Information Fusion 52 (2019) 187–205

Contents lists available at ScienceDirect

Information Fusion

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

Full Length Article

A comprehensive overview of biometric fusion

Maneet Singh [a], Richa Singh [a,*], Arun Ross [b]

[a] IIIT Delhi, India
[b] Michigan State University, USA

ARTICLE INFO

Keywords:
Biometrics
Information fusion
Multibiometrics
Soft biometrics
Continuous authentication
Privacy
Security
Cryptosystems
Spoof detection
Social networks

ABSTRACT

The performance of a biometric system that relies on a single biometric modality (e.g., fingerprints only) is often stymied by various factors such as poor data quality or limited scalability. Multibiometric systems utilize the principle of *fusion* to combine information from multiple sources in order to improve recognition accuracy whilst addressing some of the limitations of single-biometric systems. The past two decades have witnessed the development of a large number of biometric fusion schemes. This paper presents an overview of biometric fusion with specific focus on three questions: *what* to fuse, *when* to fuse, and *how* to fuse. A comprehensive review of techniques incorporating ancillary information in the biometric recognition pipeline is also presented. In this regard, the following topics are discussed: (i) incorporating data quality in the biometric recognition pipeline; (ii) combining soft biometric attributes with primary biometric identifiers; (iii) utilizing contextual information to improve biometric recognition accuracy; and (iv) performing continuous authentication using ancillary information. In addition, the use of information fusion principles for presentation attack detection and multibiometric cryptosystems is also discussed. Finally, some of the research challenges in biometric fusion are enumerated. The purpose of this article is to provide readers a comprehensive overview of the role of information fusion in biometrics.

1. Introduction

Biometrics refers to the automated process of recognizing an individual based on his/her physical or behavioral traits such as face, fingerprints, voice, iris, gait, or signature [1]. These traits are often referred to as biometric modalities or biometric cues. Over the past several years, a number of different biometric modalities [2–4] have been explored for use in various applications ranging from personal device access systems to border control systems [5].

The general framework of a typical biometric recognition system is summarized in Fig. 1. Here, given some input data (e,g, an image, video or signal), a typical biometric recognition system first performs segmentation or detection, which involves extracting the modality of interest from the input. This is followed by preprocessing, which involves data alignment, noise removal, or data enhancement. Features are extracted from the preprocessed data, which are then used by a classifier for biometric recognition. The recognition process may involve associating an identity with the input data (e.g., biometric identification) or determining if two instances of input data pertain to the same identity (e.g., biometric verification).

Traditionally, biometric recognition systems are *unibiometric*, which utilize a single biometric cue, and thus may encounter problems due to missing information (e.g., occluded face), poor data quality (e.g. dry fingerprint), overlap between identities (e.g., face images of twins) or limited discriminability (e.g., hand geometry). In such situations, it may be necessary to utilize multiple biometric cues in order to improve the recognition accuracy. For example, a border control system may use both face and fingerprints to establish the identity of an individual [6,7]. In some cases, a biometric cue could be used alongside traditional user-validation schemes such as passwords/passcodes to verify a user's identity. For example, many smartphone devices incorporate such a dual-factor authentication scheme [8,9]. In other applications, multiple sensors could be used to acquire the same biometric modality, thereby allowing the system to operate in different environments. For example, a face recognition system may use a visible spectrum camera as well as a near-infrared camera to image a person's face, thus facilitating biometric recognition in nighttime environment. The aforementioned examples underscore the need for effective biometric *fusion* techniques that can consolidate information from multiple sources.

The term *multibiometrics* has often been used to connote biometric fusion in the literature [10]. In order to develop a multibiometric system, one must consider the following three questions, (i) *what* to fuse, (ii) *when* to fuse, and (iii) *how* to fuse, each of which have been explored in this article.

What to fuse involves selecting the different sources of information to be combined, such as multiple algorithms or multiple modalities. *When*

* Corresponding author.
E-mail addresses: maneets@iiitd.ac.in (M. Singh), rsingh@iiitd.ac.in (R. Singh), rossarun@cse.msu.edu (A. Ross).

https://doi.org/10.1016/j.inffus.2018.12.003
Received 27 November 2018; Accepted 10 December 2018
Available online 7 February 2019

Fig. 1. General pipeline of a face recognition system. From a pattern recognition perspective, the most significant modules of a biometric system are the sensor module, the segmentation module; the feature extraction module; and the classification or decision-making module. Face image has been taken from the CMU Multi-PIE dataset [215].

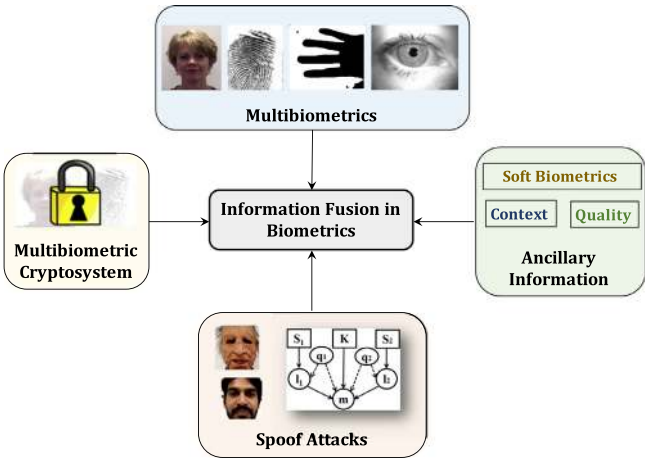


Fig. 2. Information fusion in the context of biometrics can be used to improve recognition accuracy (multibiometrics, ancillary information) or to improve security (cryptosystems, spoof detection). In some cases, non-biometric cues may also be used in the fusion framework. Images are taken from the Internet, WVU Multimodal dataset [13], and MLFP dataset [14].

to fuse is answered by analyzing the different levels of fusion, that is, the various stages in the biometric recognition pipeline at which information can be fused. *How* to fuse refers to the fusion method that is used to consolidate the multiple sources of information.

Given data from only a single modality (say face only), the performance of a recognition system can often be enhanced by incorporating some ancillary information. Incorporating details such as image quality, subject demographics, soft biometric attributes, and contextual meta-data has shown to improve the performance of recognition systems [72, 131]. While recognition performance is a major metric for evaluating biometric systems, it is important to focus on the security (and privacy) aspect of such systems as well [205]. Information fusion is seen as a viable option for securing the biometric templates in a multibiometric system. Cryptosystems based on multiple modalities have been proposed to securely store biometric templates and prevent access to the original data [11]. Biometric systems are also susceptible to spoof attacks. That is, an adversary can impersonate another person's identity by presenting a fake or altered biometric trait and gain unauthorized access. Information fusion can play a major role in the detection and deflection of such malicious activities [12]. This paper focuses on the above mentioned aspects of biometric fusion, and thus presents a survey of information fusion techniques along the lines of: (i) biometrics and ancillary information, (ii) spoof (or presentation attack) detection, and (iii) multibiometric cryptosystems.

2. Multibiometric systems

A *multibiometric system* can overcome some of the limitations of a unibiometric system by combining information from different sources in a principled manner. The utilization of multiple sources often results in improved recognition performance and enhanced system reliability, since the combined information is likely to be more distinctive to an individual compared to the information obtained from a single source.

2.1. Sources of fusion

As mentioned previously, one of the major questions for developing a multibiometric system is *what* to fuse. Fig. 3 presents the different sources of information that can be fused in a multibiometric system. Depending upon the sources of fusion, a multibiometric system can correspond to one of the following configurations: (i) multi-sensor, (ii) multi-algorithm, (iii) multi-instance, (iv) multi-sample, or (v) multi-modal.

(i) **Multi-sensor** systems combine information captured by multiple sensors for the same biometric modality. For example, a face recognition module could utilize RGB data captured using a visible spectrum camera, along with depth information captured using a 3D camera [15] or infrared data captured using an NIR camera [16–18]. Using both the images for identifying a subject would result in a multi-sensor fusion algorithm. Such systems rely

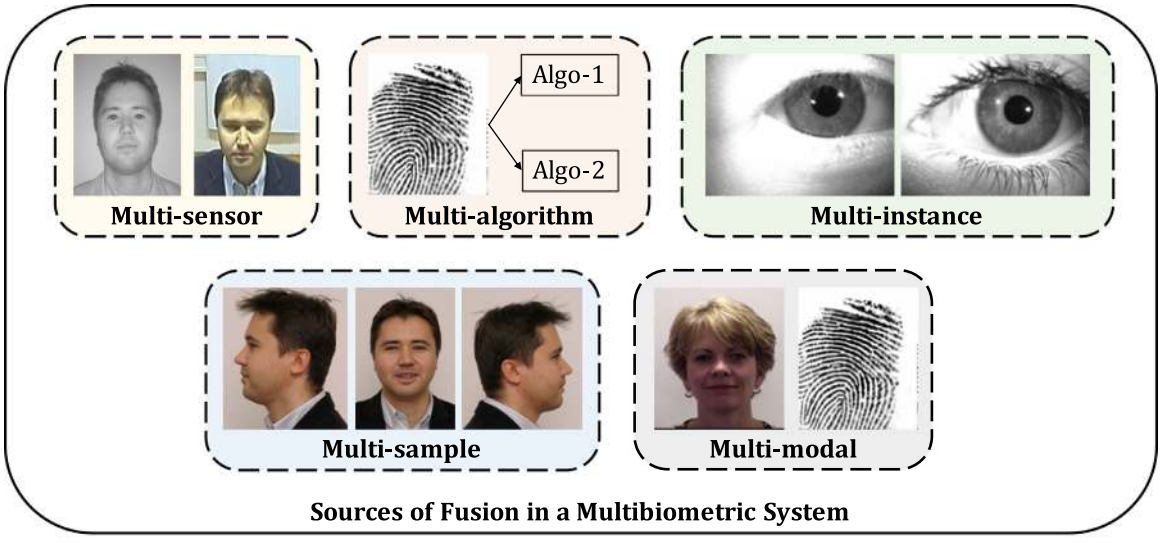


Fig. 3. Different sources of information that can be exploited by a multibiometric system. Information from multiple sensors (infrared and visible spectra) or multiple algorithms (minutiae-based and texture-based) or multiple instances (left and right irides) or multiple samples (left, frontal, and right facial profiles) or multiple modalities (face and fingerprint) can be fused. Images have been taken from the SCface dataset [216] and the WVU Multimodal dataset [13].

on a single modality for recognition; however, they capture different information from the same modality by utilizing multiple sensors [19]. Multi-sensor systems are useful in scenarios which require a different mode of capture at different times, or where discriminative information can successfully be captured by different sensors.

(ii) **Multi-algorithm** systems utilize multiple algorithms for processing an input sample. Data is captured from a biometric modality using a single sensor; however, multiple algorithms are used to process it. For example, a fingerprint recognition system could utilize both minutiae and texture features for matching fingerprints [20], or a palmprint recognition system could utilize Gabor, line, and appearance based palmprint representations for matching [21]. Such systems benefit from the advantage of extracting and utilizing different types of information from the same sample. In cases where two algorithms or feature sets provide complementary information, multi-algorithm systems can often result in improved performance.

(iii) **Multi-instance** systems capture multiple instances of the same biometric trait. In the case of iris recognition, the recognition module can utilize both left and right irides, thereby resulting in a multi-instance system [22]. Similarly, in the case of a fingerprint or palm-print recognition system, a multi-instance system can utilize data captured from the ten fingers or both palms [23,24]. Multi-instance systems may use the same feature extraction and matching methods for all instances of the biometric trait.

(iv) **Multi-sample** systems work with multiple samples of the same biometric modality, often captured with some variations. Video-based recognition models fall under this category, where, a biometric modality is captured continuously over a small period of time (e.g., several seconds long). This often results in a large number of frames containing multiplicity of information [25]. In the literature, videos have been used for performing face [25–28] and gait [29–32] recognition. This results in a multi-sample recognition system, which combines information captured across multiple video frames. Such systems are able to extract diverse information from a single biometric modality, while requiring only a single sensor.

(v) **Multi-modal** systems utilize data captured from multiple biometric cues in order to recognize a subject. A multi-modal system could utilize information captured from face, fingerprint, and iris modalities [33]; face, fingerprint, and speech modalities [34]; face and voice modalities [35,36]; ear and face modalities [37]; or even iris and periocular modalities [38,39]. Research has also focused on combining different modalities for performing speaker recognition, such as audio and lip motion [40]; audio, lip motion, and lip texture [41]; and audio, RGB, and depth information [42]. Such systems can eliminate the limitations of a particular biometric modality by having the flexibility of processing multiple modalities [43,44]. Multi-modal systems are also useful in scenarios where an individual cannot provide data for a particular biometric modality (say injured fingerprints), but can provide data pertaining to another one (say face). Fusing information from different modalities further enables extraction of distinctive features, often resulting in enhanced recognition performance [45].

Besides the aforementioned sources, non-biometric cues may also be used in the fusion process. As will be described later, information such as contextual meta-data can be used in conjunction with biometric identifiers in order to recognize an individual [46].

2.2. Levels of fusion

Fig. 4 presents the different levels at which fusion can be incorporated in a biometric pipeline, viz., (i) sensor-level, (ii) feature-level, (iii) score-level, (iv) rank-level, or (v) decision-level. Each of these levels of fusion are explained in detail below.

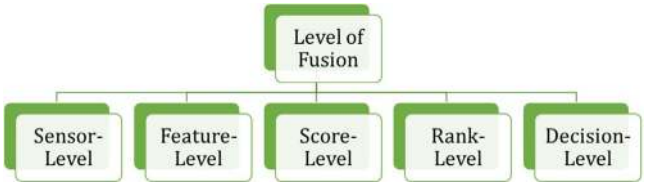


Fig. 4. Levels of fusion in a multibiometric system. These levels correspond to the various modules of a typical biometric system. See Fig. 1. While these levels of fusion are applicable to both verification and identification systems, rank-level fusion is typically applicable to only identification systems.

(i) **Sensor-level** fusion or data-level fusion generally corresponds to multi-sensor or multi-sample algorithms, where data is combined immediately after its acquisition. That is, data fusion is carried out prior to feature extraction, directly on the *raw* data [47]. In case of a face recognition module, this corresponds to direct pixel-level combination of face images captured from a camera. For example, multiple faces can be captured with pose variations such as frontal, left profile, or right profile. A mosaicing technique can be used to fuse the samples together in order to obtain a combined face representation [48]. Often a direct fusion strategy or adding pixels of two images can also be utilized [49].

(ii) **Feature-level** fusion refers to algorithms where fusion is performed on multiple features extracted from the same or different input data. This could correspond to multiple feature sets pertaining to the same biometric trait, such as textural and structural features of a face image or different features from a hand or palm-print image [50,51]. It could also correspond to features extracted from different modalities, such as face and hand images [52]. Such algorithms are often used by multibiometric cryptosystems, where features from multiple biometric sources are combined to improve security and privacy [11]. They have also been used for indexing multimodal biometric databases [53]. Feature-level fusion combines different representations in order to generate a single representation for a given individual. For example, representation learning algorithms can be used to learn a shared representation of features extracted from different modalities [54].

(iii) **Score-level** fusion corresponds to algorithms where the match scores produced by different matchers are fused together. Some of the common fusion algorithms applied at this level are mean score fusion, max score fusion, or min score fusion, where the mean, maximum, or minimum score of multiple matchers is considered as the final score [45,55–58]. Dempster-Shafer theory or probabilistic techniques such as likelihood ratio based score fusion have also been applied in the literature [59–62]. In addition, Ding and Ross [63] discuss several imputation techniques for handling missing or incomplete information in the context of score-level fusion. This is the most common type of fusion described in the literature due to the ease of accessing scores generated by commercial matchers. Most commercial matchers do not provide easy access to features or, sometimes, even the raw data.

(iv) **Rank-level** fusion is performed after comparing the input probe with the templates in the gallery set, i.e., the database. In the task of identification, where, a given probe image is compared against a gallery of images, a ranked list of matching identities is often generated by the matcher. In literature, the rank lists from multiple matchers have been fused using techniques like Borda count, logistic regression, and highest rank method [46,64–67].

In scenarios having limited access to features or match scores, rank-level fusion is often deemed effective.

(v) **Decision-level** fusion corresponds to algorithms where fusion is performed at the decision level [66,68,69]. Majority voting is one of the most common fusion algorithms applied at the decision level. Decisions taken by n matchers or classifiers are combined based on a majority vote, resulting in the final decision. Decision-level fusion has the advantage of working well with black-box systems, where only the final decisions are available [70]. This is true in the case of many commercial systems, where access to features, scores, and ranks may not be feasible.

Thus, fusion in biometrics can be invoked at different levels in the biometric recognition pipeline and can avail different sources of information. For a detailed review on the sources and levels of fusion, the reader is encouraged to refer to [10,71].

The final piece for developing a multibiometric framework is understanding *how* to fuse the diverse sources of information. The remainder of this paper presents an expansive survey of information fusion techniques applied in the context of (i) combining biometrics and ancillary information, (ii) spoof detection, and (iii) designing multibiometric cryptosystems.

3. Biometrics and ancillary information

Researchers have incorporated ancillary information in the traditional biometrics pipeline in order to improve recognition performance. Ancillary data refers to any additional information that can be provided about a particular biometric sample which might aid in the recognition process. Fig. 5 presents some of the commonly used sources of ancillary information, viz., quality estimates, soft biometric attributes, and contextual information. Ancillary information has also been used to perform continuous authentication of a subject. This section presents an overview of the literature associated with each form of ancillary information mentioned above.

3.1. Biometrics and quality

As per ISO standards (ISO/IEC 29794-1), a biometric sample is said to be of good quality if "it is suitable for automated matching". For a biometric sample, the quality is often quantified by the ease with which an image can be processed, including feature extraction and correct classification with a high confidence score. A good quality biometric sample is often associated with rich features and easy classification, whereas a poor quality sample suffers from the inherent challenge of noisy data. Bharadwaj et al. [72] present a review of biometric quality for the face, fingerprint, and iris modalities. A comprehensive survey of different quality measures proposed in the literature is presented, along with their estimation strategies and methods of incorporating quality in the biometric classification pipeline. Experimental analysis on a multimodal biometric dataset reiterated the importance of carefully selecting quality measures for enhancing recognition performance. In the literature, a sample's quality estimate has been used as ancillary information in both unibiometric and multibiometric recognition systems. As shown in Fig. 7, this is achieved by exploiting the quality information at different stages in the recognition pipeline. Fig. 6 demonstrates the inclusion of quality information for modality selection, fusion, or context switching in multi-modal recognition scenarios.

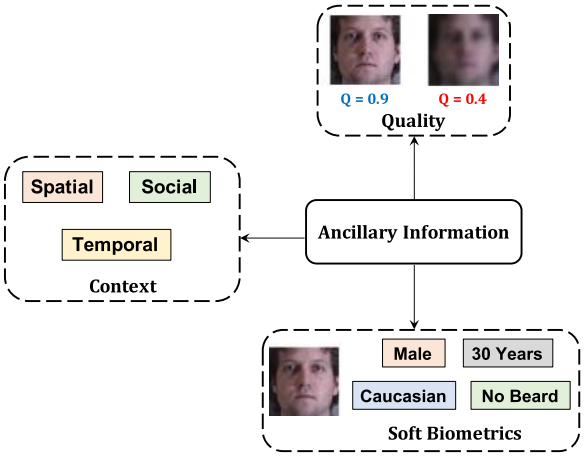


Fig. 5. Types of ancillary information that can be combined with primary biometric traits (such as faces and fingerprints) in order to improve recognition accuracy. Data quality, soft biometric attributes, and contextual information can aid in the biometric recognition process. Face image has been taken from the CMU Multi-PIE dataset [215].

In 2003, Bigun et al. [73] proposed incorporating quality information into a Bayesian statistical model for performing multimodal biometric classification. Quality is incorporated as a variance parameter such that samples with higher quality are associated with lower variance. The entire framework consists of two *supervisors*: client and impostor, that are trained for performing verification. Fierrez-Aguilar et al. [74] built upon the above architecture and presented one of the earliest works in the literature on fusing biometric quality at the score level, for performing multimodal biometric authentication. The proposed algorithm is built over a Support Vector Machine (SVM), where the training function is modified to include a quality-based cost term, thereby associating more weight with higher quality training samples. At the time of testing, the scores generated for each modality are combined in a weighted manner, where the weights are dependent on the quality scores. This ensures that samples with higher quality are given more weight when performing score level fusion of multiple modalities. Poh and Bengio [75] introduced a confidence criterion to incorporate quality when combining multiple biometric classifiers in a linear manner. Quality is considered as the derived margin, i.e. the difference between the False Acceptance Rate and the False Rejection Rate. This quality measure is integrated as an *a priori* weight when performing fusion of multiple classifiers.

In an attempt to understand the impact of fingerprint quality on different classifiers, Fierrez-Aguilar et al. [76] performed experiments using ridge-based and minutiae-based classifiers. Fingerprint images having different quality scores were used for testing and it was observed that the ridge-based system was more robust to quality variations. A weighted adaptive score fusion technique was also proposed, which combines the scores obtained from the ridge-based system (s_R) and the minutia-based system (s_M) as follows:

$$s_Q = \frac{Q}{2} s_M + \left(1 - \frac{Q}{2}\right) s_R. \tag{1}$$

Here, Q refers to the quality of the image, and s_Q refers to the final score generated by the entire framework. Consistent with their findings, as the image quality decreases, more weight is given to the score generated by the ridge-based classifier.

In 2006, Nandakumar et al. [77] proposed the use of a single quality metric for both the template (image stored in the gallery database) and the probe (image presented during verification). The authors proposed a Quality-based Product Fusion Score (QPFS) which is based upon the joint density of the match score and the quality estimated from a given gallery-probe pair:

$$QPFS(s) = \prod_{j=1}^{R} \frac{l_{j,gen}(s_j, q_j)}{l_{j,imp}(s_j, q_j)}, \tag{2}$$

where, $l_{j,gen}(s_j, q_j)$ refers to the joint density of the match score (s_j) and quality (q_j) of the j^{th} sample, and R refers to the total number of samples. In 2008, they further built upon the proposed likelihood ratio based technique by estimating the joint density using Gaussian Mixture Models [60]. They tested their approach on a bimodal system involving fingerprint and iris.

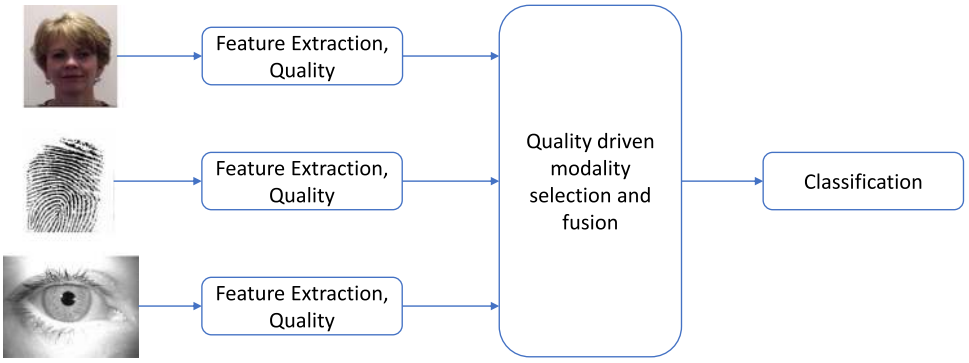


Fig. 6. Given input from multiple modalities, quality information is often used for modality selection, fusion, or context switching at run-time.

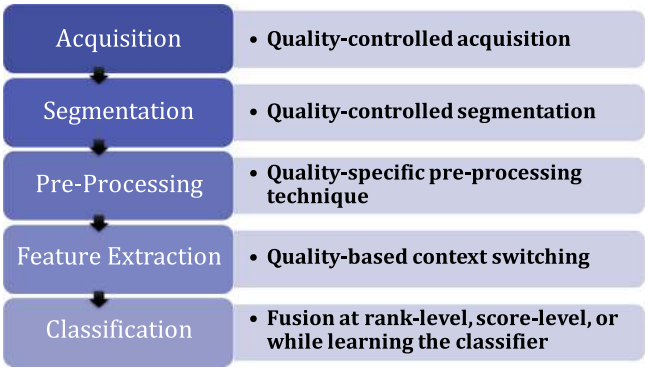


Fig. 7. Biometric sample quality has been incorporated into various modules of a biometric system. In addition, a number of fusion rules have been proposed to combine quality with features, match scores, and ranks.

Another likelihood-ratio based algorithm was proposed by Poh et al. [59], where the authors proposed incorporating both the device information and the quality information for predicting the class label of an input sample. Such a technique can be utilized in scenarios where data is collected from multiple sensors for a particular modality. Since in real world scenarios, the acquisition device can be unknown at the time of testing, a posterior probability is estimated using the quality measures. The proposed score normalization technique is thus formulated as,

$$y^{norm} = \log \frac{\sum_d p(y|C,d)p(d|q)}{\sum_d [(y|I,d)p(d|q)]}, \tag{3}$$

where, C and I refer to the client and impostor classes, respectively, and d, q corresponds to the device and quality estimate of the given sample, respectively. y refers to a vector of scores generated by different classification devices. A major limitation of this algorithm is the difficulty in keeping track of the number of devices that can be used for generating the probe images, thus limiting the applicability of the proposed model. In 2010, this model was further extended [81] and experiments were performed by incorporating quality-based score normalization as a pre-processing step in existing multi-modal fusion pipelines. Experimental analysis under different situations (known or unknown device) demonstrated the efficacy of the proposed model. Vatsa et al. [61] proposed computing the quality score of a given biometric image using Redundant Discrete Wavelet Transform (RDWT). Results were shown for the task of multimodal recognition of face and iris images. The independent match scores obtained were multiplied by the quality scores and fused using a novel 2ν-SVM fusion algorithm.

Maurer and Baker [78] presented a detailed description and analysis of a Bayesian Belief Network that was proposed earlier by them [92]. The model is built upon the motivation that "*if the match score (similarity score) of a low quality sample is high, it is extremely unlikely to be from an impostor*". This implies that simply providing a weight based on the quality of the sample while performing multi-modal fusion might negatively impact some true positive (genuine) samples. In order to address this, the authors proposed a Bayesian Belief Network where no dependence is encoded between the quality and the identity. The model is used for performing multimodal identification, or even identification of a single modality with multiple samples. The proposed architecture encodes quality in the form of *local* and *global* measures, where the global quality brings together the individual (local) multiple qualities.

Owing to the importance of multimodal biometric authentication in real world scenarios and the need for establishing a better understanding of different approaches, in 2009, Poh et al. [79] benchmarked several multimodal fusion algorithms. The algorithms were evaluated in terms of their recognition performance under different quality and cost constraints. The evaluation was carried out on the first-of-its-kind BioSecure DS2 dataset, consisting of data pertaining to several modalities, along with their quality estimates. One of the key observations of their experiment was that the fusion algorithms which incorporate derived quality measures at run-time for cross-device experiments provided better results, compared to techniques that ignore quality measures. Abaza and Ross [64] presented a Q-based Borda Count technique, which incorporates the quality estimate of the input probe image and gallery sample at the rank-level. The authors proposed modifying the Borda Count technique by incorporating the quality as a weight, in order to reduce the contribution of bad quality samples while performing rank-level fusion. The Q-based Borda Count technique can be written as:

$$R_j = \sum_{i=1}^{C} Q_{i,j} r_{i,j} \tag{4}$$

where, R_j refers to the final fused rank corresponding to the j^{th} subject of the gallery with respect to C biometric classifiers. $Q_{i,j}$ is the minimum quality value between the j^{th} user's gallery image and the i^{th} probe. $r_{i,j}$ refers to the rank assigned to the j^{th} user of the database by the i^{th} classifier. It was observed that including the sample's quality in a multiplicative manner ensures less weight to the low quality samples. Since the fusion techniques are applied at the rank-level, they can easily be incorporated into any existing multimodal classification system. Vatsa et al. [80] proposed a quality-augmented fusion technique of level-2 and level-3 features of fingerprints, based on the Dezert-Smarandache (DSm) theory of paradoxical reasoning. The authors address the scenario of missing information of low quality fingerprint images captured in real world scenarios. Quality scores are computed using the RDWT technique described earlier [61], followed by the extraction of level-2 and level-3 features of the given fingerprint image. These features are then augmented by the quality measure and fused using the DSm theory to obtain the final match score.

Tong et al. [82] built upon the model proposed by Maurer and Baker [78] and proposed a Bayesian Belief Network which incorporates the quality estimates of the probe and gallery images. They emphasized that while it is assumed that all gallery images would have higher quality, it must nevertheless be incorporated when performing identification,

Table 1
A brief summary of techniques incorporating biometric sample quality in the biometric recognition pipeline.

Year	Authors	Description
2003	Bigun et al. [73]	Quality is incorporated in a Bayesian model as a variance parameter
2005	Fierrez-Aguilar et al. [74]	SVM-based model which incorporates quality in its cost function
2005	Poh and Bengio [75]	Margin based quality measure is incorporated as an *a priori* weight for score fusion
2005	Fierrez-Aguilar et al. [76]	Weighted score fusion technique for fingerprint recognition using different features
2006	Nandakumar et al. [77]	Quality-based Product Fusion Score (QPFS) utilizing quality score for gallery and probe pair
2007	Poh et al. [59]	Score normalization technique incorporating device and quality information
2007	Vatsa et al. [61]	RDWT-based quality utilized for multimodal recognition using 2ν-SVM fusion algorithm
2008	Nandakumar et al. [60]	Built over QPFS model by incorporating Gaussian Mixture Models
2008	Maurer and Baker [78]	Bayesian Belief Network modeling *local* and *global* quality
2009	Poh et al. [79]	Benchmarked existing multimodal fusion algorithms under varying quality and cost
2009	Abaza and Ross [64]	Q-based Borda Count technique incorporating quality estimate at rank-level fusion
2009	Vatsa et al. [80]	Quality-augmented fusion technique of level-2 and level-3 fingerprint features
2010	Poh et al. [81]	Incorporates score normalization [59] as pre-processing for different multimodal pipelines
2010	Tong et al. [82]	Bayesian Belief Network which incorporates quality estimates of probe and gallery images
2010	Vatsa et al. [83]	Quality driven image and score level fusion of iris images, followed by probabilistic SVM
2010	Vatsa et al. [84]	Quality driven dynamic context switching for classifier or fusion selection
2012	Zhou et al. [85]	Eye recognition is performed using quality estimates of segmented regions
2013	Rattani et al. [86]	Incorporating sensor influence on image quality and match scores using a graphical model
2015	Bhardwaj et al. [87]	*QFuse*: An online learning framework utilizing quality based context switching
2015	Huang et al. [88]	Adaptive Biomodal Sparse Representation based Classification - quality with Sparse Coding
2016	Ding et al. [89]	Multiple Bayesian Belief Models for fusing match scores and quality values
2016	Muramatsu et al. [90]	View Transformation Model with quality-based score normalization
2017	Liu et al. [91]	Quality Aware Network (QAN): Fuse predicted quality score in a CNN

along with the quality of the probe sample. The causal relationships between the quality of samples, the decision (match/non-match), and the scores are modeled via a probabilistic graphical structure.

Working with a single modality, Vatsa et al. [83] improved the performance of iris recognition by calculating RGB channel-based quality scores and performing fusion using the two lowest quality channels. Match scores corresponding to the fused image and the highest quality image are then provided as input to a probabilistic support vector machine for obtaining the final fused match score. The proposed framework based on image-level and score-level fusion was shown to achieve improved performance. Further, they also presented a classification framework which utilized the quality vector for performing context switching to dynamically select the appropriate fusion or classification algorithm at run-time [84]. In 2012, Zhou et al. [85] proposed a quality driven technique for performing eye recognition. The algorithm involved segmenting the eye into iris and sclera, followed by computing their quality estimates independently. Independent models are trained on all three regions of the eye and, based on the quality of the three regions, a single region is selected; classification is then performed using only this selected region. Ding et al. [89] used Bayesian graphical models to understand the impact of various variables on image quality and vice-versa, and developed techniques to fuse quality with match scores and liveness values in a fingerprint verification system.

Bharadwaj et al. [87] proposed *QFuse*, an online learning framework incorporating quality based context switching for multimodal recognition. Multiple quality metrics are estimated for a given gallery and probe pair, which are provided as input to an ensemble of SVMs for choosing between unimodal classification or fusion-based classification. The incorporation of online learning in the QFuse framework makes it more applicable and usable in real world scenarios. Huang et al. [88] proposed Adaptive Bimodal Sparse-Representation based Classification (ABSRC) for performing feature fusion based on the quality of two samples. A two-step framework is proposed, where initially independent dictionaries are learned for both the modalities. Based on these dictionaries, the feature vector and Sparse Coding Error (SCE) are calculated for the samples, which are then used as quality measures. The SCE is used to generate weights, based on which the two feature vectors of different modalities are concatenated. A similar approach is followed for the learned dictionaries as well, where the dictionaries are simply concatenated based on the weights obtained. The second stage dictionary is used for performing classification. Muramatsu et al. [90] proposed a view transformation model which incorporates quality measure for cross-view gait recognition. Quality values are used to develop a score normalization framework for recognizing gait samples from different views.

It can thus be observed that quality based fusion has been performed at the feature level, score level, and decision level, by different algorithms. Research began with several probabilistic models being proposed for incorporating quality in the recognition framework; however, recent advances have led to the development of representation learning based models incorporating quality estimates in the learning stage as well. Quality information has also been utilized to create context switching based algorithms which select an algorithm for processing the input sample based on its quality. It is interesting to observe from Table 1 that research at the intersection of quality and biometrics has seen some decline in recent literature. Recently, Liu et al. [91] proposed Quality Aware Networks (QANs), for performing set-to-set matching, with application in person re-identification. QAN is built over Convolutional Neural Networks (CNNs), and strengthens our hypothesis that representation learning techniques including deep learning could benefit from fusing quality information in the recognition pipeline. Improved performance might be attained for challenging problems when matching images across scenarios such as cross-resolution face recognition or cross-sensor fingerprint matching.

3.2. Primary biometrics and soft biometrics

Soft biometrics refer to the "*characteristics that provide some information about the user, but lack the distinctiveness and permanence to sufficiently differentiate two individuals*" [93]. Fig. 9 illustrates examples of soft biometric attributes that have been extracted from primary biometric modalities. Examples include gender, ethnicity, age, stride length, weight, eye color, hair color, clothing, and facial accessories. While soft biometric traits are not discriminative enough to uniquely identify a subject, they have often been used in conjunction with primary biometric modalities to complement their performance [94]. One way to utilize soft biometric traits is by utilizing the extracted information - such as gender, ethnicity, skin color, or hair color - to reduce the search space for a given probe sample. Another commonly used technique is to extract soft biometric features and fuse them with the primary biometric trait in order to enhance identification (Fig. 8). In some cases, soft biometric attributes can be gleaned from low-resolution biometric data

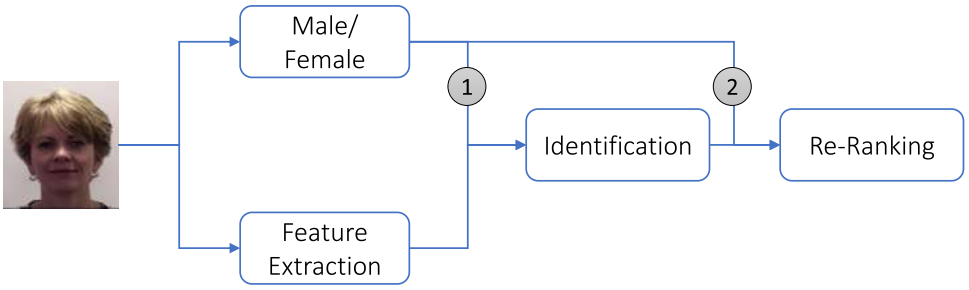


Fig. 8. Example of soft biometric fusion in a general biometric recognition pipeline. (i) Soft biometric information can be fused with primary biometric features for classification, or (ii) soft biometric information can be used for re-ranking the identification list obtained from the primary biometric traits.

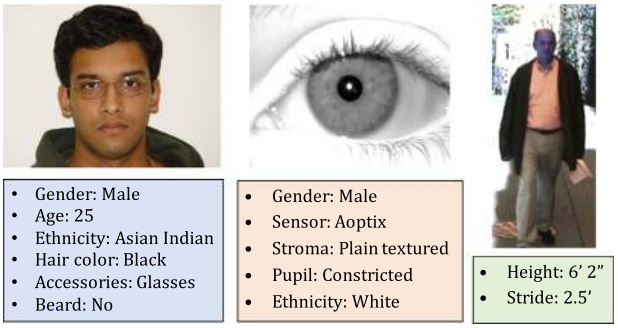


Fig. 9. Examples of soft biometric information evident in the face, iris, and gait modalities. In some cases, manually labeled or annotated soft biometric attributes are used. In other cases, these attributes are automatically extracted.

[95]. A review of techniques for extracting and using soft biometrics, especially in the case of face recognition, are provided in [94,96–99]. Table 2 presents examples of recent algorithms that utilize soft biometric information for recognition.

One of the initial approaches involving soft biometrics (referred to as 'soft measure') for identification was presented by Heckathorn et al. [100]. The authors proposed using attributes such as scars, birthmarks, tattoos, eye color, ethnicity, and gender, along with five biometric measures of height, forearm length, and wrist width, for identifying a given subject. The model was shown to be useful in scenarios where biometric scanners are unavailable and there is a need for maintaining anonymity by eliminating the storage of biometric photographs. The model was built upon the concept of "*interchangeability of indicators*", which states that "*indicators of low accuracy can produce, in combination, a highly accurate indicator*". Jain et al. [93] presented one of the first papers that explored the possibility of fusing soft biometric attributes with primary biometric traits for enhancing the recognition performance of an automated system. A probabilistic model based on the Bayes rule was used for combining the scores generated by the soft biometric and primary biometric systems. Experiments were performed on a fingerprint dataset of 160 subjects, with gender, ethnicity, and height as soft biometrics. It was observed that the utilization of additional information enhances the recognition performance by almost 6%. A similar model was later used by Guo et al. [101] for analyzing the effect of race, gender, height, and weight on cross-age face recognition. Benchmark dataset and results were also provided for the said problem.

Shortly after demonstrating the viability of fusing soft biometrics with a unimodal biometric system, Zewail et al. [102] demonstrated the effectiveness of incorporating the iris color in a multi-modal biometric system consisting of fingerprint and iris. Fusion was performed either via a weighted average of the scores or via a Parzen Classifier. Jain et al. [103] proposed utilizing the Bayesian model presented earlier [93] for combining the soft biometric attributes of gender, height, and ethnicity with fingerprint and face independently, as well as in a multi-modal configuration. Experimental results demonstrated the benefit of combining soft biometrics with both unimodal and multi-modal identification systems.

In 2006, Ailisto et al. [104] analyzed the effect of including weight and fat percentage in a fingerprint recognition system. Multiple fusion techniques were explored at the decision level: AND, OR, and weighted sum. Score level fusion was also analyzed with the help of multilayer perceptrons and SVMs. Experiments across different fusion levels reiterated the benefit of using soft biometric attributes in conjunction

Table 2
Examples of algorithms utilizing soft biometric information in conjunction with primary biometric modalities for person recognition.

Year	Authors	Description
2001	Heckathorn et al. [100]	First study demonstrating the effectiveness of soft biometrics for recognition
2004	Jain et al. [93]	Bayes rule based model for fingerprint recognition with soft and hard biometrics
2004	Zewail et al. [102]	Utilized iris color for performing multimodal recognition of fingerprint and iris
2004	Jain et al. [103]	Utilized gender, height, and ethnicity for fingerprint and face identification
2006	Ailisto et al. [104]	Incorporated weight and fat percentage for fingerprint recognition
2009	Marcialis et al. [105]	Proposed *minority groups* to reduce the false rejection rate using soft biometrics
2009	Abreu et al. [106]	Feature selection using soft biometrics
2010	Moustakas et al. [107]	User height and stride for supplementing gait recognition
2010	Park and Jain [108]	Combined facial marks with an existing face recognition algorithm
2010	Guo et al. [101]	Analyzed effect of race, gender, height, and weight for cross-age recognition
2011	Scheirer et al. [109]	Bayesian Attribute Networks for using descriptive attributes for face recognition
2011	Abreu et al. [110]	Proposed three methods for fusing soft biometric information with primary biometrics
2014	Tome et al. [111]	Evaluated effect of soft biometrics for recognition from a distance
2015	Tome et al. [112]	Fusion of continuous and discrete soft biometric traits for face recognition
2017	Mittal et al. [113]	Utilized soft biometrics for re-ordering the rank list of a face recognition model
2017	Hu et al. [114]	Tensor-based fusion of face recognition features and face attribute features
2017	Schumann and Stiefelhagen [115]	Weighted fusion of attribute prediction and face features for person re-identification
2018	Kazemi et al. [116]	Attribute centered loss for CNNs: match digital faces with sketch-attribute pairs
2018	Liu et al. [117]	Attribute guided triplet loss for heterogeneous face matching

with primary biometrics. Extending to other modalities, Moustakas et al. [107] proposed utilizing the user height and stride length information for supplementing gait recognition. A probabilistic framework was used for this purpose. In 2010, Park and Jain [108] used demographic information (gender and ethnicity) and facial marks (scars, moles, and freckles) to generate a 50-bin histogram. A soft biometric matcher was then created, the score of which was fused with that of a face matcher. This combination was observed to improve biometric performance. Scheirer et al. [109] utilized Bayesian Attribute Networks for combining multiple descriptive attributes for face identification. Descriptive attributes refer to both soft biometric traits and some non-biometric attributes as well. A noisy-OR formulation was presented that demonstrated an improvement of over 32% when compared to a face recognition algorithm. In 2014, Tome et al. [111] evaluated the effect of soft biometrics on the performance of face recognition when capturing data at varying distances from the camera. A number of soft biometric attributes were considered, which were grouped into three categories: *global, body,* and *head*. Score level fusion was then used to combine soft biometric information with face matchers. Sum rule, adaptive switch fusion rule, and a weighted fusion rule were explored, where the benefit of incorporating soft biometrics was especially significant when performing recognition at larger distances.

Soft biometric traits have also been used in other ways. In 2009, Marcialis et al. [105] demonstrated that using soft biometrics such as ethnicity and hair color with a face recognition system can help reduce the False Rejection Rate (FRR) of some users, without significantly affecting the False Acceptance Rate (FAR). A probabilistic framework was presented to predict whether an input face image belonged to a particular user, based on the extracted set of biometric features and the presence of certain soft biometric attributes. Since some soft biometric attributes (e.g., a specific hair color) are associated with only a small number of users (and not with others), it is possible to use such attributes to improve the recognition accuracy. Abreu et al. [106] utilized soft biometric attributes for feature selection, and augmented primary biometric features with the extracted soft biometric features. Experiments in the context of signature biometrics demonstrated that utilizing soft biometric traits increases identification accuracy. Further, the authors evaluated three methods for fusion: majority-based fusion, sum-based fusion, and a sensitivity-based negotiation model [110]. In 2017, Mittal et al. [113] proposed using soft biometrics such as gender, ethnicity, and skin color as a way to re-order the ranked identity list generated by a face matcher. The authors demonstrated improved performance on the problem of composite sketch recognition, where the proposed technique outperformed other algorithms for the said task. In 2018, Swearingen and Ross [118] used a label propagation scheme to deduce missing soft biometric labels (viz., gender and ethnicity) by combining face data with demographic data in a graph-like structure.

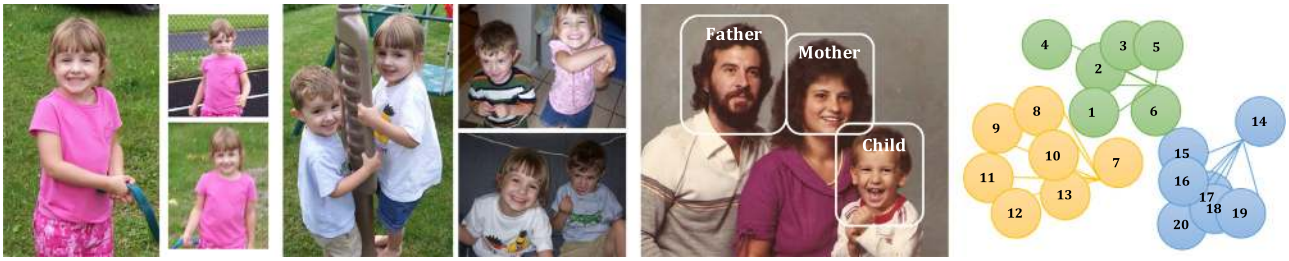
With the increased focus on deep learning techniques, some recent algorithms have incorporated soft biometric information into the deep learning pipeline. The availability of large-scale datasets with attribute information has further facilitated research in this direction [119]. Hu et al. [114] proposed a tensor-fusion based framework for combining face recognition and face attribute features, resulting in a Gated Two-stream Neural Network. Schumann and Stiefelhagen [115] proposed learning *attribute-complementary* features for person re-identification. An attribute classifier is trained for different attributes such as *male, long hair*, and *sunglasses*, followed by a person recognition network. Extracted attributes are provided as input to the recognition module, while learning weights for each attribute, in order to control their influence. Kazemi et al. [116] proposed an attribute-centered loss for training Deep Coupled Convolutional Neural Networks for matching digital face images against forensic sketches. Attribute information about the forensic sketch is used in a pair-wise fashion to learn a shared latent space consisting of several distinct centroids. For matching heterogeneous face images, soft biometric information has been incorporated in the triplet loss function [117].

The effectiveness of combining soft biometrics with primary biometric traits in improving recognition performance can thus be observed across different studies. It is interesting to note that while initial research began in the domain of fingerprint recognition, the effect of soft biometrics is now prominent across the face, iris, and gait modalities as well. Majority of the research, however, has focused on incorporating soft biometrics in a unibiometric system. While some studies have demonstrated improvement in the case of multimodal systems also, dedicated research in this direction is necessary to yield improved performance. Most of the techniques assume prior knowledge about soft biometric attributes at the time of recognition. Developing an integrated system capable of predicting soft biometric information from the input data followed by its fusion in the biometric recognition pipeline might result in more practically deployable systems.

3.3. Biometrics and contextual attributes

Generally, context is used *"to imply acceptable co-occurrence of various parts or attributes of an object or face"* [46]. Contextual attributes have been used as ancillary information along with biometric features in an attempt to aid identification performance. As shown in Fig. 10, different kinds of contextual information have been used for enhancing the recognition performance as well as for automatically tagging images in albums or family photographs. In early work, context was established in terms of temporal, spatial, and even social information. *Temporal* refers to the time of capture of images, *spatial* corresponds to the location where the image was captured, and *social* refers to some information regarding the individuals present in the images or the photographer. However, the notion of context has evolved over time. Table 3 presents examples of techniques incorporating context in the biometric recognition pipeline.

One of the initial algorithms incorporating contextual information to aid face recognition was proposed by Zhang et al. [122]. The authors utilized extended face region (specifically, the clothes) as ancillary information for face recognition. A semi-automated model was proposed that performs face tagging in family photos. The model presents



(a) Clothes Recognition (b) Co-occurrence of People (c) Family/Kin Information (d) Social Network of 20 People

Fig. 10. Examples of contextual information that can be incorporated in order to enhance the biometric recognition performance. Images are taken from the *Images of Groups* [120] and *Gallagher Collection Person* [121] datasets.

Table 3
Examples of techniques incorporating contextual information (e.g., temporal, spatial, social) in the biometric recognition pipeline.

Year	Authors	Description
2003	Zhang et al. [122]	Probabilistic Bayesian framework which incorporates clothing for face tagging
2005	Davis et al. [123]	Incorporated temporal, spatial, and social metadata for face recognition
2006	Song and Leung [124]	Novel clothes recognition algorithm fused with face recognition using spectral clustering
2007	Anguelov et al. [125]	Markov Random Field based technique for fusing clothing and facial features
2008	Gallagher et al. [120]	Clothing segmentation algorithm using graph cuts fused with facial regions for recognition
2008	Stone et al. [126]	Utilized social media networks for automated tagging of images
2009	Kapoor et al. [127]	Incorporate logical contextual constraints into active learning for tagging photographs
2010	Wang et al. [128]	Utilize social familial context for aiding face recognition
2011	Scheirer et al. [109]	Bayesian weighting approach incorporating soft biometric traits and other attributes
2012	Chen et al. [129]	Graph-based technique for predicting pair-wise relationships to improve recognition
2014	Bharadwaj et al. [46]	Social context based re-ranking algorithm using association rules
2014	Hochreiter et al. [130]	Structural Support Vector Machine incorporates album based personal and social costs
2015	Bhardwaj et al. [131]	Fusion of recognition scores obtained from a social graph and face recognition algorithm
2016	Li et al. [132]	Multi-level contextual information for person, photo, and photo-group is used to aid recognition
2017	Kohli et al. [133]	Fusion of kinship verification and face verification scores
2017	Nambiar et al. [134]	Context-specific score-level fusion for gait recognition
2017	Li et al. [135]	Person recognition in photo album using relation and scene context
2018	Sivasankaran et al. [136]	Incorporated context for continuous authentication with multiple classifiers
2018	Sankaran et al. [137]	Siamese architecture incorporating metadata for face recognition
2018	Sultana et al. [138]	Fusion of face and ear biometrics with social network information

a candidate list of potential subjects, out of which the user is asked to select the correct identity. The proposed framework is built over a probabilistic Bayesian framework which works with both facial and contextual features. Davis et al. [123] also proposed a semi-automated model for performing face tagging in photographs. The authors incorporated temporal, spatial, as well as social meta-data to aid in face recognition. Here, temporal refers to the exact time the photo was taken as per the cellular network; spatial refers to the Cell ID from the cellular network and location from Bluetooth-connected GPS receivers; and social refers to the identity of the photographer, the sender and recipients (if any) of the photo, and those who were co-present when the photo was taken (sensed via Bluetooth MAC addresses mapped to usernames). A specific logger was designed and installed on cell-phones to track the aforementioned meta-data. Sparse-Factor Analysis (SFA) was used to perform face recognition using a combination of facial features and contextual meta-data. Experimental evaluation conveyed that utilizing contextual information improves the performance of face recognition compared to using either of the information independently.

In 2006, Song and Leung [124] proposed a model which fused clothing information with face recognition results in order to perform improved person identification. A novel 'clothes recognition' algorithm was proposed, the results of which were integrated into a spectral clustering algorithm for person recognition. Logic-based constraints, such as requiring different individuals in a photograph to correspond to different identities, were also enforced by the clustering algorithm. The authors show that the performance of the clustering algorithm for face-based recognition improves when clothing information is provided and logic-based constraints are imposed.

In 2007, Anguelov et al. [125] proposed a Markov Random Field (MRF) based technique which combines clothing and facial features in order to generate a probabilistic model for predicting the identity corresponding to a given face image. Temporal context in terms of time stamps were used to create *events* based on the clothing of individuals. Different features were used for encoding the color and texture of the clothes, and the Loopy Belief Propagation (LBP) method was used for performing MRF inference. Gallagher et al. [120] proposed a clothing segmentation algorithm based on graph cuts. Features were extracted from the facial and clothing regions, and a probabilistic model was trained to perform person recognition. The authors proposed using the algorithm in the case of consumer image collections, where the number of individuals in an image are known and some individuals have already been labeled by the user. Thus, the task of the model is to detect and label the remaining individuals.

Most of the work until 2008 focused on utilizing spatial and temporal information - such as timestamps of images, geographical location, co-occurrence of individuals and personal clothing - to incorporate contextual information for aiding face recognition in group photo collections. Motivated by the large-scale availability of meta-data on social media sites such as Facebook, Stone et al. [126] proposed a technique to utilize contextual information for complementing face recognition algorithms and automatically tagging face images. The authors collected images and meta-data from a fixed set of Facebook users. The tagged images were then used to train a Conditional Random Field (CRF) algorithm built upon pairwise links between faces observed in photographs. The authors observed improved face recognition performance when incorporating contextual information in the proposed model.

Kapoor et al. [127] proposed the incorporation of logical contextual constraints into the paradigm of active learning to tag group photographs. The framework was presented for performing face tagging on personal photo and video collections using *match* and *non-match* constraints based on prior information. Further, Wang et al. [128] presented a unique formulation for incorporating social familial context into a face recognition pipeline. The authors consider the scenario where weak labels are provided for a given image, and which need to be assigned to each face present in the image. Familial social relationships, such as "mother-child" or "siblings", are used to infer the relative positioning of the face images associated with their corresponding labels. A graphical model is trained for each individual that utilizes facial features as well as features that reflect social relationships. Experimental results on datasets containing consumer images convey the efficacy of the proposed method.

Scheirer et al. [109] proposed incorporating soft biometric traits, descriptive attributes, and contextual information for face recognition. A novel Bayesian weighting approach was proposed which assigns weights to the scores obtained for all the samples in the database based on the attribute network of the gallery images and attributes/context based features extracted from the probe. Chen et al. [129] proposed a graph-based technique for predicting the pair-wise relationship between two faces in a group photograph. The proposed model generates graphs and sub-graphs, in order to understand the social relationships between people, from a given set of group photographs. The authors also propose Bag of Face subGraph (BoFG) which is based on the co-occurrence of individuals in different photographs. For a given test image, the BoFG is calculated and classification is performed based on a Naive Bayes classifier. The authors present improved performance, compared to other

techniques utilizing only descriptive visual features for performing the same task.

Bharadwaj et al. [46] proposed a social context based re-ranking algorithm for improving the classification performance of any classifier by incorporating context based rules. Association rule mining is used for inferring associations between individuals in group photographs. Multiple rules are generated and utilized to obtain context based scores. At the time of testing, these scores are combined with the normalized scores obtained from the classifier, in order to re-rank the results provided by the classifier. Hochreiter et al. [130] also proposed a technique to incorporate album based costs in a recognition framework. Two types of costs, personal and social, are included in the optimization of a structural SVM, in order to include contextual information obtained from albums of photographs.

Another algorithm for updating the rankings obtained from an existing face recognition system was proposed by Bhardwaj et al. [131]. The proposed technique utilizes a social graph (created from training images), where each node is treated as a subject, in order to learn the contextual information between the subjects. For a given group photo, the face recognition scores returned from a traditional face recognition system are combined with those obtained from the social graph, in order to perform context-aided face recognition. Recently, Li et al. [132] proposed a novel framework for utilizing multi-level contextual information at the person, photo, and photo group levels. At the person level, the algorithm utilizes contextual information related to clothes and body appearance, while in photographs of groups, a joint distribution of identities as well as meta-data is used to guide the recognition task. The authors present a framework consisting of SVMs and Conditional Random Fields to incorporate the aforementioned levels of contextual information in the recognition pipeline.

In 2017, Kohli et al. [133] proposed incorporating kinship verification scores as contextual information in the face verification pipeline. A deep learning based framework was used for kinship verification, followed by a score-level fusion with face verification via the product of likelihood ratio and SVM-based approaches. Recently, context information has been incorporated into a classifier ensemble for person re-identification or continuous authentication [134,136]. Sankaran et al. [137] proposed a Siamese convolutional neural network which utilized meta-data of face images such as *yaw, pitch*, and *face size* to enhance face recognition. Sultana et al. [138] proposed incorporating social behavioral information extracted from online social networks in a multi-modal system based on face and ear recognition. Scores of different modalities were fused at the score-level in order to obtain the final decision.

It is interesting to note the progression of what constitutes as contextual information across time. While initial research began with incorporating clothing related information in the recognition pipeline, researchers have now started utilizing social network graphs as well. Temporal information, such as the time of capture, remains an important feature for categorization of photographs into *events* in the case of tagging multiple images. Logical constraints, such as ensuring that different faces in a photograph belong to different individuals, are also often utilized by such algorithms. With the advent of social media, and easy availability of related meta-data, a majority of recent techniques have focused primarily on social networks to aid in the recognition process. Such algorithms implicitly assume an active social media presence, thereby restricting their usability. A combination of contextual information derived from social media and traditional approaches could further enhance recognition performance and improve response time for a query.

3.4. Continuous authentication

In some high security applications, it is necessary for access to be restricted to specific individuals. In such scenarios, there is often a need for authenticating the identity of an individual multiple times. For instance, when accessing confidential data over a length of time using a device, an individual may have to be *continuously* authenticated to ensure that an unauthorized adversary does not view the data during the transaction. It is, therefore, not sufficient for the individual's identity to be authenticated only at the beginning of the session - authentication has to occur at periodic intervals during the entire session (Fig. 11). Depending upon the task at hand, it might be difficult to obtain continuous data pertaining to a single biometric modality. Therefore, multiple cues are needed to facilitate user authentication in such scenarios. One of the initial papers on continuous authentication using multimodal biometrics was by Altinok and Turk [139]. The authors proposed a temporal integration technique for performing continuous authentication using multiple biometric cues - face, voice, and fingerprint. A Bayes classifier was used for combining the normalized match scores across the 3 channels, i.e., the 3 biometric cues. This was followed by using a temporal integration method to generate an expected score distribution and an estimated uncertainty of the distribution. Estimates were calculated as a function of the previous observation and the current time, in order to encode the temporal dependence between the observations.

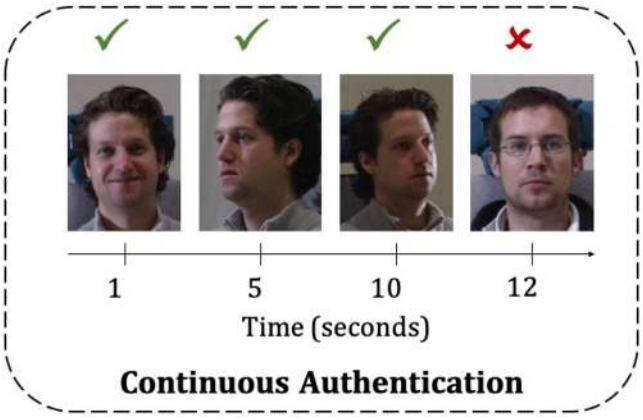


Fig. 11. In a continuous authentication system, the identity of the user is verified at regular intervals. In this illustration, the user is rejected access at t=12 seconds since the system was (correctly) unable to confirm his identity. Face images have been taken from the CMU Multi-PIE dataset [215].

In 2007, Sim et al. [140] developed a *holistic fusion* method built over Hidden Markov Models for integrating evidence from the face and fingerprint modalities over time. Another interesting technique for handling multiple modalities across time was presented by Azzini et al. [141]. The authors proposed a fuzzy controller based model which performed decision level fusion of multiple modalities. The model was built over a *trust* parameter based on which the fuzzy controller decided whether to perform authentication using a single modality or via fusion of multiple modalities. When the *trust* value goes beyond a pre-defined threshold for all scenarios, the user is logged off the system. Kwang et al. [142] performed a study on the usability of continuous authentication systems in real life. A study was performed on 58 participants wherein they were required to perform certain tasks on a Windows machine equipped with a Continuous Biometrics Authentication System (CBAS). The authors concluded in favor of using a CBAS despite having substantial system overhead.

Most of the research until 2010 focused on fusing primary biometric traits, such as fingerprint and face, for performing continuous authentication. Depending upon the system at hand and acquisition environment, obtaining a good quality face or fingerprint template *continuously* might not be a viable assumption. One can expect pose and illumination variations, incomplete or no capture, or even forced repeated co-operation from users. Inspired by these observations, Niinuma et al. [143] proposed utilizing soft biometric traits for performing continuous authentication. The proposed framework combines

Table 4
A brief summary of techniques used for continuous authentication in a multimodal setting.

Year	Authors	Description
2003	Altinok and Turk [139]	Bayes classifier based technique for temporal integration of multiple modalities across time
2007	Sim et al. [140]	Holistic fusion method built over Hidden Markov Models
2008	Azzini et al. [141]	Fuzzy controller based model for dynamically switching between modalities
2009	Kwang et al. [142]	Study on usability of a continuous authentication based system with 58 participants
2010	Niinuma et al. [143]	Combines soft biometric traits of face and clothing color with PCA based face recognition
2010	Shi et al. [144]	Learns user profile based on mobile phone usage habits using Gaussian Mixture Models
2013	Frank et al. [145]	Proposes a set of 30 behavioral touch features for smart-phone authentication
2016	Sitová et al. [146]	Hand Movement, Orientation, and Grasp (HMOG) for smartphones
2017	Peng et al. [147]	*GlassGuard*: Continuous authentication system for Google Glass using touch and voice features
2017	Fenu et al. [148]	Multimodal fusion technique for face, voice, touch, mouse, and keystroke based features
2018	Kumar et al. [149]	Score-level fusion of one-class classifiers
2018	Shen et al. [150]	Feature-level combination of multi-motion sensor behavior

face color and clothing color with PCA based face recognition. Continuous authentication is performed using the soft biometric traits, and face recognition is used for template enrollment and re-authentication. Score level fusion is utilized for continuous authentication, which is governed by a fixed threshold; a re-login is requested when authentication fails. Shi et al. [144] proposed a framework for learning user profiles based on their usage habits on mobile phones. Information such as call logs, SMS logs, browsing habits, and GPS location are used to generate features for learning each user's profile. In real time, each event generates a score value, based on which the identity of the user is either confirmed or refuted. The authors also evaluate the robustness of their method to different types of adversarial attacks.

In 2013, Frank et al. [145] proposed using a set of 30 behavioral biometric features for performing continuous authentication on smartphones. User interaction with the touch screen of their smartphone is analyzed to develop a set of different strokes that are classified using the KNN and SVM classifiers. A combination of hand movement, orientation, and grasp (HMOG) features have also been used to continuously authenticate a user on a smartphone [146]. Data collected from the accelerometer, gyroscope, and magnetometer is combined to perform unobtrusive authentication. Patel et al. [151] presented a survey on continuous authentication focusing on both unimodal and multimodal techniques. The authors discuss the progress in the field along with existing methods and related challenges. With the development in technology, the domain of continuous authentication has further expanded to incorporate wearable devices such as Google Glass as well. Peng et al. [147] proposed *GlassGuard* which fuses decision scores obtained from touch gestures and voice commands in a probabilistic manner to authenticate the user's identity. Researchers have also worked on continuous authentication for e-learning platforms [148], where a multimodal fusion technique utilizing face, voice, touch, mouse, and keystroke is proposed. Independent verification scores are calculated for each modality which are then combined via a score fusion mechanism. In order to reduce the computational cost, verification is performed at predefined time intervals.

Table 4 presents a brief summary of algorithms proposed for continuous authentication. It can be observed that the field of continuous authentication has evolved tremendously over the past decade. Researchers have attempted to address the problem in desktops, laptops, smart phones, and even wearable devices. While there exists several sophisticated algorithms in the literature, a major challenge is the lack of publicly available large datasets for the given problem. Computational efficiency in continuous authentication further remains as one of the major challenges requiring dedicated attention.

4. Biometric fusion and presentation attack detection

A biometric system is vulnerable to a number of attacks [153]. One such attack is referred to as a *presentation attack* where an adversary presents a fake or altered biometric trait to the sensor with the intention of spoofing someone else's trait; or creating a new virtual identity based on the presented trait; or obfuscating their own trait. Detecting such attacks is essential in improving the security and integrity of the biometric system. In earlier literature, *presentation attack* was synonymous with *spoof attack* and, therefore, the terms *spoof detection* and *anti-spoofing* have been used in connection with presentation attack detection (PAD).

The task of spoof detection has also been viewed as *liveness* detection, especially in the initial research revolving around fingerprint recognition [167]. Liveness detection involves predicting whether a given sample is *live/bonafide* or not, that is, whether the input is captured from a human being or is a synthetically generated artifact. Liveness or spoof detection modules can be integrated into a biometric recognition pipeline in order to create systems robust to attacks [165,168,169]. Marasco et al. [12] presented different frameworks for integrating a spoof detector with a fingerprint recognition module. The authors evaluate different techniques including sequential methods, classifier-based fusion, and a Bayesian Belief Network (BBN) which explicitly models the relationship between the spoof detection scores and biometric match scores. Further, Ding and Ross [89] explored multiple BBN architectures for modeling the influence of liveness scores on match scores (and vice-versa) and used these architectures for fusing the two scores.

In the literature, the task of spoof detection has generally been handled independently for different biometric modalities. Researchers have focused on analyzing the effect of different *attacks* on biometric systems for various modalities including face, fingerprint, iris, vein-pattern, hand geometry, and speech. Recent surveys on anti-spoofing algorithms for these modalities can be found in [170–174]. A major section of research utilizes a combination of texture or quality based features for the given task. Most of these techniques involve feature level fusion of different descriptors, where features are first concatenated and then input to a classifier (often a SVM) for spoof detection [175–180].

Wen et al. [154] proposed using Image Distortion Analysis (IDA) for identifying spoofed face images. An ensemble of SVMs is developed based on four features: specular reflection, blurriness, chromatic moment, and color diversity. The authors argue that different spoof attacks might be easily identified by different features and, therefore, learned a separate SVM for each feature. Score level fusion schemes using the min-rule and the sum-rule were used for taking the final decision. Raghavendra et al. [155] proposed using Light Field Camera (LFC) for performing facial spoof detection. As opposed to regular cameras, LFCs can be used to render an image with variations in focus and depth. This characteristic enabled the authors to observe distinct differences between the real and spoofed images. Specifically, for presentation attacks, where the input sensor is presented with a print-out of another person's biometric sample, feature level concatenation of estimated variations in focus resulted in improved performance. Arashloo et al. [156] proposed fusing Multiscale Binarized Statistical Image Features on Three Orthogonal Planes (MBSIF-TOP) and Multiscale Local Phase Quantization on Three Orthogonal Planes (MLPQ-TOP) for performing spoof detection. Fusion

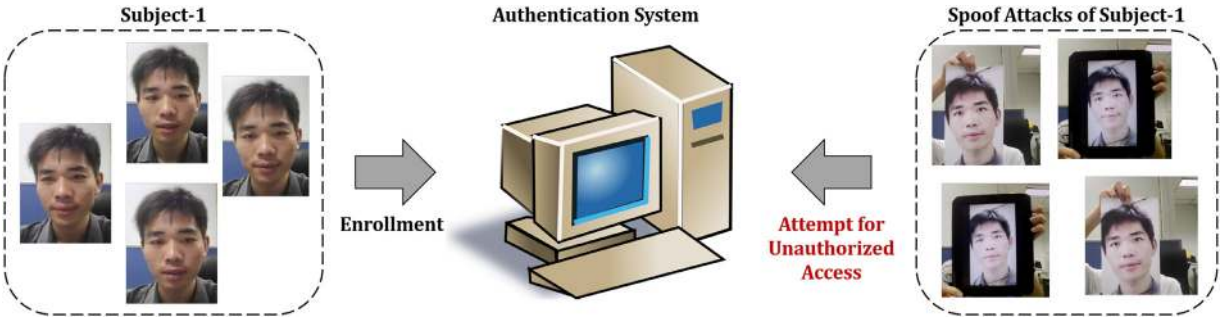


Fig. 12. Presentation attacks on a face recognition system where an adversary attempts to spoof the identity of Subject-1. Face images have been taken from the CASIA-FASD database [152].

was performed by a kernel fusion approach, termed as Spectral Regression Kernel Discriminant Analysis (SR-KDA).

In 2016, Boulkenafet et al. [157] proposed using local texture features of both the luminance and chrominance channels for performing facial spoof detection. Features extracted using Co-Occurrence of Adjacent Local Binary Patterns (CoALBP) and Local Phase Quantization (LPQ) were concatenated in the HSV and YCbCr color spaces. This was followed by classification using a SVM. The proposed technique achieved state-of-the-art results on three datasets, thereby demonstrating the benefit of combining texture features from different color spaces. Patel et al. [158] proposed concatenating color moments and Local Binary Patterns (LBP) from a given face image for performing spoof detection. An input RGB image was converted into the HSV space for calculating color moments, and the concatenated feature vector was input to a SVM for classification. Siddiqui et al. [159] proposed a multi-feature face spoof detection algorithm consisting of a multi-scale configuration of LBP and Histogram of Oriented Optical Flow (HOOF) features classified using a SVM. Experiments were performed using spoofed and bonafide video samples, wherein intra-feature and inter-feature fusion was performed at the score level. In [181], the authors designed a novel deep learning architecture that fused a CNN with a RNN in order to extract pseudo-depth images and a remote photoplethysmography (RPPG) signal from an input face video. The extracted information were then fused for face anti-spoofing.

Recently, Toosi et al. [161] presented a comparative study with ten feature descriptors for the task of fingerprint spoof detection. The authors experimented with different fusion strategies to achieve improved performance. The authors also proposed *SpiderNet*, a two-stage deep learning architecture to learn independent and combined features for different feature inputs in order to identify spoofed images. Korshunov and Marcel [162] studied the impact of score fusion for performing presentation attack detection in case of voice biometrics. The authors used eight state-of-the-art algorithms to understand the effect of mean, logistic regression, and polynomial logistic regression fusion methods. The authors also provided open source implementations of the detection algorithms, fusion techniques, and evaluation framework. A novel framework for fusing electrocardiogram (ECG) recognition with fingerprint spoof detection was proposed by Komeili et al. [163]. Two classifiers - one for ECG verification and the other for fingerprint spoof detection - were trained independently and score-level fusion was performed using weighted sum, product, and maximum fusion rules. Yadav et al. [164] presented a framework for iris spoof detection, where features from a deep learning algorithm, VGG, were fused with texture based RDWT + Haralick features. The features were concatenated and input to a neural network for performing iris spoof detection. Deep learning based CNNs have also been shown to perform well for fingerprint spoof detection [166]. Here, minutiae detection is performed on an input fingerprint image, followed by patch generation and alignment. A CNN architecture is used to generate the liveness score for each patch, followed by score-level fusion. Ding and Ross [160] proposed the use of an ensemble of one-class classifiers in order to handle the problem of limited spoof samples during training as well as address the need to develop methods for detecting previously unseen spoofs in the context of fingerprints. Each one-class classifier was based on a simple texture descriptor and was trained predominantly on bonafide fingerprint samples. Fusion of these multiple classifiers was observed to increase the robustness of the spoof detector on novel spoof fabrication materials.

It is interesting to observe that while most of the techniques do not utilize modality-specific information for performing spoof detection, none of the papers have demonstrated results across different biometric modalities (Table 5). Most algorithms are evaluated on controlled data collected in a laboratory environment, which often does not simulate the real world well. It is thus essential to develop datasets that better imitate real world scenarios in order to develop robust algorithms capable of improving state-of-the-art performance and demonstrating better generalization abilities. The Liveness Detection Competition Series[1] corresponds to a series of spoof detection challenges conducted regularly since 2009. The competition series aims at evaluating and benchmarking anti-spoofing algorithms for the tasks of iris and fingerprint spoof detection. With the development of multi-modal systems, algorithms must also be developed to handle spoof detection for multi-modal systems, especially when it might be easier to spoof one modality compared to the other (see [168,182]).

Recently, the related area of adversarial detection has garnered substantial attention, especially in the domain of deep learning [183,184]. It has been shown that adversarial samples can be created by adding small perceptible or imperceptible perturbations to the input images, which can then be used to *fool* a recognition system [185,186]. The presence of an adversarial detection module often results in a more robust recognition system that is reasonably immune to such adversarial attacks [187,188]. Most of the research in the area of adversarial detection utilizing fusion has focused on the intermediate representations obtained from the learned networks. Li and Li [189] proposed using the convolutional filter outputs of a CNN model for detecting adversarial samples. Different filter outputs are used to compute statistics for a given input which are then provided to a classifier cascade for adversarial detection. Product rule fusion is applied on the scores returned by the classifiers. Goswami et al. [188,190] proposed learning the difference between the mean unperturbed features and representations extracted from the adversarial samples. Features are extracted from multiple intermediate layers of a CNN model, followed by a SVM model for adversarial detection. The authors proposed *selective dropout* for handling adversarial samples by mitigating the effect of the adversary. Tao et al. [191] proposed detecting adversarial face samples by combining attribute information in a traditional face recognition system. As can be observed, limited research has focused on utilizing information fusion techniques for adversarial detection. As described earlier, multiple

[1] http://livdet.org/competitions.php

Table 5
A summary of some techniques that use fusion for either spoof (i.e., presentation attack) detection or for combining the anti-spoofing module of a biometric system with the biometric matcher itself.

Year	Authors	Description
2012	Marasco et al. [12]	Different frameworks for integrating a spoof detection module with a recognition system
2015	Wen et al. [154]	Ensemble of SVMs on reflection, blurriness, chromatic moment, and color diversity
2015	Raghavendra et al. [155]	Feature level concatenation with Light Field Camera based features
2015	Arashloo et al. [156]	Fused MBSIF-TOP and MLPQ-TOP using SR-KDA
2016	Ding et al. [89]	Bayesian Belief Networks for fusing match scores with liveness scores
2016	Boulkenafet et al. [157]	CoALBP and LPQ features in HSV and YCbCr colour space
2016	Patel et al. [158]	Concatenation of color moments and LBP features
2016	Siddiqui et al. [159]	Inter-feature and intra-feature score-level fusion of multi-scale LBP and HOOF features
2016	Ding and Ross [160]	Fusion of multiple one-class SVMs to improve generalizability of a fingerprint spoof detector
2017	Toosi et al. [161]	Comparative study of different fusion techniques on ten fingerprint features
2017	Korshunov and Marcel [162]	Studies impact of score fusion on presentation attack detection for voice
2018	Komeili et al. [163]	Fusion of ECG recognition and fingerprint spoof detection
2018	Yadav et al. [164]	Fusion of (VGG features+PCA) with (RDWT + Haralick) features and neural network
2018	Sajjad et al. [165]	Two-tier authentication system for recognition and spoof detection
2018	Chugh et al. [166]	CNN based spoof detection on fingerprint patches

details can be extracted from a biometric sample, such as soft biometric attributes or quality scores, thereby rendering it rich in information. Thus, the inclusion of such fusion methods in the adversarial detection module could result in enhanced performance, resulting in more robust recognition systems.

5. Multibiometric cryptosystems

Data used by a biometric system can be encrypted using strong cryptographic techniques in order to secure them against external attacks. In addition, biometric matching has to be performed in the encrypted domain to obviate the need to decrypt the data, thereby preventing an adversary from viewing the original data at any time. In multibiometric systems, where data from multiple biometric sources are available, each data piece has to be encrypted. Such systems consisting of multiple biometric data sources, along with some cryptographic technique for securing the data, are termed as Multibiometric Cryptosystems. Table 6 presents a brief summary of techniques proposed for multibiometric cryptosystems. Rathgeb and Busch [195] present a comprehensive review of the work done in the field of securing biometric templates, for both multibiometric as well as unibiometric systems. In their work, the concepts of securing a biometric template as well as a multibiometric template are explained in detail, along with a discussion of a theoretical framework for multibiometric cryptosystems.

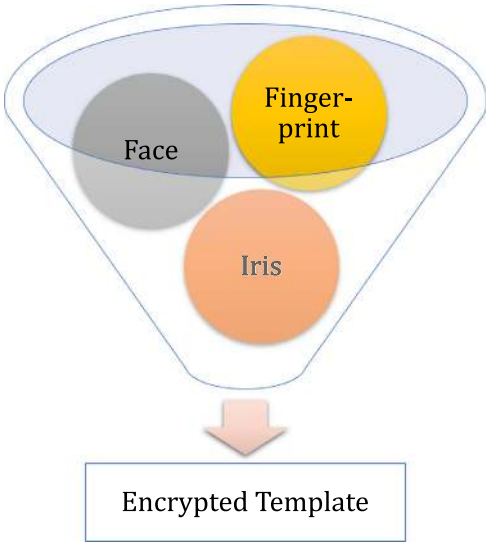


Fig. 13. Multibiometric cryptosystems utilize multiple biometric modalities and perform encryption in order to secure the biometric data of a subject.

In order to secure multibiometric systems, Sutcu et al. [192] proposed the fusion of face and fingerprint features, followed by a *secure sketch* construct for securing the fused samples, thereby making it difficult to reconstruct the original samples from the encrypted features. Minutiae based features were extracted from fingerprints, and Singular Value Decomposition (SVD) based features were used for the face modality. Nandakumar and Jain [23] proposed a fuzzy vault framework for securing multibiometric templates consisting of fingerprint and iris. Experiments were performed using multiple impressions of the same biometric modality (fingerprint), multiple instances of a biometric modality (two index fingers), and data from multiple biometric modalities (fingerprint and iris). In the proposed technique, all features were represented as elements of a Galois Field, GF. Separate techniques were presented for feature extraction from fingerprint and iris, which were then fused at the feature level, and secured via a fuzzy vault technique. Camlikaya et al. [193] proposed a template fusion technique for the fingerprint and voice modalities. The proposed algorithm strengthened the security of the multibiometric system by encoding the minutiae features obtained from the fingerprints within the voice feature vector. The authors also motivated the chosen modalities by emphasizing the desirable cancelable property of spoken words being used as a password.

In 2009, Fu et al. [194] proposed several multibiometric cryptosystem models. One model for performing fusion at the biometric level was proposed, while three models were presented for performing fusion at the cryptographic level. While no experimental evaluation was performed for the proposed models, however, an in-depth analysis of the algorithms, comparisons, and discussions were presented by the authors. Nagar et al. [11] proposed a feature-level fusion based multibiometric cryptosystem for performing recognition using features from multiple biometric modalities. A single secure sketch was generated from multiple features (of different modalities) based on two biometric cryptosystems - fuzzy vault and fuzzy commitment. A detailed experimental analysis was conducted on datasets pertaining to three modalities: face, iris, and fingerprints.

Rathgeb and Busch [22] proposed a Bloom filter based transformation technique for performing iris identification. The algorithm fused the features obtained from the left and right irides and obscured the information present in each instance independently, thus securing it against external attacks. Li et al. [197] proposed a new method for performing security analysis on multibiometric cryptosystems, based on a combination of principles from information-theory and computational security. The authors also proposed a decision-level fusion based multibiometric cryptosystem for performing fingerprint recognition. A two stage encryption was performed on the extracted

Table 6
A few examples of techniques that have been used for multibiometric cryptosystems.

Year	Authors	Description
2007	Sutcu et al. [192]	*Secure Sketch* construct for protecting face and fingerprint templates
2008	Nandakumar and Jain [23]	Fuzzy vault framework for securing multibiometric templates
2008	Camlikaya et al. [193]	Encodes minutiae features (fingerprint) within a voice feature vector
2009	Fu et al. [194]	Models for performing fusion at the cryptographic level
2012	Nagar et al. [11]	Fusion of fuzzy vault and fuzzy commitment to generate a single secure sketch
2012	Rathgeb et al. [195]	Comprehensive survey of techniques used for securing biometric templates
2014	Rathgeb and Busch [22]	Bloom filter based technique for fusing multiple features
2014	Chin et al. [196]	Random tiling and equi-probable 2^N discretization scheme for fingerprint and palmprint
2015	Li et al. [197]	Proposed technique performs security analysis on multibiometric cryptosystems
2016	Kumar and Kumar [198]	Combination of BHC encoding and hash code computation followed by cell array storage

features, followed by decision-level fusion to obtain the identity of the given sample. Kumar and Kumar [198] proposed a cell array based multibiometric cryptosystem. Bose Chaudhuri Hocquenghem (BCH) encoding and hash code computation was performed on the biometric modalities. The data was stored in the form of two cell arrays such that the hash code is distributed in the first cell array and the key is scattered across the second. Furthermore, two models were proposed, one for decision-level fusion and another for feature-level fusion. Experimental analysis depicted the superiority of decision-level fusion over feature-level fusion for the proposed multibiometric cryptosystem.

We observe that the security aspect of multibiometric systems has garnered dedicated attention. Initially, researchers applied the techniques prevalent in unibiometric systems on the fused feature vectors of different modalities. Techniques such as generation of secure sketches and fuzzy vault constructs have now been well explored. Several novel techniques have also been proposed with the aim of being more effective in terms of security and computation. In order to increase the robustness of such systems and enhance their real world applicability, research has also focused on proposing new metrics for the evaluation of the models. It is interesting to note that most of the techniques proposed in the literature for multibiometric security have focused on hand-crafted features, with a limited focus on representation learning based algorithms. We believe that this is bound to change, given the increasing interest in utilizing deep neural networks for addressing the problem of multibiometric security.

6. Research challenges and future directions

Biometric fusion has witnessed significant advancements over the past two decades in terms of algorithm development, sources of information being fused, application domains and operational data collected. The literature review in Sections 2–5 suggests that research in biometric fusion has primarily focused on combining multiple sources of information for different problems and designing new fusion algorithms. The questions of *what, when*, and *how* to fuse are important for the development of a biometric fusion system, and need to be answered during the algorithm design. However, in order to develop efficient real world biometric fusion systems, they also require efficient implementation and domain adaptation to account for changes in sensor technology, environment, target population, etc. In order to be practically deployable, we believe the following are some important research topics that require more attention and focused research efforts.

(i) **Portability of Multibiometric Solutions:** Most fusion algorithms have a number of tunable parameters. For example, even the simple sum rule for score-level fusion requires the estimation of score normalization parameters and the weight vector. Automatically deducing these parameters for different applications is not an easy task. Even learning-based methods for automatically deducing fusion parameters are vulnerable to biases in the training data. Thus, directly transferring the fusion module from one application to another may not be viable in practical systems. The problem is further exacerbated due to several variations, for instance, differences in sensors, environment, and population across applications. This raises the issue of portability. How can one design robust fusion systems that can be easily ported across applications?
Domain adaptation and transfer learning have been touted to be effective paradigms for adapting machine learning methods to new application domains [199–201]. Research at the intersection of biometric fusion and domain adaptation can have two potential directions: (a) utilizing biometric fusion for domain adaptation, and (b) incorporating domain adaptation in existing multibiometric systems for cross-domain matching. Both these directions have real world applicability with widespread impact in terms of (a) addressing the long standing problem of cross-domain matching, or (b) updating existing multibiometric systems to handle data emerging from fundamentally different distributions.

(ii) **Designing Adaptive and Dynamic Fusion Systems:** In real world applications, multibiometric systems often have to operate on large-scale data captured using multiple sensors across different geographical regions from a diverse heterogeneous population (e.g., national ID card program in India). Further, the requirements of an application and the nature of its data may change over time. In the literature, techniques such as online learning or co-training have shown to improve the performance of unibiometric recognition systems by updating them *on the fly* [201,202]. However, online learning based techniques are yet to be explored for multibiometric systems. This is an open area of research, i.e., designing fusion methods that continually evolve over time to accommodate changes in system requirements as well as variations in data distribution.
A pertinent problem is the issue of template update, i.e., modifying the stored biometric data of an individual in order to account for intra-class variations [203,204]. Aging and physical ailment can modify the biometric trait of an individual thereby requiring the enrolled data to be periodically updated. Updating the multibiometric templates of a subject over time can be an arduous task and may inadvertently result in identity creep where an impostor can exploit the template update mechanism to take over the identity of an enrolled subject. An adaptive fusion system should be able to discourage such attacks while still accounting for inevitable changes in data distribution that occur over time.

(iii) **Multibiometric Security and Privacy:** Research in soft biometrics has established the possibility of deducing additional information about an individual (e.g., age, gender, ethnicity, health condition, and genetic disorders) from their biometric data or template [94]. While this information can be used to improve recognition accuracy, it can also be deemed to violate the privacy of the subject and, in some cases, can be used for profiling an individual. The availability of multibiometric data corresponding to multiple biometric traits will only increase concerns about compromising the privacy of subjects. It is, therefore, necessary to impart security and privacy to the stored templates. In addition,

there must be legislative guarantees that prevent the data from being used beyond the purposes for which it was intended at the time of enrollment.

Recent work in differential privacy in the context of a single biometric modality [205,206] could potentially be extended to multibiometric templates. However, *guaranteeing* privacy may not be an easy task especially due to the advent of powerful deep learning techniques that can be leveraged for gleaning ancillary information about a subject that was previously not thought to be possible [207]. Another related challenge has to do with the retention of recognition accuracy while imparting security and privacy. In many cases, the use of privacy preserving or security enhancing schemes results in a degradation in recognition accuracy. Balancing security and privacy with matching accuracy is, therefore, an important challenge that needs to be judiciously resolved. Recent work in homomorphic encryption could potentially be appropriated for this purpose [208]. The use of non-biometric cues may also be needed to facilitate privacy and enhance security [209].

The principle of "signal mixing" is also being used to impart security and privacy to biometric data. Othman and Ross [47,210] describe a mixing scheme where an input fingerprint image is mixed with another fingerprint (e.g., from a different finger) in order to produce a new mixed image, that obscures the identity of the original fingerprint. This can be viewed as a data-level fusion approach. The researchers also developed a method to fuse two distinct modalities, viz., fingerprint and iris, at the image level [211].

(iv) **Resolving Conflicts Between Information Sources:** The availability of multiple biometric sources and, consequently, multiple pieces of biometric evidence, is not always beneficial. In some cases, the individual biometric sources can offer conflicting decisions about the identity of a subject. For example, in a bimodal identification system, the face and fingerprint modalities may generate a completely different list of ranked identities; or, in a multimodal verification system, half of the component classifiers might confirm the claimed identity, while the other half might refute the claimed identity. In such scenarios, it is necessary to have a principled way to generate a decision. To address this, it may be necessary to re-acquire the biometric traits of an individual and recompute the decisions. Another possibility is to consider only the outputs of the most reliable sources. However, the reliability of a source (e.g., a matcher) will depend upon multiple factors including the quality of the data and the baseline performance of the matcher itself. Pragmatic methods are needed to handle such operationally relevant situations. Another related problem is the uncertainty associated with the decisions rendered by individual matchers in a multibiometric framework. Incorporating these uncertainty (or confidence) values in the decision architecture would be essential.

(v) **Predicting Scalability of Multibiometric Systems:** A number of models have been developed to predict the scalability of a biometric system relying on a single modality [212]. Such prediction models are needed to evaluate the suitability of a given biometric system for an anticipated large-scale application. Developing such prediction models for multibiometric systems is not easy since these systems rely on multiple sources of information and, therefore, the performance of each source has to be first modeled and then combined with the models associated with the other sources [213,214]. Alternately, the entire multibiometric system can be characterized using a single model. In either case, the degrees of freedom to be considered can be intractable. Thus, predicting the scalability of a multibiometric system requires much more research, and effective models are needed to characterize the complex relationship between the individual sources.

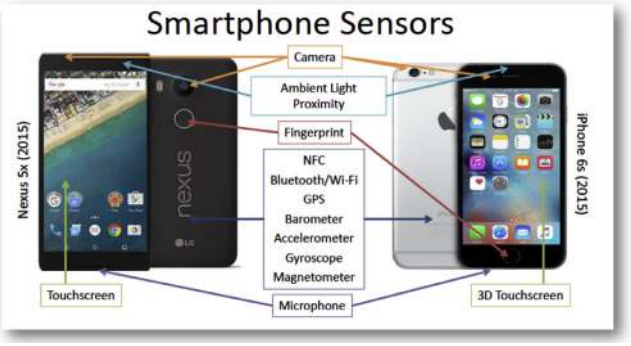


Fig. 14. Smartphones are equipped with a number of sensors. Data from these sensors can be combined in a judicious manner to perform multimodal user authentication. However, a number of research issues have to be resolved when designing such a solution. ©Debayan Deb.

(vi) **Sensor Configuration in Multimodal Systems:** One of the understudied problems in multimodal biometrics is the placement of sensors in the data acquisition module to maximize recognition performance while minimizing user inconvenience. Consider a multibiometric kiosk that identifies individuals based on their gait, face, and fingerprint modalities. As the subject approaches the kiosk, the system uses the gait information from a distance to retrieve a list of potential identities. When the subject is reasonably close to the kiosk, the face image is used to further narrow down the list of matching identities and possibly even determining the exact identity of the subject. The fingerprint sensor is only invoked if the gait and face modalities are unable to uniquely identify the subject. Such a system can arguably improve the throughput of the biometric system. However, camera placement and configuration would be a critical issue in this application. More research is needed to model user behavior in such applications and suitably adjust the placement and position of sensors to ensure that the correct identity can be rapidly determined with limited inconvenience to the user.

(vii) **Multimodal Solutions for Compact Personal Devices:** With the increasing use of mobile smartphones and wearable devices, and the need for reliably establishing identity in such compact systems, there is a tremendous opportunity to develop novel biometric sensors. Further, these personal devices are already equipped with a large number of sensors (GPS, accelerometer, gyroscope, magnetometer, microphone, NFC, and heart rate monitors) whose data can be used to identify a subject or to verify an identity in a continuous authentication scheme (Fig. 14). However, principled methods are needed to parse through this heterogeneous data and distill a compact representation that can be used for personal authentication. As described earlier, a number of continuous authentication methods have been developed for personal devices. But these methods typically pre-define the sensor modalities to be used for authentication purposes. Further, they are vulnerable to changes in a subject's behavior and cannot be easily ported from one device to another. Thus, there is a need to develop robust schemes for extracting a distinct and generalizable "signature" of a subject from the massive amounts of diverse data being generated by devices such as smartphones.

In summary, while tremendous advances have been made in the field of biometric fusion, it is now time to translate these advancements into operational systems. This provides an unprecedented opportunity for researchers to develop multibiometric solutions that are (a) practically feasible; (b) user friendly; (c) ergonomically tenable; (d) amenable to increased subject throughput; (e) scalable to large heterogeneous populations; (f) compliant with security and privacy requirements; and (g)

robust to changes in the environment, population, sensors, etc. Such solutions can impact a number of application domains including consumer electronics, banking, autonomous vehicles, robotics, health and medicine, e-commerce, law enforcement, welfare disbursement, border security, national ID cards, cybersecurity and e-voting.

Acknowledgments

M. Singh and R. Singh are partially supported through the Infosys CAI at IIIT-Delhi. Ross was supported by the US National Science Foundation under Grant Numbers 1618518 and 1617466 during the writing of this article.

References

[1] A. Jain, A. Ross, K. Nandakumar, Introduction to Biometrics: A Textbook, Springer, US, 2011.
[2] A.K. Jain, S.Z. Li, Handbook of face Recognition, Springer-Verlag, London, 2011.
[3] M.J. Burge, K. Bowyer, Handbook of iris Recognition, Springer-Verlag, London, 2013.
[4] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar, Handbook of Fingerprint Recognition, Springer-Verlag, London, 2009.
[5] A.K. Jain, K. Nandakumar, A. Ross, 50 Years of biometric research: accomplishments, challenges, and opportunities, Pattern Recognit. Lett. 79 (2016) 80–105.
[6] Spain airports implement a multi-biometric solution, (https://www.secureidnews.com/news-item/spain-airports-implement-a-multi-biometric-solution/).
[7] Multi-biometrics: the future of border control, (https://www.idemia.com/news/multi-biometrics-future-border-control-2016-04-21).
[8] V. Draluk, F. Goldfain, J.-W. Maarse, Multilevel passcode authentication, US Patent Number US8806610B2 (2014).
[9] [In-Depth Look] Samsungs Biometric Technologies Bring Added Security and Convenience to the Galaxy S8, (https://tinyurl.com/mkxsbcz).
[10] A. Ross, K. Nandakumar, A.K. Jain, Handbook of multibiometrics, Springer Publishers, 2006.
[11] A. Nagar, K. Nandakumar, A.K. Jain, Multibiometric cryptosystems based on feature-level fusion, IEEE Trans. Inf. Forensics Secur. 7 (1) (2012) 255–268.
[12] E. Marasco, Y. Ding, A. Ross, Combining match scores with liveness values in a fingerprint verification system, in: IEEE Conference on Biometrics: Theory, Applications and Systems, 2012, pp. 418–425.
[13] S. Crihalmeanu, A. Ross, S. Schuckers, L. Hornak, A protocol for multibiometric data acquisition, storage and dissemination, Technical Report, WVU, Lane Department of Computer Science and Electrical Engineering, 2007.
[14] A. Agarwal, D. Yadav, N. Kohli, R. Singh, M. Vatsa, A. Noore, Face presentation attack with latex masks in multispectral videos, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 275–283.
[15] G. Goswami, M. Vatsa, R. Singh, RGB-D Face recognition with texture and attribute features, IEEE Trans. Inf. Forensics Secur. 9 (10) (2014) 1629–1640.
[16] N. Cvejic, S.G. Nikolov, H.D. Knowles, A. Loza, A. Achim, D.R. Bull, C.N. Canagarajah, The effect of pixel-level fusion on object tracking in multi-sensor surveillance video, in: IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[17] R. Singh, M. Vatsa, A. Noore, Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition, Pattern Recognit. 41 (3) (2008) 880–893.
[18] R. Singh, M. Vatsa, A. Noore, Hierarchical fusion of multi-spectral face images for improved recognition performance, Inf. Fusion 9 (2) (2008) 200–210.
[19] T. Bourlai (Ed.), Face Recognition Across the Imaging Spectrum, Springer International Publishing, 2016.
[20] A. Ross, A. Jain, J. Reisman, A hybrid fingerprint matcher, Pattern Recognit. 36 (7) (2003) 1661–1673.
[21] A. Kumar, D. Zhang, Personal authentication using multiple palmprint representation, Pattern Recognit. 38 (10) (2005) 1695–1704.
[22] C. Rathgeb, C. Busch, Cancelable multi-biometrics: mixing iris-codes based on adaptive bloom filters, Comput. Secur. 42 (2014) 1–12.
[23] K. Nandakumar, A.K. Jain, Multibiometric template security using fuzzy vault, in: IEEE International Conference on Biometrics: Theory, Applications and Systems, 2008.
[24] A. Uhl, P. Wild, Single-sensor multi-instance fingerprint and eigenfinger recognition using weighted score combination methods, Int. J. Biom. 1 (4) (2009) 442–462.
[25] U. Park, A.K. Jain, A. Ross, Face recognition in video: Adaptive fusion of multiple matchers, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2007, pp. 1–8.
[26] G. Goswami, M. Vatsa, R. Singh, Face verification via learned representation on feature-rich video frames, IEEE Trans. Inf. Forensics Secur. 12 (7) (2017) 1686–1698.
[27] H.S. Bhatt, R. Singh, M. Vatsa, On recognizing faces in videos using clustering-based re-ranking and fusion, IEEE Trans. Inf. Forensics Secur. 9 (7) (2014) 1056–1068.
[28] S. Zhou, V. Krueger, R. Chellappa, Probabilistic recognition of human faces from video, Comput. Vision Image Understanding 91 (1) (2003) 214–245.
[29] L. Wang, H. Ning, T. Tan, W. Hu, Fusion of static and dynamic body biometrics for gait recognition, IEEE Trans. Circuits Syst. Video Technol. 14 (2) (2004) 149–158.
[30] M. Hu, Y. Wang, Z. Zhang, D. Zhang, J.J. Little, Incremental learning for video-based gait recognition with LBP flow, IEEE Trans. Cybern. 43 (1) (2013) 77–89.
[31] M.S. Nixon, T. Tan, R. Chellappa, Human identification based on gait, 4, Springer, Boston, MA, 2010.
[32] A. Kale, A. Sundaresan, A.N. Rajagopalan, N.P. Cuntoor, A.K. Roy-Chowdhury, V. Kruger, R. Chellappa, Identification of humans using gait, IEEE Trans. Image Process. 13 (9) (2004) 1163–1173.
[33] G. Goswami, P. Mittal, A. Majumdar, M. Vatsa, R. Singh, Group sparse representation based classification for multi-feature multimodal biometrics, Inf. Fusion 32 (2016) 3–12.
[34] A.K. Jain, L. Hong, Y. Kulkarni, A multimodal biometric system using fingerprint, face and speech, in: International Conference on Audio-and Video-based Biometric Person Authentication, 1999, pp. 182–187.
[35] S. Ben-Yacoub, Y. Abdeljaoued, E. Mayoraz, Fusion of face and speech data for person identity verification, IEEE Trans. Neural Netw. 10 (5) (1999) 1065–1074.
[36] A. Chowdhury, Y. Atoum, L. Tran, X. Liu, A. Ross, MSU-AVIS dataset: fusing face and voice modalities for biometric recognition in indoor surveillance videos, in: In Proceeding of International Conference on Pattern Recognition, Beijing, China, 2018.
[37] K. Chang, K.W. Bowyer, S. Sarkar, B. Victor, Comparison and combination of ear and face images in appearance-based biometrics, IEEE Trans. Pattern Anal. Mach. Intell. 25 (9) (2003) 1160–1165.
[38] D.L. Woodard, S. Pundlik, P. Miller, R. Jillela, A. Ross, On the fusion of periocular and iris biometrics in non-ideal imagery, in: International Conference on Pattern Recognition, 2010, pp. 201–204.
[39] Q. Zhang, H. Li, Z. Sun, T. Tan, Deep feature fusion for iris and periocular biometrics on mobile devices, IEEE Trans. Inf. Forensics Secur. 13 (11) (2018) 2897–2912.
[40] T. Wark, S. Sridharan, Adaptive fusion of speech and lip information for robust speaker identification, Digit. Signal Process. 11 (3) (2001) 169–186.
[41] H. Çetingül, E. Erzin, Y. Yemez, A. Tekalp, Multimodal speaker/speech recognition using lip motion, lip texture and audio, Signal Processing 86 (12) (2006) 3549–3558.
[42] V.P. Minotto, C.R. Jung, B. Lee, Multimodal multi-channel on-line speaker diarization using sensor fusion through SVM, IEEE Trans. Multimedia 17 (10) (2015) 1694–1705.
[43] J. Fierrez, A. Morales, R. Vera-Rodriguez, D. Camacho, Multiple classifiers in biometrics. part 2: trends and challenges, Inf. Fusion 44 (2018) 103–112.
[44] A. Kumar, A. Kumar, Adaptive management of multimodal biometrics fusion using ant colony optimization, Inf. Fusion 32 (2016) 49–63.
[45] A. Ross, A. Jain, Information fusion in biometrics, Pattern Recognit. Lett. 24 (13) (2003) 2115–2125.
[46] S. Bharadwaj, M. Vatsa, R. Singh, Aiding face recognition with social context association rule based re-ranking, in: IEEE International Joint Conference on Biometrics, 2014.
[47] A. Othman, A. Ross, On mixing fingerprints, IEEE Trans. Inf. Forens. Secur. 8 (1) (2013) 260–267.
[48] R. Singh, M. Vatsa, A. Ross, A. Noore, A mosaicing scheme for pose-invariant face recognition, IEEE Trans. Syst., Man, Cybernet., Part B 37 (5) (2007) 1212–1225.
[49] A. Ross, S. Shah, J. Shah, Image versus feature mosaicing: a case study in fingerprints, in: Proceedings of SPIE Conference on Biometric Technology for Human Identification, 6202, Orlando, USA, 2006, pp. 1–12.
[50] A. Kong, D. Zhang, M. Kamel, Palmprint identification using feature-level fusion, Pattern Recognit. 39 (3) (2006) 478–487.
[51] A. Kumar, D. Zhang, Personal recognition using hand shape and texture, IEEE Trans. Image Process. 15 (8) (2006) 2454–2461.
[52] A.A. Ross, R. Govindarajan, Feature level fusion of hand and face biometrics, in: Proceedings of SPIE, 2005.
[53] A. Gyaourova, A. Ross, Index codes for multibiometric pattern retrieval, IEEE Trans. Inf. Forens. Secur. 7 (2) (2012) 518–529.
[54] D. Yi, Z. Lei, S.Z. Li, Shared representation learning for heterogenous face recognition, in: IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, 2015, pp. 1–7.
[55] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, Pattern Recognit. 38 (12) (2005) 2270–2285.
[56] N. Poh, S. Bengio, Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication, Pattern Recognit. 39 (2) (2006) 223–233.
[57] M. He, S.-J. Horng, P. Fan, R.-S. Run, R.-J. Chen, J.-L. Lai, M.K. Khan, K.O. Sentosa, Performance evaluation of score level fusion in multimodal biometric systems, Pattern Recognit. 43 (5) (2010) 1789–1800.
[58] M.B. Ylmaz, B. Yankolu, Score level fusion of classifiers in off-line signature verification, Inf. Fusion 32 (2016) 109–119.
[59] N. Poh, J. Kittler, T. Bourlai, Improving biometric device interoperability by likelihood ratio-based quality dependent score normalization, in: IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007, pp. 1–5.
[60] K. Nandakumar, Y. Chen, S.C. Dass, A. Jain, Likelihood ratio-based biometric score fusion, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2008) 342–347.
[61] M. Vatsa, R. Singh, A. Noore, Integrating image quality in 2ν-SVM biometric match score fusion, Int. J. Neural Syst. 17 (05) (2007) 343–351.
[62] N. Poh, J. Kittler, A unified framework for biometric expert fusion incorporating quality measures, IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) (2012) 3–18.
[63] Y. Ding, A. Ross, A comparison of imputation methods for handling missing scores in biometric fusion, Pattern Recognit 45 (3) (2012) 919–933.
[64] A. Abaza, A. Ross, Quality based rank-level fusion in multibiometric systems, in: IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2009.
[65] T.K. Ho, J.J. Hull, S.N. Srihari, Decision combination in multiple classifier systems, IEEE Trans Pattern Anal Mach Intell 16 (1) (1994) 66–75.

[66] M.M. Monwar, M.L. Gavrilova, Multimodal biometric system using rank-level fusion approach, IEEE Trans. Syst., Man, Cybernet., Part B 39 (4) (2009) 867–878.
[67] A. Kumar, S. Shekhar, Personal identification using multibiometrics rank-level fusion, IEEE Trans. Syst., Man, Cybernet., Part C (Applications and Reviews) 41 (5) (2011) 743–752.
[68] K. Veeramachaneni, L. Osadciw, A. Ross, N. Srinivas, Decision-level fusion strategies for correlated biometric classifiers, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2008.
[69] S. Prabhakar, A.K. Jain, Decision-level fusion in fingerprint verification, Pattern Recognit. 35 (4) (2002) 861–874.
[70] M. Indovina, U. Uludag, R. Snelick, A. Mink, A. Jain, Multimodal biometric authentication methods: a COTS approach, Workshop Multimodal User Authentication (2003) 99–106.
[71] A. Ross, N. Poh, Multibiometric systems: overview, case studies, and open issues, in: Handbook of Remote Biometrics, Springer Publishers, 2009, pp. 273–292.
[72] S. Bharadwaj, M. Vatsa, R. Singh, Biometric quality: a review of fingerprint, iris, and face, EURASIP J. Image Video Process. (1) (2014) 34.
[73] J. Bigun, J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Multimodal biometric authentication using quality signals in mobile communications, in: International Conference on Image Analysis and Processing, 2003, pp. 2–11.
[74] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, J. Bigun, Discriminative multimodal biometric authentication based on quality measures, Pattern Recognit. 38 (5) (2005) 777–779.
[75] N. Poh, S. Bengio, Improving fusion with margin-derived confidence in biometric authentication tasks, in: International Conference on Audio- and Video-Based Biometric Person Authentication, 2005, pp. 474–483.
[76] J. Fierrez-Aguilar, Y. Chen, J. Ortega-Garcia, A.K. Jain, Incorporating image quality in multi-algorithm fingerprint verification, in: International Conference on Biometrics, 2005, pp. 213–220.
[77] K. Nandakumar, Y. Chen, A.K. Jain, S.C. Dass, Quality-based score level fusion in multibiometric systems, in: International Conference on Pattern Recognition, 2006, pp. 473–476.
[78] D.E. Maurer, J.P. Baker, Fusing multimodal biometrics with quality estimates via a bayesian belief network, Pattern Recognit 41 (3) (2008) 821–832.
[79] N. Poh, T. Bourlai, J. Kittler, L. Allano, F. Alonso-Fernandez, O. Ambekar, J. Baker, B. Dorizzi, O. Fatukasi, J. Fierrez, H. Ganster, J. Ortega-Garcia, D. Maurer, A.A. Salah, T. Scheidat, C. Vielhauer, Benchmarking quality-dependent and cost-sensitive score-level multimodal biometric fusion algorithms, IEEE Trans. Inf. Forensics Secur. 4 (4) (2009) 849–866.
[80] M. Vatsa, R. Singh, A. Noore, M.M. Houck, Quality-augmented fusion of level-2 and level-3 fingerprint information using DSm theory, Int. J. Approximate Reason. 50 (1) (2009) 51–61.
[81] N. Poh, J. Kittler, T. Bourlai, Quality-based score normalization with device qualitative information for multimodal biometric fusion, IEEE Trans. Syst., Man, Cybernet.- Part A 40 (3) (2010) 539–554.
[82] Y. Tong, F.W. Wheeler, X. Liu, Improving biometric identification through quality-based face and fingerprint biometric fusion, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2010, pp. 53–60.
[83] M. Vatsa, R. Singh, A. Ross, A. Noore, Quality-based fusion for multichannel iris recognition, in: International Conference on Pattern Recognition, 2010, pp. 1314–1317.
[84] M. Vatsa, R. Singh, A. Noore, A. Ross, On the dynamic selection of biometric fusion algorithms, IEEE Trans. Inf. Forensics Secur. 5 (3) (2010) 470–479.
[85] Z. Zhou, E.Y. Du, C. Belcher, N.L. Thomas, E.J. Delp, Quality fusion based multimodal eye recognition, in: IEEE International Conference on Systems, Man, and Cybernetics, 2012, pp. 1297–1302.
[86] A. Rattani, N. Poh, A. Ross, A bayesian approach for modeling sensor influence on quality, liveness and match score values in fingerprint verification, in: IEEE International Workshop on Information Forensics and Security, 2013, pp. 37–42.
[87] S. Bharadwaj, H.S. Bhatt, R. Singh, M. Vatsa, A. Noore, Qfuse: online learning framework for adaptive biometric system, Pattern Recognit. 48 (11) (2015) 3428–3439.
[88] Z. Huang, Y. Liu, X. Li, J. Li, An adaptive bimodal recognition framework using sparse coding for face and ear, Pattern Recognit Lett. 53 (2015) 69–76.
[89] Y. Ding, A. Rattani, A. Ross, Bayesian belief models for integrating match scores with liveness and quality measures in a fingerprint verification system, in: International Conference on Biometrics, 2016, pp. 1–8.
[90] D. Muramatsu, Y. Makihara, Y. Yagi, View transformation model incorporating quality measures for cross-view gait recognition, IEEE Trans Cybern. 46 (7) (2016) 1602–1615.
[91] Y. Liu, J. Yan, W. Ouyang, Quality aware network for set to set recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4694–4703.
[92] J.P. Baker, D.E. Maurer, Fusion of biometric data with quality estimates via a bayesian belief network, in: Biometric Symposium, 2005, pp. 21–22.
[93] A.K. Jain, S.C. Dass, K. Nandakumar, Can soft biometric traits assist user recognition? in: Biometric Technology for Human Identification, 2004, pp. 561–573.
[94] A. Dantcheva, P. Elia, A. Ross, What else does your biometric data reveal? a survey on soft biometrics, IEEE Trans. Inf. Forensics Secur. 11 (3) (2016) 441–467.
[95] M. Singh, S. Nagpal, R. Singh, M. Vatsa, Class representative autoencoder for low resolution multi-spectral gender classification, in: International Joint Conference on Neural Networks, 2017, pp. 1026–1033.
[96] E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez, F. Alonso-Fernandez, Facial soft biometrics for recognition in the wild: recent works, annotation, and COTS evaluation, IEEE Trans. Inf. Forensics Secur. 13 (8) (2018) 2001–2014.
[97] B.H. Guo, M.S. Nixon, J.N. Carter, Fusion analysis of soft biometrics for recognition at a distance, in: IEEE International Conference on Identity, Security, and Behavior Analysis, 2018.
[98] H. Zhang, J.R. Beveridge, B.A. Draper, P.J. Phillips, On the effectiveness of soft biometrics for increasing face verification rates, Comput. Vision Image Understand. 137 (2015) 50–62.
[99] D.A. Reid, S. Samangooei, C. Chen, M.S. Nixon, A. Ross, Soft biometrics for surveillance: an overview, in: Handbook of statistics, 2013, pp. 327–352.
[100] D.D. Heckathorn, R.S. Broadhead, B. Sergeyev, A methodology for reducing respondent duplication and impersonation in samples of hidden populations, J. Drug Issues 31 (2) (2001) 543–564.
[101] G. Guo, G. Mu, K. Ricanek, Cross-age face recognition on a very large database: the performance versus age intervals and improvement using soft biometric traits, in: International Conference on Pattern Recognition, 2010, pp. 3392–3395.
[102] R. Zewail, A. Elsafi, M. Saeb, N. Hamdy, Soft and hard biometrics fusion for improved identity verification, Midwest Symposium on Circuits and Systems, 2004.
[103] A.K. Jain, K. Nandakumar, X. Lu, U. Park, Integrating faces, fingerprints, and soft biometric traits for user recognition, in: International Workshop on Biometric Authentication, 2004, pp. 259–269.
[104] H. Ailisto, E. Vildjiounaite, M. Lindholm, S.-M. Mäkelä, J. Peltola, Soft biometrics - combining body weight and fat measurements with fingerprint biometrics, Pattern Recognit. Lett. 27 (5) (2006) 325–334.
[105] G.L. Marcialis, F. Roli, D. Muntoni, Group-specific face verification using soft biometrics, J. Visual Languag. Comput. 20 (2) (2009) 101–109.
[106] M. Abreu, M. Fairhurst, Improving identity prediction in signature-based unimodal systems using soft biometrics, in: European Workshop on Biometrics and Identity Management, 2009, pp. 348–356.
[107] K. Moustakas, D. Tzovaras, G. Stavropoulos, Gait recognition using geometric features and soft biometrics, IEEE Signal Process. Lett. 17 (4) (2010) 367–370.
[108] U. Park, A.K. Jain, Face matching and retrieval using soft biometrics, IEEE Trans. Inf. Forensics Secur. 5 (3) (2010) 406–415.
[109] W.J. Scheirer, N. Kumar, K. Ricanek, P.N. Belhumeur, T.E. Boult, Fusing with context: a bayesian approach to combining descriptive attributes, in: International Joint Conference on Biometrics, 2011.
[110] M.C. Da Costa Abreu, M. Fairhurst, Enhancing identity prediction using a novel approach to combining hard-and soft-biometric information, IEEE Trans. Syst., Man, Cybern., Part C 41 (5) (2011) 599–607.
[111] P. Tome, J. Fierrez, R. Vera-Rodriguez, M.S. Nixon, Soft biometrics and their application in person recognition at a distance, IEEE Trans. Inf. Forensics Secur. 9 (3) (2014) 464–475.
[112] P. Tome, R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, Facial soft biometric features for forensic face recognition, Forensic Sci. Int. 257 (2015) 271–284.
[113] P. Mittal, A. Jain, G. Goswami, M. Vatsa, R. Singh, Composite sketch recognition using saliency and attribute feedback, Inf. Fusion 33 (2017) 86–99.
[114] G. Hu, Y. Hua, Y. Yuan, Z. Zhang, Z. Lu, S.S. Mukherjee, T.M. Hospedales, N.M. Robertson, Y. Yang, Attribute-enhanced face recognition with neural tensor fusion networks., in: International Conference on Computer Vision, 2017, pp. 3764–3773.
[115] A. Schumann, R. Stiefelhagen, Person re-identification by deep learning attribute–complementary information, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1435–1443.
[116] H. Kazemi, S. Soleymani, A. Dabouei, M. Iranmanesh, N.M. Nasrabadi, Attribute-centered loss for soft-biometrics guided face sketch-photo recognition, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 499–507.
[117] D. Liu, N. Wang, C. Peng, J. Li, X. Gao, Deep attribute guided representation for heterogeneous face recognition, in: International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization, 2018, pp. 835–841.
[118] T. Swearingen, A. Ross, Label propagation approach for predicting missing biographic labels in face-based biometric records, IET Biom. 7 (1) (2018) 71–80.
[119] Z. Liu, P. Luo, X. Wang, X. Tang, Deep learning face attributes in the wild, in: International Conference on Computer Vision, 2015.
[120] A.C. Gallagher, T. Chen, Clothing cosegmentation for recognizing people, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[121] A. Gallagher, T. Chen, Understanding images of groups of people, in: International Conference on Computer Vision and Pattern Recognition, 2009.
[122] L. Zhang, L. Chen, M. Li, H. Zhang, Automated annotation of human faces in family albums, in: ACM International Conference on Multimedia, 2003, pp. 355–358.
[123] M. Davis, M. Smith, J. Canny, N. Good, S. King, R. Janakiraman, Towards context-aware face recognition, in: ACM International Conference on Multimedia, 2005, pp. 483–486.
[124] Y. Song, T. Leung, Context-aided human recognition–clustering, in: European Conference on Computer Vision, 2006, pp. 382–395.
[125] D. Anguelov, K. c. Lee, S.B. Gokturk, B. Sumengen, Contextual identity recognition in personal photo albums, in: IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[126] Z. Stone, T.E. Zickler, T. Darrell, Autotagging facebook: social network context improves photo annotation, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2008.
[127] A. Kapoor, G. Hua, A. Akbarzadeh, S. Baker, Which faces to tag: adding prior constraints into active learning, in: International Conference on Computer Vision, 2009, pp. 1058–1065.
[128] G. Wang, A. Gallagher, J. Luo, D. Forsyth, Seeing people in social context: recognizing people and social relationships, in: European Conference on Computer Vision, 2010, pp. 169–182.

[129] Y.-Y. Chen, W.H. Hsu, H.-Y.M. Liao, Discovering informative social subgraphs and predicting pairwise relationships from group photos, in: International Conference on Multimedia, 2012, pp. 669–678.
[130] J. Hochreiter, Z. Han, S.Z. Masood, S. Fonte, M. Tappen, Exploring album structure for face recognition in online social networks, Image Vis. Comput. 32 (10) (2014) 751–760.
[131] R. Bhardwaj, G. Goswami, R. Singh, M. Vatsa, Harnessing social context for improved face recognition, in: International Conference on Biometrics, 2015, pp. 121–126.
[132] H. Li, J. Brandt, Z. Lin, X. Shen, G. Hua, A multi-level contextual model for person recognition in photo albums, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1297–1305.
[133] N. Kohli, M. Vatsa, R. Singh, A. Noore, A. Majumdar, Hierarchical representation learning for kinship verification, IEEE Trans. Image Process. 26 (1) (2017) 289–302.
[134] A. Nambiar, A. Bernardino, J.C. Nascimento, A. Fred, Context-aware person re-identification in the wild via fusion of gait and anthropometric features, in: IEEE International Conference on Automatic Face Gesture Recognition, 2017, pp. 973–980.
[135] Y. Li, G. Lin, B. Zhuang, L. Liu, C. Shen, A. van den Hengel, Sequential person recognition in photo albums with a recurrent network, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5660–5668.
[136] D. Sivasankaran, M. Ragab, T. Sim, Y. Zick, Context-aware fusion for continuous biometric authentication, in: International Conference on Biometrics, 2018, pp. 233–240.
[137] N. Sankaran, S. Tulyakov, S. Setlur, V. Govindaraju, Metadata-based feature aggregation network for face recognition, in: International Conference on Biometrics, 2018, pp. 118–123.
[138] M. Sultana, P.P. Paul, M.L. Gavrilova, Social behavioral information fusion in multimodal biometrics, IEEE Trans. Syst. Man, Cybern. 48 (12) (2018) 2176–2187.
[139] A. Altinok, M. Turk, Temporal integration for continuous multimodal biometrics, Workshop on Multimodal User Authentication, 2003.
[140] T. Sim, S. Zhang, R. Janakiraman, S. Kumar, Continuous verification using multimodal biometrics, IEEE Trans. Pattern Anal. Mach. Intell. 29 (4) (2007) 687–700.
[141] A. Azzini, S. Marrara, R. Sassi, F. Scotti, A fuzzy approach to multimodal biometric continuous authentication, Fuzzy Optim. Decis. Making 7 (3) (2008) 243–256.
[142] G. Kwang, R.H. Yap, T. Sim, R. Ramnath, An usability study of continuous biometrics authentication, in: International Conference on Biometrics, 2009, pp. 828–837.
[143] K. Niinuma, U. Park, A.K. Jain, Soft biometric traits for continuous user authentication, IEEE Trans. Inf. Forensics Secur. 5 (4) (2010) 771–780.
[144] E. Shi, Y. Niu, M. Jakobsson, R. Chow, Implicit authentication through learning user behavior, in: International Conference on Information Security, 2011, pp. 99–113.
[145] M. Frank, R. Biedert, E. Ma, I. Martinovic, D. Song, Touchalytics: on the applicability of touchscreen input as a behavioral biometric for continuous authentication, IEEE Trans. Inf. Forensics Secur. 8 (1) (2013) 136–148.
[146] Z. Sitová, J. šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, K.S. Balagani, HMOG: new behavioral biometric features for continuous authentication of smartphone users, IEEE Trans. Inf. Forensics Secur. 11 (5) (2016) 877–892.
[147] G. Peng, G. Zhou, D.T. Nguyen, X. Qi, Q. Yang, S. Wang, Continuous authentication with touch behavioral biometrics and voice on wearable glasses, IEEE Trans. Hum. Mach. Syst. 47 (3) (2017) 404–416.
[148] G. Fenu, M. Marras, L. Boratto, A multi-biometric system for continuous student authentication in e-learning platforms, Pattern Recognit. Lett. 113 (2018) 83–92.
[149] R. Kumar, P.P. Kundu, V.V. Phoha, Continuous authentication using one-class classifiers and their fusion, in: IEEE International Conference on Identity, Security, and Behavior Analysis, 2018.
[150] C. Shen, Y. Li, Y. Chen, X. Guan, R.A. Maxion, Performance analysis of multi-motion sensor behavior for active smartphone authentication, IEEE Trans. Inf. Forensics Secur. 13 (1) (2018) 48–62.
[151] V.M. Patel, R. Chellappa, D. Chandra, B. Barbello, Continuous user authentication on mobile devices: recent progress and remaining challenges, IEEE Signal Process Mag 33 (4) (2016) 49–61.
[152] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S.Z. Li, A face antispoofing database with diverse attacks, in: IAPR International Conference on Biometrics, 2012, pp. 26–31.
[153] N. Ratha, J. Connell, R. Bolle, Enhancing security and privacy of biometric-based authentication systems, IBM Syst. J. 40 (2001) 614–634.
[154] D. Wen, H. Han, A.K. Jain, Face spoof detection with image distortion analysis, IEEE Trans. Inf. Forensics Secur. 10 (4) (2015) 746–761.
[155] R. Raghavendra, K.B. Raja, C. Busch, Presentation attack detection for face recognition using light field camera, IEEE Trans. Image Process. 24 (3) (2015) 1060–1075.
[156] S.R. Arashloo, J. Kittler, W. Christmas, Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features, IEEE Trans. Inf. Forensics Secur. 10 (11) (2015) 2396–2407.
[157] Z. Boulkenafet, J. Komulainen, A. Hadid, Face spoofing detection using colour texture analysis, IEEE Trans. Inf. Forensics Secur. 11 (8) (2016) 1818–1830.
[158] K. Patel, H. Han, A.K. Jain, Secure face unlock: spoof detection on smartphones, IEEE Trans. Inf. Forensics Secur. 11 (10) (2016) 2268–2283.
[159] T.A. Siddiqui, S. Bharadwaj, T.I. Dhamecha, A. Agarwal, M. Vatsa, R. Singh, N. Ratha, Face anti-spoofing with multifeature videolet aggregation, in: International Conference on Pattern Recognition, 2016, pp. 1035–1040.
[160] Y. Ding, A. Ross, An ensemble of one-class SVMs for fingerprint spoof detection across different fabrication materials, in: IEEE International Workshop on Information Forensics and Security, 2016, pp. 1–6.
[161] A. Toosi, A. Bottino, S. Cumani, P. Negri, P.L. Sottile, Feature fusion for fingerprint liveness detection: a comparative study, IEEE Access 5 (2017) 23695–23709.
[162] P. Korshunov, S. Marcel, Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations, IEEE J. Sel. Top. Signal Process. 11 (4) (2017) 695–705.
[163] M. Komeili, N. Armanfard, D. Hatzinakos, Liveness detection and automatic template updating using fusion of ecg and fingerprint, IEEE Trans. Inf. Forensics Secur. 13 (7) (2018) 1810–1822.
[164] D. Yadav, N. Kohli, A. Agarwal, M. Vatsa, R. Singh, A. Noore, Fusion of handcrafted and deep learning features for large-scale multiple iris presentation attack detection, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018.
[165] M. Sajjad, S. Khan, T. Hussain, K. Muhammad, A.K. Sangaiah, A. Castiglione, C. Esposito, S.W. Baik, Cnn-based anti-spoofing two-tier multi-factor authentication system, Pattern Recognit. Lett. (2018), doi:10.1016/j.patrec.2018.02.015.
[166] T. Chugh, K. Cao, A.K. Jain, Fingerprint spoof buster: use of minutiae-centered patches, IEEE Trans. Inf. Forensics Secur. 13 (9) (2018) 2190–2202.
[167] L. Ghiani, D.A. Yambay, V. Mura, G.L. Marcialis, F. Roli, S.A. Schuckers, Review of the fingerprint liveness detection (LivDet) competition series: 2009 to 2015, Image Vis Comput 58 (2017) 110–128.
[168] E. Marasco, P. Johnson, C. Sansone, S. Schuckers, Increase the security of multibiometric systems by incorporating a spoofing detection algorithm in the fusion mechanism, in: International Workshop on Multiple Classifier Systems, 2011, pp. 309–318.
[169] I. Chingovska, A. Anjos, S. Marcel, Anti-spoofing in action: joint operation with a verification system, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, pp. 98–104.
[170] J. Galbally, S. Marcel, J. Fierrez, Biometric antispoofing methods: a survey in face recognition, IEEE Access 2 (2014) 1530–1552.
[171] E. Marasco, A. Ross, A survey on anti-spoofing schemes for fingerprints, ACM Comput. Surv. (2014) 1–35.
[172] R. Ramachandra, C. Busch, Presentation attack detection methods for face recognition systems: a comprehensive survey, ACM Comput. Surv. 50 (1) (2017) 8:1–8:37.
[173] J. Galbally, M. Gomez-Barrero, A review of iris anti-spoofing, in: International Conference on Biometrics and Forensics, 2016.
[174] Z. Wu, N. Evans, T. Kinnunen, J. Yamagishi, F. Alegre, H. Li, Spoofing and countermeasures for speaker verification: a survey, Speech. Commun. 66 (2015) 130–153.
[175] J. Määttä, A. Hadid, M. Pietikäinen, Face spoofing detection from single images using micro-texture analysis, in: International Joint Conference on Biometrics, 2011.
[176] D. Gragnaniello, G. Poggi, C. Sansone, L. Verdoliva, Fingerprint liveness detection based on weber local image descriptor, in: IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, 2013, pp. 46–50.
[177] S. Bharadwaj, T.I. Dhamecha, M. Vatsa, R. Singh, Computationally efficient face spoofing detection with motion magnification, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, pp. 105–110.
[178] A. Agarwal, R. Singh, M. Vatsa, Face anti-spoofing using haralick features, in: IEEE International Conference on Biometrics Theory, Applications and Systems, 2016.
[179] J. Galbally, S. Marcel, Face anti-spoofing based on general image quality assessment, in: International Conference on Pattern Recognition, 2014, pp. 1173–1178.
[180] K. Patel, H. Han, A.K. Jain, G. Ott, Live face video vs. spoof face video: use of moiré patterns to detect replay video attacks, in: International Conference on Biometrics, 2015, pp. 98–105.
[181] Y. Liu, A. Jourabloo, X. Liu, Learning deep models for face anti-spoofing: binary or auxiliary supervision, In Proceeding of IEEE Computer Vision and Pattern Recognition, 2018.
[182] R.N. Rodrigues, N. Kamat, V. Govindaraju, Evaluation of biometric spoofing in a multimodal system, in: IEEE International Conference on Biometrics: Theory, Applications and Systems, 2010, pp. 1–5.
[183] N. Akhtar, A. Mian, Threat of adversarial attacks on deep learning in computer vision: a survey, IEEE Access 6 (2018) 14410–14430.
[184] B. Biggio, G. Fumera, P. Russu, L. Didaci, F. Roli, Adversarial biometric recognition: a review on biometric system security from the adversarial machine-learning perspective, IEEE Signal Process Mag. 32 (5) (2015) 31–41.
[185] S.-M. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, P. Frossard, Universal adversarial perturbations, in: IEEE International Computer Vision and Pattern Recognition, 2017, pp. 86–94.
[186] S.-M. Moosavi-Dezfooli, A. Fawzi, P. Frossard, Deepfool: a simple and accurate method to fool deep neural networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2574–2582.
[187] A. Agarwal, R. Singh, M. Vatsa, N. Ratha, Are image agnostic universal adversarial perturbations for face recognition difficult to detect, IEEE Int. Conf. Biom. (2018).
[188] G. Goswami, A. Agarwal, N. Ratha, R. Singh, M. Vatsa, Detecting and mitigating adversarial perturbations for robust face recognition, Int. J. Comput. Vis. (2019).
[189] X. Li, F. Li, Adversarial examples detection in deep networks with convolutional filter statistics, in: IEEE International Conference on Computer Vision, 2017, pp. 5775–5783.
[190] G. Goswami, N. Ratha, A. Agarwal, R. Singh, M. Vatsa, Unravelling robustness of deep learning based face recognition against adversarial attacks, in: AAAI Conference on Artificial Intelligence, 2018.
[191] G. Tao, S. Ma, Y. Liu, X. Zhang, Attacks meet interpretability: attribute-steered detection of adversarial samples, in: Advances in Neural Information Processing Systems, 2018, pp. 7728–7739.
[192] Y. Sutcu, Q. Li, N. Memon, Secure biometric templates from fingerprint-face features, in: IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[193] E. Camlikaya, A. Kholmatov, B. Yanikoglu, Multi-biometric templates using fingerprint and voice, in: Proceedings of SPIE, 2008.

[194] B. Fu, S.X. Yang, J. Li, D. Hu, Multibiometric cryptosystem: model structure and performance analysis, IEEE Trans. Inf. Forensics Secur. 4 (4) (2009) 867–882.
[195] C. Rathgeb, C. Busch, Multi-biometric template protection: issues and challenges, New Trends and Developments in Biometrics, 2012.
[196] Y. Chin, T. Ong, A. Teoh, K. Goh, Integrated biometrics template protection technique based on fingerprint and palmprint feature-level fusion, Inf. Fusion 18 (2014) 161–174.
[197] C. Li, J. Hu, J. Pieprzyk, W. Susilo, A new biocryptosystem-oriented security analysis framework and implementation of multibiometric cryptosystems based on decision level fusion, IEEE Trans. Inf. Forensics Secur. 10 (6) (2015) 1193–1206.
[198] A. Kumar, A. Kumar, A cell-array-based multibiometric cryptosystem, IEEE Access 4 (2016) 15–25.
[199] V.M. Patel, R. Gopalan, R. Li, R. Chellappa, Visual domain adaptation: a survey of recent advances, IEEE Signal Process. Mag. 32 (3) (2015) 53–69.
[200] S.J. Pan, Q. Yang, et al., A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (10) (2010) 1345–1359.
[201] H.S. Bhatt, R. Singh, M. Vatsa, N.K. Ratha, Improving cross-resolution face matching using ensemble-based co-transfer learning, IEEE Trans. Image Process. 23 (12) (2014) 5654–5669.
[202] R. Singh, M. Vatsa, A. Ross, A. Noore, Biometric classifier update using online learning: a case study in near infrared face verification, Image Vis. Comput. 28 (7) (2010) 1098–1105.
[203] U. Uludag, A. Ross, A. Jain, Biometric template selection and update: a case study in fingerprints, Pattern Recognit. 37 (7) (2004) 1533–1542.
[204] A. Rattani, B. Freni, G.L. Marcialis, F. Roli, Template update methods in adaptive biometric systems: a critical review, in: International Conference on Biometrics, 2009, pp. 847–856.
[205] S. Chhabra, R. Singh, M. Vatsa, G. Gupta, Anonymizing k facial attributes via adversarial perturbations, in: International Joint Conference on Artificial Intelligence, 2018, pp. 656–662.
[206] V. Mirjalili, S. Raschka, A. Ross, Gender privacy: an ensemble of semi adversarial networks for confounding arbitrary gender classifiers, in: IEEE International Conference on Biometrics: Theory, Applications and Systems, 2018.
[207] A. Martinez-Monseny, D. Cuadras, M. Bolasell, J. Muchart, C. Arjona, M. Borregan, A. Algrabli, R. Montero, R. Artuch, R. Velázquez-Fragua, A. Macaya, C. Pérez-Cerdá, B. Pérez-Dueñas, B. Pérez, M. Serrano, From gestalt to gene: early predictive dysmorphic features of pmm2-cdg, J. Med. Genet. (2018).
[208] M. Gomez-Barrero, E. Maiorana, J. Galbally, P. Campisi, J. Fierrez, Multi-biometric template protection based on homomorphic encryption, Pattern Recognit. 67 (2017) 149–163.
[209] A. Bhargav-Spantzel, A.C. Squicciarini, S. Modi, M. Young, E. Bertino, S.J. Elliott, Privacy preserving multi-factor authentication with biometrics, Technical Report, Center for Education and Research in Information Assurance and Security, Purdue University, 2007.
[210] A. Ross, A. Othman, Mixing fingerprints for template security and privacy, in: European Signal Processing Conference, 2011, pp. 554–558.
[211] A. Othman, A. Ross, Fingerprint + Iris = IrisPrint, in: Proc. of SPIE Biometric and Surveillance Technology for Human and Activity Identification, 2015, pp. 1–13.
[212] N.A. Schmid, J.A. O'Sullivan, Performance prediction methodology for biometric systems using a large deviations approach, IEEE Trans. Signal Process. 52 (10) (2004) 3036–3045.
[213] R. Wang, B. Bhanu, Performance prediction for multimodal biometrics, in: International Conference on Pattern Recognition, 3, 2006, pp. 586–589.
[214] K. Nandakumar, A. Ross, A.K. Jain, Biometric fusion: Does modeling correlation really matter? in: Proceedings of the IEEE International Conference on Biometrics: Theory, Applications and Systems, 2009, pp. 271–276.
[215] R. Gross, I. Matthews, J. Cohn, T. Kanade, S. Baker, Multi-pie, Image Vis. Comput. 28 (5) (2010) 807–813.
[216] M. Grgic, K. Delac, S. Grgic, SCface–surveillance cameras face database, Multimedia Tools Appl. 51 (3) (2011) 863–879.