ASIAN JOURNAL OF ELECTRICAL AND ELECTRONIC ENGINEERING
Vol. 4 No. 2 2024
E-ISSN: 2785-8189

A Systematic Review on Facial Detection and Recognition: Limitations and Opportunities

Mohammad Amir Khan*, Hasan Firdaus Bin Mohd Zaki, Azhar Bin Mohd Ibrahim, S.M Afzal Hoq and Amir Akramin bin Shafie

Department of Mechatronics, International Islamic University of Malaysia, Kuala Lumpur, Malaysia

*Corresponding author: hasanzaki@iium.edu.my

(Received: 2 October 2024; Accepted: 4 October 2024)

Abstract— Face recognition technology is a biometric tool that identifies people by facial characteristics. Individuals collect the facial photography, which is then automatically processed by picture recognition software. Face detection and recognition have several potential applications in various departments like security, education, healthcare, etc. Therefore, the fundamentals and methods of broad facial detection and recognition have been discussed in this article. Owing to the outbreak of the pandemic, people are now required to wear masks so that the spreading of the coronavirus is prevented, which makes it challenging to monitor sizable crowds of mask-wearing individuals. Face masks have higher interclass similarities and interclass variability because they cover a significant portion of the face, fooling face recognition systems' facial verification process. Thus, this paper has also discussed various aspects of masked face recognition.

Keywords: *Detection, Recognition, Masked, Facial, Artificial Intelligence, Algorithms.*

1. INTRODUCTION

Face recognition is one of the most frequently used applications of picture analysis. A subset of the visual pattern recognition problem is face recognition. Humans constantly recognize visual patterns; our eyes are how we take in visual data. The brain interprets this information as meaningful concepts. Whether a picture or a movie, it is a matrix of many pixels for a computer. The computer should determine what idea every piece of data in data is referring to. A primary classification issue in visual model recognition exists here [1]. To do face recognition, it is crucial to identify the face's owner in the portion of the data that all machines understand as the face. It's a subdivision issue. Globally, the COVID-19 coronavirus disease has tremendously impacted people's daily lives. The World Health Organization has advised individuals to wear face masks in all public areas to stop the spread of COVID-19 [2]. It can be challenging to check people in public places for face masks manually. Furthermore, using face masks renders conventional face recognition methods, generally created for exposed faces, ineffectual. Thus, there is a pressing need to create a reliable system that can identify various people while donning a face mask and identify those who are not.

Although it is starting to be used in other sectors like logistics, retail, smartphone, transportation, education, real estate, government management, entertainment advertising, network information security, and others, face recognition is most frequently used in attendance, access control, security, and finance [3]. Face recognition can be used in security to identify offenders and alert authorities to potentially dangerous circumstances. Since artificial intelligence technology has advanced significantly, we now need recognition technology that is more precise, adaptable, and quick. Face detection is the first step in all facial analysis algorithms, including face alignment, face modeling, face relighting, face identification, face verification, authentication, head pose tracking, and recognition of facial emotions, gender, and age, among others [4].

This paper thoroughly analyzes face detection and the numerous face detection techniques in section 2. Masked face detection and various data sets used for face detection are also discussed in this section. In addition, challenges and applications of Face Detection have also been presented. After this, systems for generic facial recognition are examined in section 3. The development stages of facial recognition, the approaches taken for facial recognition, and the evaluation criteria of a facial recognition system are discussed in this section. Section 4 presents a detailed analysis of Masked Facial Recognition Systems. Recent advancements in Masked Facial

DOI: https://doi.org/10.69955/ajoeee.2024.v4i2.70

Recognition have been presented in section 5. The authors also provide a gap and some limitations of the current systems at the end of this section. Lastly, section 6 concludes the article.

2. FACE DETECTION

The initial step in face recognition is face detection. Face detection locates and measures a human face in a digital image. In the digital image, all additional objects like bodies, trees, and buildings are disregarded in favor of identifying facial characteristics. It might be considered a specific example of object-class identification, where the task is locating and figuring out the dimensions of every object in an image that fits into a particular class [5]. There are certain restrictions in advanced unconstrained situations, although general face detection algorithms can be applied to some extent. A vast area being obscured makes it challenging to detect occluded faces in unrestricted contexts because intraclass similarity and intraclass variation increase [6]. Utilizing adaptive technologies, numerous strategies are developed to tackle the issue successfully.

When faces are obscured, the current face detection models occasionally perform less accurately. They developed unique methods for obscured face detection to solve this issue. Due to intraclass similarity and intraclass variability, occluded object detection in unconstrained environments can be difficult when a significant portion of the environment is obscured. Numerous methods are used for obscured object detection. In a lab setting or interior situations, specialized methods like MTCNN, Sckit-Image, and Haar cascades can produce good results for detecting obscured faces. The convolutional correlational filter effectively addresses the issue—the same problem benefits from binary segmentation [7].

2.1 Masked Facial Detection Methods

We mainly concentrate on masked facial detection techniques in this area. Various Masked Facial Detection techniques are depicted in Figure 1. All methods are categorized using two kinds based on their characteristics: manually created feature-based methods and neural network-based methods. Standard methods also frequently include custom feature-based techniques. Depending on how many detectors are used, they can be further divided into two categories: single-detector and multiple-detector methods [8]. Most detectors use the AdaBoost algorithm [9]. Various detectors, such as the face detector, facial mask detector, nose and mouth detector, eye detector, and mouth and nose detector, are chosen or combined. Several researchers are interested in neural network-based techniques. The approaches can be divided into three groups based on how many stages they have: single-stage methods, two-stage methods, and multistage methods. They are primarily implemented for single-stage procedures through transfer learning of object detection algorithms. YOLO series methods, such as YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, and equivalent micro variants, are an example [10].

Regarding the utilization of neural networks, two-stage approaches can be further separated into three categories: neural network + neural network, neural network + hand-crafted feature, and hand-crafted feature + neural network. Face region classification and face region pre-detection are the two components of two-stage techniques. The first component is used to identify potential facial locations, while the second half is utilized to categorize the circumstances surrounding mask use. More sophisticated processing stages or employing several models in multistage approaches result in higher computation costs [11].

2.2 Data sets

Many datasets are proposed by academics worldwide to train detection or classification methods to observe the conditions of wearing masks. These models will be used in edge nodes or monitoring systems. Detailed descriptions and explanations of some of these datasets are provided in this section.

First, we will provide a previous dataset on masked face detection. 2017 Ge et al. [12] suggested the extensive MAFA dataset. It was asserted to be the largest mask-wearing dataset existing before 2017. Thirty-five thousand eight hundred six masked faces can be found among 30 811 online photographs that make up MAFA. The dataset is probable to have occluded faces since it includes a variety of masks, such as single-colored manufactured objects, hands, neckerchiefs, hair, and medical masks. Six attributes are assigned to the dataset: face orientation, eye position, mask location, face location, occlusion degree, and type of mask. It delivers appropriate samples and considers roughly 60 situations with masked faces. Many occlusions, meanwhile, are ineffective at shielding patients from infection threats.

The "MaskedFace-Net Image Dataset (MFNID)" was created by Cabani et al. [13] to produce simulated correct/incorrect masked faces. Candidate face detection, facial landmarks detection, mask-to-face mapping, and manual image filtering are the four phases that make up the system. Roy et al. [14] used photos from the Internet to compile the Moxa3K dataset. There are 3000 photos in total. The dataset considers boundary conditions carefully; for instance, if a handkerchief covers a face, it will be in the "mask" class. Many diverse samples are included in Moxa3K, including blurry, rotating, crowded, and lighting-related samples. All the facial areas are annotated in Pascal VOC format "LabelImg" and YOLO format, with 9161 faces and 2015 masked faces included. As a result, it gives researchers more options for how to train their machine-learning models. The robustness of masked facial detectors should increase in this circumstance.

By assembling photos from the publicly accessible datasets FFHQ [15], LFW [16], CelebA [17], YouTube videos, and the Internet, Eyiokur et al. [18] presented an Unconstrained Face Mask Dataset (UFMD). UFMD is a complicated collection that encompasses ethnicity, age, gender, and indoor and outdoor conditions, thanks to these publicly available photos. To increase the robustness of masked face detectors, UFMD also considers many head posture variations. The UFMD consists of 21 316 photographs divided into three classes: 10 618 images with faces that are masked, 10 698 images not masked, and 500 images with the wrong masks on. The website, according to the writers, will be accessible soon. Singh et al. manually created seven thousand five hundred images—5191 training photos, 1599 validation images, and 710 test images [19]. These pictures are from MAFA [12] and Wider Face [20]. The two classes "face" and "face-mask" in Singh's dataset are used to train a model to assess whether someone is wearing a mask. The extent of the crowing can be evaluated using the detection results. Annotations are supplied for bounding boxes. Table 1 provides a summary of these datasets.

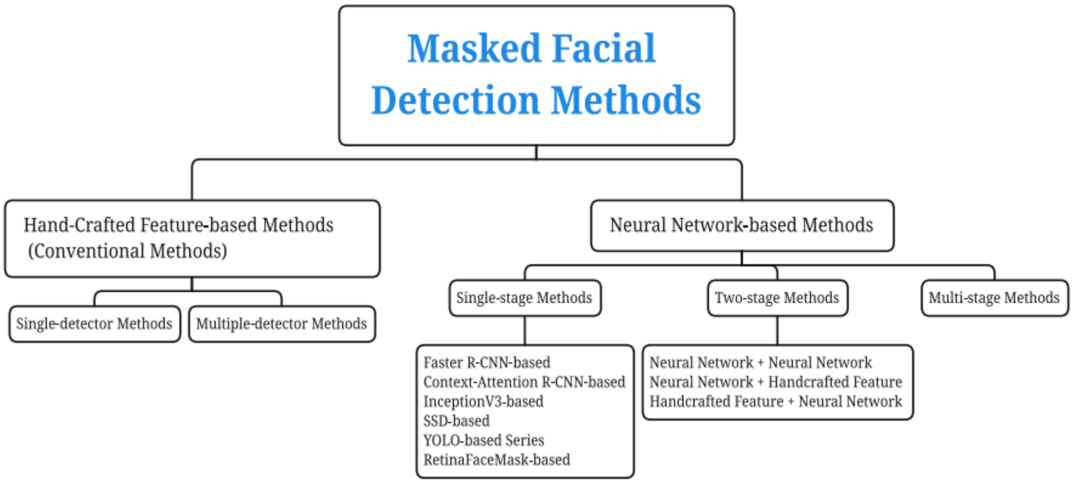

Fig. 1. Masked Facial Detection Methods [11]

Table 1. Summary of the Datasets

Dataset name	**Main characteristics**	**Image reality**	**Image number**	**Category**	**Masks number**	**Scale**	**Head pose**	**Scene**	**Annotation**	**Open**
[12]	The pictures are all taken from the Internet. For each face region, six attributes are manually annotated.	Real	30811	Multiple mask types	35806 masked faces	Medium -large	Various	Complex	Yes	Yes
[13]	Face images are from FFHQ	Simulated	137016	Two	Sixty-seven thousand one hundred ninety-three faces with correct masks. 69823 faces with incorrect masks	Large	Frontal	Simple	No	Yes
[14]	The images are from the Kaggle	Real	3000	Two	9161 faces without masks.	Small Mediu	Various	Complex	Yes	Yes

Dataset name	Main characteristics	Image reality	Image number	Category	Masks number	Scale	Head pose	Scene	Annotation	Open
	data set, which has images captured in Russia, Italy, China, and India during a pandemic				3051 masked faces	m Large				
[18]	The pictures come from YouTube videos, FFHQ, LFW, CelebA, and other online sources.	Real	21316	Three	10698 faces without masks,10618 correct masked, 500 incorrect	Large	Frontal to Profile	Medium	Yes	Soon Open
[19]	The data set includes MAFA, WIDER, and FACE.	Real	7500	Two	5191 training images, 1599 validation images, 710 testing images	Small Medium Large	Various	Complex	Yes	Yes

2.3 *Challenges in Face Detection*

Face detection problems lead to a deterioration in face detection accuracy and rate of detection. Complex backgrounds, excessive faces in photographs, strange expressions, illuminations, low resolution, and face occlusion (hiding the face by any object) are a few of these difficulties. The variables considered are skin tone, altitude, proximity, hat, scarf, hand, hair, and any other thing, Background complexity, excessive faces, the image's orientation, etc. An example of this can be seen in Figure 2 [21].

Figure 2. Challenges in face detection [21].

2.4 *Applications of Face Detection*

Gender classification, document and access control, biometric attendance, human-computer interface, face detection for autofocus in some modern digital cameras, photography, facial feature extraction, and marketing are some examples of applications for face detection. A webcam can be built into a television and detect any face that walks by. The approach next establishes the age range, gender, and race of the face. A sequence of advertisements specific to the indicated race, gender, and age can be played after collecting the data. [22].

3. GENERIC FACIAL RECOGNITION SYSTEMS

People started looking into ways to program machines to detect faces in the 1950s. Face geometry was primarily used for face recognition, as applied research in face recognition engineering began in 1964. The process below can create a face recognition system [23]. Identifying a specific area of a picture as a face is known as face localization or detection. This method has many applications, including face tracking, location estimation, and compression. Face normalization is one of the most important considerations when using a vector of geometrical features [24]. They must be somehow normalized to make the retrieved features independent of the face's position, scale, and rotation in the image plane. The next step in the feature extraction process is getting hold of pertinent facial characteristics from the data. The feature extraction procedure needs to be efficient regarding memory and processing time.

The feature extraction method includes dimension reduction, feature extraction, and feature selection. Any pattern recognition algorithm has to conduct dimensionality reduction [25]. The number of test images, features, and the complexity of the classifier all affect how well it performs. After feature extraction, feature selection is frequently carried out [26]. Features are first extracted from the face photos to identify the best subset of characteristics. Lastly, the face should be recognized by the system. The system would expose an identity from a database while performing an identification task. An accuracy metric, a classification algorithm, and a comparison method are used in this stage. Figure 3 shows the structure of the generic Face Recognition system.

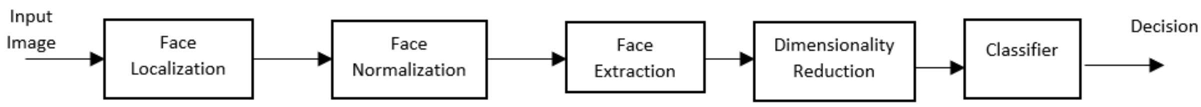

Figure 3. Generic Recognition

3.1 *Development stages of Facial recognition*

3.1.1 *Early algorithm stage*

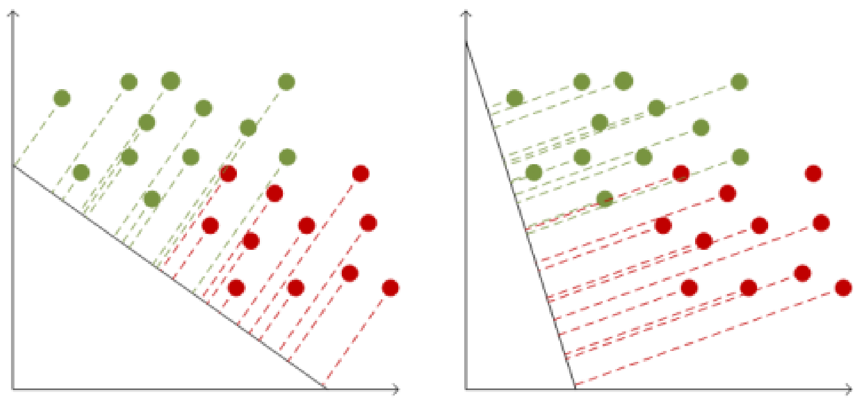

Figure 4. Comparison between PCA and LDA [29]

PCA- The principal component analysis (PCA) algorithm is the one most frequently used to reduce the dimensionality of data. PCA uses feature face extraction in algorithms for facial recognition. In 1991, MIT Media Laboratory's Turk and Pentland developed principal component analysis for facial recognition [27].
LDA - can employ linear discriminate analysis (LDA) for face recognition datasets with labels. It is applied in the classification process. As shown in Figure 4, While PCA requires that the data variance after dimensionality reduction be as significant as possible so that the data can be divided as widely as possible, LDA requires that the variance within the same category of data groups after projection be as small as possible and the variance between groups to be as large as possible [28].

3.1.2 *Artificial features and classifier stage*

SVM- Vapnik and Cortes introduced the support vector machine (SVM) in 1995. A method known as a support vector machine was explicitly created for high-dimensional, small-sample facial recognition problems. It is a classifier that was created using the generalized portrait algorithm. Due to its superior text classification performance, it quickly gained popularity as a machine learning technology [30]. AdaBoost- Schapire made the initial suggestion for boosting the algorithm. It is used to find faces. Any learning algorithm can be more accurate by using a boosting method. The fundamental concept is to combine various classifiers into a better final classifier by a few straightforward principles to improve overall performance [31].

SMALL SAMPLES - The term "small sample problem" alludes to the fact that there aren't enough face recognition training examples available, which prevents most face recognition algorithms from performing as well as they should. Numerous studies have been conducted to store picture information, maintain the association between samples, mitigate the effects of noise, and improve the face recognition effect more efficiently. To address the issue of limited sample sizes, Howland et al. devised a method that combined linear discriminant analysis with unspecified singular value decomposition (GSVD) [32].

NEURAL NETWORKS - An algorithm known as a neural network was created to replicate the human brain for face recognition. Face recognition has grown to be one of the research areas in the field of neural networks as one of the most crucial biometric identification techniques. Figure 5 depicts a typical neural network structure [33].

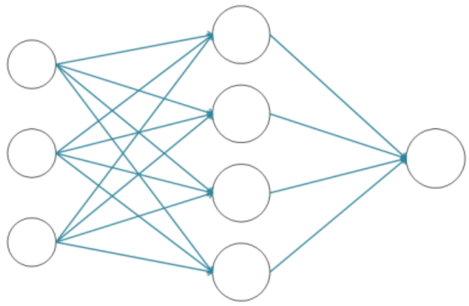

Figure 5. Typical Neural Network Structure [29]

3.1.3 Deep Learning

A subset of machine learning is deep learning. Without the requirement for feature extraction stages, deep learning may automatically identify the features required for categorization during the training phase. This is done to make network learning acquire more valuable features for face recognition. Deep learning has drastically changed the face recognition industry. Deep learning, frequently used in face recognition, can be broken down into the following categories, as shown in Figure 6. By integrating local perception areas, shared weights, and downsampling of face images, CNN optimizes the model structure using the data's locality and other variables. The complexity of the face's shape and texture contributes to the challenge of facial recognition. The authors in [34] presented a deep nonlinear face shape extraction approach from coarse to fine (coarse-to-fine auto-encoder networks, CFAN) further to enhance the nonlinear regression capacity of the algorithm to acquire hardiness to changes such as form.

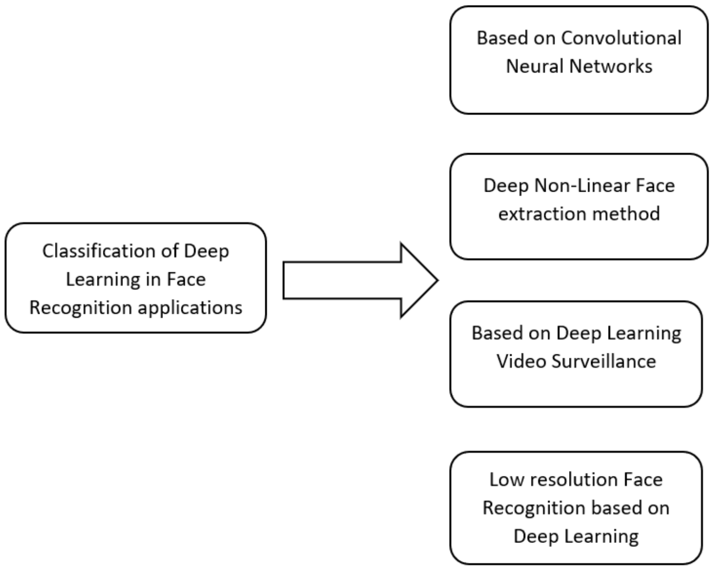

Figure 6. Deep Learning in Face Recognition

For video search and surveillance, it's crucial to identify people reliably and rapidly in the footage. Schofield et al. devised a deep convolution neural network technique to examine animal behavior that could automatically recognize, track, and record human faces in the video [35]. Fourthly, deep learning-based low-resolution face recognition. The gathered face photos in practical applications frequently alter their position, and the limited image resolution causes a quick fall in face image identification performance. To enhance low-resolution face recognition, the most cutting-edge supervised discriminant learning approach was adopted [36], and the generative confrontation network pre-training method and the entire convolution structure were included.

3.2 Approaches for Facial Recognition

3.2.1 Feature Based approaches

Feature-based face recognition selects a few features to solely identify people using prior knowledge or local traits of faces. Local features chosen from facial photos include the eyes, nose, mouth, chin, and head contour. Topological graphs are used to express relationships between features [37]. A straightforward deterministic graph-matching algorithm takes advantage of the fundamental structure to extract recognizable faces from a database. Elastic Bunch Graph Matching is a technique that matches the gallery set, which serves as the model face graph, to the probe set, which serves as the input face graphs, to identify faces. The idea of nodes is fundamental to Elastic Bunch Graph Matching. A particular face feature point represents each node in the input face graph. For instance, one node might represent an eye, another might represent a nose, and so on to describe the other elements of the face. As shown in Figure 7, a graph-like data structure fitted to the face's shape is created by connecting the nodes for the input face graph [38].

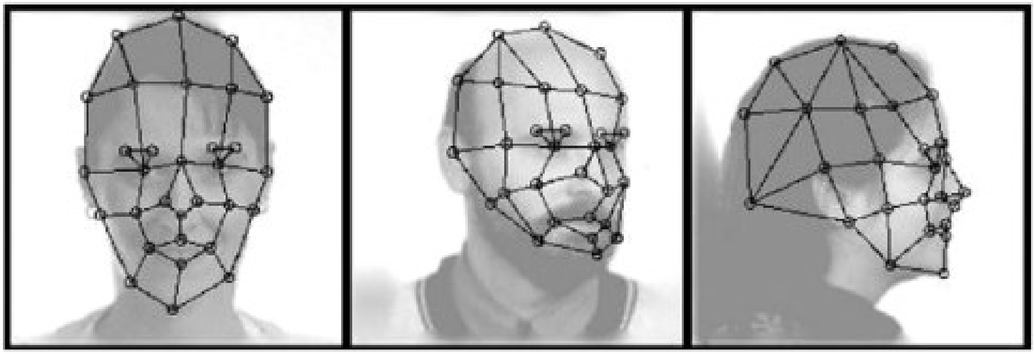

Figure 7. Face Topological graph [38]

3.2.2 Appearance Based approaches

Many academic fields, including biometrics, pattern recognition, and computer vision, have paid attention to appearance-based face identification approaches. Remarkably, two classifications—holistic and hybrid approaches—are implied. The entire face region is used as the initial input in the holistic approach to a recognition system. Additionally, it tries to capture the best possible representation of a face image as a whole and to use statistical patterns in pixel intensity variations [39]. Eigenimages, which come from principal component analysis, are one of the most often utilized representations of the facial region. The second group includes hybrid strategies; just as the human perception system uses local features and the entire face region to identify a face, a machine identification system should do the same.

3.2.3 Soft Computing based approaches

Soft computing tools can also be used for face recognition. Soft computing techniques like neural networks, fuzzy logic, and GA are commonly used. In face recognition, artificial neural networks are a standard tool [40]. They have been applied to categorization and pattern recognition. Because neural networks are nonlinear, their usage is enticing. Therefore, the feature extraction stage might be more efficient than linear KarhunenLoeve methods. The way a neural network is designed is crucial for effective recognition. It is highly dependent on the program being utilized. Higher recognition accuracy is provided by a face recognition system created using fuzzy logic and neural networks [41].

Artificial neural networks, fuzzy logic, and genetic algorithms (GAs) are examples of soft computing approaches that have become significant analytical methodologies in computer vision research. An artificial neural network is an effective technique for resolving the nonlinearity imposed by various restrictions. Fuzzy logic is also utilized to simulate human perception and thought. It is commonly known that analysis using fuzzy sets and fuzzy logic, as well as precise cognition, contribute to the efficacy of the human brain. Fundamental application limitations are almost always accompanied by uncertainty, a typical pattern recognition issue. Fuzzy logic-based analysis has demonstrated significantly improved results in pattern identification. GA is a potent optimization and search algorithm that is based on the notion of natural selection [42]. GA is effective at speeding up computation for a massive heap of space. Because recognizing faces from an ample heap space takes a lot of time, a GA-based strategy is employed to identify the unknown image quickly. GA is used when a user has insufficient time to provide findings without checking each database face. Faster face recognition will be possible with feature extraction and GA. Table 2 represents the summary of these approaches.

Table 2. The summary of these approaches

S.no	Approach	Description
1.	Feature-Based	selects a few features from faces based on past knowledge or local factors to identify people uniquely.
2.	Appearance Based	It is claimed that there are two categories: holistic and hybrid approaches.
3.	Soft Computing Based	We utilize soft computing methods like fuzzy logic, neural networks, and GA.

3.2.4 Algorithms for Face Recognition

Appearance-based or model-based algorithms can be used for facial recognition. The portrayal of the face in these techniques sets them apart. A face is represented using appearance-based approaches as a collection of raw-intensity images. Considered a high-dimensional vector, an image. A feature space is typically derived from the image distribution using statistical methods. This approach can be categorized as linear or nonlinear; linear appearance-based approaches reduce dimensions linearly. The complexity of nonlinear appearance algorithms is higher [1]. A nonlinear manifold is approximated by linear subspace analysis.

On the other hand, the model-based method [43] aims to model a human face. The calculated model's parameters identify the image when introducing the new sample. Model-based methods might be 2D or 3D in nature. These models are frequently morphable, allowing the classification of faces even in position alterations. Elastic Bunch Graph Matching or 3D Morphable Models are two examples of this methodology [44].

3.3 Evaluation Criteria for Facial Recognition

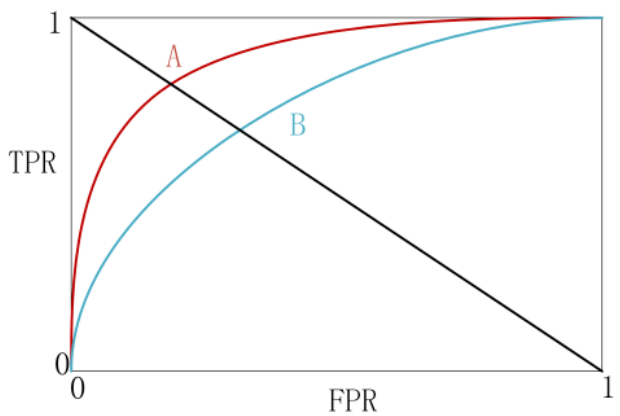

Figure 8. TPR FPR curve for two classifiers [29]

Three crucial metrics for assessing the effectiveness of the face recognition algorithm are accuracy (ACC), the ROC curve, and the area under the curve (AUC) value [45]. The ACC index is frequently used in facial recognition tasks. If M photographs are correctly identified out of the total N images in the testing set, The definition of ACC is as follows.

$$ACC = M/N \tag{1}$$

The algorithm's performance improves as the ACC value increases. In the face recognition task, the ROC first calculates the distance measurement or the similarity between photos. Then, it completes the recognition following the threshold to identify whether two photographs (also known as sample pairs) came from the same person [46]. The abscissa of the ROC curve represents the false positive rate (FPR), and the recall rate or valid positive rate is defined by the ordinate (TPR). The definitions of FPR and TPR are given as follows.

$$TPR = TP/(TP + FN) = FP/(FP + TN) \tag{2}$$

TP stands for the cheerful sample pair that the model incorrectly predicted, FN for the positive sample pair that the model incorrectly predicted, TN for the adverse sample pair that the model incorrectly predicted, and FP for the adverse sample pair that the model incorrectly predicted. Different TPR and FPR values can be obtained by adjusting different thresholds, and ROC curves can be produced (https://blog.csdn.net/). The red and blue curves in Figure 8 represent the TPR FPR curves of two distinct classifiers. The curve's point corresponding to a threshold value is depicted as the ROC curve. The method performs better when the ROC curve is near the upper left corner. In other words, it can still attain a high recall rate even with a meager incorrect recognition rate. AUC value, or the region below the ROC curve, is a scalar to assess the model's strengths. The higher the AUC value, the better the algorithm's performance (https://blog.csdn.net/) [29].

4. MASKED FACIAL RECOGNITION

The most challenging facial occlusion difficulty is recognizing masked faces since they obscure a significant portion of the frontal face, often 60%, which includes rich features like the nose and lips. Face masks deceive face recognition systems' facial verification processes by obscuring a significant portion of the face, resulting in higher inter-class similarities and interclass variance [47].

Matching a masked face with unmasked or masked faces is the goal of the masked face identification issue. When mask-wearing became crucial in containing the COVID-19 pandemic's transmission of the virus, this novel and complex study topic gained significant importance. MFR can support face recognition systems' ability to adapt to real-world situations, which is essential for maintaining public safety [48]. However, the researchers were forced to consider alternate solutions because the many mask types in this procedure made it impossible for the typical facial recognition systems to function correctly. Mask-wearing has become customary during this Covid-19 outbreak, and sadly, it appears that this will continue owing to new mutations. As a result, it is now difficult to deal with this issue efficiently and adequately [49]. The necessity for large-scale masked face datasets for training is revealed by deep learning being a promising contender in solving this complex problem. However, because these datasets are scarce and insufficient, researchers have begun using face-masking technologies to create artificially masked face datasets from existing large-scale face datasets. Masked face datasets can be made using two methods [50]. Collecting actual facial photos of people wearing masks is the first strategy. This strategy, however, takes a lot of effort and money. It is challenging to create an extensive image dataset. Finding the anchor points in the facial photos and adding artificial masks is a different strategy [51].

5. A RECENT INVESTIGATION IN THE FIELD OF MASKED FACIAL RECOGNITION

To address the critical biometric issue of masked face recognition—particularly in light of the continuing COVID-19 pandemic—the work in [52] presents a novel attention-guided deep CNN model. We suggest a dual-branch training method that incorporates our Upper Patch Attention (UPA) module as an auxiliary attention branch into the trunk CNN and directs the backbone network to improve the feature extraction capability in the non-occluded facial regions while maintaining the overall perception of a whole face image.

The Deepmasknet framework for precise face mask detection and masked facial recognition has been given in [53]. An internal unified MDMFR dataset has also been created to evaluate the effectiveness of the suggested strategy. To test face mask detection and masked facial recognition, we have assembled a sizable and varied collection of face photos. The accuracy of 93.33% for masked facial recognition and 100% for face mask detection have demonstrated our Deepmasknet model's superiority to existing methods. Additionally, test results on our MDMFR dataset and the three standard Kaggle datasets have confirmed the proposed Deepmasknet model's robustness for face mask detection and masked facial recognition under a variety of conditions, including variations in face angles, lightning conditions, gender, skin tone, age, and others.

A deep transfer learning model has been constructed [54] to recognize various forms of face masks. The suggested process has two key phasesData preprocessing (conversion of image format, ROI identification, and normalization) and data augmentation are included in the first step of the recognition process, which uses a deep transfer learning neural network (rotation, width, and height shifting, shearing, horizontal flipping, and pixel filling).

The work in [55] suggests and puts into practice a face mask detector that can determine whether a person's face is adequately protected by a mask in settings where COVID-19 regulations are in effect. This model used a two-stage detector construction. The model was trained via transmission learning with the MobileNetV2 framework on a dataset of 7514 photos. To extract the area of interest (ROI), a Haar cascade-based feature extractor and classification technique were used. This led to identifying the sections, which were then forwarded to the learning classification method to obtain the desired result. According to the results of the experiments, training accuracy was 98 percent, and testing accuracy was 97.45 percent.

The depth, azimuth, and elevation geometric properties of a 3D face are used in [56] to depict the face. The intrinsic benefits of 3D faces improve the 3D masked face recognition network's stability and viability. For the 3D masked face recognition network to fully utilize the depth, azimuth, and elevation information in differentiating face identities, a facial geometry extractor is also proposed to emphasize discriminative facial geometric traits. The DAE image stores depth, azimuth, and elevation data in three channels. This study suggests using FGE to draw attention to critical facial geometric elements on each channel. Figure 9 depicts the FGE's organizational structure. The DAE image's depth, azimuth, and elevation channel components correspond to a different convolution kernel. A single convolution kernel handles one channel, and one convolution kernel convolves a single channel. Each output feature map from one-to-one convolution emphasizes the associated geometric features on the face.

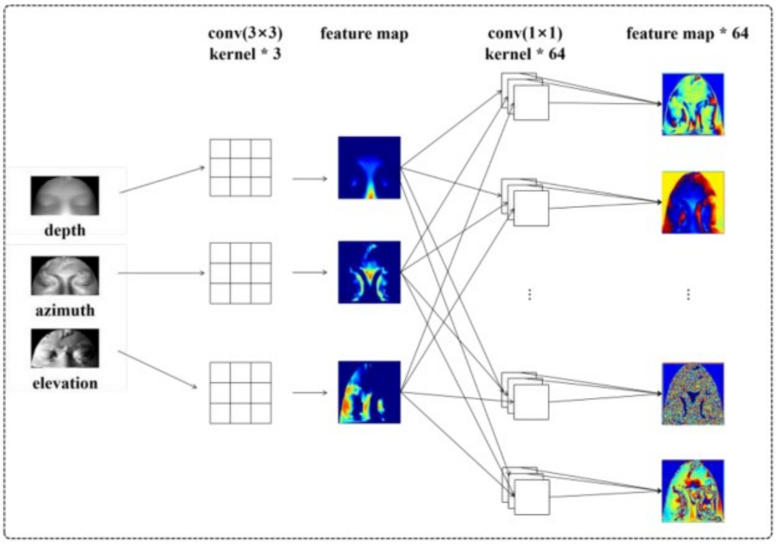

Figure 9. The structure of the Facial Geometry Extractor (FGE) [56]

In [57], the authors have put forth a method that enables us to keep using some facial recognition techniques for faces covered by safety masks, like those used for COVID-19. Our method comprises the preliminary processing of a picture with a hidden section. This preprocessing is based on determining the hue of the skin before processing. The recommended method entails first identifying the occluded and non-occluded facial regions and then distinguishing faces from skin regions. We validated and verified our method using FEI database face photos by introducing safety masks. We applied our strategy to enhance the functionality of the traditional Eigenfaces technology, which, when used directly, does not work on faces with protective masks. The research in [58] offers a straightforward method for achieving this goal using essential machine learning technologies, including TensorFlow, Keras, OpenCV, and Scikit-Learn. The proposed method correctly identifies the face in the picture or video and assesses whether it is wearing a mask. It can distinguish a face and a mask in motion and a video as a surveillance task performance. The method achieves superb accuracy. The authors investigate the best parameter settings for the Convolutional Neural Network model (CNN) to precisely detect the presence of masks without producing over-fitting.

The necessity for contactless biometrics during the COVID-19 pandemic has aided in the growth of masked face recognition. Due to the vast differences between entire faces and masked faces, it is challenging to obtain strong face representations when training a masked face recognition model. These two challenges must be addressed. The research in [56] suggests using non-masked face photos to train a 3D-masked face recognition network to address the first problem. This work uses the depth, azimuth, and elevation geometric properties of 3D faces to represent the face in the second problem. The intrinsic benefits of 3D faces improve the 3D masked face recognition network's stability and viability. A facial geometry extractor is also suggested to draw attention to distinguishing facial geometric aspects. The 3D masked face recognition network may fully utilize the depth, azimuth, and elevation information in differentiating face identities.

Table 4. Summary of recent work

Reference	Methodology	Result
[52]	A novel attention-guided deep CNN model is suggested.	The efficiency of the strategy has been extensively tested on synthetic and real-masked face datasets.
[53]	A Deepmasknet framework has been created for precise face mask detection and masked facial recognition.	The accuracy of 93.33% for masked facial recognition and 100% for face mask detection have demonstrated that our Deepmasknet model is better than existing methods.
[54]	A deep transfer learning model recognizes various types of face masks.	According to experimental data, the deep residual networks (ResNet101v2 and ResNet152v2) offer the best performance with the highest accuracy and negligible loss.
[55]	A face mask detector that can determine whether a person's face is covered adequately at locations where COVID-19 laws are enforced was created using a two-stage detector structure.	According to the results of the experiments, training accuracy was 98 percent, and testing accuracy was 97.45 percent.
[57]	The suggested method first identifies the occluded and non-occluded portions of the face and then distinguishes faces from skin regions.	The obtained experimental findings partially back the method the authors recommend for identifying face wear
[58]	These include TensorFlow, Keras, OpenCV, and Scikit-Learn, which are fundamental machine learning technologies.	The method can be applied to identify a face and mask in motion and video. The technique achieves superb accuracy.
[56]	The depth, azimuth, and elevation of the 3D face are used in this study's representation of the face's geometry. A facial geometry extractor is suggested to draw attention to distinguishing facial geometric traits.	The experimental results on four open 3D face datasets demonstrate that the proposed 3D masked face recognition network enhances the masked face recognition accuracy, proving that it is feasible to train the masked face recognition model using non-masked face photos.
[59]	The suggested approach is based on the FaceNet framework. It aims to improve the performance of both scenarios with and without a mask by altering the face recognition model currently in use.	The outcome displays a remarkable accuracy of 99.2% on a scenario with faces wearing masks.
[60]	A modified convolutional neural network creates a system for masked face recognition-based attendance (CNN).	The trial outcomes demonstrate the suggested CNN's high efficacy (i.e., 98.92%) in identifying masked faces for recording attendance.

In [59], a method for identifying human faces while wearing a mask is suggested. The lower third of the human face is obscured and cannot be utilized in the facial recognition learning process. Therefore, the proposed method is designed to recognize human faces on any facial features that are accessible, which may vary depending on whether a mask is worn or not—design, method, and strategy. The suggested approach is based on the FaceNet framework. It aims to improve the performance of both scenarios with and without a mask by altering the face recognition model currently in use. Then, simulated masked-face images are computed on top of the original face photos to aid in the learning process of face recognition.

Additionally, feature heat maps are created to show the bulk of the facial features that are important for identifying people wearing masks. In [60], a modified convolutional neural network creates a masked facial recognition-based attendance system (CNN). To accomplish this, CNN's original Softmax classifier is replaced with a Support Vector Machine. In a 5-fold cross-validation, the improved CNN's performance in identifying faces with masks was compared to that of other CNNs. The trial outcomes demonstrate the suggested CNN's high efficacy (i.e., 98.92%) in identifying masked faces for recording attendance.

Compared to cutting-edge automatic FR solutions, the research in [61] offers a joint evaluation and in-depth analysis of the face verification performance of human specialists. This incorporates four automatic recognition solutions and a thorough review by human experts. A series of key takeaways on various facets of the association between the verification behavior of humans and machines are presented at the study's conclusion. Table 4 provides a summary of the recent work.

5.1 Limitations and Gap

1. Mask-wearing conditions are ignored by some approaches, which merely identify the masked and non-mask facial classes.
2. It is common knowledge that inappropriate mask use will not stop the propagation of COVID-19. The detection of mask-wearing situations was only possible using a few methods. As a result, more algorithms should be tested to identify situations where people are putting masks over their faces.
3. There still needs to be standardized evaluation for many masked facial detection methods, even though some literature evaluates several methods. Several methodologies might be used on different platforms.
4. The outcomes presented in the original literature merely enable conceptual comparisons for readers. Making a fair judgment is difficult. Quite a few techniques can be used to produce a good performance.
5. Several approaches need more specific information regarding their cost-effectiveness and operating environment.
6. Running time is a significant measurement statistic in practical applications.
7. Using current ways to maintain performance with lightweight equipment is a challenge.

6. CONCLUSION

Face recognition technology has made enormous strides with the advancement of science and technology, but there is still potential for improvement in terms of practical use. The authors have reviewed a series of masked face detection methods. In addition to this, various challenges faced in facial detection have been highlighted. Apart from facial detection, this paper presents an overview of the recent work done in Facial Recognition systems. Considering the current pandemic outbreak, the authors have tried to pay special attention to masked facial detection and recognition. After carrying out a detailed review, it is safe to say that although great efforts are being taken in the field of facial detection and recognition or masked facial detection and recognition, detecting the proper mask-wearing conditions is still a concern.

REFERENCES

[1] T. Serre, G. Kreiman, M. Kouh, C. Cadieu, U. Knoblich, and T. Poggio, "A quantitative theory of immediate visual recognition," Prog. Brain Res., vol. 165, pp. 33–56, Jan. 2007. https://doi.org/10.1016/S0079-6123(06)65004-8

[2] N. Khan and S. Fahad, "Critical Review of the Present Situation of Corona Virus in China," SSRN Electron. J., Feb. 2020. https://doi.org/10.2139/SSRN.3543177

[3] P. Kaur, K. Krishan, S. K. Sharma, and T. Kanchan, "Facial-recognition algorithms: A literature review," https://doi.org/10.1177/0025802419893168, vol. 60, no. 2, pp. 131–139, Jan. 2020. https://doi.org/10.1177/0025802419893168

[4] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," Int. J. Comput. Vis., vol. 57, no. 2, pp. 137–154, May 2004. https://doi.org/10.1023/B: VISI.0000013087.49260.fb

[5] M. K. Hasan, M. S. Ahsan, Abdullah-Al-Mamun, S. H. S. Newaz, and G. M. Lee, "Human Face Detection Techniques: A Comprehensive Review and Future Research Directions," Electron. 2021, Vol. 10, Page 2354, vol. 10 no. 19, p. 2354, Sep. 2021. https://doi.org/10.3390/electronics10192354

[6] R. Ranjan et al., "A Fast and Accurate System for Face Detection, Identification, and Verification," IEEE Trans. Biometrics, Behav. Identity Sci., vol. 1, no. 2, pp. 82–96, Apr. 2019. https://doi.org/10.1109/TBIOM.2019.2908436.

[7] S. Hemathilaka and A. Aponso, "A Comprehensive Study on Occlusion Invariant Face Recognition under Face Mask Occlusion," Jan. 2022. https://doi.org/10.5121/csit.2021.111804

[8] S. Stein and G. A. Fink, "A new method for combined face detection and identification using interest point descriptors," 2011 IEEE Int. Conf. Autom. Face Gesture Recognition. Work. FG 2011, pp. 519–524, 2011. https://doi.org/ 10.1109/FG.2011.5771452

[9] H. Zeng, F. Qin, and K. Lin, "An Optimized Face Detection Based on Adaboost Algorithm," Proc. 2018 Int. Conf. Inf. Syst. Comput. Aided Educ. ICISCAE 2018, pp. 375–378, Mar. 2019. https://doi.org/ 10.1109/ICISCAE.2018.8666925

[10] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLO-face: a real-time face detector," Vis. Comput., vol. 37, no. 4, pp. 805–813, Apr. 2021. https://doi.org/10.1007/S00371-020-01831-7

[11] B. Wang, J. Zheng, and C. L. P. Chen, "A Survey on Masked Facial Detection Methods and Datasets for Fighting Against COVID-19," IEEE Trans. Artif. Intell., vol. 3, no. 3, pp. 323–343, 2022. https://doi.org/10.1109/TAI.2021.3139058

[12] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with LLE-CNNs," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, pp. 426–434, 2017. https://doi.org/10.1109/CVPR.2017.53

[13] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "MaskedFace-Net – A dataset of correctly/incorrectly masked face images in the context of COVID-19," Smart Heal., vol. 19, p. 100144, Mar. 2021. https://doi.org/10.1016/J.SMHL.2020.100144

[14] B. Roy, S. Nandy, · Debojit Ghosh, · Debarghya Dutta, · Pritam Biswas, and T. Das, "MOXA: A Deep Learning Based Unmanned Approach for Real-Time Monitoring of People Wearing Medical Masks," Trans. Indian Natl. Acad. Eng. 2020 53, vol. 5, no. 3, pp. 509–518, Jul. 2020. https://doi.org/10.1007/S41403-020-00157-z

[15] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks." pp. 4401–4410, 2019. https://doi.org/10.1109/CVPR.2019.00453

[16] G. B. Huang and E. Learned-Miller, "Labeled Faces in the Wild: Updates and New Reporting Procedures," Accessed: Dec. 21, 2022. [Online]. Available: http://vis-www.cs.umass.edu/.

[17] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep Learning Face Attributes in the Wild." pp. 3730–3738, 2015, Accessed: Dec. 21, 2022. [Online]. Available: http://personal.ie.cuhk.edu.hk/.

[18] F. I. Eyiokur, H. K. Ekenel, and A. Waibel, "Unconstrained face mask and face-hand interaction datasets: building a computer vision system to help prevent the transmission of COVID-19," Signal, Image Video Process., pp. 1–8, Jul. 2022. https://doi.org/10.1007/s11760-022-02308-x

[19] S. Singh, U. Ahuja, M. Kumar, K. Kumar, and M. Sachdeva, "Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment," Multimed. Tools Appl., vol. 80, no. 13, pp. 19753–19768, May 2021. https://doi.org/10.1007/s11042-021-10711-8

[20] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "WIDER FACE: A Face Detection Benchmark." pp. 5525–5533, 2016. https://doi.org/10.1109/CVPR.2016.596

[21] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," Artif. Intell. Rev., vol. 52, no. 2, pp. 927–948, 2019. https://doi.org/10.1007/s10462-018-9650-2

[22] D. S. Brar, A. Kumar, Pallavi, U. Mittal, and P. Rana, "Face Detection for Real World Application," Proc. 2021 2nd Int. Conf. Intell. Eng. Manag. ICIEM 2021, pp. 239–242, Apr. 2021, https://doi.org/10.1109/ICIEM51511.2021.9445287

[23] A. Review, "Past, Present, and Future of Face Recognition: A Review," 2020.

[24] K. Nagano et al., "Deep face normalization," ACM Trans. Graph., vol. 38, no. 6, Nov. 2019. https://doi.org/10.1145/3355089.3356568

[25] S. Ayesha, M. K. Hanif, and R. Talib, "Overview and comparative study of dimensionality reduction techniques for high dimensional data," Inf. Fusion, vol. 59, pp. 44–58, Jul. 2020. https://doi.org/10.1016/J.INFFUS.2020.01.005

[26] B. Venkatesh and J. Anuradha, "A review of Feature Selection and its methods," Cybern. Inf. Technol., vol. 19, no. 1, pp. 3–26, 2019, https://doi.org/10.2478/CAIT-2019-0001

[27] A. Alahmadi, M. Hussain, H. A. Aboalsamh, and M. Zuair, "PCAPooL: unsupervised feature learning for face recognition using PCA, LBP, and pyramid pooling," Pattern Anal. Appl., vol. 23, no. 2, pp. 673–682, May 2020. https://doi.org/10.1007/s10044-019-00818-y

[28] A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang, "A hybrid improved kernel LDA and PNN algorithm for efficient face recognition," Neurocomputing, vol. 393, pp. 214–222, Jun. 2020. https://doi.org/10.1016/j.neucom.2019.01.117

[29] L. Li, X. Mu, S. Li, and H. Peng, "A Review of Face Recognition Technology," IEEE Access, vol. 8, pp. 139110–139120, 2020. https://doi.org/10.1109/ACCESS.2020.3011028

[30] P. VenkateswarLal, G. R. Nitta, and A. Prasad, "Ensemble of texture and shape descriptors using support vector machine classification for face recognition," J. Ambient Intell. Humaniz. Comput., pp. 1–8, Apr. 2019. https://doi.org/10.1007/S12652-019-01192-7

[31] K. Sowmya, "Facial Recognition for Automated Attendance System Using Ada Boost Algorithm."

[32] F. Xu, J. Gao, and X. Pan, "Cow Face Recognition for a Small Sample Based on Siamese DB Capsule Network," IEEE Access, vol. 10, pp. 63189–63198, 2022. https://doi.org/10.1109/ACCESS.2022.3182806

[33] S. Almabdy and L. Elrefaei, "Deep Convolutional Neural Network-Based Approaches for Face Recognition," Appl. Sci. 2019, Vol. 9, Page 4397, vol. 9, no. 20, p. 4397, Oct. 2019. https://doi.org/ 10.3390/APP9204397

[34] J. Zhang, S. Shan, M. Kan, and X. Chen, "Coarse-to-Fine Auto-encoder Networks (CFAN) for real-time face alignment," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8690 LNCS, no. PART 2, pp. 1–16, 2014. https://doi.org/10.1007/978-3-319-10605-2_1

[35] D. Schofield et al., "Chimpanzee face recognition from videos in the wild using deep learning," Sci. Adv., vol. 5, no. 9, Sep. 2019. https://doi.org/10.1126/sciadv.aaw0736

[36] P. Li, L. Prieto, D. Mery, and P. J. Flynn, "On Low-Resolution Face Recognition in the Wild: Comparisons and New Techniques," IEEE Trans. Inf. Forensics Secur., vol. 14, no. 8, pp. 2000–2012, Aug. 2019. https://doi.org/10.1109/TIFS.2018.2890812

[37] S. H. Oh, G. W. Kim, and K. S. Lim, "Compact deep learned feature-based face recognition for Visual Internet of Things," J. Supercomput., vol. 74, no. 12, pp. 6729–6741, Dec. 2018. https://doi.org/10.1007/S11227-017-2198-0

[38] L. C. Jain, U. Halici, I. Hayashi, S. B. Lee, and S. Tsutsui, "Intelligent Biometric Techniques in Fingerprint and Face Recognition," Intell. Biometric Tech. Fingerpr. Face Recognition., no. July 1997, pp. 1–463, 2022, https://doi.org/10.1201/9780203750520

[39] P. Payal and M. M. Goyani, "A comprehensive study on face recognition: methods and challenges," https://doi.org/10.1080/13682199.2020.1738741, vol. 68, no. 2, pp. 114–127, Feb. 2020. https://doi.org/10.1080/13682199.2020.1738741

[40] K. Amit, "Artificial Intelligence and Soft Computing : Behavioral and Cognitive Modeling of the Human Brain," Artif. Intell. Soft Comput., Oct. 2018. https://doi.org/10.1201/9781315219738

[41] Y. I. Daradkeh, I. Tvoroshenko, V. Gorokhovatskyi, L. A. Latiff, and N. Ahmad, "Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic," IEEE Access, vol. 9, pp. 13417–13428, 2021. https://doi.org/10.1109/ACCESS.2021.3051625

[42] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on the genetic algorithm: past, present, and future," Multimed. Tools Appl., vol. 80, no. 5, pp. 8091–8126, Feb. 2021. https://doi.org/10.1007/s11042-020-10139-6

[43] J. Li, T. Qiu, C. Wen, K. Xie, and F. Q. Wen, "Robust Face Recognition Using the Deep C2D-CNN Model Based on Decision-Level Fusion," Sensors 2018, Vol. 18, Page 2080, vol. 18, no. 7, p. 2080, Jun. 2018. https://doi.org/10.3390/S18072080

[44] M. P. Beham and S. M. M. Roomi, "A review of face recognition methods," Int. J. Pattern Recognit. Artif. Intell., vol. 27, no. 4, pp. 1–35, 2013. https://doi.org/10.1142/S0218001413560053

[45] J. J. Howard, L. R. Rabbitt, and Y. B. Sirotin, "Human-algorithm teaming in face recognition: How algorithm outcomes cognitively bias human decision-making," PLoS One, vol. 15, no. 8, p. e0237855, Aug. 2020. https://doi.org/10.1371/journal.pone.0237855

[46] P. Grother, M. Ngan, K. Hanaoka, W. Ross, and W. Copan, "Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects," 2019. https://doi.org/10.6028/NIST.IR.8280

[47] M. Alexiou, I. P. Ktistakis, and G. Goodman, "Towards a Masked Face Recognition Algorithm: A Novel Rule Based Hybrid Algorithm," 6th South-East Eur. Des. Autom. Comput. Eng. Comput. Networks Soc. Media Conf. SEEDA-CECNSM 2021, 2021. https://doi.org/10.1109/SEEDA-CECNSM53056.2021.9566244

[48] Y. Martínez-Díaz, H. Méndez-Vázquez, L. S. Luevano, M. Nicolás-Díaz, L. Chang, and M. González-Mendoza, “Towards Accurate and Lightweight Masked Face Recognition: An Experimental Evaluation,” IEEE Access, vol. 10, pp. 7341–7353, 2022. https://doi.org/10.1109/ACCESS.2021.3135255

[49] H. Du, H. Shi, Y. Liu, D. Zeng, and T. Mei, "Towards NIR-VIS Masked Face Recognition," IEEE Signal Process. Lett., vol. 28, pp. 768–772, 2021. https://doi.org/10.1109/LSP.2021.3071663

[50] K. Wang et al., "Mask Aware Network for Masked Face Recognition in the Wild," Proc. IEEE Int. Conf. Comput. Vis., vol. 2021-Octob, pp. 1456–1461, 2021. https://doi.org/10.1109/ICCVW54120.2021.00168

[51] J. Yu, X. Hao, Z. Cui, P. He, and T. Liu, "Boosting Fairness for Masked Face Recognition," Proc. IEEE Int. Conf. Comput. Vis., vol. 2021-Octob, pp. 1531–1540, 2021. https://doi.org/10.1109/ICCVW54120.2021.00178

[52] Y. Zhang, X. Wang, M. S. Shakeel, H. Wan, and W. Kang, "Learning upper patch attention using dual-branch training strategy for masked face recognition," Pattern Recognit., vol. 126, 2022. https://doi.org/10.1016/j.patcog.2022.108522

[53] N. Ullah, A. Javed, M. Ali Ghazanfar, A. Alsufyani, and S. Bourouis, "A novel DeepMaskNet model for face mask detection and masked facial recognition," J. King Saud Univ. - Comput. Inf. Sci., Volume 34, Issue 10, Part B, 2022. https://doi.org/10.1016/j.jksuci.2021.12.017

[54] R. Mar-Cupido, V. García, G. Rivera, and J. S. Sánchez, "Deep transfer learning for the recognition of types of face masks as a core measure to prevent the transmission of COVID-19," Appl. Soft Comput., vol. 125, p. 109207, 2022. https://doi.org/10.1016/j.asoc.2022.109207

[55] M. Gupta, G. Chaudhary, D. Bansal, and S. Pandey, "DTLMV2—A real-time deep transfer learning mask classifier for overcrowded spaces," Appl. Soft Comput., vol. 127, no. June 2020, p. 109313, 2022. https://doi.org/10.1016/j.asoc.2022.109313

[56] Y. Wang, Z. Yang, Z. Zhang, H. Zang, Q. Zhu, and S. Zhan, "Masked Face Recognition with 3D Facial Geometric Attributes," ACM Int. Conf. Proceeding Ser., pp. 13–21, 2022. https://doi.org/10.1145/3529446.3529449

[57] F. Ennaama, K. Benhida, and S. Essalki, "Proposed approach to improve facial recognition techniques for occluded faces by Covid-19 mask protection," IFAC-PapersOnLine, vol. 55, no. 12, pp. 456–461, 2022. https://doi.org/10.1016/j.ifacol.2022.07.354

[58] G. Kaur et al., "Face mask recognition system using CNN model," Neurosci. Informatics, vol. 2, no. 3, p. 100035, 2022. https://doi.org/10.1016/j.neuri.2021.100035

[59] W. Moungsouy, T. Tawanbunjerd, N. Liamsomboon, and W. Kusakunniran, "Face recognition under mask-wearing based on residual inception networks," Appl. Comput. Informatics, 2022. https://doi.org/10.1108/ACI-09-2021-0256

[60] Y. Bong and G. Lee, "International Journal on Robotics, Automation and Sciences," vol. 3, pp. 33–41, 2021. https://doi.org/10.33093/ijoras.2021.3.6

[61] N. Damer, F. Boutros, M. Süßmilch, M. Fang, F. Kirchbuchner, and A. Kuijper, "Masked face recognition: Human versus machine," IET Biometrics, vol. 11, no. 5, pp. 512–528, 2022. https://doi.org/10.1049/bme2.12077