Multimedia Tools and Applications (2023) 82:1669–1748
https://doi.org/10.1007/s11042-022-13248-6

A comprehensive survey on techniques to handle face identity threats: challenges and opportunities

Mayank Kumar Rusia[1] · **Dushyant Kumar Singh**[1]

Received: 9 February 2021 / Revised: 3 February 2022 / Accepted: 15 May 2022 /
Published online: 10 June 2022

Abstract

The human face is considered the prime entity in recognizing a person's identity in our society. Henceforth, the importance of face recognition systems is growing higher for many applications. Facial recognition systems are in huge demand, next to fingerprint-based systems. Face-biometric has a highly dominant role in various applications such as border surveillance, forensic investigations, crime detection, access management systems, information security, and many more. Facial recognition systems deliver highly meticulous results in every of these application domains. However, the face identity threats are evenly growing at the same rate and posing severe concerns on the use of face-biometrics. This paper significantly explores all types of face recognition techniques, their accountable challenges, and threats to face-biometric-based identity recognition. This survey paper proposes a novel taxonomy to represent potential face identity threats. These threats are described, considering their impact on the facial recognition system. State-of-the-art approaches available in the literature are discussed here to mitigate the impact of the identified threats. This paper provides a comparative analysis of countermeasure techniques focusing on their performance on different face datasets for each identified threat. This paper also highlights the characteristics of the benchmark face datasets representing unconstrained scenarios. In addition, we also discuss research gaps and future opportunities to tackle the facial identity threats for the information of researchers and readers.

Keywords Biometrics · Face recognition · Authentication · Computer vision · Machine learning · Deep learning · Image processing

1 Introduction

Nowadays, face biometric-based recognition has become a serious requirement in authentication systems for public safety and security. The rich facial structure and non-intrusive

✉ Mayank Kumar Rusia
mayank.qip18@mnnit.ac.in

Dushyant Kumar Singh
dushyant@mnnit.ac.in

[1] CSED, MNNIT Allahabad, Prayagraj, Uttar Pradesh, INDIA

property have attracted more attention from communities of researchers than other biometric traits such as fingerprint, iris, palm, and more. Face biometric-based sensors use an individual's physical appearance (characteristics) to distinguish one person from others [17]. In the recent past, the face biometric-based recognition systems have demonstrated remarkable developments in intelligent surveillance, financial systems, security monitoring, forensic investigations, the civil aviation industry, and other areas due to their convenience and reliability. Face detection is the first essential step in any face recognition activity. Humans can quickly and easily identify individuals through their visual and mental abilities. Nevertheless, it can become a vital problem if the computer performs the same task, especially for unconstrained real-time scenarios [129]. Thus we need an automated system capable of performing face detection and recognition tasks efficiently and intelligently in a real-time environment. The face recognition process involves locating the face region, aligning the detected face region, extracting the discriminant features, and finally classifying the face, which ultimately determines a person's identity. Figure 1 presents the complete process of face recognition. The key terms utilized in face recognition are described below:

- **Face detection:** A process of selecting a region of interest (i.e., face) from the input image or the video sequence.
- **Face image preprocessing:** The acquired face image cannot be directly considered as final input for face recognition as this might contain some additional unwanted information such as ear, neck, dressing accessories, and jewellery. This extra information may vary every time and adds the probability of getting erroneous features for further processing. The other reason for preprocessing is to enhance the quality of the captured image through various image processing methods such as alignment, normalization, standardization, and noise removal.
- **Feature extraction:** Feature extraction is a process of extracting essential and important characteristics of the object of interest, which finally is transformed as a one-dimensional vector, typically in maximum applications. The feature of an object may include color, texture, and shape.
- **Pattern matching/Feature classification:** Pattern matching compares the input image or video sequence to the stored database image (i.e., template) and generates a similarity

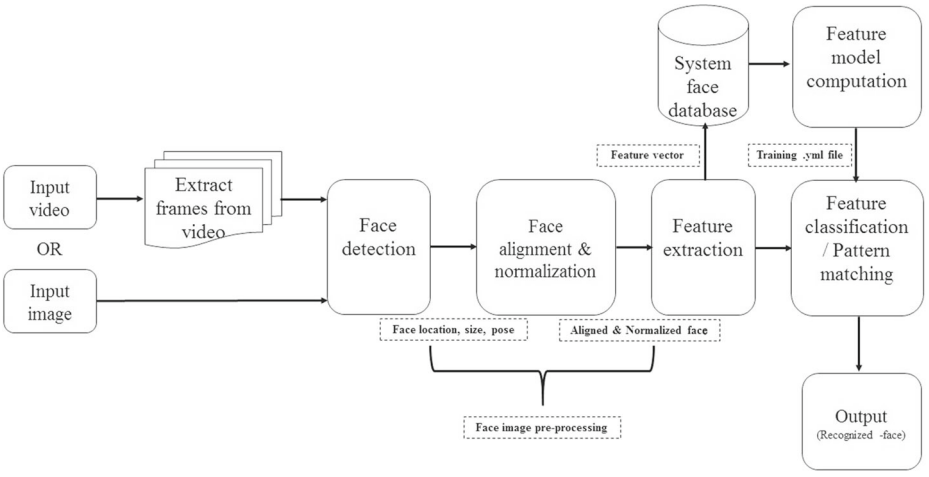

Fig. 1 Face recognition process

score. The extracted features (patterns) are used to validate a person's identity based on a similarity score. In the same way, the threshold value is utilized to classify a claimed identity as being acceptable or not.

The rapid evolution of facial recognition systems in real-time applications has raised new concerns about their ability to resist identity threats, particularly in non-intrusive applications. For instance, automatic border control systems [80] of the aviation industry, where an automated face recognition system is installed at each entry and exit point to authenticate the identity of individuals without any security personnel. The face is a traditionally exposed (i.e., visible) part of the human body and is easily accessible in photos and videos through various social media such as Facebook, WhatsApp, Instagram, and others. Therefore, the fraudsters can easily steal, misuse, and modify these available identities for any illegal activity as it is almost impossible to keep this biometric trait secret. Face-based identity threats are intended to fool facial recognition systems by presenting facial biometric artifacts [110] in place of the real person's face identity. Such face identity loss is a major challenge and brings a significant change in the rapid development of a real-time applications equipped with face recognition system. The contributions of some of the earlier generic surveys and state-of-the-art research work to address the challenges of facial recognition systems are displayed in Table 1.

A careful study of this evolutionary research domain through various research and review articles reveals the following motivating observations:

- Most state-of-the-art research and review articles represent one or more of the three challenges associated with the face recognition system. None of the surveys have provided a comprehensive classification of face identity threats.
- No review article provides comparative studies covering all potential face identity threats together and their truly adaptive solution aspects.
- The Global Research Committee has paid less attention to these critical threats that can substantially impair the performance of the face recognition system.

Considering these persuasive observations, we provide a comprehensive and critical survey of the various identity threats in face recognition. This survey paper provides imperative attention to the most clamorous real-time face identity threat such as face spoofing, partial face occlusion, face pose variation, facial expression variation, and illumination. However, the scope of this survey also touches upon contemporary research, which includes other challenging factors for facial recognition, such as plastic surgery, aging, gender, camera viewpoints, noise, and cluttered environmental effects. We mainly emphasize on various methods developed and used to reduce the impact of face identity threats. Our contributions and finding are summarized in this paper as follows:

- We present a comprehensive assortment of face recognition algorithms, summarizing features and recent development for each category.
- We propose a novel taxonomy of potential face identity threats identified in face recognition.
- We explore various scenarios in this survey such as direct spoofing (plastic surgery, make-up), indirect spoofing (photo attack, video attack, mask attack), Zero Effort Imposter (identical twins and face morphing), other factors (occlusion, expression, aging, race, gender, pose, illumination, low resolution, cluttered background, camera orientation), and modularity impact on face recognition over the last decade.

Table 1 Summary of state-of-the-art surveys addressing facial recognition challenges

Reference	FRM	TFIT	FDD	Face Identity Threats															
				DS		IDS		ZEI		Other Factors									
				PS	MU	2D	3D	IT	Mo	Occ	Exp	Ag	R	G	Po	ill	LR	CB	CO
Tamilselv M., et al. [157]	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Jia S., et al. [68]	N	N	Y	N	N	N	Y	N	N	N	N	N	N	N	N	N	N	N	N
Guo G, et al. [55]	Y	N	Y	N	Y	N	N	N	N	Y	Y	N	N	N	Y	Y	Y	N	N
Revina IM, et al. [125]	Y	N	Y	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N	N
Minaee S, et al. [104]	Y	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Sawant MM, et al. [134]	Y	N	Y	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N
Zhou S, et al. [190]	Y	N	Y	N	N	N	Y	N	N	Y	Y	N	N	N	Y	N	N	N	N
Mortezaie Z, et al. [107]	Y	N	Y	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N
Bowyer KW, et al. [22]	Y	N	Y	N	N	N	N	N	Y	N	Y	N	N	N	N	N	N	N	N
Soltanpour S, et al. [150]	Y	N	N	N	N	N	Y	N	N	N	N	N	N	N	N	N	N	N	N
Mahmood Z, et al.[97]	Y	N	Y	N	N	N	N	N	N	Y	Y	N	N	N	Y	Y	Y	N	N
Abate AF, et al. [1]	Y	N	Y	N	N	N	N	N	N	Y	N	Y	N	N	Y	Y	N	N	N
This Survey	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**

FRM-Face Recognition Methods, TFIT- Taxonomy of Facial identity Threats, FDD- Face Dataset Details, DS- Direct Spoofing, IDS-Indirect Spoofing, ZEI- Zero Effort Imposter, PS-Plastic Surgery, MU-Makeup, 2D-Two-dimensional Photo/Video Attack, 3D-Three-dimensional Mask Attack, IT-Identical Twins, Mo-Morphing, Occ-Occlusion, Exp-Expression, Ag-Aging, R-Race, G-Gender, Po-Pose, ill-Illumination, LR-Low Resolution, CB-Cluttered Background, CO-Camera Orientation, Y-Yes, N-No

- We provide a concise description of the various state-of-the-art techniques that can be applied to minimize face recognition threats. Furthermore, this survey includes a tabular comparison of these techniques, which would help in understanding the handling of face identity threats.
- We also provide a comprehensive comparison of various state-of-the-art countermeasures techniques from contemporary insight research conducted to reduce the impact of identified face recognition threats.
- We also highlight the various available face datasets for each identified category of face identity threats, including static images, video sequences, and heterogeneous data.
- We point out some research opportunities and solution aspects for readers and researchers to address these challenges efficiently.

The structure of this paper (as shown in Fig. 2) is organized in the following way: Section 1 gives a brief introduction of face detection and recognition along with the motivations behind this work. We also provided a tabular summary of earlier generic surveys with a clear indication of the main contribution of this survey, as shown in Table 1. Section 2 contains a classification of traditional face recognition techniques with a summary. Section 3 proposes a new hierarchy of potential face identity threats. Section 4 describes various techniques for mitigating identity threats based on features extraction, classification, dimensionality reduction, and neural network-based algorithms with tabular comparison. Section 5 represents an extensive tabular analysis of the various approaches proposed to mitigate the face identity threats. We also highlight a detailed description of the various available face datasets in this section. In addition, we provide a summary and remarkable point for best methodologies after each table with their pros and cons. Section 6 discusses each challenge along with some significant opportunities for researchers and readers. Finally, the conclusions followed by the future work are summarized in Section 7 of this paper.

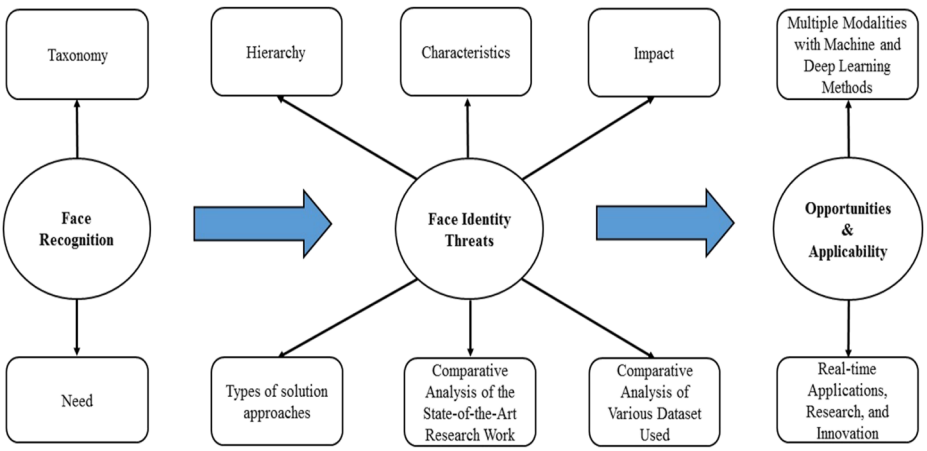

Fig. 2 The structured workflow of the paper

2 Classification of face recognition approaches

Face recognition is a computer vision technique that enables a computer to infer a person's identity from an image. Face recognition is sometimes referred to as face identification and verification. Both the terms identification and verification are similar, but the application aspects of these two are different. Face verification means authorizing someone based on one-to-one mapping, reflecting the perception of "Is this person X?". For instance, smartphones can be locked or unlocked through genuine face biometric. In face identification (recognition), the system looks at the person's identity from database images to find a match for that person. It refers to a one-to-many relationship. The notion of this identification system is "who is this person?". Examples of face identification can be a surveillance system [147] and an attendance monitoring system [129]. In this section, we summarize the traditional to recent trends in face recognition.

Research related to face recognition began in the early 1970s, but since late 1998, this research domain has been witnessing rapid improvement. The face identity of an individual can be realized in two steps, the first is the detection of face location (region of interest), and the second is classification of the detected face region. In the early 1990s, face localization [89] was identified through face numerology, where each component of the face (i.e., eyes, nose, mouth, chin, and more) is assigned a specific numeric value. Various methods such as the annotated point, the distance between the annotated points, and the angle joining these points are considered to locate these face regions in an image. Inspired by facial numerology, the researchers identified some universal face shapes to distinguish facial features from other body parts, such as the oval, long, round, square, diamond, and heart-shaped.

Since the last decade, several approaches have been introduced following the invention of the deep learning approach with efficient computation power and ample memory space for storing large databases. Figure 3 represents a broad classification of all the traditional approaches to date. We are considering the maximum of them concisely as these approaches

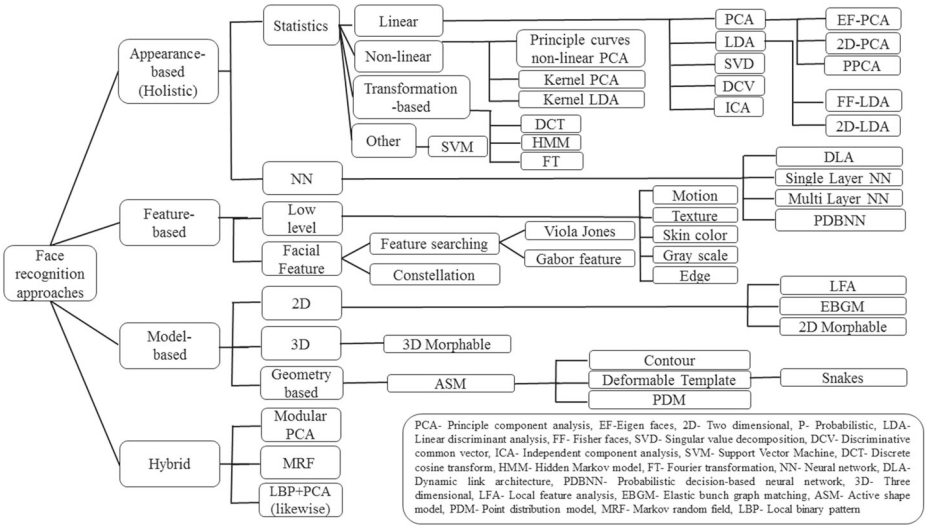

Fig. 3 Classification of face recognition approaches

are still alive and used extensively. Researchers mainly refer to these traditional approaches to provide an efficient solution for a wide variety of real-world problems.

Traditional face recognition approaches are categorized into four types: appearance-based, feature-based, model-based, and hybrid approaches.

2.1 Appearance-based approaches

Appearance-based approaches [26, 36, 59, 70, 112, 120, 144, 185] rely entirely on the facial appearance. These holistic techniques focus on facial geometry that includes linear, non-linear, transformation-based components without impairing facial information. These approaches are simple and fast to deploy for real-world problems.Appearance-based approaches can be divided into two subcategories: linear and non-linear. The linear approach consists of a one-dimensional function to map the independent variable with the dependent variable, while the non-linear approach involves multidimensionality. Non-linearity is represented by the kernel functions (i.e., redial-basis and Gaussian distribution) to map between higher-dimensional spaces to low-dimensional space. Appearance-based approaches are more prone to unconstrained pose, expression, low illumination, and cluttered backgrounds. With the recent advances in machine learning in terms of computational power, a remarkable change has been observed in the learning process concerning the time and cost to solve complex problems. The most common machine learning models are neural networks and support vector machines.

2.2 Features-based approaches

In contrast to appearance-based approaches, feature-based approaches are more promising for finding a face in an image using low, middle, and high-level feature extraction [7, 56, 73, 82, 98, 109, 141, 145, 149]. Low-level features include skin color, edges (i.e., intensity change), texture, and size. Skin color is the key feature to distinguish the human face (i.e., visible part of the body) from other objects. Edges can be measured by changing the intensity of the pixel values. The texture is a feature used to represent the spatial distribution of different color spaces and their intensity values in an image. Low-level facial features [87] include universal facial attributes (eyes, nose, mouth, chin, and jawline) to locate the face regions in the entire image. Various techniques exist for extracting facial features, such as Viola-Jones, Gabor features.

2.3 Model-based approaches

The model-based approach [10, 18, 23, 72, 82, 153, 164] is an automatic feature detection approach intended to generate a set of unique patterns to correlate it with query samples. This approach can be used well if the prior knowledge of the various facial features and the distance between edges and corners is known. Model-based approaches can be categorized into two-dimensional, three-dimensional, and geometry-based models. Elastic Bunch Graph Matching (EBGM) deals with the shape and texture of a face image proposed by Wiskott et al. [5]. The two-dimensional morphable face model has been proposed by Cootes et al. [53] to understand the variations in face architecture. A geometry-based active shape model utilizes the local image to correct the shape of the features. The ASM model comprises three segments: contour, deformable templates, and point distribution model. The contour

was first introduced in 1987 by Kass et al. [74] to represent an object's outer shape (i.e., boundary line). The deformable template is proposed by Yuille et al. [179] that contains elementary information about the face. The point distribution model (PDM) represents the valid interpretations of face features, which encounter non-rigid features such as lip tracking. The three-dimensional morphable model is used in the analysis of 3D shapes of the human face.

2.4 Hybrid-based approaches

The hybrid approach [43, 48, 60, 142, 167] is a random combination of more than one technique to effectively enhance system performance. This method is complex and difficult to implement in practice because it processes multiple techniques simultaneously. Turk and Pentland [162] were the first to propose a face recognition technique based on the approximation of a face to distance map via global minima. The weights of the eigenvalues are represented through the discriminating features for each face sample. Distance from the free space (DFFS) approximates the face area.

3 Identity threats in face recognition

The human face is the interpersonal identity of individuals in communication [106]. Identity is a term related to socio psychological theories, applied to the conceptualization and analysis of the human face [152]. A person's identity can be analyzed and identified through their looks (i.e., face), qualities, and expressions. The face identity threat is a possible danger that might exploit a vulnerability to breach security and cause potential harm. The threat of such identity loss makes the task of facial recognition quite formidable. Numerous algorithms have been proposed or optimized by the Biometric Research Group [66] for real-time face recognition.

Nevertheless, the vulnerability and challenges in the face recognition systems have received scant attention in the literature, which this article sums up. The absence of such analysis in the literature made the intention of this survey be empirically focused, which this article efficiently accomplishes. After carefully perusing various reviews and research papers, the authors found that there could be two main categories for face identity threats: intentional and unintentional. In the intentional threat, the real person's identity is spoofed by fraudsters to fool the verification device and perform certain illegal activities. In contrast, the original identity carries out the unintentional threat with or without the intent to deceive the verification system. Figure 4 represents the proposed taxonomy for potential face identity threats.

3.1 Intentional threats (spoofing identity used)

Intentional threats are activities carried out by an imposter to deliberately enforce the identity of a real person (i.e., impersonation) to mislead the verification systems. Here, the unauthorized person uses their own identity to commit such fraud. More specifically, identity spoofing refers to gaining someone else's privilege unlawfully to access an authorized person's face rights through any alternative means. This category can be considered in two types: the first is direct spoofing, and the second is indirect spoofing.

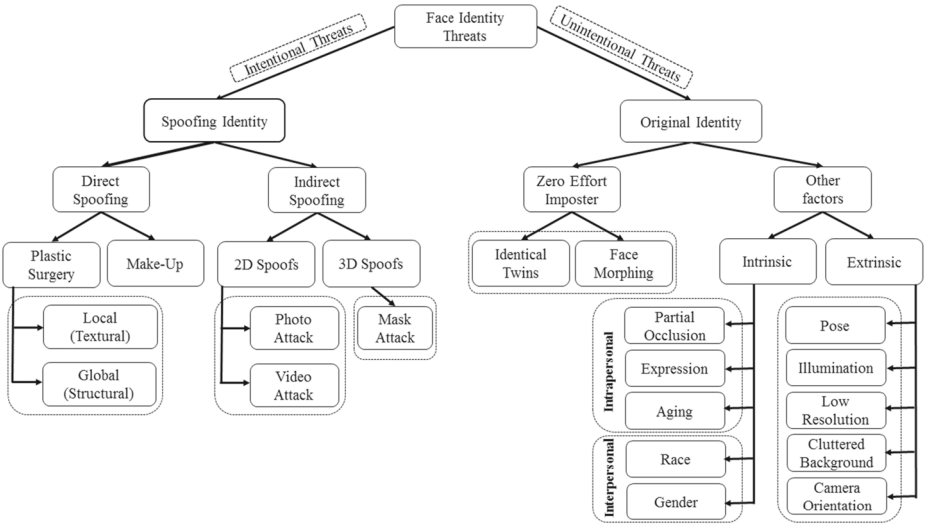

Fig. 4 Proposed hierarchy for potential face identity threats

3.1.1 Direct spoofing

Direct spoofing [77, 83] involves producing synthetic or manipulated identities in place of actual ones. This artificial or spurious identity is the cause of various artifacts that can modify the appearance of the face. More specifically, direct spoofing refers to the situation where one contaminant tries to personate the identity of others through various synthetic or medical treatment practices. These synthetic or medical treatment processes can be temporary, long-lasting, and permanent. However, concealing a person's identity (impersonation) implies that the imposter wants to execute the illegal activity, particularly where authentication is required. Table 3 of Section 5 compares the state-of-the-art countermeasures techniques, focusing on methodology, the datasets, and the technique's performance. Direct spoofing can be of two types based on permanent and temporal artifacts, i.e., plastic surgery and the makeup. Figure 5 depicts the scenarios for analyzing the impact of plastic surgery and makeup.

(i) Plastic surgery Plastic surgery [90] is a part of medical science used to restore and repair a person's facial identity. These plastic surgeries are performed to improve the appearance of specific facial features that are damaged due to an accident, birth disorder, some disease infections, aging effects, burn, and any other facial feature's discrepancies. Major categories of facial feature surgery include: Nasal surgery (rhinoplasty), face enhancement surgery (rhytidectomy), eyelid surgery (blepharoplasty), chin surgery (genioplasty), lip augmented surgery (liposuction), brow lift surgery, and cheekbones implants surgery. A person with a firm intention of committing fraud and evading law enforcement by impersonating another identity can benefit from such plastic surgery without the fear of being recognized. Nowadays, plastic surgery is more common in practice due to affordability and modernization, leading to a drop in the performance of the face recognition system. Figure 5 (a) depicts the plastic surgery scenario. Plastic surgery can be divided into local (textural) surgery and global (structural) surgery.

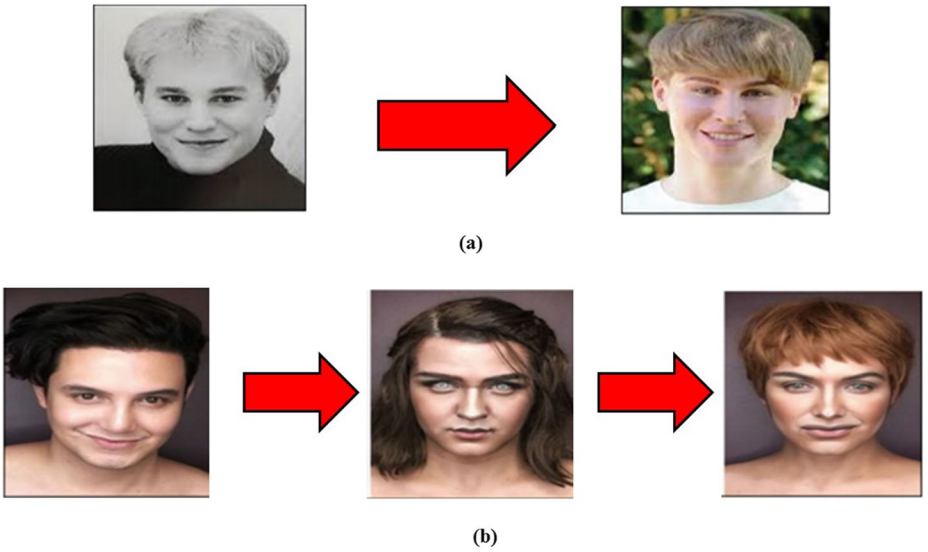

Fig. 5 Some scenarios of direct spoofing [100, 131](a) Plastic surgery, (b) Makeup attack

Local (textural) plastic surgery [35] is a minor surgery used to correct certain anomalies of facial appearance. Localized texture features may be the forehead, eyelids, nose, cheek, chin, and jawline. A small change in any of these imperative face features causes a definite change in the geometric distance of neighboring features that can result in a reasonably distinct or unseen face. Similarly, removing blemishes, warts, moles, and acne from the face can eventually confuse verification tools. Local plastic surgery is often misused by criminals to elude law enforcement and security controls. This local (textural) plastic surgery is sometimes referred to as cosmetic surgery.

Global (structural)) Plastic surgery [155] is used to reconstruct functional damage and disorders (defects) in the structure of facial features. These deforming defects include facial cuts and burns operated by a team of specialist surgeons. In this irreversible process of global surgery, facial identity (geometric structure) is modified for a long life that can never be regained. Global plastic surgery-based problems may allow an imposter to proceed without fear of being identified by verification tools.

(ii) Makeup Makeup [119] can be considered a second category of direct spoofing (i.e., temporal). Makeup is the process of altering a person's appearance (i.e., looks) for a period of time. This transformation process can substantially alter the appearance of the face, leading to the failure of the face recognition system, especially in the case of older people. There are various makeup patterns available in the real world, such as light makeup and heavy makeup depending on an individual's purpose, durability, and skin tone. Figure 5 illustrates the scenarios for analyzing the impact of plastic surgery and makeup.

The keyword Light makeup does not represent any quantitative information about the beauty product applied to the face. Instead, it refers to qualitative measures. The difference in the face's texture before and after applying makeup is called the notion of light makeup. A small or negligible difference in facial texture before and after makeup may not be easily

recognized as it coincides with the skin tone, which enhances some of the facial ingredients without any exposure. Unlike light makeup, heavy makeup is visible on the face (see Fig. 5(b)) due to a mismatched complexion with the material applied on the facial skin. Heavy makeup can aesthetically affect perceptual changes in facial appearance through widely accented eyes and dark lips, resulting in a low detection rate for the authentic look. The importer uses certain facial grooming products to create complexity and impediments for the facial recognition system.

In order to mislead the verification system, the duplicate actors used disguised faces to pose as original actors. The imposters may use additional props such as hats, rings, and more to maintain similarity with the real person. This type of threat is temporary. However, it can cause considerable damage to the face recognition system.

3.1.2 Indirect spoofing

Indirect spoofing (presentation attack) [61] does not directly modify real face identity, although artifacts such as digital photos, printed photos, facial videos, and three-dimensional masks that mimic the original identity are used to fool the device. Indirect spoofing can be of two types: two-dimensional and three-dimensional. A comparative analysis of the state-of-the-art techniques that reduce the impact of indirect spoofing is shown in Table 4 of Section 5. Figure 6 represents some real-world scenarios involving indirect spoofing attacks.

(i) 2D Spoofs In 2D spoofing [116], the imposter uses the two-dimensional props form such as printed photos, mobile photos, paper photos, and videos to gain illegal access to the system. The printed photo, scanned photo, and replayed video scenarios for the indirect spoofing attack are shown in Fig. 6. These digital photos and videos are easily accessible to fraudsters through social media such as Facebook, WhatsApp and others. The imposter displays these artifacts before the verification tool to gain illegal access to the real identity.

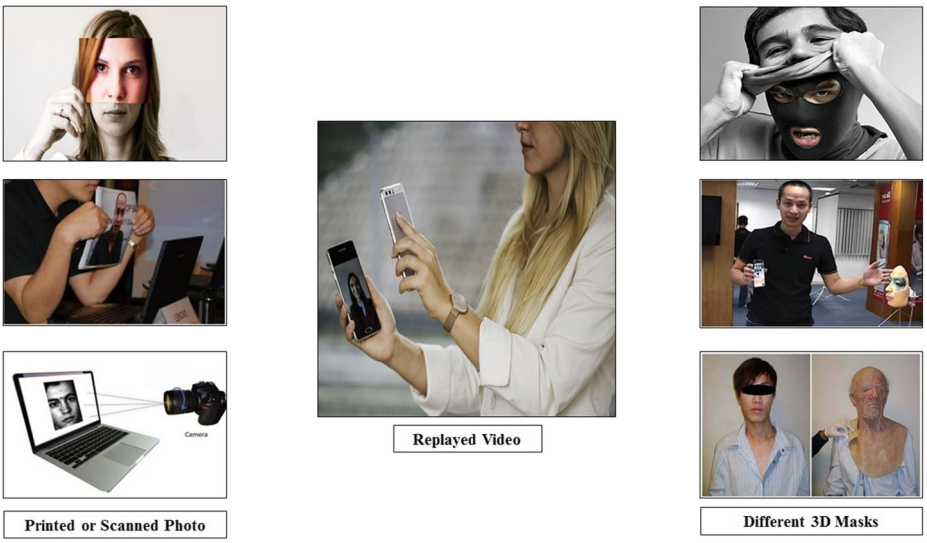

Fig. 6 Scenarios of indirect spoofing [68, 100]

In practice, two-dimensional spoofing is mostly adopted by imposters due to the simple, inexpensive process.

(ii) 3D Spoofs Three-dimensional spoofing [68] is an act of circumventing a facial recognition system by using synthetic 3D face masks to mimic the identity of a genuine person at the time of verification, as depicted in Fig. 6 for various 3D masks. However, this attempt requires technical prowess and high-quality synthetic material-based 3D masks that are too expensive for nefarious users to portray a real person and deceive someone. Nowadays, bank robberies and other big scams are carried out on the basis of similar three-dimensional spoofing that can breach the authenticity of face recognition to large extent

3.2 Unintentional threats (original identity used)

The act of misleading the face recognition system without any prejudicial intent using the original identity is known as an unintentional threat. Sometimes, the face verification system fails to recognize the identity even after considering the actual enrolment (i.e., the registered user) due to various factors such as different poses, expressions, partial occlusions, illuminations, and many more. However, this threat also supports cases where a legitimate user is involved under an agreement to bypass the verification device, such as identical twins and face morphing. The unintentional threats can be divided into zero effort imposter and, intrinsic and extrinsic factors. These categories are briefly discussed sequentially in subsequent sub-sections.

3.2.1 Zero effort imposter

As the name suggests, zero effort imposter [136] is an act to deceive the system with zero effort. Sometimes, passive imposter benefits from the limitation of the biological and genetic structure of the human face consisting of similarities in facial features. This threat inadvertently allows fake identities to access the resource due to extensive resemblance with the original biometric captured at the time of enrolment. Two types of scenarios caused these threats: identical twins and face morphing. State-of-the-art approaches addressing identical twins and morphing threats are analyzed comparatively in Table 5 of Section 5.

(i) Identical twins Identical twins [105] are siblings who have the same date of birth. The term identical refers to the uniqueness in both people's appearance and behavioral characteristics due to similar environmental conditions during the birth process. The passive imposter (i.e., zero effort fraudsters) takes the first person's face-biometric from the twins at the time of enrolment and accomplishes the verification through the second person's face-biometric from the twins by taking the benefit of the doubt as both the persons have a facial similarity. This act is also known as interchange the identity.

(ii) Face morphing Face Morphing [137] is a severe security breach for the face recognition system. Face morphing is a technique for reproducing a new face identity by combining two or more face identities with an almost identical facial appearance to facilitate illegal access. More specifically, the facial identity of two similar-looking individuals is merged in the ratio that each identity can access the resource that contributed at the time of enrolment (i.e., registration). Face morphing can be performed using the benefit of the doubt considering the minimum confidence (i.e., threshold) value. However, if the threshold is set extremely high and the merged identities do not have similar facial features, the verification system may

fail to recognize the individual or be rejected. Figure 7 shows the scenario of combining multiple identities and authenticating both to gain illegal access (i.e., false positive). In the recent past, many researchers in the biometric domain have shown their interest in this sparking field, i.e., face morphing detection.

3.2.2 Other factors

The face recognition literature confirms that verification systems may be inadequate or ineffective even after the actual identity is not spoofed (impersonated). Factors responsible for the decline in recognition rate without being impersonated fall under this threat, which can be further categorized into two levels: intrinsic factors and extrinsic factors. Intrinsic factors include scenarios where physical information of a face is lost or is insufficient to classify it as a face. The intrinsic factor includes two subcategories: intrapersonal and interpersonal. Figures 8(a) and 8(b) depict intrapersonal intrinsic factors for partial occlusion of the face and different facial expression scenarios, respectively. However, the interpersonal scenarios-related samples are not represented here due to unavailability of benchmark datasets for this category. Extrinsic factors are responsible for the alteration in facial appearance due to variations in face poses, illumination (lighting effects), and camera viewpoints. The low resolution and cluttered background are also significant factors that fall under this category. Figure 8 (c) shows the extrinsic factor-based threats scenarios for different poses. Subsequent subsections provide a concise detail for each of these threats in sequential order.

(i) Intrinsic factors Intrinsic factors are responsible for modifying the human face's physical characteristics resulting in insufficient information to represent it as a face accurately. A comparative analysis of state-of-the-art countermeasure techniques capable of identifying such threats and providing appropriate solutions is shown in Table 6 of Section 5.

Intrapersonal identity threat is a condition where a person's facial identity is modified to such an extent that it cannot be easily recognized. For instance, facial information is not sufficient to identify the person in partial occlusion due to certain artifacts such as masks and own body parts, which hide essential facial information. Other scenarios for intrapersonal identity threat may be different facial expressions and aging factors that may impair actual facial location.

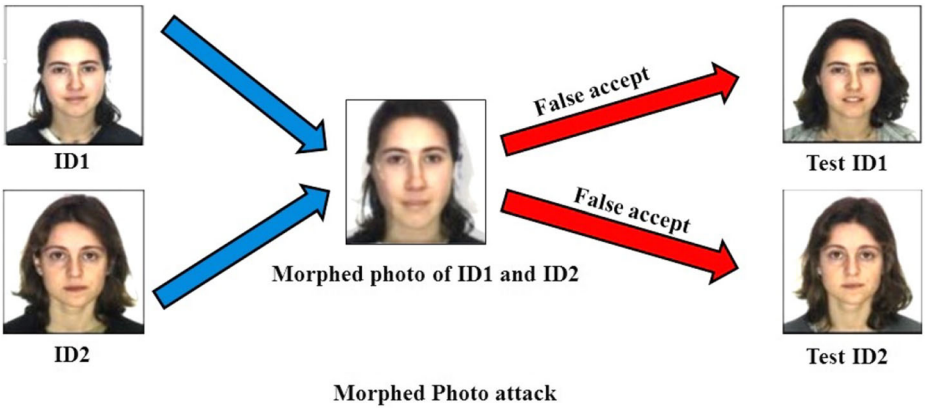

Fig. 7 Process of attempting morphed face attack [100, 135]

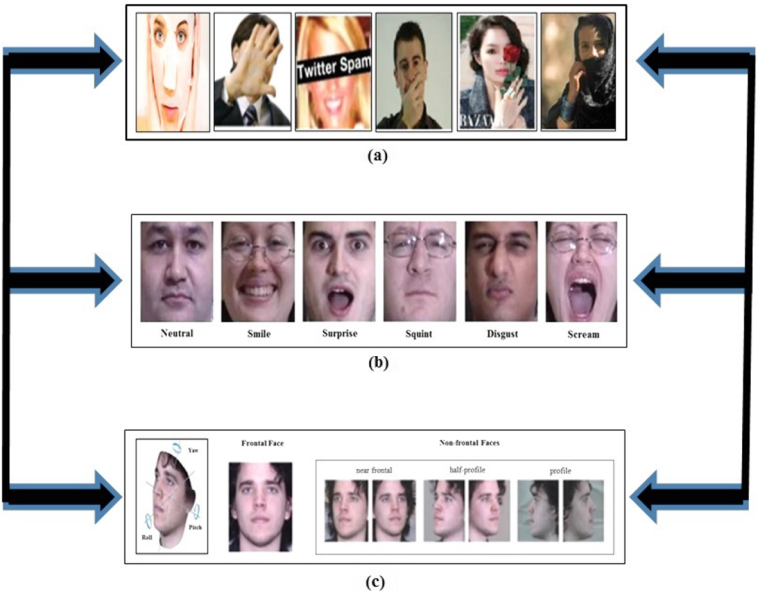

Fig. 8 Other scenarios of face identity threats [41, 67, 86, 189] (a) and (b) Intrinsic factors (c) Extrinsic factor

– **Partial occlusion:**
Partial occlusion [143] plays a vital role in deteriorating performance in face recognition systems under unconstrained conditions. Partial occlusion includes all scenarios where some part of the face is hide using specific object, such as sunglasses, COVID-19 face masks, paper strips, tattoos, scarves, and own body parts as shown in Fig. 8 (a)). The face recognition system works well when all facial features such as eyes, nose, mouth, and jawline are visible and positioned in frontal view in the input image. The de-occlusion process poses a significant challenge to preserve the actual identity of the individual, which may reduce the facial recognition rate. Partial occlusion of the face consisting of two props (i.e., scarf and sunglasses) has attracted the intense attention of researchers and has been extensively investigated. However, we also provide some insights for face masks-based occlusion, which received exceptional attention from the Biometric Research Committee during the recent COVID-19 pandemic in state-of-the-art research.
– **Expression:**
Facial expressions [99, 128] are symbolic notations of a human's emotion. The facial expression reflects the biological and neurological activity in the brain at that time, represented through various facial muscles' moments. However, these facial expressions could deteriorate the natural geometry of the face, resulting in poor recognition of the face. Face recognition process is solely based on facial features with a specific structure. A small change in facial features or muscle contraction makes a significant difference in the recognition process. Figure 8(b)) depicts six facial expressions such as neutral, smile, surprise, squint, disgust, and scream.
– **Aging:**
Aging [175] refers to a gradual transformation in facial features over time, i.e., from an early age to the onset of adulthood. These facial transformations are dependent on

various factors such as skin tissues (fatty or thin), bones, lifestyle, social and economic status, diet, and disease. The gradual changes in facial skin can be measured through wrinkles, freckles, and sagging in different people and groups based on age, sex, and race. The demand for accurate estimation and prediction of a person's age is increasing daily, especially for real-world applications such as dermatology, cosmetics, missing child investigation, and many more.

Interpersonal identity threat is a case where face appearance is analyzed among different peoples or groups based on their gender and race.

- **Gender:**
  Gender Identification [126] through facial features is a significant challenge to face recognition systems. Gender recognition may be an easy task for humans, but predicting a person's gender electronically or using software to face images poses a severe concern for face verification systems.
- **Race:**
  The term human race [78] classifies different groups of people based on their inherited behavioral and biogenetic (i.e., physical) differences. Discriminating parameters can be skin color, hair color, facial features, and eye formation based on the physical and psychological characteristics of the human. Races have distinctive characteristics of people from different continental aggregates (i.e., geographical locations) that reflect distinct attitudes and beliefs with various cultural interventions.

(ii) Extrinsic factors Extrinsic factors include non-cooperative subjects and unconstrained conditions such as pose variations, low resolution, poor illuminations, cluttered background (noise), and different camera orientations resulting in poor recognition rate due to alteration in the facial appearance. A comparative analysis of state-of-the-art countermeasure techniques capable of identifying such threats and providing appropriate solutions is shown in Table 7 of Section 5.

- **Pose:**
  A pose containing a non-frontal face causes a severe threat to preserving a person's true identity, significantly reducing facial recognition rates. Profile images or half-profile images are occasionally provided as test samples to match them with enrolled frontal face images, failing identification. State-of-the-art research discusses three possible variations in face poses, namely yaw (horizontal), pitch (vertical), and rolls (longitudinal or angular displacement) (see Fig. 8 (c)).
- **Illumination:** The extrinsic factor that significantly affects the face recognition system is the variation in the illumination. Several reasons can be responsible for this threat, such as samples captured from far away, background lighting effects from another lighting device, day-night vision impact, shadow, and many more.
- **Low resolution and noise:**
  Low-resolution is a significant challenge for face recognition systems, particularly when mapping low-resolution and high-resolution images. Resolution represents the total number of pixels concerning the width and height of an image. A small image has few pixels, which are sometimes insufficient to discriminate important features resulting in a poor recognition rate. Frames (i.e., images) captured with surveillance CCTV cameras that are too far from the object may capture blurry, noisy, or poor quality images resulting in failing recognition.

Noise can be added at any stage that causes considerable damage to the pixels concerned. Several image pre-processing operations such as enhancement and filtering can be applied to segregate the noise from the test samples.

– **Cluttered background:**
The cluttered and complex background constitutes a significant threat in face recognition. The face recognition systems are best suited for ideal (i.e., constant) background-related problems. However, the system's performance can inevitably decrease under the interference of cluttered background scenes such as motion, color, texture, overlapping of objects, and environmental conditions (i.e., dust, haze, cloud, and rain). The researchers preferred highly constrained and strongly perception-based simulated environments for face recognition tasks in earlier times. However, real-time scenarios have not been adequately addressed in recent past research. Conventional face recognition techniques such as contour-based and receptive field-based techniques are not convenient for effective facial recognition because the contour includes additional pixels, while the receptive field has standard dimensions that may omit some significant facial features. These extra pixels in the background cause a significant disturbance for facial recognition system.
– **Camera orientation:**
Camera position greatly influences face recognition and surveillance system [147]. Some of the significant adjustment features that may affect the performance of the face recognition system can be divided into specification-based and physical-based adjustments. Specification-based adjustments include the configuration of the camera, such as the model, lens quality, shutter speed, aperture, and resolution for capturing the images. The physical adjustments include the camera's location, the distance of an object from the camera, alignment (i.e., direction), height, and other factors. All the mentioned parameters including camera's position should be effectively maintained. Deployment of multiple cameras at a predetermined location and aggregation of multiple observations with different functions can solve this problem. However, this approach is not feasible for time, maintenance cost, and unpredictable environmental concerns. An additional light projection-based method can be used for the transformation invariant adversarial pattern generation.

4 Countermeasure techniques against face identity threats

The face biometric-based recognition system has attracted the attention of researchers, although it still suffers from various face identity threats to real-world problems. Several state-of-the-art countermeasures (i.e., anti-spoofing) techniques have been investigated to mitigate the impact of face identity threats over the last decade. This paper briefly introduces the basic functioning of some robust anti-spoofing techniques based on dimensionality reduction, feature extraction, classification, and neural network. A comparative study of various countermeasure techniques, including their characteristics, purpose, advantages, and disadvantages, is described in Table 2.

4.1 Dimensionality reduction techniques

Dimensionality reduction is the most popular method in machine learning. This technique reduces the feature dimensions (i.e., the number of independent variables) by considering only those essential variables that are accountable for discriminating the features. Given the

Table 2 Comparative analysis of countermeasure techniques

Techniques	Purpose	Characteristics	Applicability	Advantages	Disadvantages
PCA [19, 27, 32, 34, 88, 149]	DR , FE	Unsupervised, maximize inter-class distance	PA, textural/structural PS, POcc	Noise reduction	Poor result for large dataset, outliers, higher variance
LDA [15, 19, 49, 65, 164, 186]	DR , FE	Supervised, maximize inter-class, minimize intra-class distance	PS, MU, PA, and ill	Reliable and efficient method	Assumption based method that can affect the results
SVD [62, 70, 188]	DR	Matrix decomposition, pattern-based solution, generalization of Eigen face	ill, CB	Optimized information using few coefficients	Slow, expensive singularity problem
DCV [4]	DR , CLS	Variation of Fisher's LDA with small sample size	POcc, FS, Exp	Reduces singularity and small sample size problem	Handling of large matrices, complex method
Kernel PCA [170]	DR , FE	Transformation of non-linear patterns into linearly separable high-dimensional space	FR, handwriting recognition	Dealing with non-linear distribution-based unconstrained problem	Longer computation time, over fitting issue
kernel LDA [180]	DR	Non-parametric method, allows efficient computation of Fisher discriminant	FR, Exp	No assumption required for input distribution	Small sample size problem

Table 2 (continued)

Techniques	Purpose	Characteristics	Applicability	Advantages	Disadvantages
DCT [6, 111, 117, 120, 174]	DR , CLS	Transformation-based holistic method used to represent the sum of sinusoids for different magnitudes and frequencies	FMo, textural and structural PS, ill	Fast, provides constant matrix, preserve energy	Quantization is required
LBP [29, 40, 46, 84, 96, 123, 176]	FE	Image texture-based analysis through spatial structure, mathematically proven	FSD, Ag, G	Robust, efficient for illu-t mination, time, cos	Large false positive
HOGs [5, 31, 159, 186] [33]	FDe	Two main parameters, i.e., gradients direction and its magnitude	FSD, FMo, G, and illumination variations	Robust to variable lighting conditions	High dimensional feature space, cost, large datasets
SIFT [8, 9, 71, 142, 171]	FDe	Local features detection	PS, IT, POcc, LR, CB	Transformation invariance (S, Ro), efficient for Omni-directional	The complexity and run time
SURF [19, 60, 113, 144]	FDe	Extracts salient features (S, Orientation, ill)	POcc, LR	Eliminate the undesired motion found in videos, higher efficiency	Illumination variations issue

Table 2 (continued)

Techniques	Purpose	Characteristics	Applicability	Advantages	Disadvantages
Gabor Wavelet [8, 26, 34, 56, 82]	FE, MM	Biological inspired features, scale and orientation based features	FSD, IT, Exp	Invariance to shift, rotate and illumination change	Large memory, cost, and higher dimensionality issue
Viola- Jones [13, 25, 71, 108, 129, 149]	FE , CLS	Robust and generalized technique for face recognition	LR, G	large features, fast, best for low-resolution images	Frontal face images required, sensitive to lighting conditions
Skin Color Modelling [76, 79, 146]	FE (Low-level)	Y parameter in YIQ, YUV and YCbCr shows the luminance, and other two for chrominance. Hue, saturation, and intensity contain the color depth, purity, and brightness, respectively	FD, MU, POcc	Depth color information, fast processing in controlled environment	Not suitable for unconstrained condition, performance is dependent on the color-model used
SVM [10, 13, 20, 25, 27, 51, 135, 138, 151, 184]	CLS	Multi-class classifier, support vectors, decision boundary, and kernel discriminative classifier	Structural PS, FSD, FMo, POcc, and Exp	Handle noise, less chance of over fitting, real valued features	High computational cost

Table 2 (continued)

Techniques	Purpose	Characteristics	Applicability	Advantages	Disadvantages
K-NN algorithm [4, 15, 40, 176]	CLS	An alternate of SVM, unsupervised clustering-based	FMo, Exp	Suitable to find out the loss/error estimation	Not fit for large dataset, long process time
HMM [112, 132]	CLS	A generative classifier focused on sequence of symbol emitted by system underlying random walk between states	Pattern recognition, classification, and structure analysis	Strong statistical foundation	Not suitable for higher order correlation
SLNN [2]	FE , CLS	Human brain oriented feed forward neural network consist of two layer architecture	POcc	Easy setup and less computation	Separable data is desired, cannot deal complex nonlinear problems efficiently
MLNN [119, 156]	Automatic FE , CLS	At least one hidden layer is required including input and output layer	FSD, Exp, medical diagnosis	Easily tackle complex problem	Heavy computation, large space, long time

Table 2 (continued)

Techniques	Purpose	Characteristics	Applicability	Advantages	Disadvantages
CNNs [52, 123, 124, 154, 167, 177, 178, 183, 184]	Image-based FE , CLS	A deep learning technique, which takes image data as input	FSD, Po, Exp, Occ, LR	Supports transfer learning by sharing the pre-trained weights, fast	Layers interpretations is not clear, complicated hidden layer mechanism
Euclidean-DMC [9, 13, 45, 112]	CLS , LF	Distance between two data sample (lets p and q) for n-dimensional feature space	PS, G, LR, Occ	An effective method to find uniqueness	Assumes in prior for misplacing of data points
Manhattan- DMC [82]	CLS , LF	Distance between two data samples measured along the axes at right angle	FR, video surveillance, Crime monitoring, Occ	This method has robustness to outliers	Generates large value for two similar images that represents the dissimilarity
Chi-Square- DMC [176]	CLS	Investigates the difference between what actually find in study (observed), and what is expected to find (hypothesis)	Histogram matching, Exp	Suitable for comparing different histograms, easy computation and interpretation	Requires data is in numeric form to deal with higher degree of categories
Cosine Similarity-DMC [45, 111, 117]	CLS, SE	Similarity index between two different vectors, Cosine angular represents the product of two vectors with direction	PS, FR and camera orientation	Good accuracy, vectors are used to measure similarity, direction and angular displacement	If two vectors lies on the same line than the cosine value will be 1, and the similarity value will be 0

DR-Dimesionality Reduction, FE- Feature Extraction, FDe- Feature Descriptor, FR- Face Recognition, FD- Face Detection, FS- Face Spoofing Detection, CLS-Classification, PA- Presentation Attack, PS- Plastic Surgery, MU- Makeup, MM-Multiple Modularity, LF- Loss Function, IT- Identical Twins, FMo- Face Morphing, POcc-Partial Occlusion, Exp- Expression, Ag-Aging, R- Race, G- Gender, Po- Pose, ill- Illumination variations, LR- Low Resolution, CB- Cluttered Background, CO-Camera Orientation, S-Scaling, Ro-Rotation, DMC-Distance Metric Classifier, SE- Similarity Evaluation

scope of this paper, some of the most prominent methods of dimensionality reduction are discussed here.

Principal component analysis [19, 27, 32, 34, 88, 149] is an unsupervised, non-parametric, and statistical analysis-based algorithm that was first proposed by Sirovich and Kirby [102] in the 1980s. is used for various tasks such as dimensionality reduction, feature extraction, and classification for given distributions. For instance, if we have m independent variables, PCA uses only the p variable, where p≤m is used for feature extraction. PCA reduces the data from n-dimension to k-dimension and computes the covariance matrix as shown in (1).

$$\sum = \frac{1}{M}\sum_{i=1}^{n}(x^i)(x^i)^T \tag{1}$$

The PCA method is quite helpful for significant variance and noisy data-based problems. The covariance problem can be easily handled by two-dimensional principal component analysis.

Linear discriminating analysis [15, 19, 49, 65, 164, 186] is a supervised learning-based algorithm. LDA is used to minimize the dimensionality of the feature space, maximize the two-class distance, and at the same time minimize the intra-class variance of the class. The LDA process is consists of three steps: the first is to calculate the distance value by the mean of different classes, also termed class variance as shown in (2).

$$s_b = \sum_{i=1}^{n} N_i \cdot (\bar{x_l} - x_i) \cdot (\bar{x_l} - x_i)^T \tag{2}$$

The second step determines the distance between a sample and the mean value, known as intra-class variance as represented in (3).

$$s_w = \sum_{i=1}^{n}(N_i - 1) \cdot s_t = \sum_{i=1}^{n}\sum_{j=1}^{N_i}(x_{i,j} - \bar{x_l}) \cdot (x_{i,j} - \bar{x_l})^T \tag{3}$$

The third step is to apply dimensionality reduction, where an intraclass variance is minimized, and the interclass variance is maximized. To do this, the fisher's criteria P must satisfy the following (4):

$$P_{lda} = \arg\max_{P}\left|\frac{P^T \cdot s_b \cdot P}{P^T \cdot s_w \cdot P}\right| \tag{4}$$

Fisher Face LDA is an extension of LDA, which provides robust solutions for identifying changes in illumination and various emotions.

Singular value decomposition [62, 70, 188] is an effective method for facial recognition related problems including low illumination and cluttered backgrounds scenarios. SVD can be deployed to extract the covariance matrix based on the statistical model to analyze data efficiently. The eigenvalues and eigenvectors can be easily calculated through matrix diagonalization as shown in (5).

$$A_{m\times n} = U_{m\times m} S_{m\times n} V^T_{n\times n} \tag{5}$$

Where,A is any matrix of order $m\times n$, U is the square matrix (i.e., column vector), S is a diagonal matrix of order $m\times n$, and V^T can only exist, if the eigenvectors are linearly independent. Now the two singular vectors, U and V are arbitarily taken for orthogonal symmetric and square matrix such as AA^T and A^TA. Both these matrices have the same rank, say R and eigenvalues, either zero or positive. U and V must satisfy the following

condition of the (6).

$$U^T U = V^T V = I(IdentityMatrix) \tag{6}$$

Discriminative common vector [4] extracts critical information from training samples of different classes, known as common vectors. The DCV process involves three steps: the first is to calculate the null spaces in the interclass matrix. The second is to assign the sample from each class. The third is to calculate the class variance.

Kernel PCA [170] is an unsupervised technique that involves a kernel trick to represent the non-linear geometry of face images. The kernel PCA includes the following steps for mapping the high-dimensional feature spaces to input (i.e., face images).

- Calculate dot product of the image matrix with the help of kernel function.
- Find the eigenvector of subspace for the covariance matrix and then apply normalization.
- Now, check the test point on the eigenvector using the kernel function.

Sometimes two-dimensional PCA poses a multi-dimensionality problem for estimating the covariance matrix. Therefore, a kernel trick (i.e., K2DPCA) is utilized to eradicate the non-linearity by transforming the high-dimensional data into low-dimensional feature space. Liu et al. [91] proposed a weighted kernel PCA method that provides promising results to recognize facial features.

Kernel LDA [180] is a non-linear expansion of the LDA method used for dimensionality reduction. The kernel is used to represent the non-linear distribution of features.

Discrete cosine transformation [6, 111, 117, 120, 174] is a transformation-based holistic method representing the sum of two-dimensional sinusoids for the given images consisting of different magnitudes and frequencies. The DCT technique is primarily utilized for analysis of textural and structural plastic surgery, face morphing, illumination, camera orientation, and image compression. The JPEG lossy compression is the best example based on DCT techniques. The working of DCT is similar to discrete Fourier transformation as represented in (7).

$$B_{p,q} = \alpha_p \cdot \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{m,n} \cdot \cos\left(\frac{\pi(2m+1)p}{2M}\right) \cdot \cos\left(\frac{\pi(2n+1)q}{2N}\right) \tag{7}$$

where, $M \times N$ is the input image, $A_{m,n}$ is the intensity of the pixel in row m and column n. (7) must satisfy the condition $0 \leq p \leq M-1$ and $0 \leq q \leq N-1$ for all inputs. α_p and α_q are the normalization factors as represented in (8) and (9), respectively.

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}} & for, p = 0 \\ \sqrt{\frac{2}{M}} & otherwise, 1 \leq p \leq M-1 \end{cases} \tag{8}$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}} & for, q = 0 \\ \sqrt{\frac{2}{N}} & otherwise, 1 \leq q \leq N-1 \end{cases} \tag{9}$$

The B_{pq} is known as the DCT coefficient.

4.2 Feature extraction technique

Feature extraction is an essential face recognition process, where the discriminating features are extracted from the input data and mapped with the stored data. Large-scale research work has been done in the field of features extraction. This survey paper discusses

robust state-of-the-art feature extraction techniques including most sophisticated feature descriptors along with a tabular comparison of their advantages and disadvantages for face recognition problems (see Table 2).

Local Binary Pattern [21, 29, 40, 46, 84, 96, 123, 176] is a robust and extensively used feature extraction method, especially in face recognition. Different variants of LBP methods are proposed and optimized to address various facial identity threats in state-of-the-art research. Figure 9 demonstrates the LBP's process for a grayscale image.

LBP is computationally efficient concerning the time and cost, which makes it most interesting and useful. However, the illumination effect for a simple kind of LBP is monotonic.

Histograms of oriented gradients (HOGs) [5, 31, 159, 186] were proposed by Lowe in 2004 [33]. HOG is a widely used feature descriptor for object detection in the field of computer vision. This method represents the histogram for several oriented gradients. A histogram is the distribution of intensity over an image, whereas a gradient or slope is a kernel used to detect edges and other discriminable features of an object or a face. The term oriented refers to the motion of gradients along with directions. Generally, HOG calculates the two main parameters, i.e., gradients direction and magnitude. The term magnitude refers to the abrupt change in intensity. The HOG computation includes some processing steps as follows:

– HOG can be applied to any image containing a fixed aspect ratio. However, the magnitude and orientation of the horizontal and vertical gradient can be calculated using (10) and (11):

$$Magnitude\ G_{(x,y)} = \sqrt[2]{\left(G_x^2 + G_y^2\right)} \tag{10}$$

$$\theta_{(x,y)} = \arctan\left(\frac{G_y}{G_x}\right) \tag{11}$$

Where, x and y represents the change in intensity concerning vertical and horizontal direction, respectively.

– HOG shows promising results for 8×8 image size, thus a gradient for RGB channel window with two parameters (i.e., magnitude and direction) are considered here to simplify the calculation. Thus, $8\times8\times3 = 192$ will be the total pixel value, and the total number of bins will be $8\times8\times2 = 128$.

– HOG is calculated as a one-dimensional vector with nine bins (i.e., elements) with a range of 0 to 160. The contribution of each bins' response can be plotted on the

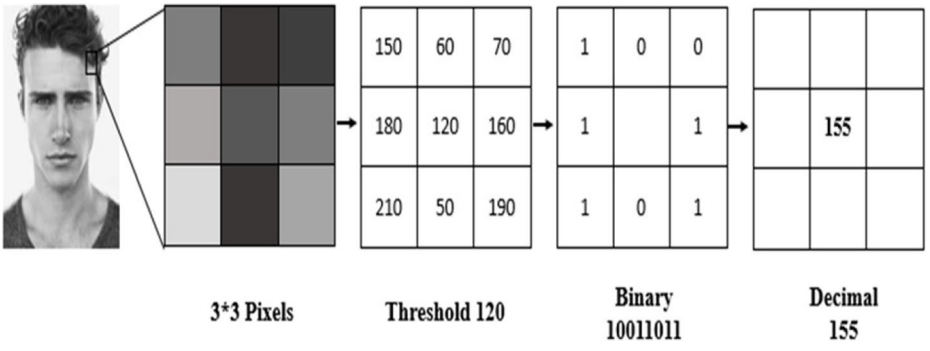

Fig. 9 LBP operation

histogram graph, and finally the gradient can be normalized by applying the L2-norm to reduce the lighting effect.

Scale-invariant feature transform [8, 9, 71, 142, 171] is a local feature descriptor technique for image-based problems. The SIFT model showed efficient processing of omni-directional images with different transformation invariance such as scaling, rotation, and more. Figure 10 depicts the logic to separate the key-point descriptors from image gradients.

Speed up robust features [19, 60, 113, 144] is a fast and efficient feature descriptor for extracting salient features with invariance to scale, orientation, and illumination changes for face detection. Sometimes SURF is useful for eliminating undesired motion found in videos using its features extracting capabilities.

Gabor wavelet transformation [8, 26, 34, 56, 82] includes a biologically inspired features extraction capability for face detection. The Gabor wavelet features consist of two parameters: scale and orientation. Gabor wavelet performs tremendously, especially when modularity is used. Sometimes wavelets can be used as an activation function for deep learning neural networks.

Voila jones [13, 25, 71, 108, 129, 149] is a simple, fast, and robust method proposed by Paul Viola & Michael Jones in 2001 for real-time face detection and recognition. The Viola-Jones method provides faster computation due to integral images and represents efficient results for low-resolution (i.e., small-sized) images. The following steps include in viola-Jones process as follows:

- **Haar-Like features** [129, 149] is a fast, robust, and efficient classifier consists of different grayscale patterns (approx. 162000 features) of black and white pixels that represent a high resemblance to human facial features. Figure 11 pictorially represents the most common Haar-like features: edge features, line features, center-surround features, and four-rectangular features.

  The haar classifier matches their features with the human face's grayscale shades near the eye, lip, and eyebrow to distinguish the face and a non-face class.
- **Integral image** [75] is used to calculate the shaded region of the detected face. Figure 12 shows a straightforward calculation of the region of interest (i.e., shaded pixels) for any image using the four associated values. (12) demonstrates the process of calculating any region from the image. Here, we consider the shaded region of Fig. 12

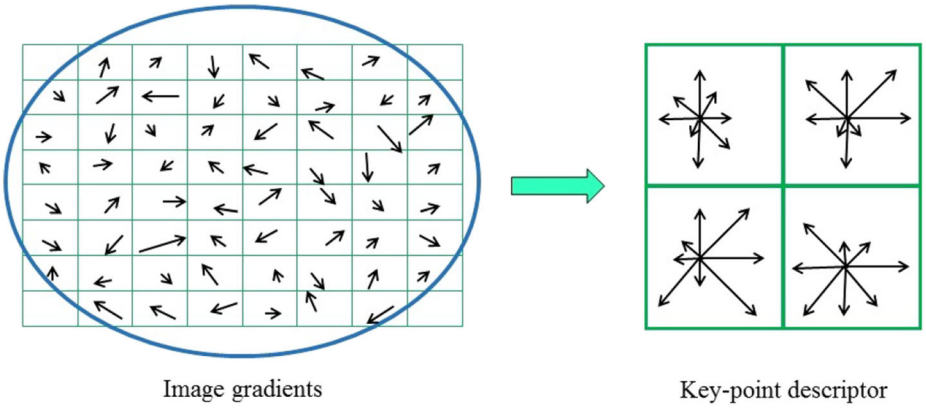

Fig. 10 Key-point descriptor calculation in scale invariant feature transform (SIFT) [47]

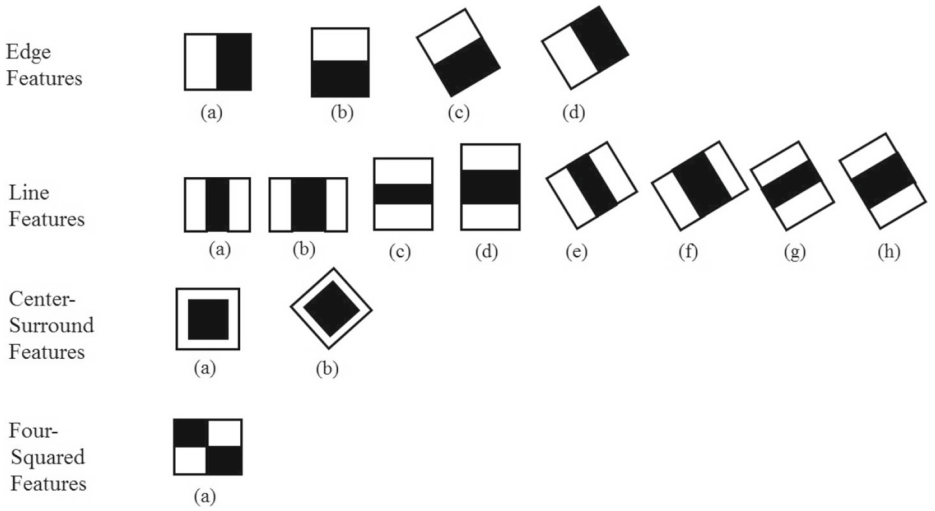

Fig. 11 Pictorial representation of most common Haar-like features

for better understanding the concept.

$$Sum = I(A) + I(D) - I(B) - I(C) \tag{12}$$

Where, A and B are the points from left to right at top position and C and D are the bottom ends points of the shaded cell regions. The sum denotes the total pixel values of the shaded area as shown in left side image.

– **Adaboosting** [25, 71] is used to train the Haar-like feature to obtain a robust learned model. Adaboost creates a strong classifier by combining multiple weak classifiers together consisting of specific facial features. The (13) represents the sum of feature maps consists of boosting coficient and training samples for all inputs.

$$F(x) = \alpha_1\ f(x_1) + \alpha_2\ f(x_2) + \alpha_3\ f(x_3) \ldots\ldots\ldots + \alpha_n\ f(x_n) \tag{13}$$

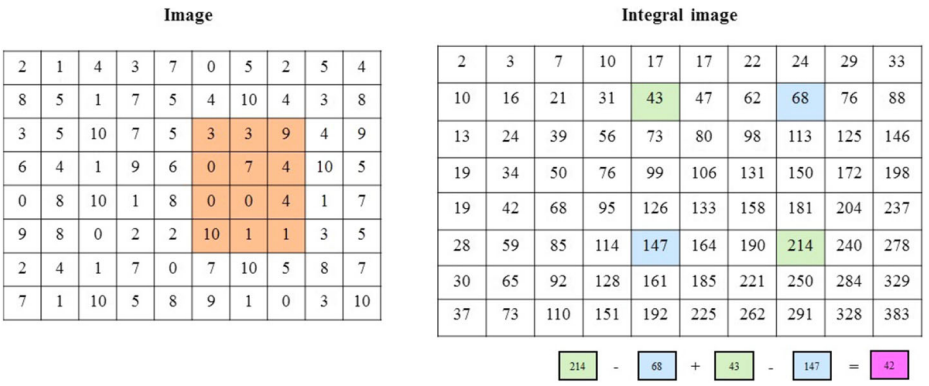

Fig. 12 Calculation to obtain the shaded region using integral image

The sum of all such weighted feature maps can be determined using (14).

$$F_t(x) = \sum_{t=1}^{T} f_t(x) \tag{14}$$

Where α_i, is the boosting factor and $f(x_i)$ is the training samples for i=1, 2, 3, ... n, and $F_t(x)$ is the summation of all such functions for t=1, 2, 3, ... $T-1, T$.

– **Cascading** is a process of separating strong classifier from other weak classifiers. The classifiers can be evaluated on the basis of matching feature's performance applied to the images to be tested. The cascading of the classifier is shown in (15):

$$E_t(x) = \sum_{t=1}^{\infty} E\left[f_{t-1} \cdot (x_i) + \alpha_t \cdot h(x_i)\right] \tag{15}$$

Skin color is a special feature of the human face to differentiate the face from other parts of the body. State-of-the-art research investigated various skin color models, although we consider the most prominent skin color models [76, 79, 146], such as RGB, CMY, YUV, YC_bC_r, HSV, and CIE-XYZ. The RGB model is comprised of three primary colors, namely, Red, Green, and Blue. RGB color is the most reliable and convenient for the human visual system. The CMY (or CMYK) model is used interchangeably with the RGB model. The YUV model includes two parameters: the first is the Y channel (i.e., luminance), responsible for detecting of colors, and the second parameter is the combination of U and V to represent the chrominance values. The YUV is used in television for video broadcasting, e.g., PAL or NTSC. The YC_bC_r model [95] is the most prominent model used to represent the colors for digital TV due to its ability to tackle complex scenarios. Here the Y parameter represents the luminance part, and the other two combined parameters, C_b and C_r, show the chrominance part of the color. The YC_bC_r model has a mathematical connection with the well-known RGB model, The parameter Y can be derived from (16).

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \tag{16}$$

Whereas, the chrominance parameters for blue and red color can be calculated by the (17) and (18), respectively.

$$C_b = (B - Y) \tag{17}$$

$$C_r = (R - Y) \tag{18}$$

The HSV model [14] includes Hue, Saturation, and Value parameters. In this model, Hue comprises color depth information, saturation defines the strength of the whiteness, and the value represents the count of intensity (i.e., brightness). Sometimes, HSV is also referred to as HSB or HSI [11] model, where H and S parameters are identical to the HSV model, whereas B in HSB stands for brightness factor, same as parameter 'I' in HSI stands for intensity value. HSV is useful for face spoofing threats as it can extract depth information. H, S, and V parameters can be determined from the (19)-(21), respectively.

$$H = \cos^{-1}\left(\frac{0.5 \times [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-G) \times (R-B)}}\right) \tag{19}$$

$$S = 1 - 3\frac{Min(R, G, B)}{R + G + +B} \tag{20}$$

$$V = \frac{(R + G + +B)}{3} \tag{21}$$

CIE-XYZ has three parameters such as X, Y, and Z, which have a relation with R, G, and B parameters of the RGB color model as shown in the following (22)–(24).

$$X = 0.490186 \times R + 0.309879 \times G + 0.199934 \times B \tag{22}$$

$$Y = 0.177015 \times R + 0.812324 \times G + 0.010660 \times B \tag{23}$$

$$Z = 0.010077 \times G + 0.989922 \times B \tag{24}$$

This model is suitable for color matching purposes as it can efficiently measure even small discriminable changes in the colors.

4.3 Classification technique

Various classifiers have been proposed and discussed in state-of-the-art research works. However, this section only provides contemporary insights for techniques that are primarily used to tackle face identity threats such as support vector machine, k- nearest neighbor, Bayesian network, hidden Markov model, and distance-based classifiers. The comparative analysis for all classification-based countermeasure techniques is shown in Table 2.

Support vector machine [10, 13, 20, 25, 27, 51, 135, 138, 151, 184] is a robust multi-class classifier primarily used for face recognition problems with machine learning and deep learning architectures. SVM identifies support vectors for each class best separated from the decision boundary using either of the two non-linear kernel functions, i.e., RBF and Gaussian. Figure 13 shows the typical flow of the SVM process to calculate the support vectors. The best separating line that classifies different regions is also known as a hyperplane.

The kernel can be a linear, polynomial, radial basis, and simple a sigmoid function. The non-linearity can be handled efficiently by the kernel trick in SVM as indicated in (25).

$$Z = e^{-y(x^2+y^2)} \tag{25}$$

In SVM, the radial basis function can be implemented using (26) and (27), and the Gaussian kernel can be implemented using (28) as shown below.

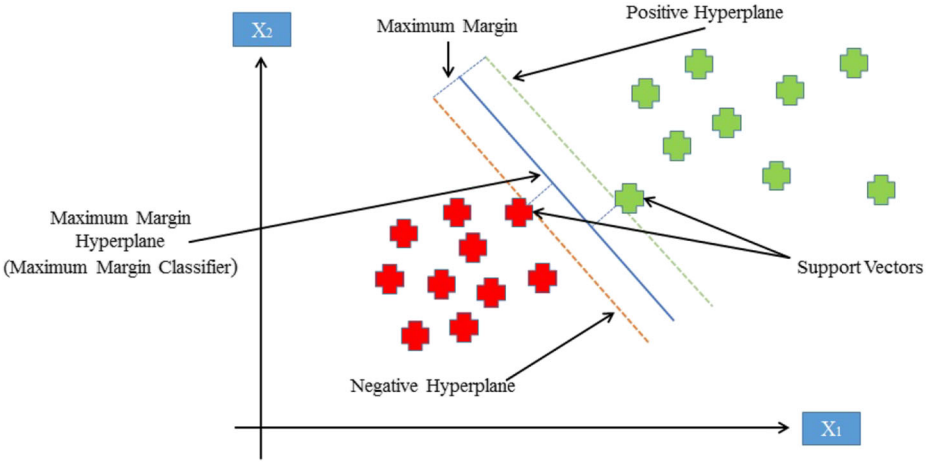

Fig. 13 SVM process

– **Radial basis function (RBF):**

$$D_{(x,y)} = \sqrt{(x^2 + y^2)} \tag{26}$$

$$f(x) = \sum_{i}^{N} \alpha_i \cdot y_i \exp\left(-\frac{\| x - x_i \|^2}{2\sigma^2}\right) + b \tag{27}$$

– **Gaussian Kernel:**

$$K(x, y) = \exp\left(-\frac{\| x - y \|^2}{2\sigma^2}\right) + b \tag{28}$$

where, (x, y) is the coordinate of the centre.

K-nearest neighbors [4, 15, 40, 176] can be considered as a replacement for SVM, especially where brute force effort is not reliable over a long period of time. The K-NN algorithm compares the nearest neighbors and based on that inference, it counts the votes and classifies the decision.

Bayesian network [173] is a statistical and probability analysis-based model widely used to solve real-world problems. The outcome of the Bayes theorem is not always perfect but almost trustable as it is associated with the persuasive logic of statistical analysis. The purpose of Bayes' theorem is to determine the most likely hypothesis from the fundamental knowledge extract from the given input data and their prior probability for different theories. Bayes' theorem is shown in (29):

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)} \tag{29}$$

Where, $P(H)$= Prior probability of hypothesis H, $P(D)$= Prior probability of training data D, $P(H|D)$= Prior probability of hypothesis H given D (Posterior Density), $P(D|H)$= Prior probability of training data D given hypothesis H (Likelihood of D given H).

Hidden markov model [112, 132] is designed to focus on sequences of emitted or lost regions of the face rather than on the entire face. These emitted (or lost) regions of the face precisely fix the occluded (i.e., covered) part of the face when compared with a full face. This model considers the five standard face attributes such as forehead, eyes, nose, mouth, and chin to normalize the holistic face region in accordance with sequence-wise contextual facial grammar. Figure 14 represents a complete framework of HMM with a neural network for recognizing facial expressions based on face grammar.

Distance metrics classifiers are mainly used in many machine learning and deep learning techniques. These methods utilize the correlation between different data features to predict the class. The most popular distance metrics classifiers are Euclidean distance, Manhattan distance, Chi-square distance, and cosine similarity.

Euclidean distance [9, 13, 45, 112] is an efficient and effective method for measuring the distance between two data samples for any order of dimensions. Euclidean distance can be represented by (30).

$$D_{Euclidean} = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \tag{30}$$

Where, p and q are the samples for n-dimensional feature space, and D is the distance of these data samples. $\sum$ is the sum of the calculation for each category.

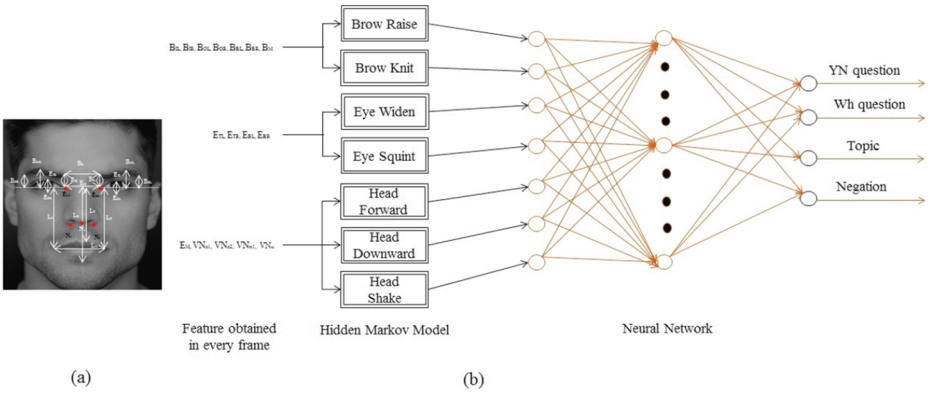

Fig. 14 Framework for recognizing facial expressions [187] (a). Distance feature (b). Combination of HMM and NN

Manhattan distance [82] is the distance between two data samples that can be measured at right angle to the axes. This method is also referred as City Block Distance or D4 method. The distance between two points $A(p_1, q_1)$ and $B(p_2, q_2)$, can be measured through the (31).

$$D_{Manhattan} = \sqrt{\sum_{i=1}^{n} |p_i - q_i|} \tag{31}$$

Where, n is the total number of samples and let's suppose, i and q are the two data points for i=q=1 and i=q= 2. It can be increased for any number of data samples. $\sum$ is the summation of such categories.

Chi-square distance [176] method investigates the difference between what is actually found in the study (i.e., observed data), and what is expected to be found find (i.e., hypothesis) for each considerable data point. The Chi-square method is best suited for comparing different histogram patterns in the face recognition domain because of its easy computation and interpretation. The chi-square distance can be calculated as shown in (32).

$$D_{Chi-square} = \chi^2 = \frac{1}{2}\sum_{i=1}^{n} \frac{(X_i - Y_i)^2}{(X_i + Y_i)} \tag{32}$$

Where, χ^2 is the measure of chi-sqaure, X_i and Y_i are the observed data samples. $\sum$ is the sum for each category.

Cosine similarity measure [45, 111, 117] is the similarity index between two different vectors. The cosine represents the product of two vectors concerning the angular direction, where these vectors are pointing. The cosine similarity score can be measured by using (33).

$$D_{Cosinesimilarity} = \cos\theta = \frac{p \cdot q}{\| p \| \cdot \| q \|} = \frac{\sum_{i=1}^{n} p_i \cdot q_i}{\sqrt{\sum_{i=1}^{n} {p_i}^2} \cdot \sqrt{\sum_{i=1}^{n} {q_i}^2}} \tag{33}$$

Where, $\| p \|$and $\| q \|$ are the length for vector p and vector q, respectively. $\frac{p}{\|p\|}$ and $\frac{q}{\|q\|}$ are the normalized vectors. $\sum$ is the sum for each category. The term $\sum_{i=1}^{n} p_i \cdot q_i$ represents the productive sum of two vectors for each category.

4.4 Neural networks-based techniques

A neural network [52, 94, 109, 112, 124, 144] is a set of algorithms that learns from the data itself and subsequently, identify discriminating patterns (relationships) to classify these data. A neural network is a replica of the human brain, consisting of various components such as neurons, dendrites, axons, and soma that are span millions. The smallest unit of the neural network is the perceptron (two-layer artificial neural network), first introduced by Rosenblatt in 1957. Neural network-based techniques are the primarily preferred in state-of-the-art research to solve real-world problems, including the countermeasures for facial identity threats. A neural network is an appropriate option for small and large database-related problems. The structure of a simple neuron is depicted in Fig. 15.

Single-layer neural networks [2] aim to understand the psychology of the human brain using a simple type of feed-forward neural network. This method takes the inputs in numeric values referred to as weights to generate output based on specific functions. Single-layer neural networks include a simple architecture with low computational cost to solve real-world problems. The architecture of a single-layer neural network is depicted in Fig. 16.

Multi-layer neural network (MLNN) [119, 156] is another neural network consisting of more than two layers, including at least one hidden layer. Two or more multi-layer neural networks can create a new and faster model known as the ensemble model [26], which is used to efficiently perform various complex tasks, such as medical diagnosis, face spoofing, and more. However, it requires more computation time and cost.

Convolutional neural networks (CNNs) [52, 123, 124, 154, 167, 177, 178, 183, 184] is a powerful feature extraction and classification technique. The convolution operation is performed to find a correlation among different pixels associated with that image. Figure 17 illustrates the complete process of CNN, including feature vector, multiple feature maps, pooling, flattening, fully connected layer, and finally, the classification.

CNN supports transfer learning by sharing the weights of pre-trained models like ImageNet. There are different types of CNN variants that make it robust such as AlexNet, VGG16 [16, 101], VGG19 [139], ResNet50 [92], InceptionV3 [30], InceptionResNetV2 [37], [122], MobileNet [39], and deep residual networks (DenseNet121 [166], DenseNet169 [182], DenseNet201 [169]). Other modular variants of convolutional neural network such as scattered CNN [177], Multitask- CNN [111], and CNN with RNN [67] are also analyzed and compared here. The CNN-based model [84] for video attack detection provides error rates of only 1.3% for the replay-attack, and 2.5% for the CASIA-FA datasets. The face-based presentation attack detection method provides the best error rate of only 0.3% with

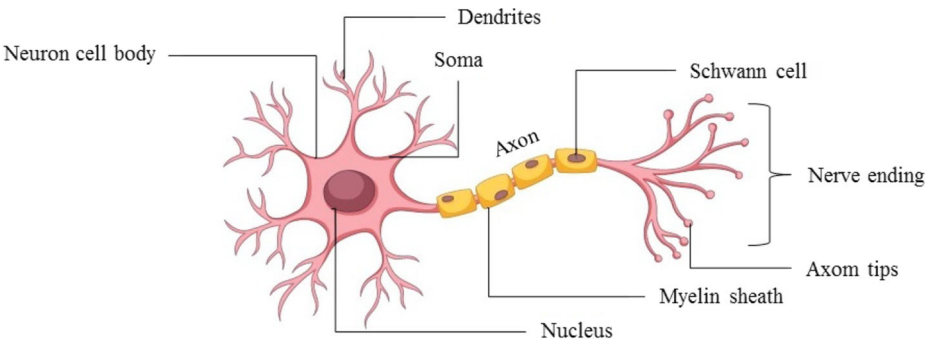

Fig. 15 The structure of neurons

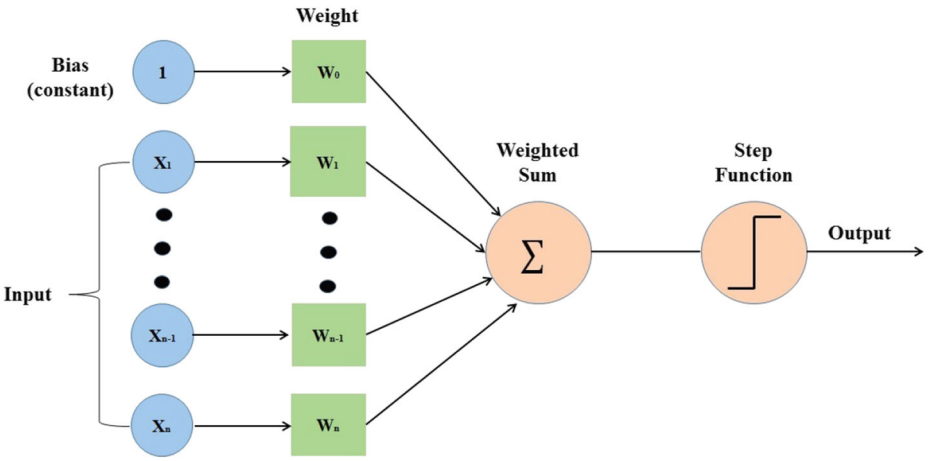

Fig. 16 Single layer neural network

the wide multi-channel presentation attack (WMCA) dataset for a three-dimensional silicon mask using a CNN. Therefore, the computation cost of TCDCN is very less. Table 2 provides a comparative analysis of countermeasure techniques to handle various facial identity threats.

Summary and remarks

Table 2 provides a detailed comparative analysis of various countermeasure techniques based on dimensionality reduction, feature extraction-based techniques, classification-based techniques, and neural network-based techniques to reduce the impact of various challenges associated with face recognition systems. Some methods can be used for multiple purposes.

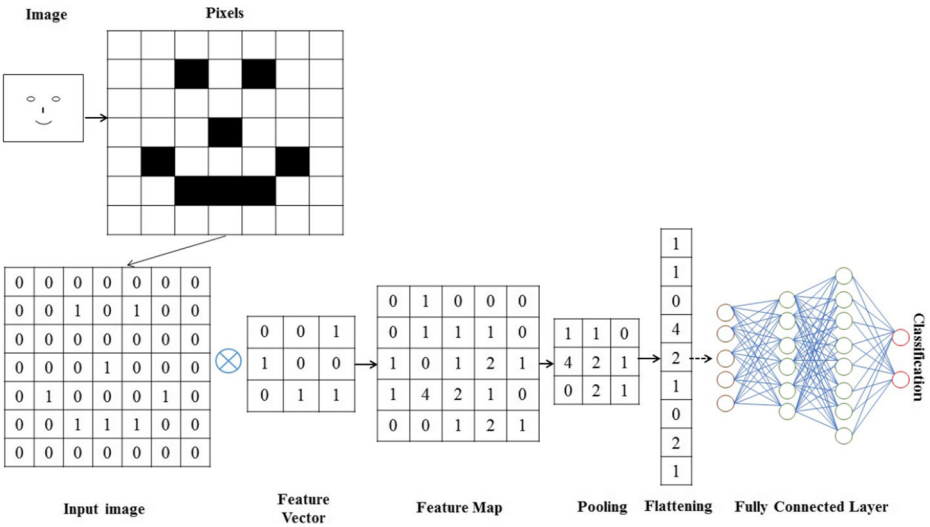

Fig. 17 The process of Convolution Neural Network

The fourth column of Table 2 clearly indicates the applicability of these countermeasure techniques on identified threats.

5 Comparative analysis of state-of-the-art approaches

In the recent past, various state-of-the-art approaches have been proposed and optimized to provide an efficient solution to the above-mentioned face identity threats. This paper provides a detailed comparative study of different state-of-the-art approaches for these identified threats based on the significance, relevance, and research interest. This paper divides the complete analysis of countermeasure techniques for facial identity threats into four subsections: direct spoofing, indirect spoofing, zero effort imposter, and other factors. Each sub-section is composed of tabular comparison followed by a summary on the best methodologies based on the performance measures. We also indicate some remarks on the merits and demerits of the techniques. Table 3 to Table 7 represents the countermeasure techniques proposed to reduce the impact of these identified facial threats. Table 8 presents the comparative studies of state-of-the-art approaches involving modular (i.e., multiple) face identity threats. Table 9 provides a detailed description of the various available benchmark face datasets to address these face identity threats. The abbreviations and their meanings are mentioned at the bottom of each table. However, repeated abbreviations are explained only once.

5.1 Comparison and discussion on direct spoofing

We have categorized direct spoofing countermeasure techniques into two sub-categories: plastic surgery and makeup. The tabular comparison take into account various factors such as the concept, the methodology used, the dataset used, the performance measures, and the limitation. Table 3 represents a comparative analysis of state-of-the-art techniques to reduce the impact of direct spoofing of the face.This analysis of countermeasure techniques includes three sub-sections: textural plastic surgery, structural plastic surgery, and makeup-based spoofing.

Summary and Remarks

Few research articles are published in this particular area of research, which includes holistic face region-based [36], ocular information-based [71], and depth information-based (i.e., heterogeneous sketches) [19] countermeasure methods. However, this survey paper only includes some of the best methodologies with their pros and cons to overcome these threats.

The Geometric-based analysis [45] and periocular biometric-based analysis [130] outperforms the other recent methods. The average recognition rate for local and global plastic surgery is 78.5% and 76.1%, respectively. The fusion of feature-based and texture-based methods [6] represents an accuracy of 91%. Singh R, et al. [149] represents a hierarchy for all possible plastic surgeries on face regions performed to date globally. The research article [71] represents 87.4% rank-1 accuracy for plastic surgery.

A shape, texture, and skin color analysis-based methods for detecting facial makeup with a detection rate of 93.5% have been proposed by Chen C., et al. [25]. Ueda S., et al. [163] represents the two different scenarios for evaluating disguised faces (i.e., impersonation and obfuscation) with explicit noting for light makeup and heavy makeup. Transfer learning-based method (named Style Transfer) with the cycle-consistent generative adversarial network (i.e., cycle-GAN) is proposed by Chang H., et al. [24] to find out the

Table 3 Comparative analysis of techniques to mitigate face-based direct spoofing

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
PS (T)	[45]	Analysis of pre-geometric and pose geometric changes through GFRPS and minimum distance classifier technique	GFRPS, MO, ED, CS, and NN classifier	PSFD	IR for local- 78.5%, and global 76.1%, GFRPS achieved 79.80% IR (Rank-1)	Hard to detect each appearance change
	[6]	A fusion of feature-based GIST and texture-based LBP methods for plastic surgery images considering edges, corners	FE- GIST, LBP, CLS-cosine distance metrics	PSFD with 1012(pre and post-surgery) (506 subjects)	VerA 91% (max)	The time complexity is not measured
	[130]	Facial marks are identified using HOG to evaluate the pre-surgery and post-surgery impact on face	Laplacian of Gaussian, HOG, SURF, Sobel, and CED	PSFD	SURF outperform others EER-42%, RR- 99.8% for FMD	Not suitable for critical plastic surgery cases
PS (T & St)	[149]	An analytical aspect is reviewed for FR system after PS with 900 individual face databases.	Polar Gabor NN transform(GNN),PCA, CLBP SURF, LFA, FDA	PSDB	GNN outperform others with 53.7% IA (Rank-1)	Sensitivity and privacy issues, not suitable to find geometric changes
	[71]	The face and ocular information-based method for face identification is proposed. A review on various surgery approaches including information of commercial software is also discussed.	FE- VJ, SIFT- LBP, Identification-Cumulative Match Characteristic	PSFD, ocular dataset	face and ocular fusion accuracy is 87.4% (Rank-1)	Low-resolution, variations in scale and expression, duplicates

Table 3 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[36]	FAce Recognition against Occlusions (FARO) with expression variations divides the face into multiple regions and Partitioned Iterated Function System (PIFS) process them on the basis of codes.	FE- PCA, LDA, FARO and FACE, SFA FDA, LBP CLS- SVM, k-NN	AR-Faces	GNN algorithm's performance among all local and global PS process.	Dataset is synthesized from benchmark DB that can affect results for real-world problem.
	[109]	A survey on state-of-the-art techniques analyzing the performance for facial plastic surgery is presented.	PCA, FDA, Geometric Features, LFA, LBP, SURF Gabor NN, PSO, PIFS, and SSIM	Public face surgery dataset	IR for GPROF method is up to 90% (rank-1)	Un-trustable high accuracy is achieved for different altered probe and gallery
	[192]	Overview of PS based FR techniques considering the relevance, applications, and surgeon's recommendation for patient is discussed.	NA	2878 frontal images with 9 genetic disorder (36 points annotation)	Apple iPhone-X frequently updates the user's face print information.	The nonlinear alterations are appeared in facial landmarks
	[19]	A review on contemporary research to investigate the interaction between facial PS and FR software.	FE - Circular LBP, SURF, EV-SIFT, CLS- PCA, FDA, LDA	PSFD and facial pathology.	IA in range of 15- 99%	Security, ethical, and non-linearity

Table 3 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[163]	Three categories of disguise make-up (i.e., No, Light and Heavy) are investigated with considering false positive and false negative.	A three-factor (i.e., no, light and heavy) repeated-measures ANOVA is used.	24 Japanese women have participated for this reserach	RR for no makeup- high light makeup-medium heavy makeup- very low	Unable to recognize the heavy make-up.
	[34]	A non-permanent technique involving a face altering mechanism for investigating the make-up is deployed here.	Pre-processing-DoG, FE-Gabor wavelets, LBP, Projection generation-Verilook Face Toolkit, DR- PCA, LDA	YMU, VMU DB	EER range 6.50% (LBP) to 10.85% (Verilook), EER for LGBP in YMU and VMU are 15.89% (no makeup), 5.42% (full makeup)	Only female subjects are taken into consideration.
	[25]	An algorithm to classify the make-up in input image using shape, texture and color information for only considering eye and mouth features is presented.	Color space (RGB, Lab, HSV), Adaboost, SVM, GMM, LBP, DFT	YMU (151 subjects, 600 images) MIW (125 subjects, 154 images)	DR 93.5% (with 1% FPR), accuracies with SVM - 91.20±0.56, with Adaboost- 89.94±1.60, Overall-95.45%.	Only female subjects are considered in the database, thus not a generalized method.

Table 3 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[26]	A sampling patch-based ensemble learning method to classify the image, before and after the makeup.	FE- LGGP, HGORM, DS-LBP, SRS-LDA, CLS-CRC and SRC	YMU	EERs and GARs for COTS-1, COTS-2 and COTS-3 are 12.04%, 7.69%, 9.18%, and 48.86%, 76.15%, 58.48%, respectively.	Only female subjects are considered, thus not generalized.
	[24]	A novel GAN-based unsupervised method for cycle-consistent asymmetric functions is deployed using reciprocal functions in forward (i.e., encode style transfer) and backward (i.e., destroy style) direction.	CycleGAN	Self-created dataset from YMU tutorial video	Significant results on transfer makeup styles to people having different skin-tones, original-tone with preserved identity	The performance degrades with heavy make-up.
	[148]	Disguised faces in the wild dataset is proposed with the evaluation of impersonation in three levels of difficulties (i.e., Impersonate, obfuscation, overall)	Analysis on Impersonation & obfuscation attack.	DFW, AEFRL, MiRAFace, UMDNets	IR for AEFRL of 96.80%, and Obfuscation rate for MiRA face 90.65%, and best Overall rate is 90.62% for MiRA face. (best accuracy)	The created dataset is not suitable for all types of presentation attack.

PS- Plastic Surgery, T- Textural, St- Structural, GFRPS- Geometrical Face Recognition after Plastic Surgery, Mo- Morphological operation, PSDB- Plastic Surgery Face Database, CS- Cosine similarity, NN- Neural Network, ED- Euclidean Distance, FE- Feature Extraction, CLS-Classification, EER-Equal Error rate, RR- Recognition rate, FMD-Face Mask Detection, CED- Canny Edge Detector, DoG- Difference of Gaussian, CLBP- Circular Local Binary Pattern,IA- Identification Accuracy, GNN- Neural Network-based Gabor Transform, SFA- Split Face Architecture, YMU- YouTube MakeUp,VMU- Synthetic Virtual Makeup,LGGP- Local Gradient Gabor Pattern, HGORMHistogram of Gabor Ordinal Ratio Measures, DS- Densely Sampled, CRC-Collaborativebased Representation Classifiers, SRC- Sparse-based Representation Classifiers, PSO- Particle Swarm Optimization, PIFS- Partitioned Iterated Function System, SSIM-Structural Similarity Image Maps

optimized result with understanding the impact of makeup on faces in face recognition. Among other practical approaches, the patch-batch ensemble learning method [26] and the non-permanent face altering approach [34] provide significant results specific to this problem.

An arduous study of state-of-the-art research for plastic surgery and make-up based face identity confirms that researchers have paid less attention to this area. The main reason for this issue is the unavailability of appropriate benchmark datasets with non-linear variations. Plastic surgery-based datasets are difficult to analyze, especially for evaluating real-world scenarios, because of the various legal safety and security concern of the public. Thus, no state-of-the-art approach has provided an acceptable level of identification for plastic surgery-based problems. However, Singh M., et al. [148] introduced a new benchmark dataset in 2018 for disguised faces, named Disguised Faces in the Wild (i.e., DFW). This dataset provides an impersonation rate of up to 96.80%.

5.2 Comparison and discussion on indirect spoofing

We have categorized indirect spoofing countermeasure techniques into two sub-categories: two-dimensional spoofs and three-dimensional spoofs. Table 4 shows a comparative analysis of the state-of-the-art techniques to reduce the impact of indirect spoofing which again consists of photo attack, video attack, and mask attack. Extensive research work has been done in the field of photo attacks and video attacks, as these types of spoofing are easy to implement and inexpensive in terms of cost. In contrast, mask spoofing (attacks) are costly and hard to implement practically.

Summary and remarks

Spoofed faces typically have some image distortion reflecting the poor quality of the input image compared to the original acquisition. The color texture [27], and luminance parameters [20] are significant factors to distinguish these distortions from the real acquisition. Choi J., et al. [29] investigated the thermal, infrared, and visual imaging-based samples for the experiments. The authors obtained the best recognition rate with these samples for identifying the spoofed faces. The moiré digital grid patterns [51] also perform better for detecting presentation attacks. The HOG descriptors with light field disparity outperform other methods for classifying spoofed face and genuine face [138]. The kernel-based method [178] also shows efficient results for replay attack (i.e., video attacks) problems. Silicon mask detection-based technique [52] represents best results (i.e., 0.3% error rate only) among the state-of-the-art approaches to detect 3D face mask. In reference [121], the most prominent anti-spoofing methods are critically analysed.

5.3 Comparison and discussion on zero effort imposter

Zero effort imposter threat constitutes two categories the first is identical twins and the second is a face morphing. However, these threats have not been resolved till date due to some facts such as unavailability of datasets, biological and genetic reasons of medical science. Table 5 demonstrate a comparative analysis of the state-of-the-art techniques representing the recent development to mitigate the impact of zero effort imposter threats.

Summary and remarks

The identical twins-based problem poses a significant threat to the existing face recognition systems. No research has proved the significant consequences of facial recognition for identical twin problems in practical situations to date. However, few researchers indicate some alternatives to address this issue. Phillips P. J. et al. [118] presented a covariate

Table 4 Comparative analysis of techniques to mitigate face-based indirect spoofing

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
PA	[29]	A new partial least squares-discriminant analysis method to reduce the modularity impact in visible and thermal images	LBP, HOG, Gabor filter and PLSDA	(X1 Collection) UND dataset	Thermal-visible Facial IR 49.9%.	Modality gap is more, less efficient to night vision
	[51]	A novel overlapping digital grid method to detect the presentation attack by utilizing Moiré pattern is proposed.	FE- Moiré Pattern digital grid, CLS- SVM (RBF kernel)	MIT-CBCL, Yale, Caltech	This method provides good results for low pixel ratio.	Not suitable with illumination variance problem
	[20]	A non-intrusive software-based method to evaluate the luminance color information from the face image.	Disparity, CoALBP, LBP, LPQ, BSIF, SID, SVM with histogram	CASIA-FASD, Replay Attack, MSU-MFSD	EER For given DB (in %)- Video- 3.2, 0, 3.5, Image- 2.1, 0.4, 4.9, respectively.	The color-texture analysis is not generalized.
	[138]	The Light field disparity-based face PAD techniques are reviewed with introducing a new HOG descriptor.	HOG, SVM	NUAA, Print attack, Replay Attack, CASIA	ACER- Laptop 0.05%, Mobile 0.02% paper -0.5%	Unavailability of appropriate Light-field DB, focus-based poor results
	[27]	A CNN-based method with color-texture features (RI-LBP) is introduced here.	FE- ResNet-18 and rotation invariant -LBP), CLS- SVM, DR- PCA	NUAA Replay CASIA-FASD, MSU-MFSD	ERR for NUAA, Replay, CASIA, MSU are 0.5, 2.3, 4.4, 3.1 respectively.	The results on cross-DB must be generalized.

Table 4 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[124]	A new database SCFASD is introduced with disparity layer-based supervised CNN classifier to efficiently detect face liveness.	Dynamic disparity maps and CNN	Stereo face anti-spoofing DB for (printed photo, mobile, tablet)	Overall APCER = 0.86±0.80	More generalization is expected for real-time applications.
	[56]	The adaptive fusion of each classifier scores is performed on multimodal biometrics (face, finger, and iris) considering the concurrent (boosted) and discordant (suppressed) Modalities.	Hough Transform, Gabor, minutiae features, Gaussian kernel function, Min-max, threshold	Chimeric multimodal databases	Average accuracy 99.5%, EER 0.5%	It doesn't support a dynamic environment
	[123]	DL-based facial PAD method using perturbation with pre-processing is proposed.	CNN, LBP	CASIA, Idiap Replay-Attack, OULU-NPU DB.	ACER (in %)- 3.89 (Oulu NPU), 0.23 (CASIA and Idiap), 0.97 (MSU-USSA)	Other handcrafted method also has to be evaluated.
VA	[84]	A LBP and CNN-based methods are comparatively evaluated for face spoofing.	LBP and CNN	Replay Attack and CASIA-FA	ACER- Replay Attack 1.3%, CASIA-FA 2.5%	Slow speed recognition, sensitive to noise
	[151]	Three discriminative (SPMT, SSD, TFBD) representations for face PAD is performed.	FE- SPMT, SSD, TFBD, CLS- SVM	Replay Attack and CASIA-FA	ACER- Replay Attack 0%, CASIA-FD- 2.58%	Sensitive to stereo type, binocular camera is needed
	[178]	A deep architecture-based FLD to prevent video spoofing attack is proposed here.	FE-CNN and generalized multiple kernel learning (GMKL),CLS- SVM	Replay Attack and CASIA-FA	Replay Attack 0% CASIA-FA 2.58%	Edge enhancement, texture differences are not considered.

Table 4 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[69]	Several face PAD techniques are reviewed to generalize the mobile-based authentication for various cross-databases.	Mobile spoofing datasets	MSU-USSA, Replay Mobile, Oulu NPU	ACER (in %)- 8.33 (Oulu NPU), 0.26 (Replay mobile), 0.97 (MSU-USSA)	The biased DB has to be more practical and generalized.
MA	[121]	A comprehensive review for the state-of-the-art anti spoofing techniques are investigated here.	Challenge response, blink detection	Replay Attack, 3D mask Attack, Print Attack DB	APCER for Print Wrap and display are 5.27%, 1.21%, 0.71%, respectively	H/W: costly, overhead, S/W: device dependent, Not generalized
	[52]	A multi-channel CNN method with a new WMCA dataset consisting of 2D and 3D attack for Impersonation and Obfuscation condition is introduced here.	CNN with a novel Wide Multi-Channel PAD	WMCA dataset with thermal, near-infrared, color, and depth information.	ACER of 0.3%	Unseen attack protocol evaluation for problems is not efficient.
	[49]	An image-quality assessment-based fast non-intrusive method with motion cues is proposed for face spoofing detection.	LDA, SVM	Replay attack, Replay mobile, 3D MAD	HTER-0.024% (Replay attack)	This process is not suitable for large amount of database.

PA-Photo attack, VA- video attack, MA- Mask attack, PAD- Presentation attack detection, FLD- face liveness detection, SPMT- spatial pyramid coding micro-texture, TFBD- template face matched binocular depth, SSD- single shot multiBox detector,IR- Identification Rate, APCER-Attack presentation classification error rate, ACER- Average classification error rate, HTER- Half total error rate, WMCPA-Wide Multi-Channel Presentation Attack dataset

Table 5 Comparative analysis of techniques to handle zero effort imposter threats

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
ZEI (IT)	[118]	A covariate analysis is performed to detect the identical twins. Multiple biometric are considered for evaluations of twin pairs.	Multiple Biometric Evaluations (MBE).	Identical Twins pair image DB (Same day-126, after one year-24 (all in pairs))	90% confidence value from 0.01-0.04 EER.	Unable to distinguish outside light, smiling-neutral expression, gender and age.
	[164]	A face similarity-based recognition of identical twins including various facial expressions is investigated here.	1-Iterative Closest Point (ICP), SIFT, 2- Shape index, Extended- LBP, 3-ED, 4- UR3D (Wavelet + LDA)	3D TEC DB (3D scans of 107 pairs of twins, having Neutral & Smile facial expressions	3D face recognition accuracy for twins exceeds 90%	The modality of facial similarity with expression variation makes a challenging issue.
	[82]	The twins' pairs are identified through analysis of face-based aging features. Here, noticeable parameters like brow furrows, laugh lines, and forehead lines are considered despite wrinkles.	Gabor wavelets, Modified-ASM, LFDA, FDA, Locality Preserving Projection (LPP), Manhattan distance	UND-twins DB.	A person having smile expression without wearing glasses obtained accuracy of 96.67% on 240 twin pairs.	Only smile facial expression is evaluated
	[153]	An automatic gradient-based method to recognize facial marks of Identical twins on multi-scale parameters (bright and dark regions) is proposed here.	Manual annotation, Gaussian pyramid, ASM face contour, Bipartite Graph Matching, FRST	Twins Days Festival in Twinsburg, Ohio. facial marks as a biometric signature	It is observed that face mark detector is slow while manual annotation method is quite complex and time taken process.	Only face mark is considered, thus not generalized.

Table 5 (continued)

Type/ Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[114]	The identical twins pairs are evaluated for various scenarios such as gender, illumination, age, and expressions using seven different FR methods.	Multiple Biometric Evaluations (MBE).	Twins Days festival [142] in Twinsburg, Ohio in August 2009 and August 2010.	EER (in %) 0.2 for the Studio-Studio, 1.1 for ambient conditions. EER≤0.1(M), EER 4.1-17.4 for others.	It is very challenging to distinguish identical twins in un-ideal condition.
ZEI (FMo)	[113]	A 2D surface reflectance and 3D shape-based FR technique consists of 0, 30, and 60 degree viewpoints is proposed here.	Laser scanning, 3D shape, 2D surface reflectance, P2-norm	Self-created dataset with 60 volunteers from the University of Texas at Dallas	The data patterns for female category are comparatively easy to understand and interpret.	It provides viewpoint for maximum of 60 degree rotation, thus not good for profile images>60.
	[120]	A novel collaborative method for micro-textural feature extraction is proposed to apply face averaging, and face morphing operation for real-life scenarios.	LBP, BSIF Image-Gradient magnitude and LPQ, DCT, (HSV, YCbCr)	Bonafied image DB, synthesized morphed and averaged DB	FRS Vulnerability for averaging and morphing are 90.33 and 83.62, respectively with IAPMR. EER for averaging and morphing are 9.48 and 2.93, respectively (in %)	The non-correlating ethnicity is the major drawback of this technique.
	[135]	A deep review on conceptual categorization of face morphing techniques with their assessment metrics.	Free-Form Deformation (FFD), SVM, optimization of warping, VGG-19, BSIF, LBP, HOG, k-NN	Automated in-house generated DB from real-world attack scenarios, FEI DB	NA (no significant result obtained, It requires more research in future)	Quality, overfitting benchmark database are the major issues

BSIF- Binarized Statistical Image Features, FMo- Face Morphing, FRST- Fast Radial Symmetry Transform, DB-Database, ASM- Active Shape Model

analysis-based method, which provides better recognition results. The impact of various facial expressions and illuminations on identical twins are investigated by Paone JR., et al [114] to achieving better results. The aging factors such as the erosion ratio of skin tone and wrinkles are investigated to achieve effective results for similar threats [82]. The research article [153] indicates that every human face has at least one unique facial marking that distinguishes it from other people. These facial marks can be moles, freckles, birthmarks, patches, and scars depending on ethnicity, race, and various genetic (hormonal) changes. Furthermore, the literature confirms that these facial marks are particularly helpful in effectively discriminating the features for identical twins.

5.4 Comparison and discussion on other factors (intrinsic threats)

We have categorized intrinsic threats related countermeasure techniques into two subcategories: Intrapersonal and interpersonal. The intrapersonal factors that becomes barrier to face recognition can again be classified into partial occlusion, various facial expression, and aging. In the same way, the interpersonal factors can also be divided into race and gender. The threat of race is a genuine problem that currently needs to be solved. Therefore, satisfactory results are still awaited. Face-based gender identity threats is also unsolved issue till date due to various factors such as absence of adequate and benchmark dataset, sensitivity and social exclusion. Table 6 represents a comparative analysis among various interinsic factors.

Summary and remarks

Table 6 represents a comparative analysis of various recent countermeasure methods to tackle intrinsic facial identity threats such as partial occlusion, facial expression, aging, race, and gender. Lu Xh, et al. [94] implemented a local and global feature-based method to detect partial occlusion of the face. In contrast, Wan J., et al. [165] analysed the appearance-based method. Lian Y., et al. [88] proposed infrared and thermal image data-based technique methods to present a better recognition rate for the partial occlusion.

The nuclear-norm-based approach [42] obtains accuracy up to 93.1%. However, this method is limited to showing poor results with illumination variations. A single-shot-multibox detector method [191] achieves an average precision of up to 95.46% for all categories of partial occlusions. Anzar S., et al. [8] introduced a new technique that includes two sophisticated features descriptors: bi-orthogonal discrete wavelet and SIFT. This method achieves the best accuracy of up to 90.5% for randomly applied partial occlusions. Reference [32] critically reviewed state-of-the-art approaches detecting partial face occlusion. This study concludes that the partial occlusion detection rate for sunglass prop can be improved up to 94.2%. The gravitational search-oriented Restricted Boltzmann Machine (RBM) with the SURF feature-based method [144] provides the highest accuracy of partial occlusion up to 98.72%. Extensive research work has been done in this evolutionary field of facial expression recognition. The literature study confirms that a total of seven noticeable facial expressions such as angry, disgust, fear, happy, neutral, sad, and surprise have been studied and investigated by researchers worldwide due to global uniqueness. State-of-the-art research includes shape-based approach [177], kernel-based approach [10], real-time video-based approach [117], and short-time annotation-based approach [63]. Most research works represent expression-wise accuracy rather than each expression. The illumination invariance-based facial expression method [40] achieves up to 99.82% accuracy for the NVIE dataset with infrared light. However, micro-expressions-based research such as blinking eyelids is not adequately addressed and still needs to be explored by research communities. Aging is also a significant factor that affects the automatic face recognition

Table 6 Comparative analysis of techniques to mitigate the impact of intrinsic threats

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
Occlusion	[158]	Local distribution based occlusion detection for palm print-based uniform patterns.	ULBP (Uniform-LBP), NN, Thresholding	PolyU2D, CASIA, IIITDM palm-print DB	Up to 36% of occlusion is detected	Not tested with other biometric traits.
	[32]	A review on state-of-the-art approaches to detect occlusion in 3D face and subsequently the restoration strategies.	Radial curve, PCA, Template machine	Bosphorus,UMB GavabDB EURECOM KinectFace DB	RR for face occlusion (hairs)- 98.0%, glasses-94.2%, overall- 96.3 %	Not suitable for Uncontrolled conditions
	[42]	Characterization of the occluded and corrupted region using nuclear norm to get dictionary with best results.	SRC, NNB matrix regression, NNAODL	AR, Extended Yale- B DB, LFW	Accuracy 93.1%	Unable to handle non-monotonic illumination and noise problems.
	[165]	A cascaded deep generative regression model for an occlusion-free face alignment with GAN (locating occlusion), DRM(enhancement), Cascading(facial landmarks).	GAN, DRM, and Cascading	OCFW, COFW, 300W, and AFLW	EER 5.72% for COFW, 6.97% for 300W, 1.80% for AFLW.	Fusion of multiple processes can degrade the performance of occlusion.
	[88]	PCA and infrared thermal-based method to detect the occlusion on face regions is proposed here.	PCA, Infrared thermal imaging, BPNN, SVM	Self-created sample data	Highest face recognition rate 95%	Recognition rate is dependent on kernel parameter.

Table 6 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[94]	The local and global face features are extracted using high filtering function in the presence of different light source to analyze the halo effect on input images.	Recurrent Neural Network	AR face dataset	Sunglasses occlusion-98.45-88.4, Scarf shield-95.37-62.3 (in%)	Complex algorithm, which consumes more computation time
	[144]	Wavelet and SURF-based occlusion detection method utilizing gravitational search algorithm for feature extraction and recognition.	Wavelet, SURF, holo-entropy, DNN-RBM FE-appearance-based	Derf's accumulation (video dataset) and HMDB51	98.72% Accuracy	Complex architecture, computation overhead.
	[191]	A deep SSD-based occlusion detection method to find the target location of the face. It can deal with sunglasses, face mask, hats, and other accessories on face.	SSD (Single Shot Multi-Box Detector)	Seven types of self-built dataset for face occlusion	The average precision has reached to 95.46% for all categories (mAP).	No benchmark dataset is available.
	[8]	This paper proposed BDW-SIFT for partial occlusion. Wavelet SIFT covers 15% to 60% occlusion of the face.	BDW-SIFT (scale invariant feature transform with discrete wavelets), DWT	MUCT dataset	Testing accuracy for occlusion-Vertical- 87%, Horizontal 86%, diagonal 46.5%, Random 90.5%	Due to mass key points the time consuming process
Expression	[63]	User dependent unsupervised approach referred as PADMA, to find affective states from spontaneous facial expressions is proposed here.	PADMA using AMIL, Clustering	CMU Multi-PIE ,UNBC-MSARV	User dependent 58.5%, 59.2% User independent 71.3%, 25.7%	Occurrence of unknown affect in face gestures fails and not detected.

Table 6 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[40]	A uniform- LBP method is proposed for features extraction with Legendre moments, while k-NN is used for classification.	uLBP- (feature extraction), Legendre- feature vector, k-NN with L1 Norm Classification	IRIS and NVIE Database	RR and accuracy for IRIS 90.63 and 98.83 (visible), 99.83 and 99.98 (infrared). For NVIE 99.46, 99.97 (visible) and 99.82, 99.97(infrared).	Performance reduces for visible images having low illumination.
	[117]	An unsupervised framework for spontaneous expression recognition is proposed with preserved discriminative information of the input (i.e., video).	HOOF, MBH, UAM, SEV, Cosine, LDA, PLDA, SVM	BP4D, AFEW	Accuracy (Expression vector + SVM) for BP4D-81.3%, AFEW-74.1%	Experiments involve trimmed clipping videos, thus real-time testing result might be poor.
	[187]	An automated machine-based FER system including a dataset development algorithm is proposed with a detailed state-of-the-art review on occlusion effect.	Sparse and reconstruction approach, SVM, DBN, CNN	JAFFE, CK, CAM3D, BU-3DFD	CK+ database shows highest accuracy among all	Hard to detect face if occluded in group of faces
	[177]	A CNN-based method to identify the start to end points (onset, apex, and offset) based on face shape for each expression is proposed.	Scattering CNN	BU-3DFE, BP-4D	Precision 81.35%, Recall 87.19%	Not efficient for 3D and multi model domain
	[184]	An optimized method to recognize the relevant facial expressions for an input image is investigated.	CNN, SVM	BU-3DFE, SFEW	BU-3DFE- 1.72%,, SFEW-1.11%	Not suitable for extent variance in pose.

Table 6 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[132]	A non-posed image acquisition method, for evoking natural expressions through strong influence of the subjects in various scenarios: playing video games, during interviews, and watching emotional video is proposed.	Appearance based and feature based analysis, Static and dynamic classifiers (BN, HMM)	UT-Dallas FEEDTUM, NVIE, AM-FED	Expression of players of Video game is performed better for emotion diagnosis	Spontaneous expression intensity estimation for uncontrolled profile image is not done
	[103]	A supervised transfer learning method to recognize various facial expressions for multiple subjects (simultaneously) in a single frame is presented.	CNN (MobileNet) with Softmax and center loss function.	JAFFE and CK+	JAFFE (95.24%)and CK+ (96.92%).	Only frontal faces are taken into account
	[12]	A novel software for emotion recognition using webcam data is introduced based on FURIA algorithm and unordered Fuzzy rule induction.	FURIA algorithm with unordered fuzzy rule induction	Cohn-Kanade AU-coded expression extended database (CKplus)	Overall average accuracy of 83.2% (α) level	This method unable to detect the micro expressions
	[10]	An optimized kernel-based SVM method is introduced to classify the various facial expressions.	RBF kernel. SVM, Dimensionality reduction	JAFFE CK	Cross validation result 94.3%	Not able to detect non-basic (despite universal)expressions
Aging	[115]	A 3D shape and wrinkle texture-based aging invariance method is proposed to investigate 3D view-invariant face models.	3D Aging, PCA, AAM,m RBF	2D face aging DB, FG-NET, MORPH, BROWNS using FaceVACS	IA (before& after) aging are 26.4, 37.4 (FGNET), 57.8, 66.4, (MORPH), 15.6, 28.1 (BROWNS)	Automation of Age invariant FR system is hard to implement due to noise, textural info

Table 6 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[101]	A novel Wrinkle Oriented AAM (WO-AAM) is proposed with new channel to analyze wrinkles, empiric joint probability density by Kernel Density estimation, than synthesize new plausible wrinkles.	PCA,GAN, Recurrent face aging (RFA), VGG-16, CNN	Self-built 400 A Caucasian women, age (43-85 years) with average age-69	10-years aging & rejuvenating period, the estimation of age can be added by 4.9 and be subtracted by 4 years, respectively.	Result influenced by exposure of sunlight, alcohol consumption, dark spot.
Race	[57]	Other Race Effect (ORE) in FR analyze the different races considering two measures: assessed the quality (and quantity) of social interaction and the time measure (i.e., spent in Western country) with least information of contact.	Race contact questionnaire in study phase and recognition phase.	23 Chinese and 25 Caucasian participant with variability in contact, effect, inversion decrement for other-race, and cross-race	The overall model was significant (2,47) = 30.99; $p<0.0001$ and explained 38.2 % of the sample variance.	This method is not generalized and doesn't perform well for other continental location.
	[160]	A detailed study of the Race condition, includes participants training to recognize African, American (or Hispanic) through their faces at individual level and classify them at on the basis of race.	Electroencephalogram (EEG), Behavioral training	24 students of 18-29 years from university of Victoria, created DB	The fine-grained visual discrimination supported by N250 that reflects the formation of perceptual representations.	Adult face recognition for other race is not inflexible, intuitive experience is meaningless.

Table 6 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
Gender	[13]	A novel geomteric curve feature-based method is proposed for gender classification. The saptial features are extracted through circular and radial sets using Euclidean distance. Then, Adaboost algorithm for minimal feature selection.	FE-LBP, Multi-scale local normal patterns (MS-LNPs), Euclidean distance for Geometric (circular/radial) curve feature extraction, FS- Adaboost, CLS- SVM	FRGCv2 dataset	Face recognition rate for a rank-1 is 98% and a gender classification rate is 86%.	It needs a large dataset with annotation, variety in race could affect the results
	[96]	An analysis and evaluation of hormone replacement therapy for classification of gender is presented. The full face is considered for detecting the significant regions. The fusion of texture-based periocular features with patch-based LBP reveals more prominent results.	HRT, Texture-based face matchers, LBP, HOG, and patch-based local binary patterns (p-LBP)	A DB >1.2 M face image from You Tube with HRT for gender transformation from several months to three years.	The evaluated periocular-fused patch-based face matcher outperforms PittPatt SDK v5.2.2 by 76.83% and Cognetic FaceVACS v8.5 by 56.23% for rank-1 accuracy.	Performance is highly dependent on final alteration after effective medical treatment.

NNAODL-Nuclear Norm based adapted Occlusion Dictionary Learning, POcc-Partial Occlusion, PADMA- Personalized Affect Detection with Minimal Annotation, SRC-Sparse Representation Coding, NNMR-Nuclear norm based matrix regression, DRM-Deep Regression Model, BDW- Biorthogonal Discrete Wavelet, AMIL-Association-based Multiple Instance Learning, HOOF-Histograms of oriented optical flow, AAM- Active appearance model, GAN-Generative Adversarial Networks

system. Due to various deformities occurring on the face over time, such as facial hair discoloration, wrinkles, freckles, and sagging, the verification tool can not detect the same person, who enrolled their face biometric a few years ago (say 20 years), resulting in face recognition fails. The wrinkle-oriented active appearance model [177] provides significant results for predicting age over time. However, this research threat has not received enough attention compared to other face identity threats.

The race is area of research involves only analytical observation, but no experimental evidence exists to date as benchmark datasets are not available, and subjects of this race must be from different geographical locations. Gender recognition based on facial features is also lacking due to various reasons such as the non-availability of adequate and benchmark datasets, sensitivity and social exclusion, country-wise laws and regulations, legal recognition for gender change and hormone replacement policy, and facial surgery. The significant parts of gender classification include men and women. The physical structure (i.e., body) and appearance (i.e., face) of men and women differ broadly. Sexual dimorphism is a medical term used to represent males' and females' body shape and morphology. Recent studies show that men and women may differ based on soft tissues, i.e., skin and muscles, and hard tissues like skeletal or skull elements. The race is the most ignorant field in state-of-the-art research due to the unavailability of benchmark datasets till now. Thus, no experimental evidence exists for this significant challenge, although few analytical analysis-based studies for different geographical locations appear in the literature.

5.5 Comparison and discussion on other factors (extrinsic threats)

We have categorized extrinsic threats related countermeasure techniques into five subcategories: pose variation, illumination variation, low resolution, cluttered background, and camera orientation. Table 7 represents a detailed comparison of various state-of-the-art techniques used to minimize the impact of extrinsic factors.

Summary and remarks

Face pose variation is a significant challenge for face recognition systems. Several methods have been proposed to overcome this problem. A thermal image-based method [103] provides the best match score of up to 83% for blood vessel-based analysis. The structural and contextual feature-based method [183] considering Euclidean loss function achieves a mean error of 3.26% and standard deviation of 0.83% for the AR database. Li J., et al. [86] introduced facial attributes prediction model and Wu Y., et al. [172] proposed fiducial facial landmark detection model to identify different poses efficiently. Deep architecture-based methods perform extremely well, especially for the problem of different poses. The deformable FaceNet-based approach achieves a recognition accuracy of 82.11% on the MEGA face challenge dataset.

Several recent works, raising the low-illumination and variable lighting conditions have been introduced. However, no work provides excellent results when dealing with lighting variations. Zhang G., et al.[186] proposed a gradient-based method for an illumination insensitive solution. This method obtains a recognition rate of 99.83% for the PIE database and 98.96% for the Yale B databases. The adaptive homomorphic method for eight local directional patterns [46] obtains a recognition rate of 99.45% with CMU-PIE databases and 96.67% with Yale B databases. Other methods, such as spectral analysis [21] and color-image classification method [141], also achieve good recognition results.

Singular value decomposition [70] and thermal imaging-based approach [108] outperform other recent methods. Various filters segregate the noise factors from the input

Table 7 Comparative analysis of techniques to reduce the impact of extrinsic factors

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
Pose	[23]	Thermal images-based method to recognize the face-biometric using contour and morphology with blood vessel network.	PCA, Bayesian Network	Synthesized DB for multi-pose (thermal face), UMD database	The best matching score is 83%	Fake vascular contours contribute in matching process with poor results
	[18]	A detailed review on various recent methodology and taxonomy under varying face poses is presented.	Low level, motion, shape, 3D, CLM, CQF, AAM	AR, LPFW	CLM,CQF,AAM shows better results among other SOTA approaches	Not effective for heavy occlusion and varying illumination condition.
	[183]	A contextually discriminative feature and structural loss function-based deep approach to detect various face poses.	CNN, Structural, contextual, Euclidian loss	LFW and Net, UMD face	For AR database, Mean error 3.26%, Standard Deviation 0.83%	Not provide good result for yaw displacement.
	[133]	A functional regression solution for the least square problem is introduced to predict shape displacement.	iCCR Algorithm, cascade regression, Monte-Carlo sampling	300-VW dataset	20 times faster, real face tracking as compared to other recent approaches.	Not efficient with Pose variance, illumination, expression.
	[38]	A novel metric learning approach to reduce synthesized variation for single training image. In addition, a multi-depth extended mode of genetic elastic model is developed to handle illumination variations.	3D Multi-depth generic elastic model in association of extension (3D-EGEM), Linear regression	Multi-PIE database	This method obtains average accuracy of 99.3% with Multi-PIE database.	It works on single training image, thus generalization for deep learning.

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[86]	An end-to-end pipeline-based AFFAIR method is proposed to achieve three tasks: learning global transformation, identifying the face location, and merging of local and global features to get robust attribute.	AFFAIR	CelebA, LFWA, MTFL	86.55% Average accuracy among gender, smile, glass and pose	Fixed number of facial point is considering.
	[172]	A review on facial LMD approaches consists of holistic (global facial shape and appearance), CLM- (local appearance), regression-(implicitly capturing of facial shape and appearance).	Holistic, constrained local model (CLM) regression based method.	BioID, AR, Extended Yale-B, FERET, CK/CK+, Multi-PIE, XM2VTSDB	Regression based modal represents the fast and efficient performance among others.	Poor results in extreme head pose, occlusion, strong illumination.
	[58]	A CNN-based DFN model is proposed for recognize the face pose variation. Here, a DCL, ICL, and loss functions are implemented to reduce the intra-class feature variation.	FE-DFN, loss function-DCL and ICL for displacement and identity consistency loss	DFN, MF1, Face scrub dataset	Identification accuracy of DFN on MEGA face challenge 1 is 82.11%	If the pose of the face is more than 60% then it shows poor results.
	[50]	Geometric projection and DL-based coarse-to-fine method is proposed for face pose estimation (i.e., yaw, pitch and roll)	CNN InceptionResNetV2, Geometric Projection	BiWi pointing'4, unconstrained DB AFLW	Classification result for BIWi, Pointing'04 and AFLW datasets are 97.50%, 82.45%, 93.25%, respectively	Errors in some extreme poses are large, results to big deviation

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
illumination	[186]	A theoretical analysis-based novel method to extract illumination insensitive features is introduced under Gradient faces on uncontrolled and natural lighting condition.	Histogram equalization, log-transform, low-curvature image simplifier PCA, LDAMSR, SQI, LTV, Gradient- faces	PIE DB (68 subjects), Yale-B (10 subjects), Outdoor DB (132 subjects)	RR in outdoor and natural light condition for PIE DB, Yale B DB are 99.83% (68 subj), 98.96% (10 subj), and 95.61%, respectively.	Illuminance at each point is considered as smooth, thus not generalized with real practice.
	[21]	Intra-spectral and cross-spectral FR is investigated through SWIR, MWIR, and NIR standoff distances in controlled and uncontrolled scenarios.	FR using PCA, PCA + LDA, BIC, LBP and LTP, DoG	SWIR, MWIR, NIR	SWIR-100%, MWIR-90%. NIR- 80% identification rate	Uncontrolled cross-spectral matching is the main challenge
	[46]	An adaptive harmonic filtering-based method is proposed by utilizing filter stretching and Kirsh compass iin all eight local directions to create illumination invariance.	Low- dimensional linear subspaces, HE, gamma intensity correction, Self-quotient image (SQI). AH-ELDP	CMU-PIE, Yale B, Extended Yale B	RR of 99.45% (CMU-PIE), 96.67% (Yale B) and 84.42% for Extended Yale B face images by considering single image per subject.	Constructing a linear subspace and requiring several sample images for training.
	[141]	The SIFT and state-of-the-art FR methods are analysed based on their performance for hyper spectral images.	LBP, Gabor wavelets, HOG, SVM and SIFT	PolyU-HSFD, CMU-HSFD	The SIFT method outperforms others recent methods for illumination issues.	This method has generalization issue.

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[62]	A logarithm high frequence-based SVD method is proposed to generate face using frequency interoretation. A local-region based nearest negihbor method is deployed to combine discriminative weights (DWs) and Gaussian weights (GWs).	HF-SVD,AHFSVD, DWLNN,GWLNN, FLNN, H& LSVD, SQI, LTV, S& L-LTV, Log-DCT, LBP, TT, Gradient-face, Weber-face, and MSLDE, bipolar sigmoid function	Yale B, CMU PIE, LFW, and self-built driver face databases.	Recognition rate (in %) on the Yale B face DB - DWLNN and GWLNN with best RR 98.10%, 98.73%. average RR for GWLNN-99.97, H&LSVD-GWLNN-99.94, and for drive face DB GWLNN-average RR is 73.89	H&L-SVD is a complex illumination model, GWLNN- is not good for unequal light in small regions
	[176]	A novel mathematically proved method referred as pixel-wise AWFGT is proposed. The LBP feature is separated feature from the weber face to reduce the impact of illumination variation.	AWFGT, intensity transformation without blurring using gamma correction, LBP, k-NN, chi-square	Yale B, CMU-PIE	Recognition rate for Yale B- 99.55%, CMU-PIE-96.63%	It performs on pixel wise operation that shows more time consumption.
LR	[108]	A fast, robust, appearance, and geometric information-based method is proposed to accurately detect low- resolution images using thermal images.	Haar features, Adaboost, Rotation invariant Gaussian distribution, LBP, BRIEF, and SURF	Thermal/visible dataset (X1- Collection) from UND, IRIS Face DB.	Automatic extraction from an Inter-Pupil Distance = 24, 64×64 pixels thermal image. BRIEF signature provides accurate and fast FR	A problem like pose variation is unsolved using this method.

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[70]	Hallucination and recognition-based method with SVD is proposed to handle the low resolution-based input face.	PCA, SVD, ED, Simultaneous Face Hallucination for Verification/ identification (SHV/SHI).	LFW DB, AR	Average PSNR and SSIM for proposed SHV = 22.72, 0.6627, and for SHI= 22.83, 0.6685	It is assumed that two similar faces can have the same local-pixel structure.
	[15]	ICA I (linear face images-original) and ICA II (noisy images) (column vector) architecture are optimized to show the effectiveness of model using five classifiers for five separate benchmark face datasets.	Log-ICA (I & II), LDA, SVM, K-NN, DT, RF	IRIS, FERET, CMU-PIE, USTC-NVIE, Yale, CK, JAFFE Dataset	Except Yale database, log-ICA-II and LDA achieve 59.3%, highest accuracy 89.33%- normal, 85.82% for thermal images.	This method is not suitable for occluded face images.
	[9]	A novel noise robust-SIFT feature descriptor is proposed. The proposed method with two benchmark dataset JAFFE and ORL represents the remarkable performance over existing approaches for face recognition.	SIFT, Laplacian of Gaussian (LoG), Difference of Gaussian (DoG), Euclidean distance	JAFFE and ORL face databases	The noise-robust SIFT technique obtained RR of 88.85% and 91.2% for JAFFE and ORL DB respectively.	The pixel-wise operation is performed, thus its time consuming

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[28]	A preserved slack block-diagonal-based method to show dynamic target structure matrix is proposed. A noise-robust dictionary learning algorithm with two layers (i.e., Laplacian and Gaussian) is utilized by SBD structure represented as SBD^2L.	SBD, SBD^2L,VGG16	AR, Extended Yale B, CMU PIE, Labeled Faces	SBD^2L model achieves the highest RR (worst case is still as high as 60.9%) under different numbers of dictionary atoms.	If numbers of dictionary atoms are too large then recognition result will be low.
	[181]	A CNN-based novel technique to resolve low resolution problem is proposed, which consists of five layers mappings with fourteen high resolution face layers involving non-linear transformation.	DCNN- VGGNet, Back propagation, Optimization- SGD (Stochastic gradient descent)	FERET, LFW, and MBGC datasets	FERET (6×6, 12×12) - 81.4%, 92.1%, LFW (8×8)- 76.3%, MBGC(12×12)- 68.64%, overall 5% improvement in LR.	This performance of this model gradually degrades, if we have very low size images
CB	[93]	The methods that can distinguish face images from sketches involving cluttered backgrounds, noise and deformed images are investigated here. A full CNN (i.e., pFCN) method consists of two stages, first is preprocessing and sketch synthesis and second is feature extraction is investigated.	pFCNN. L1 loss function	Public face sketch DB, Cross DB, CUHK Face Sketch DB(CUFS), AR DB, XM2VTS DB	The average SSIM value for L1-pfCN is 61.78 (for CUHK student dataset). RSLCR is 56.10 (for CUFS dataset), pfCN + RSCLR is 48.04 (for cross dataset)	More complex background or heavy noise can affect the SSIM value.

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[127]	A large benchmark video dataset named Extended Tripura University Video Dataset (ETUVD), consists of complex atmospheric condition for motion objects is introduced.	Bayesian Strategy, Filtering, Histogram Equalization, Learning Strategy	Self- created video dataset ETUVD comprises 147 video clips (each 2-5min long)	This dataset provide more efficient results over 26 other classification method and 04 Deep learning based methods.	Weather degradation may affect the results.
	[168]	Where-What Networks (WWNs)-based technique to simulate the information processing pathway is proposed involving Synapse Maintenance (for background interference) and Neuron Regenesis (for improving the network) considering size, type, and location simultaneously.	WWN-7 model with Hebbian learning rule, receptive fields, update rules, PCA	Simulated scenario with face images (LFW) of 5 types, 11 sizes, and 225 Complex background locations.	For two mechanism RR 0.9960 Location error - 0.9638, size error- 1.0845.	TH angle of the faces, occlusion is not considered here. It consists high computation Complexity.
CO	[16]	Motion analysis-based optimized method with task specific camera placement is discussed to enhance object images for unconstrained or dynamic environment.	PCA, Kalmen filter(Tracking), Least square fitting	Self- created real-time videos from different camera angle	Max percent for Indoor, pedestrian and vehicle, pedestrian only, vehicle only are 99,92,95,97 (In %) respectively	Simulation of real-world environment is taken into consideration that results poor performance in real practice.

Table 7 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[111]	A transformation invariant adversarial light projections conducting real-time damage with feasibility assurance is analyzed. Experiments comprised a webcam and projector to conduct attacks (i.e., impersonation and obfuscation).	Multitask CNN-based FD and landmark estimation method for FaceNet, SphearFace Commercial face. cosine distance metric, fusion function	Two open-source, and one commercial FR DB (50 subjects in each case)	FaceNet and SphereFace is suitable for all white-box obfuscation attempts, while black-box setting succeeded 7 out of 10 attempts on commercial face.	The camera adjustment or view point is highly correlated with lighting condition.

LR- Low resolution, CB-Cluttered Background, CO-Camera Orientation, DFN - Deformable FaceNet, AFFAIR- lAndmark Free Face AttrIbute pRediction, CQF- Convex Quadratic Fitting, AWFGT- adaptive Weber face-based gamma transformation, MF1- Megaface challenge 1

samples, and results represent a low-resolution invariant outcome with better recognition rates.

A self-organizing map-based method is introduced in state-of-the-art research to identify the foreground objects with reasonable accuracy through rearranging the neurons with the highest receptive fields. The feature-based techniques, such as constellation (for statistical and probabilistic analysis) and Viola Jones (for Adaboosting and Cascading), are implemented and extensively analyzed in [168] to tackle cluttered background problems. Handling dynamic background and shadow of the image are still challenging tasks due to the absence of any feasible benchmark dataset related to static or dynamic background. However, background subtraction and frame differencing are the most reliable techniques to detect the motion in the object.

Some of the significant adjustment features that may affect the performance of the face recognition system can be divided into specification-based and physical-based. Specification-based includes the configuration of the camera, such as the model, lens quality, shutter speed, aperture, and resolution for capturing images. The physical-based features include the camera's location, the distance of an object from the camera, alignment (i.e., direction), height, and other factors. All the mentioned parameters including camera's position should be effectively maintained. Deployment of multiple cameras at a predetermined location and aggregation of multiple observations with different functions can solve this problem. However, this approach is not feasible for time, maintenance cost, and unpredictable environmental concerns. An additional light projection-based method can be used for the transformation invariant adversarial pattern generation.

5.6 Comparison and discussion on other factors (extrinsic threats)

Sometimes, more than one face identity threats occur simultaneously, thus causing a huge impact on results. Researchers noticed this observation and proposed methods to effectively resolve modular (i.e., multiple) type of face identity threats effectively. Table 8, represents a detailed comparative analysis of state-of-the-art techniques, which are used to reduce the effect of multiple identity threats from the hierarchy as depicted in Fig. 4 of Section 3. Here, we are only describing some techniques based on performance relevance in tabular form .

Summary and remarks

The face recognition system is adoptable if it is effective in an unconstrained environment. More obviously, a face recognition system must be capable enough to deal with multiple challenges at once. Some researchers proposed a robust face recognition system that achieved efficient results even when multiple face identity threats were present. Wu, Yue et al. [171] proposed an integrated and robust facial landmark detection method for occlusion and head pose estimation using SIFT feature descriptor with minimum average error for the multi-PIE datasets. Jang, Jinhyeok, et al. [67] presented a recurrent neural network-based method to distinguish various facial expressions with genders and age variations through visual effects. This method achieved an average accuracy of 91.36%. Shi Y. et al. [142] investigated various facial expressions based on variable illumination and poses. The SIFT, PCA, and SVM methods are utilized here for feature extraction, dimensionality reduction, and classification, respectively. This method receives the best recognition accuracies of 98.52% for dry faces without additional light and 96.97% for dry faces with two additional light sources. Duan, Qingyan, and Lei Zhang [43] introduced a boosting GAN (BoostGAN) method for large-pose variations and corrupted regions, considering partial, incomplete, and patch-based occlusions on the face. This method provides the best recognition rate of 99.48% for single point occlusion and 99.45% for multipoint occlusions.

Table 8 Comparative analysis of techniques to mitigate modular(multiple) face identity threats

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
Multiple FIT	[112]	Visual cues and head motion-based method for sign language prediction using HMM under partial occlusion and facial expression recognition is proposed.	FE-BN, PPCA, HMM, CLS- NN, Euclidean distance	Self-created video sequences with motion face feature, facial expression dataset	The best RR of 74% is obtained with the Gold tracks, while Bayes tracks and KLT tracks yielded 63% and 59%, respectively.	This face tracking tool is unable to detect the face in fast motions, and heavy occlusions
	[60]	Different feature extraction methods such as LBP, WLD GJD, SIFT, SURF for FR in unconstrained environment are investigated	LBP, WLD, SIFT, SURF & GJD	Equinox and UCH Thermal FF DB	The WLD method outperforms the other methods.	GJD has low RR in outdoor setups represents poor generalization.
	[171]	The facial landmark-based unified and robust methods for occlusion and head pose estimation are proposed, referred as FLMD, HPE, and FDF.	SIFT, Regression	BU4DFE, BU, COFW, Multi -PIE (with self-occlusion) databases	Average mean absolute error for HPE-4.4(BU), FLMD- 6.4 (COFW), FLMD- 3.5(Multi-PIE), Overall Normalized error-6.4 (COFW)	Cascading of multiple processes creates a big generalization issue.
	[67]	A recurrent deep learning-based method to predict facial expression, age, and gender through mimicking the human's activity using visual effects is proposed.	CNN+RNN+ spatiotemporal manifold	Facial Expression, Gender and Age dataset	91.36% overall average accuracy	It does not support video based face attribute recognition.

Table 8 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[81]	The three significant issues are identified and countermeasures are proposed for one image per subject, face occlusion, and various facial expressions.	Distance matrix, Thresholding, Affine-wrapping	ORL	NA	Modularity of occlusion with FER affects the results drastically.
	[154]	A method to detect the facial landmarks in 3D shape using features and model based technique is proposed.	Active Normal Model with CNN, 3D morphable model, cascade regression	BU-3DFE, Bosphorus, BU-4DFE	Overall Mean error is 3.33%, and Standard Deviation is 2.08%	Model free tracking and object localization are the practical problems.
	[142]	A SIFT feature extraction technique to extract regional gradient information along with SVM and PCA-based classification for facial expressions is proposed.	SIFT, SVM with PCA	Video-based CK database for FER, 15 volunteers are considered with various illumination, poses, and facial moisture level.	The average recognition accuracies are 98.52% and 94.47% (no additional light sources), and 96.97% and 95.40% (two additional light sources), respectively. Dry face recognition outperforms wet faces.	This method is not suitable for pose variation above 15% in yaw, pitch and roll direction.
	[87]	A method consisting of three tasks, first is to detect edge and color feature (histogram), second is enhancement (self-adaptive feature fusion), and third is to upgrade the object model (drifting) is proposed.	Fusion of Color, edge orientation histogram and self-adaptive Bayesian estimation	Self-created dataset	Running time results for the task 1 and 3 are 61.6 and 109.9, respectively	The obtained results were found good only when the initial template is well defined.

Table 8 (continued)

Type/Focus on	Ref.	Concept	Methodology Used	Dataset Used	Performance	Limitation
	[140]	A deep orientation-based method to estimate four different tasks (i.e., voxel, occlusion invariance, 3D mesh and landmark) is introduced with triplet loss to predict the results.	Locality Preseving Projection, GAN, Autoencoder	KinectFace, Bosphorus, and UMB-DB	This approach obtains 86.1%, 75.5, 81.3%, 83.9% accuracy for voxel, occlusion, landmark, and with 3D mesh, respectively.	Noise and micro expression analysis cannot be detected
	[43]	An efficient, effective boosting-GAN method for large-pose variations and corrupted regions (e.g., nose and eyes) is proposed with two consideration, first is that occlusion is partial, incomplete, and patch-based and the second is an encoder-decoder network (coarse face synthesis) and boosting(face generation) with an aggregation structure.	BoostGAN	Multi-PIE, LFW	RR for single point and random multipoint occlusion with 15% face pose variation are 99.48%, 99.45% respectively	The occlusion by different objects like sunglass, scarf, and mask are not considered.
	[167]	A CNN-based method named Region Attention Network is proposed to identify occluded face regions having variation in pose and facial expressions.	CNN (Region Attention Network)	AffectNet, SFEW, FER-Plus, and RAF-DB.	RAN performance Occlusion 83.63%, pose (30, 45 degree) 82.23%, 80.40%, respectively. for RAFDB 86.90%	Biased loss is calculated that is more prone to errors.

FIT- Facial Identity Threats, WLD-Weber Linear Descriptors, GJD- Gabor Jet Descriptors

However, this method is not generalized as it excludes scarves, sunglasses, and masks types of occlusions.

5.7 Comparison and discussion on various dataset used in facial identity threats

Datasets play a vital role in training the machine for finding efficient solutions to any face identity threat. There are various adequate and benchmark datasets available in the literature. This survey provides a comparative description of these significant datasets. Table 9 denotes the exclusive features of the benchmark face dataset, including dataset name, specification, applicability or attack type, references, and the source links to download these datasets. The specification (i.e., the second column) of these datasets provides further information about the color channel, pattern (image or video), image size (training and testing), number of subjects, and samples per subject. This paper organizes these datasets in a systematic order of applicability, although some datasets are modular and can be used for more than one attack. The source links given in this table are accessible at this time, although this may be changed or removed in the future.

Summary and remarks

Most machine learning and deep learning-based face recognition algorithms are trained on different face datasets to detect faces in the unconstrained environment, such as different expressions, poses, occlusions, illuminations, resolutions, and more. To further investigate the suitability and effectiveness of countermeasure techniques for the identified threats, we provide a comparative study of different benchmark face datasets, including their specifications in Table 9. The Plastic Surgery, the Muct dataset, FaceScrub, Eurecom Facial Cosmetic Dataset, MIW, and DFW datasets are primarily utilized for training the machine against a direct spoofing attack. The NUAA, Print-Attack, YMU, Oulu-NPU, Replay-Attack, MSU-MFSD, and CASIA-FASD datasets are used to train the model against indirect face spoofing attacks. The UND-Twins (2009-2010) dataset is the only reliable dataset to distinguish the twin pairs. However, no benchmark dataset is found in the literature rather a synthesized dataset is generated through various software and mobile APIs to tackle the face morphing threat efficiently. Other factors that lead to the failure of facial recognition systems are handled by the countermeasure techniques using various datasets such as JAFFE, BU-4DFE, BU-3DFE, CK/CK+, and FER-2013 for various facial expressions, FMD for partial face occlusions, LFWcrop, and FEI for pose variations. Some benchmark modular datasets such as AR, FERET, The Yale Face Database, Extended Yale B, CMU Multi-PIE, BP-4D, EURECOM KinectFace DB, and more are considered to have more than one threat.

6 Discussion

In this paper, a detailed hierarchy of different potential face identity threats is addressed and analyzed sequentially with their appropriate countermeasure approaches. Section 1 briefs the significance of the facial recognition system, the main steps of the face recognition process, characteristics, and usability. Then, the motivation for this survey paper is mentioned, which indicates how these threats can be dangerous to public safety and security systems if not adequately addressed. We also noted the innovative contribution to writing this survey, followed by the structured workflow of the paper. In Section 2, we presented a complete hierarchy of traditional face recognition approaches, including appearance-based, feature-based, model-based, and a combination of multiple approaches, i.e., hybrid-based, including their brief insights. In section 3, we proposed a new taxonomy for potential face

Table 9 Tabular representation of various datasets used in countermeasure techniques to handle face identity threats

	Specification of Dataset							
Dataset Name	Channel	Pattern (I/V)	Size (Tr/T)	Subjects	Samples	Applicable	Ref.	Source Link
JAFFE	Gray	Image	256×256	10 F	213 Total	Exp (6 basic) frontal face	[10, 103]	https://www.kasrl.org/jaffe_download.html
AR	RGB, Gray	Image	576×468	126 (70M, 56F)	3276 Total	Occ (sunglass, scarf), ill, Exp	[36, 94]	https://www2.ece.ohio-state.edu/~aleix/ARdatabase.html
BU-4DFE	RGB	Image, Video	3D	101 (58F, 43M)	60K Total	Exp (6 Basic)	[154, 171]	https://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html
BU-4DFE	RGB	Image, Video	3D	101 (58F, 43M)	60K Total	Exp (6 Basic)	[154, 171]	https://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html
FERET	Gray	Images	256×384	30	14,051 Tota 1	ill, Exp, Po, Occ variations	[15, 172]	http://www.itl.mist.gov/iad/humanid/feret/

Table 9 (continued)

	Specification of Dataset							
Dataset Name	Channel	Pattern (I/V)	Size (Tr/T)	Subjects	Samples	Applicable	Ref.	Source Link
CK/CK+	Gray	Video(325 clips)	640×480	97(123-CK+)	593 Total in CK+	Exp (6 Basic), non-Occ, frontal face, ill	[142, 172]	https://www.kaggle.com/shawon10/ck-facial-expression-detection/data
The Yale Face	Gray	Images	320×243	15 (14 M, 1 F)	165 Total	ill, Exp	[176, 186]	http://vision.ucsd.edu/datasets/yale_face_dataset_original/
Extended Yale B	Gray	Images	640 × 480	28	16128 Total	Po (9), ill(64)	[46]	http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/ExtYaleB.html
FER-2013	Gray	Images	48×48	NA	35887 Total	Exp (7 universal)	[54]	https://www.kaggle.com/msambare/fer2013
DFW	RGB	Images	400/600	1000	11157 Total	Normal, disguised, and impersonation	[148]	http://iab-rubric.org/DFW/2018.html
BOSPHORUS	RGB, Gray	Images	3D	105	4666 Total	Exp, Occ(facial hair,hand, eyeglass), Po	[32, 154]	http://bosphorus.ee.boun.edu.tr/default.aspx

Table 9 (continued)

Dataset Name	Specification of Dataset				Samples	Applicable	Ref.	Source Link
	Channel	Pattern (I/V)	Size (Tr/T)	Subjects				
UMB	RGB	2D, 3D	12, 24, 36, 48 pixels scale	143	1473 Total	4 Exp, Occ (scarf, hat, eyeglass, hand)	[32]	http://www.ivl.disco.unimib.it/minisites/umbdb/
BioID	Gray	Image	384 × 286	23	1521 Total	Synthesized set of eye positions	[172]	https://www.bioid.com/facedb/
EURECOM KFDB	RGB	Images	256×256 Depth	52 (38M, 14F)	427 images	ill, Occ, Exp (9)	[32, 140]	http://rgb-d.eurecom.fr/
EURECOM FCDB	RGB	Makeup, non-makeup	150×140 face	50	389 images	MU	[44]	http://fcd.eurecom.fr/
FaceScrub	RGB	Images	80/20	530 (265M, 265F)	1,07,818 Total	Face alignment, MU and FR	[3]	http://vintage.winklerbros.net/facescrub.html
ORL	Gray	Images	92 × 112	40	400 Total	ill, Occ, Exp	[9, 81]	http://nimfa.biolab.si/nimfa.examples.orl_images.html
NUAA	RGB	Images	640×480	15	5105 valid,7509 PAD	Face artefacts generation, PAD	[26, 138]	http://parnec.nuaa.edu.cn/_upload/tpl/02/db/731/template731/pages/xtan/NUAAImposterDB_download.html

Table 9 (continued)

	Specification of Dataset							
Dataset Name	Channel	Pattern (I/V)	Size (Tr/T)	Subjects	Samples	Applicable	Ref.	Source Link
CASIA FASD	RGB	Photo, Video	640×480, 1280×720, 1920×1050	50	600 Total 150 (RA) 450 (SA)	PA, VA	[20, 27]	https://pypi.org/project/bob.db.casia-fasd/
Print attack	RGB	Videos	320×240	50	200 samples	2BG, ill(Controlled, adverse), PAD	[138]	https://www.idiap.ch/en/dataset/printattack
MSU MFSD	RGB	Video Clips	640×480, 720×480	35	280 (P+V)	Reflection, Blurriness, PA,RA	[20, 27]	https://pythonhosted.org/bob.db.msu_mfsd_mod/
Oulu-NPU	RGB	Video Clips	720×480	55 (40M, 15 F)	360(RA), 720 (PA), 720(VA)	Print PA, Replay VA	[123]	https://www.idiap.ch/software/bob/docs/bob/bob.db.oulunpu/master/index.html
REPLAY ATTACK	RGB	Video Clips	320×240	50	1300 (300 (RA), 1000 (SA))	Printed, digital photo, Replayed video for PAD	[151, 178]	https://www.idiap.ch/en/dataset/replayattack
300-VW dataset	RGB	Video Clips	128×128	300	1 minute video (25 - 30 FPS)	Lm	[133]	https://ibug.doc.ic.ac.uk/resources/300-VW/
BP-4D	RGB + Gray	Video clips (431)	1040×1392	41(23 F, 18 M)	Total 2952 (1080)	27 AU, Occ, Po, Exp (8), 2D/3D Lm (83)	[117, 177]	http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html

Table 9 (continued)

	Specification of Dataset							
Dataset Name	Channel	Pattern (I/V)	Size (Tr/T)	Subjects	Samples	Applicable	Ref.	Source Link
CMU Multi-PIE	RGB	Images	3072×2048	337	Exp-19, ill-15	profile face, Exp, Occ, Po	[62, 63]	https://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html
Caltech Faces	Gray	Images	304×312	27	10524 Total	ill, Exp, BG	[51]	http://www.vision.caltech.edu/Image_Datasets/Caltech_10K_WebFaces/
PSFD	RGB	Images	163×131 to 288 × 496	900	Total 1800 (pre/post-surgery)	Texture and structure-based surgery	[71, 149]	http://iiitd.edu.in/iab/Image_Analysis_and_Biometrics_Group/Resources.html
Muct database	RGB	Images	480×640	50	500 Total	Texture-based Plastic surgery	[8]	http://www.milbo.org/muct/
YMU	RGB to gray	Makeup images	128×128	151	Total 600 (with and with-out makeup)	MU, disguised, spoofing, PAD	[25, 26]	http://iprobe.cse.msu.edu/dataset_detail.php?id=3&?title=YouTube_Makeup_Dataset_(YMU)

Table 9 (continued)

Dataset Name	Specification of Dataset				Samples	Applicable	Ref.	Source Link
	Channel	Pattern (I/V)	Size (Tr/T)	Subjects				
MIW	RGB	Internet image	Patch 16×16	125	154 (with/without makeup)	MU, disguised, spoofing, PAD	[85]	http://iprobe.cse.msu.edu/dataset_detail.php?id=5&?title=Makeup_in_the_Wild_(MiW)
UND-Twins	RGB	Colored image	Random	435	24050 Total	Identical twins	[82]	https://cvrl.nd.edu/projects/data/#nd-twins-2009-2010
FMD	RGB	Images	Random	Group-based	853	Partial Occ (COVID-19)	[92]	https://www.kaggle.com/andrewmvd/face-mask-detection
FEI	RGB	Images	640×480	200	2800 Total	Po, hairstyle, adorns	[161]	http://fei.edu.br/~cet/facedatabase.html
LFWcrop	RGB, Gray	Images	64×64	1680	13000 Total	Po, scale variations	[64]	https://conradsanderson.id.au/lfwcrop/

Occ-Occlusion, Po-Pose, ill-Illumination, Exp-Expression, AU–Action Unit, Lm-Landmark, BG-Background, M-Male, F-Female, P-Photo, V-Video, PA-Photo Attack, VA-Video Attack, RA- Replay-video Attack, SA- Spoofing Attack, FPS-Frames Per Second, DB-Database, N-Normal, D-Disguised, Im-Impersonator, MU-Makeup, NUAA-Nanjing University of Aeronautics and Astronautics, CASIA FASD-Chinese Academy of Science Face Anti-Spoofing Database, MSU MFSD- Michigan State University-Mobile Face Presentation Attack, YMU-YouTube makeup database, MIW-Makeup in the wild database, DFW-Disguised Faces in the Wild dataset, FMD- Face mask detection dataset, CK- Cohn-Kanade, FEI- Educational Foundation of Ignatius, LFW- Labeled Faces in the Wild, UND- Twins- University of Notre Dame Twins dataset (2009-2010), KFDB-KinectFace DB, FCDB-Facial Cosmetics DB, MUCT-Milborrw University of Cape Town

identity threats with a brief description of each threat in a systematic manner. The impact of these threats on the system's performance is primarily focused and deeply analyzed in this survey. Section 4 provides a brief overview of the various state-of-the-art countermeasure techniques developed to mitigate the effects of these threats to facial recognition. In Section 5, the over 120 research articles and review papers are explored extensively to conduct comparative analysis based on various parameters such as author's contribution, methodology used for feature extraction and classification, the dataset used, the performance measures, and the limitations.

State-of-the-art countermeasure approaches are compared in a separate table for each identified face identity threat, as shown in Fig. 4 of Section 3. A comparison of effective approaches for handling modularity-based threats in face recognition is also given in Table 8. In addition, we also highlight the exclusive information of the various available benchmark datasets in the tabular form used to handle these facial recognition threats. This section is an essential part of this survey paper. The critical summary points of this study are as follows:

- The most significant face identity threats are face spoofing, partial face occlusion, face pose variation, various facial expressions, and illumination/ lighting effects, which can drastically affect the system's performance.
- Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Support Vector Machines (SVM) are more robust machine learning methods for feature extraction and classification to address the above face identity threats.
- Local Binary Pattern (LBP), and other variants of LBP such as circular-LBP, Densely Sampled-LBP, Rotation Invariant-LBP, uniform-LBP, Co-occurrence of adjacent-LBP, Patch-based-LBP, Extended-LBP, and is more widely used for feature extraction purposes. These methods provide the best solutions for facial recognition problems.
- Other efficient methods for resolving face identity threats are Gabor-wavelet (i.e., powerful feature descriptor), k-NN (unsupervised clustering algorithm), Discrete Cosine Transform, and Euclidean Distance.
- Deep Convolutional Neural Network (DCNN) and variants such as Scattered- CNN and Multi-Task- CNN provides a robust solution. However, CNN also supports transfer learning techniques to utilize the learning weights from a pre-trained model to achieve significant results with a large dataset quickly. The best transfer learning techniques to handle face identity threats are MobileNetV2, VGG16, VGG19, InceptionResNetV2, InceptionV3, ResNet50, DenseNet121, DenseNet169, and DenseNet201.
- Benchmark dataset FERET is a multipurpose dataset containing a total of 14051 images for the occlusion, expressions, illumination, and pose variations problems.
- Other benchmark datasets include CMU-Multi-PIE (illumination and expression), Extended Yale B (Different poses and illumination), DFW (impersonation and disguised face), BU-4DFE (expression and facial landmarks), AR (facial landmarks, occlusions, illumination), FER-2013 (universal facial expressions), and BOSPHORUS (face poses, expressions, and occlusion) in the form of images and videos.
- Plastic surgery database is a single benchmark dataset consisting of 900 subjects and a total of 1800 images for pre-surgery and post-surgery. Due to safety and security concerns for the patients and the physicians, no other datasets have been created to date.

This section summarizes what we have done so far with the key highlights of the survey paper. This survey manuscript provides a concise, crisp, and efficient analysis of various facial identity threats. We hope this article will be of great help to scholars and others in

their future research work in the domains of computer vision, machine learning, and deep learning. Researchers and readers can benefit from this paper for their future research work.

7 Conclusion and future scope

For the last decade, facial recognition has been plagued by various challenges such as face spoofing, occlusion, various expressions, poses, and many more that lead to a drastic reduction in the system's performance. Finding these challenges has been a consequential and influential research topic in the field of computer vision. However, earlier generic surveys and state-of-the-art research have discussed only one or more of the three challenges associated with the face recognition approach. This survey paper proposes a new taxonomy to face-based potential identity threats, including their essential aspects, a concise and crisp detailing of countermeasure methodologies based on dimensionality reduction, feature extraction, classification, and neural network techniques. A comprehensive analysis of various countermeasure techniques is compared on the ground of concept, methodology, datasets, performance, and limitations for each identified threat category. Direct spoofing, zero effort imposter, and intrinsic interpersonal factors such as race and gender have attracted less attention from research communities. The unavailability of adequate and benchmark datasets due to security, geographical regions, laws, and social exclusion is the main reason for less research in this domain. However, these threats are rarely used in practice due to being expensive and requiring technical expertise on the imposter's end. Therefore, only theoretical and analytical observations are effectively discussed in the literature as these threats have no experimental evidence.

In contrast, indirect spoofing, and other factors (intrinsic and extrinsic factors), have attracted the primary attention of researchers due to the fact that it is simple, commonly used, and requires minimum technical skill. We found that some appearance-based machine learning approaches such as PCA, LDA, and SVM provide robust solutions to handle these face identity threats. The feature-based LBP techniques with their different variants, Gabor wavelet, k-NN, DCT, and Euclidean distance, are also preferred by various researchers. Deep architecture-based techniques (i.e., DCNN) have always been considered a fast and efficient method to deal with large databases. This paper covers various deep architectures such as Scattered- CNN and Multi-Task-CNN, including transfer learning methods such as MobileNetV2, VGGNet16, InceptionResNetV2, and others. A tabular comparison of modular (multiple) threat-based techniques to represent the impact on face recognition is also discussed. We also added a summary and remarks after each comparison table to understand the insights. Apart from these, we also highlight the exclusive details of various available benchmark face datasets needed to handle these facial recognition threats as the synthesized dataset may not reflect real-world scenarios. This survey found that the FERET dataset is a multi-purpose dataset used to provide the solution for a variety of threats such as occlusion, pose and expression variation, and illumination. The other effective datasets to deal with face identity threats could be DFW, AR, BU-4DFE, CMU- Multi-PIE, BOSPHORUS, FER-2013, and Extended Yale B databases. Further, we discuss the research opportunities to researchers and readers for their future research work. We hope that this paper will help researchers and readers to further expand research in this domain. In the future, we will endeavour to provide a robust, computationally efficient, and real-time-based solution approach to tackle identified facial identity threats.

Declarations

Conflict of Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Abate AF, Nappi M, Riccio D, Sabatino G (2007) 2d and 3d Face recognition: a survey. Pattern Recognit Lett 28(14):1885–1906
2. Agarwal V, Bhanot S (2018) Radial basis function neural network-based face recognition using firefly algorithm. Neural Comput Applic 30(8):2643–2660
3. Al Jazaery M, Guo G (2019) Automated cleaning of identity label noise in a large face dataset with quality control. IET Biometrics 9(1):25–30
4. Al-Dabagh MZN, Alhabib MM, Al-Mukhtar F (2018) Face recognition system based on kernel discriminant analysis, k-nearest neighbor and support vector machine. Int J Eng Res 5(3):335–338
5. Albiol A, Monzo D, Martin A, Sastre J, Albiol A (2008) Face recognition using hog–ebgm. Pattern Recogn Lett 29(10):1537–1543
6. Ali ASO, Sagayan V, Malik A, Aziz A (2016) Proposed face recognition system after plastic surgery. IET Comput Vis 10(5):344–350
7. Ansari M, Singh DK et al (2021) Human detection techniques for real time surveillance: a comprehensive survey. Multimed Tools Appl 80(6):8759–8808
8. Anzar S, Amrutha T (2020) Efficient wavelet based scale invariant feature transform for partial face recognition. In: AIP conference proceedings, AIP publishing LLC, vol 2222, p 030017
9. Arya K, Rajput SS, Upadhyay S (2019) Noise-robust low-resolution face recognition using sift features. In: Computational intelligence: theories, Applications and Future Directions-Volume II. Springer, pp 645–655
10. Ashir AM, Eleyan A, Akdemir B (2020) Facial expression recognition with dynamic cascaded classifier. Neural Comput Applic 32(10):6295–6309
11. Astawa I, Putra I, Sudarma IM, Hartati RS (2017) The impact of color space and intensity normalization to face detection performance. TELKOMNIKA (Telecommunication Comput Electron Control [Internet] 15(4):1894
12. Bahreini K, van der Vegt W, Westera W (2019) A fuzzy logic approach to reliable real-time recognition of facial emotions. Multimed Tools Appl 78(14):18943–18966
13. Ballihi L, Amor BB, Daoudi M, Srivastava A, Aboutajdine D (2012) Boosting 3-d-geometric features for efficient face recognition and gender classification. IEEE Trans Inf Forensics Secur 7(6):1766–1779
14. Bargshady G, Zhou X, Deo RC, Soar J, Whittaker F, Wang H (2020) The modeling of human facial pain intensity based on temporal convolutional networks trained with video frames in hsv color space. Appl Soft Comput 97:106805
15. Bhowmik MK, Saha P, Singha A, Bhattacharjee D, Dutta P (2019) Enhancement of robustness of face recognition system through reduced gaussianity in log-ica. Expert Syst Appl 116:96–107
16. Bodor R, Drenner A, Schrater P, Papanikolopoulos N (2007) Optimal camera placement for automated surveillance tasks. J Intell Robot Syst 50(3):257–295
17. Bolle RM, Connell JH, Pankanti S, Ratha NK, Senior AW (2013) Guide to biometrics. Springer Science & Business Media
18. Borude PR, Gandhe S, Dhulekar PA, Phade G (2015) Identification and tracking of facial features. Procedia Comput Sci 49:2–10
19. Bouguila J, Khochtali H (2020) Facial plastic surgery and face recognition algorithms: interaction and challenges. A scoping review and future directions. Journal of stomatology, oral and maxillofacial surgery
20. Boulkenafet Z, Komulainen J, Hadid A (2016) Face spoofing detection using colour texture analysis. IEEE Trans Inf Forensics Secur 11(8):1818–1830
21. Bourlai T, Cukic B (2012) Multi-spectral face recognition: identification of people in difficult environments. In: 2012 IEEE international conference on intelligence and security informatics, IEEE, pp 196–201
22. Bowyer KW, Chang K, Flynn P (2006) A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. Comput Vis Image Underst 101(1):1–15
23. Buddharaju P, Pavlidis IT, Tsiamyrtzis P, Bazakos M (2007) Physiology-based face recognition in the thermal infrared spectrum. IEEE Trans Pattern Anal Mach Intell 29(4):613–626

24. Chang H, Lu J, Yu F, Finkelstein A (2018) Pairedcyclegan: asymmetric style transfer for applying and removing makeup. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 40–48
25. Chen C, Dantcheva A, Ross A (2013) Automatic facial makeup detection with application in face recognition. In: 2013 International conference on biometrics (ICB), IEEE, pp 1–8
26. Chen C, Dantcheva A, Ross A (2016) An ensemble of patch-based subspaces for makeup-robust face recognition. Inf fusion 32:80–92
27. Chen FM, Wen C, Xie K, Wen FQ, Sheng GQ, Tang XG (2019) Face liveness detection: fusing colour texture feature and deep feature. IET Biometrics 8(6):369–377
28. Chen Z, Wu XJ, Yin HF, Kittler J (2020) Noise-robust dictionary learning with slack block-diagonal structure for face recognition. Pattern Recogn 100:107118
29. Choi J, Hu S, Young SS, Davis LS (2012) Thermal to visible face recognition. In: Sensing Technologies for global health, military medicine, disaster response, and environmental monitoring II; and biometric technology for human identification IX, international society for optics and photonics, vol 8371, p 83711L
30. Chowdary GJ, Punn NS, Sonbhadra SK, Agarwal S (2020) Face mask detection using transfer learning of inceptionv3. In: International conference on big data analytics. Springer, pp 81–90
31. Dadi HS, Pillutla GM (2016) Improved face recognition rate using hog features and svm classifier. IOSR J Electron Electr Commun Eng 11(04):34–44
32. Dagnes N, Vezzetti E, Marcolin F, Tornincasa S (2018) Occlusion detection and restoration techniques for 3d face recognition: a literature review. Mach Vis Appl 29(5):789–813
33. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE Computer society conference on computer vision and pattern recognition (CVPR'05), ieee, vol 1, pp 886-893
34. Dantcheva A, Chen C, Ross A (2012) Can facial cosmetics affect the matching accuracy of face recognition systems? In: 2012 IEEE Fifth international conference on biometrics: theory, applications and systems (BTAS), IEEE, pp 391–398
35. De Freitas Pereira T, Komulainen J, Anjos A, De Martino JM, Hadid A, Pietikäinen M, Marcel S (2014) Face liveness detection using dynamic texture. EURASIP J Image and Video Process 2014(1):1–15
36. De Marsico M, Nappi M, Riccio D, Wechsler H (2015) Robust face recognition after plastic surgery using region-based approaches. Pattern Recogn 48(4):1261–1276
37. Deeb A, Roy K, Edoh KD (2020) Drone-based face recognition using deep learning. In: International conference on advanced machine learning technologies and applications. Springer, pp 197–206
38. Deng W, Hu J, Wu Z, Guo J (2018) From one to many: pose-aware metric learning for single-sample face recognition. Pattern Recogn 77:426–437
39. Dey SK, Howlader A, Deb C (2021) Mobilenet mask: a multi-phase face mask detection model to prevent person-to-person transmission of sars-cov-2. In: Proceedings of international conference on trends in computational and cognitive engineering. Springer, pp 603–613
40. Dhekane M, Seal A, Khanna P (2017) Illumination and expression invariant face recognition. Int J Pattern Recognit Artif Intell 31(12):1756018
41. Ding C, Tao D (2016) A comprehensive survey on pose-invariant face recognition. ACM Trans Intell Syst Technol (TIST) 7(3):1–42
42. Du L, Hu H (2019) Nuclear norm based adapted occlusion dictionary learning for face recognition with occlusion and illumination changes. Neurocomputing 340:133–144
43. Duan Q, Zhang L (2020) Look more into occlusion: realistic face frontalization and recognition with boostgan. IEEE transactions on neural networks and learning systems
44. Eckert ML, Kose N, Dugelay JL (2013) Facial cosmetics database and impact analysis on automatic face recognition. In: 2013 IEEE 15Th International workshop on multimedia signal processing (MMSP), IEEE pp 434–439
45. El-Said SA, Abol Atta HM (2014) Geometrical face recognition after plastic surgery. Int J Comput Appl Technol 49(3-4):352–364
46. Faraji MR, Qi X (2014) Face recognition under varying illumination based on adaptive homomorphic eight local directional patterns. IET Comput Vis 9(3):390–399
47. Fassold H, Rosner J (2015) A real-time gpu implementation of the sift algorithm for large-scale video analysis tasks. In: Real-Time Image and Video Processing 2015, International Society for Optics and Photonics, vol 9400, p 940007
48. Feng Y, Yuen PC, Jain AK (2008) A hybrid approach for face template protection. In: Biometric Technology for Human Identification V, International Society for Optics and Photonics, vol 6944, p 694408

49. Fourati E, Elloumi W, Chetouani A (2020) Anti-spoofing in face recognition-based biometric authentication using image quality assessment. Multimed Tools Appl 79(1):865–889
50. Gao F, Li S, Lu S (2020) How frontal is a face? quantitative estimation of face pose based on cnn and geometric projection. Neural Comput Applic, PP 1–17
51. Garcia DC, de Queiroz RL (2015) Face-spoofing 2d-detection based on moiré-pattern analysis. IEEE Trans Inf Forensics Secur 10(4):778–786
52. George A, Mostaani Z, Geissenbuhler D, Nikisins O, Anjos A, Marcel S (2019) Biometric face presentation attack detection with multi-channel convolutional neural network. IEEE Trans Inf Forensics Secur 15:42–55
53. Gerig T, Morel-Forster A, Blumer C, Egger B, Luthi M, Schönborn S, Vetter T (2018) Morphable face models-an open framework. In: 2018 13th IEEE International conference on automatic face & gesture recognition (FG 2018), IEEE, pp 75–82
54. Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B, Cukierski W, Tang Y, Thaler D, Lee DH et al (2013) Challenges in representation learning: a report on three machine learning contests. In: International conference on neural information processing. Springer, pp 117–124
55. Guo G, Zhang N (2019) A survey on deep learning based face recognition. Comput Vis Image Underst 102805:189
56. Gupta K, Walia GS, Sharma K (2020) Quality based adaptive score fusion approach for multimodal biometric system. Appl Intell 50(4):1086–1099
57. Hancock KJ, Rhodes G (2008) Contact, configural coding and the other-race effect in face recognition. Br J Psychol 99(1):45–56
58. He M, Zhang J, Shan S, Kan M, Chen X (2019) Deformable face net: learning pose invariant feature with pose aware feature alignment for face recognition. In: 2019 14th IEEE International conference on automatic face & gesture recognition (FG 2019), IEEE, pp 1–8
59. Hermosilla G, Ruiz-del Solar J, Verschae R (2017) An enhanced representation of thermal faces for improving local appearance-based face recognition. Intell Autom Soft Comput 23(1):1–12
60. Hermosilla G, Ruiz-del Solar J, Verschae R, Correa M (2012) A comparative study of thermal face recognition methods in unconstrained environments. Pattern Recogn 45(7):2445–2459
61. Hernandez-Ortega J, Fierrez J, Morales A, Galbally J (2019) Introduction to face presentation attack detection. In: Handbook of Biometric Anti-Spoofing. Springer, pp 187–206
62. Hu C, Lu X, Ye M, Zeng W (2017) Singular value decomposition and local near neighbors for face recognition under varying illumination. Pattern Recogn 64:60–83
63. Huang MX, Ngai G, Hua KA, Chan SC, Leong HV (2015) Identifying user-specific facial affects from spontaneous expressions with minimal annotation. IEEE Trans Affect Comput 7(4):360–373
64. Huang GB, Ramesh M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. univ. Massachusetts, Amherst, MA, USA
65. Izenman AJ (2013) Linear discriminant analysis. In: Modern multivariate statistical techniques. Springer, pp 237–280
66. Jain AK, Nandakumar K, Ross A (2016) 50 years of biometric research: accomplishments, challenges, and opportunities. Pattern recognit lett 79:80–105
67. Jang J, Cho H, Kim J, Lee J, Yang S (2018) Facial attribute recognition by recurrent learning with visual fixation. IEEE Trans Cybern 49(2):616–625
68. Jia S, Guo G, Xu Z (2020) A survey on 3d mask presentation attack detection and countermeasures. Pattern Recognit 98:107032
69. Jia S, Guo G, Xu Z, Wang Q (2020) Face presentation attack detection in mobile scenarios: a comprehensive evaluation. Image and Vision Comput 93:103826
70. Jian M, Lam KM (2015) Simultaneous hallucination and recognition of low-resolution faces based on singular value decomposition. IEEE Trans Circuits Syst Video Technol 25(11):1761–1772
71. Jillela R, Ross A (2012) Mitigating effects of plastic surgery: fusing face and ocular biometrics. In: 2012 IEEE Fifth international conference on biometrics: theory, applications and systems (BTAS), IEEE, pp 402–411
72. Kak SF, Mustafa FM, Valente P (2018) A review of person recognition based on face model. Eurasian J Sci Eng 4(1):157–168
73. Karve S, Shende V, Ahmed R (2018) A comparative analysis of feature extraction techniques for face recognition. In: 2018 International conference on communication information and computing technology (ICCICT), IEEE, pp 1–6
74. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput vision 1(4):321–331
75. Képešiová Z, Kozák Š (2018) An effective face detection algorithm. In: 2018 Cybernetics & Informatics (K&I), IEEE, pp 1–6, vol 2018

76. Kolkur S, Kalbande D, Shimpi P, Bapat C, Jatakia J (2017) Human skin detection using rgb, hsv and ycbcr color models. arXiv:170802694
77. Kose N, Dugelay JL (2014) Mask spoofing in face recognition and countermeasures. Image Vis Comput 32(10):779–789
78. Krishnapriya K, Albiero V, Vangara K, King MC, Bowyer KW (2020) Issues related to face recognition accuracy varying based on race and skin tone. IEEE Trans Technol Soc 1(1):8–20
79. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. Artif Intell Rev 52(2):927–948
80. Labati RD, Genovese A, Muñoz E, Piuri V, Scotti F, Sforza G (2016) Biometric recognition in automated border control: a survey. ACM Comput Surv (CSUR) 49(2):1–39
81. Lahasan B, Lutfi SL, San-Segundo R (2019) A survey on techniques to handle face recognition challenges: occlusion, single sample per subject and expression. Artif Intell Rev 52(2):949–979
82. Le THN, Luu K, Seshadri K, Savvides M (2012) A facial aging approach to identification of identical twins. In: 2012 IEEE Fifth international conference on biometrics: theory, applications and systems (BTAS), IEEE, pp 91–98
83. Li L, Correia PL, Hadid A (2017) Face recognition under spoofing attacks: countermeasures and research directions. Iet Biometrics 7(1):3–14
84. Li L, Feng X, Xia Z, Jiang X, Hadid A (2018b) Face spoofing detection with local binary pattern network. J Vis Commun Image Represent 54:182–192
85. Li Y, Huang H, Cao J, He R, Tan T (2019) Disentangled representation learning of makeup portraits in the wild. Int J Comput Vis, pp 1–19
86. Li J, Zhao F, Feng J, Roy S, Yan S, Sim T (2018) Landmark free face attribute prediction. IEEE Trans Image Process 27(9):4651–4662
87. Li T, Zhou P, Liu H (2019) Multiple features fusion based video face tracking. Multimed Tools Appl 78(15):21963–21980
88. Lian Y, Wang Z, Yuan H, Gao L, Yu Z, Chen W, Xing Y, Xu S, Feng L (2020) Partial occlusion face recognition method based on acupoints locating through infrared thermal imaging. In: 2020 International wireless communications and mobile computing (IWCMC), IEEE, 1394–1399
89. Lin L, Zhang D, Luo P, Zuo W (2020) Face localization and enhancement. In: Human centric visual analysis with deep learning. Springer, pp 29–45
90. Liu X, Shan S, Chen X (2012) Face recognition after plastic surgery: a comprehensive study. In: Asian conference on computer vision. Springer, pp 565–576
91. Liu N, Wang H, Yau WY (2006) Face recognition with weighted kernel principal component analysis
92. Loey M, Manogaran G, Taha MHN, Khalifa NEM (2021) Fighting against covid-19: a novel deep learning model based on yolo-v2 with resnet-50 for medical face mask detection. Sustainable Cities Soc 65:102600
93. Lu D, Chen Z, Wu QJ, Zhang X (2019) Fcn based preprocessing for exemplar-based face sketch synthesis. Neurocomputing 365:113–124
94. Lu Xh, Wang Lf, Qiu Jt, Li J (2020) A local occlusion face image recognition algorithm based on the recurrent neural network. In: International conference on multimedia technology and enhanced learning. Springer, pp 159–170
95. Luo Y, Guan YP (2017) Adaptive skin detection using face location and facial structure estimation. IET Comput Vis 11(7):550–559
96. Mahalingam G, Ricanek K, Albert AM (2014) Investigating the periocular-based face recognition across gender transformation. IEEE TTrans Inf Forensics Secur 9(12):2180–2192
97. Mahmood Z, Muhammad N, Bibi N, Ali T (2017) A review on state-of-the-art face recognition approaches. Fractals 25(02):1750025
98. Makhija Y, Sharma RS (2019) Face recognition: novel comparison of various feature extraction techniques
99. Mancini C, Falciati L, Maioli C, Mirabella G (2020) Threatening facial expressions impact goal-directed actions only if task-relevant. Brain sci 10(11):794
100. Marcel S, Nixon MS, Fierrez J, Evans N (2019) Handbook of biometric anti-spoofing: presentation attack detection. Springer
101. Martin V, Seguier R, Porcheron A, Morizot F (2019) Face aging simulation with a new wrinkle oriented active appearance model. Multimed Tools Appl 78(5):6309–6327
102. Meytlis M, Sirovich L (2007) On the dimensionality of face space. IEEE Trans Pattern Anal Mach Intell 29(7):1262–1267
103. Miao Y, Dong H, Jaam JMA, Saddik AE (2019) A deep learning system for recognizing facial expression in real-time. ACM Trans Multimed Comput Commun Appl (TOMM) 15(2):1–20
104. Minaee S, Abdolrashidi A, Su H, Bennamoun M, Zhang D (2019) Biometrics recognition using deep learning: a survey. arXiv:191200271

105. Mohammed BO, Shamsuddin SM, Hasan S (2019) An overview of uni-and multi-biometric identification of identical twins. IEIE Trans Smart Process Comput 8(1):71–84
106. Moore J (2017) Performative face theory: a critical perspective on interpersonal identity work. Commun Monogr 84(2):258–276
107. Mortezaie Z, Hassanpour H (2019) Asurvey on age-invariant face recognition methods. Jordanian J Comput Inf Technol (JJCIT) 5(02):87–96
108. Mostafa E, Hammoud R, Ali A, Farag A (2013) Face recognition in low resolution thermal images. Comput Vis Image Underst 117(12):1689–1694
109. Nappi M, Ricciardi S, Tistarelli M (2016) Deceiving faces: when plastic surgery challenges face recognition. Image Vis Comput 54:71–82
110. Neal TJ, Woodard DL (2016) Surveying biometric authentication for mobile device security. J Pattern Recognit Res 1(74-110):4
111. Nguyen DL, Arora SS, Wu Y, Yang H (2020) Adversarial light projection attacks on face recognition systems: a feasibility study. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 814–815
112. Nguyen TD, Ranganath S (2008) Tracking facial features under occlusions and recognizing facial expressions in sign language. In: 2008 8th IEEE International conference on automatic face & gesture recognition, IEEE, pp 1–7
113. O'Toole AJ, Vetter T, Blanz V (1999) Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing. Vision Res 39(18):3145–3155
114. Paone JR, Flynn PJ, Philips PJ, Bowyer KW, Bruegge RWV, Grother PJ, Quinn GW, Pruitt MT, Grant JM (2014) Double trouble: Differentiating identical twins by face recognition. IEEE Trans Inf forensics Sec 9(2):285–295
115. Park U, Tong Y, Jain AK (2010) Age-invariant face recognition. IEEE Trans Pattern Anal Mach Intell 32(5):947–954
116. Peng F, Qin L, Long M (2018) Face presentation attack detection using guided scale texture. Multimed Tools Appl 77(7):8883–8909
117. Perveen N, Roy D, Mohan CK (2018) Spontaneous expression recognition using universal attribute model. IEEE Trans Image Process 27(11):5575–5584
118. Phillips PJ, Flynn PJ, Bowyer KW, Bruegge RWV, Grother PJ, Quinn GW, Pruitt M (2011) Distinguishing identical twins by face recognition. In: Face and gesture 2011, IEEE, pp 185–192
119. Prasad PS, Pathak R, Gunjan VK, Rao HR (2020) Deep learning based representation for face recognition. In: ICCCE 2019. Springer, pp 419–424
120. Raghavendra R, Raja K, Venkatesh S, Busch C (2017) Face morphing versus face averaging: vulnerability and detection. In: 2017 IEEE International Joint Conference on Biometrics (IJCB), IEEE, pp 555–563
121. Ramachandra R, Busch C (2017) Presentation attack detection methods for face recognition systems: a comprehensive survey. ACM Comput Sur (CSUR) 50(1):1–37
122. Reddy GV, Savarni CD, Mukherjee S (2020) Facial expression recognition in the wild, by fusion of deep learnt and hand-crafted features. Cogn Syst Res 62:23–34
123. Rehman YAU, Po LM, Komulainen J (2020) Enhancing deep discriminative feature maps via perturbation for face presentation attack detection. Image and Vision Comput 94:103858
124. Rehman YAU, Po LM, Liu M (2020) Slnet: stereo face liveness detection via dynamic disparity-maps and convolutional neural network. Exp Syst Appl 142:113002
125. Revina IM, Emmanuel WS (2021) A survey on human face expression recognition techniques. J King Saud Univ -Comput Inf Sci 33(6):619–628
126. Rodríguez-Gómez P, Romero-Ferreiro V, Pozo MA, Hinojosa JA, Moreno EM (2020) Facing stereotypes: erp responses to male and female faces after gender-stereotyped statements. Soc Cogn Affect Neurosci 15(9):928–940
127. Roy SD, Bhowmik MK (2020) Annotation and benchmarking of a video dataset under degraded complex atmospheric conditions and its visibility enhancement analysis for moving object detection. IEEE Trans Circuits Syst Video Technol
128. Rusia MK, Singh DK (2021) An efficient cnn approach for facial expression recognition with some measures of overfitting. Int J Inf Technol 13(6):2419–2430
129. Rusia MK, Singh DK, Ansari MA (2019) Human face identification using lbp and haar-like features for real time attendance monitoring. In: 2019 Fifth international conference on image information processing, (ICIIP), IEEE, pp 612–616
130. Sabharwal T, Gupta R (2020) Facial marks for enhancing facial recognition after plastic surgery. Int J Inf Technol, pp 1–6

131. Sabharwal T, Gupta R, Kumar R, Jha S et al (2019) Recognition of surgically altered face images: an empirical analysis on recent advances. Artif Intell Rev 52(2):1009–1040
132. Saha P, Bhattacharjee D, De BK, Nasipuri M (2019) A survey on image acquisition protocols for non-posed facial expression recognition systems. Multimed Tools Appl 78(16):23329–23368
133. Sánchez-Lozano E, Tzimiropoulos G, Martinez B, De la Torre F, Valstar M (2017) A functional regression approach to facial landmark tracking. IEEE Trans Pattern Anal Mach Intell 40(9):2037–2050
134. Sawant MM, Bhurchandi KM (2019) Age invariant face recognition: a survey on facial aging databases, techniques and effect of aging. Artif Intell Rev 52(2):981–1008
135. Scherhag U, Rathgeb C, Merkle J, Breithaupt R, Busch C (2019) Face recognition systems under morphing attacks: a survey. IEEE Access 7:23012–23026
136. Schuckers S (2016) Presentations and attacks, and spoofs, oh my. Image Vis Comput 55:26–30
137. Seibold C, Samek W, Hilsmann A, Eisert P (2017) Detection of face morphing attacks by deep learning. In: International workshop on digital watermarking. Springer, pp 107–120
138. Sepas-Moghaddam A, Pereira F, Correia PL (2018) Light field-based face presentation attack detection: reviewing, benchmarking and one step further. IEEE Trans Inf Forensics Secur 13(7):1696–1709
139. Serengil SI, Ozpinar A (2020) Lightface: a hybrid deep face recognition framework. In: 2020 Innovations in intelligent systems and applications conference (ASYU), IEEE, pp 1–5
140. Sharma S, Kumar V (2020) Voxel-based 3d occlusion-invariant face recognition using game theory and simulated annealing. Multimed Tools Appl 79(35):26517–26547
141. Sharma V, Van Gool L (2016) Image-level classification in hyperspectral images using feature descriptors, with application to face recognition. arXiv:160503428
142. Shi Y, Lv Z, Bi N, Zhang C (2019) An improved sift algorithm for robust emotion recognition under various face poses and illuminations. Neural Comput Applic, pp 1–15
143. Shinwari AR, Ayoubi M (2020) A comparative study of face recognition algorithms under occlusion. Technology 2(1):85–95
144. Shirley C, Mohan NR, Chitra B (2020) Gravitational search-based optimal deep neural network for occluded face recognition system in videos. Multidim Syst Sign Process, pp 1–27
145. Singh DK, Kushwaha DS (2016) Analysis of face feature based human detection techniques. Int J Control Theory Appl 9(22):173–180
146. Singh DK, Kushwaha DS (2016) Ilut based skin colour modelling for human detection. Indian J Sci Technol, vol 9 (32)
147. Singh DK, Paroothi S, Rusia MK, Ansari MA (2020) Human crowd detection for city wide surveillance. Procedia Computer Science 171:350–359
148. Singh M, Singh R, Vatsa M, Ratha NK, Chellappa R (2019) Recognizing disguised faces in the wild. IEEE Trans Biom Behav Identity Sci 1(2):97–108
149. Singh R, Vatsa M, Bhatt HS, Bharadwaj S, Noore A, Nooreyezdan SS (2010) Plastic surgery: a new dimension to face recognition. IEEE Trans Inf Forensics Secur 5(3):441–448
150. Soltanpour S, Boufama B, Wu QJ (2017) A survey of local feature methods for 3d face recognition. Pattern Recogn 72:391–406
151. Song X, Zhao X, Fang L, Lin T (2019) Discriminative representation combinations for accurate face spoofing detection. Pattern Recogn 85:220–231
152. Spencer-Oatey H (2007) Theories of identity and the analysis of face. J Pragmat 39(4):639–656
153. Srinivas N, Aggarwal G, Flynn PJ, Bruegge RWV (2012) Analysis of facial marks to distinguish between identical twins. IEEE Trans Inf Forensics Secur 7(5):1536–1550
154. Sun J, Huang D, Wang Y, Chen L (2019) Expression robust 3d facial landmarking via progressive coarse-to-fine tuning. ACM Trans Multimed Comput Commun Appl (TOMM) 15(1):1–23
155. Sun Y, Tistarelli M, Maltoni D (2013) Structural similarity based image quality map for face recognition across plastic surgery. In: 2013 IEEE Sixth international conference on biometrics: theory, applications and systems (BTAS), IEEE, pp 1–8
156. Surekha S (2020) Deep neural network-based human emotion recognition by computer vision. In: Advances in electrical and computer technologies. Springer, pp 453–463
157. Tamilselvi M, Karthikeyan S (2018) A literature survey in face recognition techniques. Int J Pure Appl Math 118(16):831–849
158. Tamrakar D, Khanna P (2015) Occlusion invariant palmprint recognition with ulbp histograms. Procedia Comput Sci 54:491–500
159. Tan H, Yang B, Ma Z (2013) Face recognition based on the fusion of global and local hog features of face images. IET comput vision 8(3):224–234
160. Tanaka JW, Pierce LJ (2009) The neural plasticity of other-race face recognition. Cogn Affect Behav Neurosci 9(1):122–131

161. Thomaz CE, Giraldi GA (2010) A new ranking method for principal components analysis and its application to face image analysis. Image and vision comput 28(6):902–913
162. Turk MA, Pentland AP (1991) Face recognition using eigenfaces. In: Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition, IEEE Computer Society, pp 586–587
163. Ueda S, Koyama T (2010) Influence of make-up on facial recognition. Perception 39(2):260–264
164. Vijayan V, Bowyer KW, Flynn PJ, Huang D, Chen L, Hansen M, Ocegueda O, Shah SK, Kakadiaris IA (2011) Twins 3d face recognition challenge. In: 2011 International joint conference on biometrics, (IJCB), IEEE, pp 1–7
165. Wan J, Li J, Lai Z, Du B, Zhang L (2020) Robust face alignment by cascaded regression and de-occlusion. Neural Netw 123:261–272
166. Wang Q, Guo G (2019) Benchmarking deep learning techniques for face recognition. J Vis Commun Image Represent 65:102663
167. Wang K, Peng X, Yang J, Meng D, Qiao Y (2020) Region attention networks for pose and occlusion robust facial expression recognition. IEEE Trans Image Process 29:4057–4069
168. Wang D, Wang H, Sun J, Xin J, Luo Y (2020) Face recognition in complex unconstrained environment with an enhanced wwn algorithm. J Intell Syst 30(1):18–39
169. Waseem M, Khowaja SA, Ayyasamy RK, Bashir F (2020) Face recognition for smart door lock system using hierarchical network. In: 2020 International conference on computational intelligence (ICCI), IEEE, pp 51–56
170. Winant D, Schreurs J, Suykens JA (2019) Latent space exploration using generative kernel pca. In: Artificial Intelligence and Machine Learning. Springer, pp 70–82
171. Wu Y, Gou C, Ji Q (2017) Simultaneous facial landmark detection, pose and deformation estimation under facial occlusion. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3471–3480
172. Wu Y, Ji Q (2019) Facial landmark detection: a literature survey. Int J Comput Vis 127(2):115–142
173. Xiao Q, Song R (2018) Action recognition based on hierarchical dynamic bayesian network. Multimed Tools Appl 77(6):6955–6968
174. Xie Z, Shi L, Li Y (2020) Two-stage fusion of local binary pattern and discrete cosine transform for infrared and visible face recognition. In: International conference on intelligent and interactive systems and applications. Springer, pp 967–975
175. Yang C, Lv Z (2020) Gender based face aging with cycle-consistent adversarial networks. Image Vis Comput 100:103945
176. Yang C, Wu S, Fang H, Er MJ (2019) Adaptive weber-face for robust illumination face recognition. Computing 101(6):605–619
177. Yao Y, Huang D, Yang X, Wang Y, Chen L (2018) Texture and geometry scattering representation-based facial expression recognition in 2d+ 3d videos. ACM Trans Multimed Comput Commun Appl (TOMM) 14(1s):1–23
178. Yu C, Yao C, Pei M, Jia Y (2019) Diffusion-based kernel matrix model for face liveness detection. Image Vis Comput 89:88–94
179. Yuille AL (1991) Deformable templates for face recognition. J Cogn Neurosci 3(1):59–70
180. Zafeiriou S, Tzimiropoulos G, Petrou M, Stathaki T (2012) Regularized kernel discriminant analysis with a robust kernel for face recognition and verification. IEEE Trans Neural Netw Learn Syst 23(3):526–534
181. Zangeneh E, Rahmati M, Mohsenzadeh Y (2020) Low resolution face recognition using a two-branch deep convolutional neural network architecture. Expert Syst Appl 139:112854
182. Zavan FHdB, Bellon OR, Silva L, Medioni GG (2019) Benchmarking parts based face processing in-the-wild for gender recognition and head pose estimation. Pattern Recogn Lett 123:104–110
183. Zeng J, Liu S, Li X, Mahdi DA, Wu F, Wang G (2017) Deep context-sensitive facial landmark detection with tree-structured modeling. IEEE Trans Image Process 27(5):2096–2107
184. Zhang F, Mao Q, Shen X, Zhan Y, Dong M (2018) Spatially coherent feature learning for pose-invariant facial expression recognition. ACM Trans Multimedia Comput Commun Appl (TOMM) 14(1s):1–19
185. Zhang X, Sugano Y, Fritz M, Bulling A (2017) It's written all over your face: full-face appearance-based gaze estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 51–60
186. Zhang T, Tang YY, Fang B, Shang Z, Liu X (2009) Face recognition under varying illumination using gradientfaces. IEEE Trans Image Process 18(11):2599–2606
187. Zhang L, Verma B, Tjondronegoro D, Chandran V (2018) Facial expression analysis under partial occlusion: a survey. ACM Comput Sur (CSUR) 51(2):1–49
188. Zhang G, Zou W, Zhang X, Zhao Y (2018) Singular value decomposition based virtual representation for face recognition. Multimed Tools Appl 77(6):7171–7186

189. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. ACM comput sur (CSUR) 35(4):399–458
190. Zhou S, Xiao S (2018) 3d face recognition: a survey. Hum-centric comput inf sci 8(1):1–27
191. Ziwei X, Liang Z, Jingyu P, Jinqian Z, Hongling C, Yiwen Z, Xi H, Siyuan X, Haoyang Y (2020) Face occlusion detection based on ssd algorithm. In: 2020 IEEE 10th international conference on electronics information and emergency communication (ICEIEC), IEEE, pp 362–365
192. Zuo KJ, Saun TJ, Forrest CR (2019) Facial recognition technology: a primer for plastic surgeons. Plast Reconstr Surg 143(6):1298e–1306e