeepID2 achieved
99.15% accuracy with LFW. This is a 67% improvement over its predecessor, DeepID. [59]
DeepID2+, a face recognition algorithm based on DeepID2, was proposed by Professor Tang Xiaoou's team in 2015. The dimension of fully connected layers increases to 512 layers. The LFW dataset test can achieve 99.47% recognition accuracy [53], which exceeds the recognition accuracy of the human eye.

2.6. Facenet Algorithm

Facenet is a face recognition system released by Google in 2015, which improves the loss function of Neural Networks, proposes a new loss function based on Euclidean face similarity measure, directly abandons traditional softmax classification, and triplet The loss is the loss function [8th]. The 128-dimensional vector distances given by the vectors of the penultimate layer in the structure are computed just for face recognition. The advent of
Facenet brings with it a new loss function that shows that the feature dimension can be distributed over 128 dimensions. Facenet has also achieved excellent results in practical use, and in 2015 he achieved a high face recognition accuracy of 99.63% with LFW. Therefore, it surpassed the best face recognition performance with the highest accuracy[52].

2.7. ResNet-based Face Recognition

A Residual Neural Network (ResNet) is a popular deep learning model that uses residual blocks to overcome the problem of training very deep networks. Use these blocks to skip connections and jump between levels. ResNet makes it easier to train huge networks without increasing the training error rate.

Ze Lu et al. [55] proposed a low-resolution face recognition model called the “deeply coupled ResNet model”. A ResNet-like network, a trunk network, was used to extract features common to facial photographs of different resolutions. We then used a bifurcation network to learn a joint mapping to project the image features. The proposed model was experimented on LFW (with different probe sizes) and SCface datasets (three different datasets depending on the camera distance) and showed 93.6%–98.7% and 73.0%–98.0% accuracy in face recognition. has been achieved.

Storey et al. [56] in their study, introduced the 3DPalsyNet framework for mouth movement detection and facial paralysis grading. To capture the dynamic action of the video data, they used a modified 3D CNN architecture with his ResNet backbone. The structure was tested with two sets of data. The CK+ and proprietary facial paralysis datasets achieved F1 scores of 82% for mouth movements and 88% for grading facial paralysis, respectively. This is higher than the score of 3D CNN indicating the ability to perform efficient facial analysis. The video shows the sequence.

Peng et al. [57] provided two approaches to facial recognition. The first was to convert the initial residual ResNet scale factors from hyperparameters to trainable parameters. Use a small initial value of 0.1 to ensure an initially stable training network. The second is the use of Leaky ReLU and PReLU in the Inception ResNet module to improve network performance by maximizing the use of input data. Both methods have been tested on the datasets VGGFace2, MS1MV2, IJBB, and LFW and show improved accuracy and stability of the training process.

Li et al. [58] proposed an improved facial emotion detection model by using ResNet-50 as the backbone of the network and CNN for feature extraction. BN and activation function ReLU were used to increase the convergence ability of the model. The model was tested on her dataset of 20 different subjects (700 images) with different facial expressions and age groups and was found to have good accuracy of 95.39±1.41%.

P. year &	Method/Algorithm	Accuracy

Study	Dataset used	
2021, [58]	CNN + ResNet-50 Private dataset of 20 different subjects (700 images)	95.39 ± 1.41%
2019, [57]	Modified 3D CNN + ResNet Private Facial Palsy dataset	F1 score: 88%
2019, [56]	Modified 3D CNN + ResNet CK+ dataset	F1 score: 82%
2018, [55]	Deep coupled ResNet mode LFW face database & SCface datasets	93.6 - 98.7 % 73.0 - 98.0 %

Table 6. ResNet-based methods for face recognition.

IV. Summary of datasets used by researchers for Face Recognition

Year	Dataset	Total Images/Videos Features
2012	CASIAFA SD [27]	600 (240 for training, 360 Anti-spoofing data set. Footage was shot for testing), 50 topics (12 different lighting and resolutions. videos per topic)
2012	GTIVehicl e [35]	3425 car rear images, This database contains images extracted from 3900 images are extracted footage. TTable 8. Summary of datasets used in various Face Recognition. he images cover a variety of driving from the road sequence. conditions, especially those related to weather.
2012	ReplayAtt ack [26]	1300 videos Face fake database. All videos were created by displaying a snapshot or video from the same client for at least 9 seconds, or by displaying a video from a real client trying to access the laptop via the built-in webcam.
2011	YTF [19]	3,425 videos, 1,595 The video was sourced from YouTube. There people are an average of 2.15 videos per topic. The average video length is 181.3 frames, the shortest clip is 48 frames and the longest clip is 6070 frames.
2010	FEI [23]	2800 image set Faces of people between the ages of 19 and 40 are used in the images, each distinguished by their striking appearance, hairstyle and jewelry.
2009	MBGC [42]	628 videos recordings There are 4025 frames with the left iris visible and 4013 frames with the right iris visible. Iris On Move (IOM) technology captured near-infrared video of the face.
2002 2003	CMUPIE [46]	41,368 images Each photo was taken in 13 different

2007	CASPEAL [50] LFW [18]	poses with 43 different lighting situations and 4 facial expressions. 99594 images A database of Chinese faces with large-scale images. To collect 27 photos from three planes, each person is asked to look straight ahead, up and down. The database also includes 5 facial expressions, 6 accessories and 15 lighting adjustments. 13,233 images, 1680 Its goal is to collect facial images and other persons important data for the Living People category on Wikipedia.
2001	YALEB [14]	5760 images Each subject is viewed in 576 different ways (9 poses x 64 lighting conditions). A photo was also taken with ambient (background) lighting for each person in a specific pose.
2001	Extend Yale B [14]	2414 images, 38 subjects Images were shot with different lighting and different facial expressions to achieve excellent results.
1999	XM2VT S [17]	2360 rounds, 295 people The dataset comes with manually placed perspectives for all 2360 images for better recognition.
1998	AR [12]	4000 images, 126 persons
		The image consists of different facial (56 female / 70 male) expressions with and without sunglasses or scarf (eg neutral expression, smile, anger, shout), and different lighting conditions.
1998	FERET [15]	14,126 images, 1199 It consists of a set of 1564 images and a set of individuals 365 repeating images.
1998	JAFFE [32]	213 images Ten Japanese female models posed with 7 facial expressions (6 basic facial expressions + 1 expressionless facial expression).
1998	UMIST [48]	564 images, 20 subjects Each image has different positions, from profile to full face. Subjects represent a wide range of races and genders, resulting in a more complete data set.
1997	YALE [47]	165 grayscale images Each subject has 11 pictures, one for each emotion or face composition: Center Light, Glasses, Happy, Left Light, No Glasses, Normal, Right Light, Sadness, Drowsiness, Shock and Wink.
1994	ORL [9]	400 images Different facial expressions, facial details (with or without glasses), and lighting

		conditions were used in the images. A dark, uniform background was used with the subject in a vertical position.

Table 7. Summary of datasets used in various Face Recognition.

IV. SUMMARY

However, some algorithms can be combined (such as the Principal Component Analysis and Linear discriminant analysis algorithms mentioned above) to achieve higher face detection accuracy. For traditional face recognition algorithms, deep learning and Neural Network-based face recognition algorithms are characterized by high robustness, a high degree of automation, and high recognition accuracy. Some algorithms, like the Facenet algorithm, have better face recognition capabilities than the human eye. However, there is still a lot of room for improvement in smart face recognition algorithms such as Improved Neural Network structure (DeepID) and loss algorithm (Facenet) as described above. Facial recognition currently has many development paths, including Strengthening the theoretical investigation of deep learning algorithms, strengthening the investigation of semi-supervised as well as unsupervised learning.

V. CONCLUSION

In this article, we have explored existing face recognition techniques based on various descriptor methods combined with machine learning classifiers such as SVM, deep learning, and transfer learning. We also listed common datasets used for face recognition technology. A limitation of existing facial recognition techniques is that they use datasets of images captured in specified and controlled environments and that the performance of these systems degrades under adverse conditions known as semantic adversarial attacks. , or to be degraded when downloaded from the Internet. Our study provides insight into existing facial recognition technologies for researchers wishing to conduct research in this area. A challenge for future research is to develop robust face recognition algorithms that can process low-resolution images captured in uncontrolled environments.

VI. REFERENCES

[1]. Jafri, R., & Arabnia, H.R. (2009). A Survey of Face Recognition Techniques. J. Inf. Process. Syst., 5, 41-68.

[2]. Lu, Wen-yao and Ming Yang. “Face Detection Based on Viola-Jones Algorithm Applying Composite Features.” 2019 International Conference on Robots & Intelligent System (ICRIS) (2019): 82-85.

[3]. Li, Haoxiang, Zhe L. Lin, Xiaohui Shen, Jonathan Brandt and Gang Hua. “A convolutional neural network cascade for face detection.” 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015): 5325-5334.

[4]. Chen, Haonan, Yaowu Chen, Xiang Tian and Rongxin Jiang. “A Cascade Face Spoofing Detector Based on Face Anti-Spoofing R-CNN and Improved Retinex LBP.” IEEE Access 7 (2019): 170116-170133.

[5]. X. Hu and B. Huang, "Face Detection based on SSD and CamShift," 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), 2020, pp. 2324-2328, doi: 10.1109/ITAIC49862.2020.9339094.

[6]. Journal. Overview of Face Recognition Technology[J]. Computer Knowledge and Technology, 2019,15(02):171-172.

[7]. A. Alkandari and S. J. Aljaber, "Principle Component Analysis algorithm (PCA) for image recognition," 2015 Second International Conference on Computing Technology and Information Management (ICCTIM), 2015, pp. 76-80, doi: 10.1109/ICCTIM.2015.7224596.

[8]. F. Wang, F. Xie, S. Shen, L. Huang, R. Sun and J. Le Yang, "A Novel Multiface Recognition Method With Short Training Time and Lightweight Based on ABASNet and H-Softmax," in IEEE Access, vol. 8, pp. 175370-175384, 2020, doi: 10.1109/ACCESS.2020.3026421.

[9]. Ferdinando and Andy Harter. “Parameterisation of a stochastic model for human face identification.” Proceedings of 1994 IEEE Workshop on Applications of Computer Vision (1994): 138-142.

[10]. Wang, Minghui & Song, Lili & Sun, Kaisong & Jia, Zhigang. (2020). F-2D-QPCA: A Quaternion Principal Component Analysis Method for Color Face Recognition. IEEE Access. 8. 217437-217446. 10.1109/ACCESS.2020.3041847.

[11]. M. Chen, M. Jiang, J. Sun, J. Kong, and J. Hou, "Face Recognition Based on CSGF(2D)2PCANet," in IEEE Access, vol. 6, pp. 45153-45165, 2018, doi: 10.1109/ACCESS.2018.2865425.

[12]. Martinez, Aleix. & Benavente, Robert. (1998). The AR face database. Tech. Rep. 24 CVC Technical Report.

[13]. C. -Y. Low, A. B. -J. Teoh and K. -A. Toh, "Stacking PCANet +: An Overly Simplified ConvNets Baseline for Face Recognition," in IEEE Signal Processing Letters, vol. 24, no. 11, pp. 1581-1585, Nov. 2017, doi: 10.1109/LSP.2017.2749763.

[14]. A. S. Georghiades, P. N. Belhumeur and D. J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 643-660, June 2001, doi: 10.1109/34.927464.

[15]. Phillips, P. Jonathon & Moon, Hyeonjoon & Rizvi, Syed & Rauss, Patrick. (2000). The FERET Evaluation Methodology for Face-Recognition Algorithms. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 22. 1090 - 1104. 10.1109/34.879790.

[16]. The Georgia Tech Face Database, [Online], Available: http://www.anefian.com/research/face_reco.htm.

[17]. Messer, K. & Matas, Jiri & Kittler, J. & Jonsson, K. & Luettin, Juergen & Maître, Gilbert. (2000). Xm2vtsdb: The extended m2vts database. Proc. of Audio- and Video-Based Person Authentication.

[18]. Huang, Gary & Mattar, Marwan & Berg, Tamara & Learned-Miller, Eric. (2008). Labeled Faces in the Wild: A Database forStudying Face Recognition in Unconstrained Environments. Tech. rep..

[19]. L. Wolf, T. Hassner and I. Maoz, "Face recognition in unconstrained videos with matched background similarity," CVPR 2011, 2011, pp. 529-534, doi: 10.1109/CVPR.2011.5995566.

[20]. Dalali, Shivakumar & L., Suresh. (2016). Daubechives Wavelet Based Face Recognition Using Modified LBP. Procedia Computer Science. 93. 344-350. 10.1016/j.procs.2016.07.219.

[21]. Tang, Jialin & Su, Qinglang & Su, Binghua & Fong, Simon & Cao, Wei & Gong, Xueyuan. (2020). Parallel Ensemble Learning of Convolutional Neural Networks and Local Binary Patterns for Face Recognition. Computer

Methods and Programs in Biomedicine. 197. 105622. 10.1016/j.cmpb.2020.105622.

[22]. Abdul Muqeet, Mohd & Holambe, Raghunath. (2017). Local binary patterns based on directional wavelet transform for expression and pose-invariant face recognition. Applied Computing and Informatics. 15. 10.1016/j.aci.2017.11.002.

[23]. Thomaz, Carlos & Giraldi, Gilson. (2010). A new ranking method for Principal Components Analysis and its application to face image analysis. Image Vision Comput.. 28. 902-913. 10.1016/j.imavis.2009.11.005.

[24]. Zhang, Wanling & Xiang, Shijun. (2020). Face anti-spoofing detection based on DWT-LBP-DCT features. Signal Processing: Image Communication. 89. 115990. 10.1016/j.image.2020.115990.

[25]. Shi, Liangliang & Wang, Xia & Shen, Yongliang. (2020). Research on 3D face recognition method based on LBP and SVM. Optik. 220. 165157. 10.1016/j.ijleo.2020.165157.

[26]. Chingovska, I & Anjos, A & Marcel, Sébastien. (2012). On the Effectiveness of Local Binary Patterns in Face Anti-spoofing. IEEE International Conference of the Biometrics Special Interest Group.

[27]. Zhang, Zhiwei & Yan, Junjie & Liu, Sifei & Lei, Zhen & Yi, Dong & Li, Stan. (2012). A face antispoofing database with diverse attacks. Proceedings - 2012 5th IAPR International Conference on Biometrics, ICB 2012. 26-31. 10.1109/ICB.2012.6199754.

[28]. Julius Žemgulys, Vidas Raudonis, Rytis Maskeliūnas, Robertas Damaševičius, Recognition of basketball referee signals from videos using Histogram of Oriented Gradients (HOG) and Support Vector Machine(SVM), Procedia Computer Science, Volume 130, 2018, Pages 953-960, ISSN 1877-0509, DOI: 10.1016/j.procs.2018.04.095.

[29]. Rameswari, R. & Kumar, S. & Aananth, M. & Deepak, C.. (2020). Automated access control system using face recognition. Materials Today: Proceedings. 45. 10.1016/j.matpr.2020.04.664.

[30]. Chitlangia, Aditya & Ganesan, Malathi. (2019). Handwriting Analysis based on Histogram of Oriented Gradient for Predicting Personality traits using SVM. Procedia Computer Science. 165. 384-390. 10.1016/j.procs.2020.01.034.

[31]. Lakshmi, D. & Ponnusamy, Ramesh. (2021). Facial Emotion Recognition using modified HOG and LBP features with Deep Stacked AutoEncoders. Microprocessors and Microsystems. 82. 103834. 10.1016/j.micpro.2021.103834.

[32]. Lyons, M.Kamachi and J. Gyoba, "The Japanese Female Facial Expression (JAFFE) Database," The Japanese Female Facial Expression (JAFFE) Dataset, Zenodo, DOI: 10.5281/zenodo.3451524, 1998.

[33]. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-specified Expression," Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, pp. 94- 101, DOI: 10.1109/CVPRW.2010.5543262, San Francisco, USA, 2010.

[34]. Yan, M. Yu, Y. Yu, and L. Fan, "Real-time Vehicle Detection Using Histograms of Oriented Gradients and AdaBoost Classification," Optik, vol. 127, no. 19, pp. 7941-7951, 2016.

[35]. J. Arróspide, L. Salgado and M. Nieto, "Video Analysis Based Vehicle Detection and Tracking Using an MCMC Sampling Framework," EURASIP Journal on Advances in Signal Processing, vol. 2012, Article ID: 2012 (2), DOI: 10.1186/1687-6180-2012-2, 2012.

[36]. Y.-D. Zhang, Z.-J. Yang, H.-M. Lu, X.-X. Zhou, P. Philips, Q.-M. Liu and S.-H. Wang, "Facial

Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation," IEEE Access, vol. 4, pp. 8375-8385, DOI: 10.1109/ACCESS.2016.2628407, 2016.

[37]. T. T. D. Pham and C. S. Won, "Facial Action Units for Training Convolutional Neural Networks," IEEEAccess, vol. 7, pp. 77816-77824, DOI: 10.1109/ACCESS.2019.2921241, 2019.

[38]. I. Omara, A. Hagag, S. Chaib, G. Ma, F. E. Abd El-Samie and E. Song, "A Hybrid Model Combining Learning Distance Metric and DAG Support Vector Machine for Multimodal Biometric Recognition," IEEE Access, vol. 9, pp. 4784-4796, DOI: 10.1109/ACCESS.2020.3035110, 2021.

[39]. F. Zhang and F. Wang, "Exercise Fatigue Detection Algorithm Based on Video Image Information Extraction," IEEE Access, vol. 8, pp. 199696-199709, DOI: 10.1109/ACCESS.2020.3023648, 2020.

[40]. A. R. Syafeeza, M. Khalil-Hani, S. S. Liew and R. Bakhteri, "Convolutional Neural Network for Face Recognition with Pose and Illumination Variation," International Journal of Engineering and Technology (IJET), vol. 6, no. 1, pp. 44-57, 2014.

[41]. Erfan & Mohammad, Rahmati & Mohsenzadeh, Yalda. (2017). Low Resolution Face Recognition Using a Two-Branch Deep Convolutional Neural Network Architecture. Expert Systems with Applications. 139. 10.1016/j.eswa.2019.112854.

[42]. P. J. Phillips et al., "Overview of the Multiple Biometrics Grand Challenge," Proc. of Advances in Biometrics (ICB 2009), Part of Lecture Notes in Computer Science, vol. 5558. Springer, Berlin, Heidelberg, [Online], Available: https://doi.org/10.1007/978-3-642-01793-3_72, 2021.

[43]. J. Im, S. Jeon and M. Lee, "Practical Privacy-Preserving Face Authentication for Smartphones Secure Against Malicious Clients," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 2386-2401, DOI: 10.1109/TIFS.2020.2969513, 2020.

[44]. S. Sengupta, J. Chen, C. Castillo, V. M. Patel, R. Chellappa, and D. W. Jacobs, "Frontal to Profile Face Verification in the Wild," Proc. of the IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1-9, DOI: 10.1109/WACV.2016.7477558, Lake Placid, USA, 2016.

[45]. T. Goel and R. Murugan, "Classifier for Face Recognition Based on Deep Convolutional Optimized Kernel Extreme Learning Machine," Computers & Electrical Engineering, vol. 85, Paper ID: 106640, DOI: 10.1016/j.compeleceng.2020.106640, 2020.

[46]. T. Sim, S. Baker and M. Bsat, "The CMU Pose, Illumination, and Expression Database," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pp. 1615-1618, Dec. 2003.

[47]. A. Georghiades, P. Belhumeur and D. Kriegman, "Yale Face Database," Yale University, [Online] Available: http://cvc.yale.edu/projects/yalefaces/yalefa, 1997.

[48]. F. Zhao, J. Li, L. Zhang, Z. Li, and S. Na, "Multi-view Face Recognition Using Deep Neural Networks," Future Generation Computer Systems, vol. 111, pp. 375-380, DOI: 10.1016/j.future.2020.05.002, 2020.

[49]. W. Gao et al., "The CAS-PEAL Large-scale Chinese Face Database and Baseline Evaluations," IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans, vol. 38, no. 1, pp. 149- 161, DOI: 10.1109/TSMCA.2007.909557, 2008.

[50]. Al-Shannaq and L. Elrefaei, "Age Estimation Using Specific Domain Transfer Learning,"

Jordanian Journal of Computers and Information Technology (JJCIT), vol. 06, no. 02, pp. 122-139, June 2020.
[51]. S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah, and S. U. Ali, "Facial Recognition Using Convolutional Neural Networks and Implementation on Smart Glasses," Proc. of the IEEE International Conference on Information Science and Communication Technology (ICISCT), pp. 1-6, Karachi, Pakistan, 2019.
[52]. Bao Ruidong, Zhao Min. Research Review of Face Recognition based on Convolutional Neural Network [J]. Software Guide, 2008,17(04):6-8.
[53]. Zhang Yiwei. Face Recognition Research based on Convolutional Neural Networks [D]. Jilin University,2019.
[54]. Wang Jingjing, Liu Rong, Hou Yujing. Research on Face Detection and Recognition Algorithm based on DeepID [J]. Computer Knowledge and Technology, 2008,14(17):220-221.
[55]. G. Storey, R. Jiang, S. Keogh, A. Bouridane and C. Li, "DPalsyNet: A Facial Palsy Grading and Motion Recognition Framework Using Fully 3D Convolutional Neural Networks," IEEE Access, vol. 7, pp. 121655-121664, DOI: 10.1109/ACCESS.2019.2937285, 2019.
[56]. S. Peng, H. Huang, W. Chen, L. Zhang, W. Fang, "More Trainable Inception-ResNet for Face Recognition," Neurocomputing, vol. 411, pp. 9-19, DOI: 10.1016/j.neucom.2020.05.022, 2020.
[57]. B. Li and D. Lima, "Facial Expression Recognition via ResNet-50," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 57-64, DOI: 10.1016/j.ijcce.2021.02.002, 2021.
[58]. H. Wang, D. Zhang and Z. Miao, "Fusion of LDB and HOG for Face Recognition," Proc. of the 37th IEEE Chinese Control Conf., pp. 9192-9196, DOI: 10.23919/ChiCC.2018.8483900, Wuhan, China, 2018.
[59]. Jing Chenkai, Song Tao, Zhuang Lei, Liu Gang, Wang Le, Liu Kailun. Overview of Face Recognition Technology based on Deep Convolutional Neural Network [J]. Computer Application and Software, 2008,35(01):223-231.
[60]. R. Gross, I. Matthews, J. Cohn, T. Kanade and S. Baker, "Multi-PIE," 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, 2008, pp. 1-8, doi: 10.1109/AFGR.2008.4813399.
[61]. Priyanka Shahane, Deipali Gore "A Survey on Classification Techniques to Determine Fake vs. Real Identities on Social Media Platforms," IJRDT, 2018.

Cite this article as :

Shantanu Jain, Vrushaket Chaudhari, Rushikesh Chuadhari, Tanvesh Chavan, Prof. Priyanka Shahane, "A Survey on Face Recognition Techniques in Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 6, pp. 50-66, November-December 2022. Available at doi : https://doi.org/10.32628/CSEIT228558
Journal URL : https://ijsrcseit.com/CSEIT228558