Chapter 2

Biometric-Based Human Recognition Systems: An Overview

David Palma and Pier Luca Montessoro

Abstract

With the proliferation of automated systems for reliable and highly secure human authentication and identification, the importance of technological solutions in biometrics is growing along with security awareness. Indeed, conventional authentication methodologies, consisting of knowledge-based systems that make use of something you know (e.g., username and password) and token-based systems that make use of something you have (e.g., identification card), are not able to meet the strict requirements of reliable security applications. Conversely, biometric systems make use of behavioral (extrinsic) and/or physiological (intrinsic) human characteristics, overcoming the security issues affecting the conventional methods for personal authentication. This book chapter provides an overview of the most commonly used biometric traits along with their properties, the various biometric system operating modalities as well as various security aspects related to these systems. In particular, it will be discussed the different stages involved in a biometric recognition process and further discuss various threats that can be exploited to compromise the security of a biometric system. Finally, in order to evaluate the systems' performance, metrics must be adopted. The most widely used metrics are, therefore, discussed in relation to the provided system accuracy and security, and applicability in real-world deployments.

Keywords: biometrics, authentication, identification, human traits, evaluation criteria, pattern recognition system, security, vulnerabilities

1. Introduction

This chapter stands as an introduction to the field of biometrics which is rising as an advanced layer to many user- and enterprise-centric security systems. In fact, conventional authentication methods, such as traditional passwords, have long been a weak point for security systems. Biometrics aims to answer this issue by linking proof-of-identity to our physiological traits and behavioral patterns. It is therefore important to present the concepts and primitives of performance metrics due to their impact on secure biometric systems. Thus, a brief overview is given to describe the main biometric traits along with their properties as well as the various biometric system operating modalities and the relatively known vulnerabilities. Finally, the criteria for performance evaluation have been defined to determine the system accuracy and security which are related to the applicability in real-world deployments.

2. Biometric traits

Various biometric modalities have been developed over the years making the biometric technology landscape very vibrant. Prominent examples of physiological/ biological and behavioral biometric characteristics, which have been the purpose of major real-world applications, are illustrated in **Figure 1**.

2.1 Physiological/biological (intrinsic) human characteristics

Biological biometrics make use of traits at a genetic and molecular level which may include features like DNA or blood, whilst physiological biometrics involve the individual physical traits like a fingerprint, iris, or the shape of the face. On the other hand, behavioral biometrics are based on patterns unique to each person, for example, how an individual walks, speaks, or even types on a keyboard. Some examples of biometric traits are briefly described below.

Fingerprint: Fingerprint recognition, which measures a finger's unique pattern, is one of the oldest forms of biometric identification. This trait appears as a series of dark lines and white spaces when captured from the device and it consists of a set of ridges and valleys located on the surface tips of a human finger to uniquely distinguish individuals from each other. The fingerprint features are generally categorized into—(i) macroscopic ridge flow patterns (core and delta points), (ii) minutia features (which consists of the ridge bifurcations/trifurcation and the ridge endings), and (iii) pores and ridge contour attributes (incipient ridges, pore, shape, and width). Fingerprints of identical twins are different and so are the prints on each finger of the same person [1].

Face: Facial features use the location and shape (geometry) of the face, including the distance between the eyes, the distance from the chin to the forehead, or other measures that involve eyebrows, nose, lips, and jawline [2]. This kind of recognition is

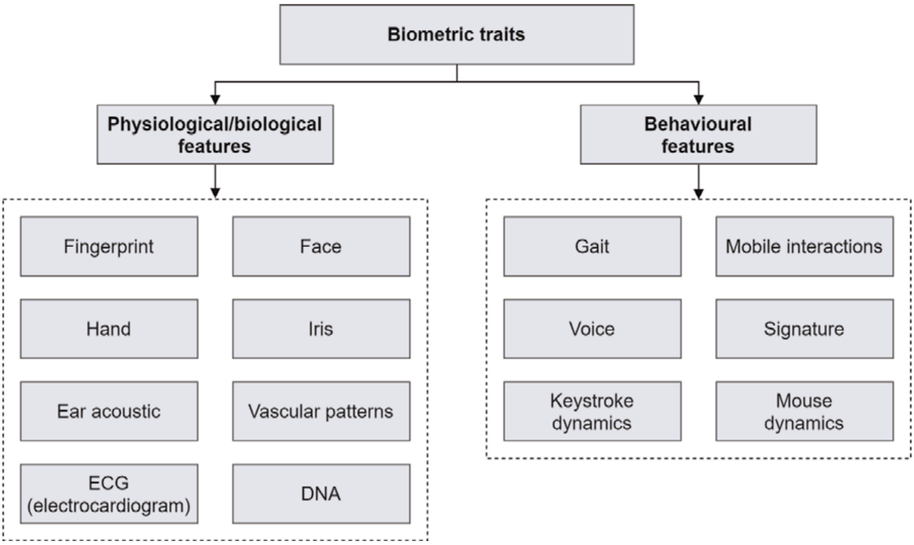

Figure 1.
Examples of physiological/biological and behavioral traits applied in biometric recognition applications.

a nonintrusive method with reasonable authentication performance in commercially available systems. However, several constraints may be imposed by the systems on how the facial images are obtained to work properly, for example, controlled illumination and background. Moreover, its susceptibility to change due to factors such as aging or expression may present a challenge [3].

Hand geometry: This trait is based on the geometric characteristics of the hand such as the length and width of fingers, their curvature, and their relative position to other features of the hand. Though once a dominant method of biometric measurement due to the requirement of the low complexity in feature extraction and low-cost imaging, modern advances in biometrics have replaced its relevance in most applications [4]. Furthermore, such a biometric trait is not known to be very distinctive and hand geometry-based recognition systems cannot be scaled up for systems requiring the identification of an individual from a large population. In addition, hand-geometry features from both hands are expected to be similar, as their anatomy is quite similar [5].

Iris: Systems based on this trait are among the most accurate biometric systems available. This human characteristic refers to the colored part in the eye that consists of thick, thread-like muscles characterized by unique folds and patterns that can be used to identify and verify the identity of humans. Furthermore, this biometric trait is stable because iris patterns do not vary during the course of a person's life and are not susceptible to loss, manipulation, or theft, making an iris recognition system robust to spoofing attacks. One interesting point worth noting is that even the two eyes in the same person have different patterns [6].

Ear acoustic: The main purpose of this kind of recognition system is to map one aspect within acoustic ear recognition, namely the performance of the ear characteristics bands and peaks. An ear signature is generated by probing the ear with inaudible sound waves which are reflected bouncing in different directions and picked up by a small microphone. The shape of the ear canal determines the acoustic transfer function which forms the basis of the signature. The recognition process is also possible, whilst the subject is on the move and caters to the protection of secrecy, which expands the applicability of this technology [7].

Vascular patterns: This biometric trait has been largely investigated for its advantages over other features. In fact, the vascular pattern of the human body is unique to every individual, even between identical twins [8], remains steady during the course of a person's life, and lies underneath the human skin ensuring confidentiality and robustness to counterfeiting, as opposed to other intrinsic and extrinsic biometric traits that are more vulnerable to spoofing, thus leading to important security and privacy concerns [9]. To acquire the network structure of blood vessels underneath the human skin, a vascular-based recognition system uses near-infrared light to reflect or transmit images of blood vessels, since they are almost invisible in normal lighting conditions [10]. The most commonly used vascular biometric solutions use hand-oriented modalities, such as finger vein, palm vein, hand dorsal vein, and wrist vein recognition, as well as eye-oriented modalities, such as retina and sclera recognition [11].

Electrocardiogram (ECG): This trait considers the human heart and body anatomic features form the shape of the ECG signal typically acquired using a few electrodes, amplifiers, filters, and a data acquisition module, and which reports the strength and timing of the electrical activity of the heart [12]. However, scientific findings to date throw doubt on the specificities of real-world application scenarios and acceptability by the potential end users, which pose several constraints and questions.

Deoxyribonucleic acid (DNA): DNA matching is based on a common molecular biology method named short tandem repeat (STR)[1] analysis, which is used to compare allele repeats at specific locations on a chromosome in DNA between two or more samples [14, 15]. DNA-based biometric recognition has been widely used in forensic science and scientific investigation due to its very high accuracy, despite the fact that identifications require tangible physical samples and cannot be done in real time.

2.2 Behavioral (extrinsic) human characteristics

Keystrokes, handwriting, gait, how a person uses a mouse, and other movements are some of the behavioral traits that a biometric system may analyze to assess the individual's identity.

Gait: This characteristic may be changeable over a large time span due to various reasons, such as weight gain [16]. Thus, it can be used in low-security applications for massive crowd surveillance as it can quickly identify people from afar based on their walking style, even harnessing the potential of a large number of surveillance cameras installed in public locations into a biometric system. In fact, such a system does not require the individuals to be cooperative, nor that they wear any special device or equipment to be recognized [17].

Mobile interactions: It is based on the unique ways in which users swipe, tap, pinch-zoom, type, or apply pressure on the touchscreen of mobile devices like tablets and phones, thus providing characteristic patterns that may be used to identify people, even considering further features deriving from on-board sensors such as GPS, gyroscope, and accelerometers [18], which can also be configured to collect data in passive mode. Therefore, mobile interactions-based biometrics focuses not so much on the outcome of the user's actions but rather on the way a user performs those actions.

Signature: Signature recognition is the most widely accepted method for documents authentication and it makes use of shorter handwriting probes compared to text-independent writer recognition methods, but it requires to write the same sign every time. A signature authentication scheme can be categorized into two methods—(i) off-line or static (the signature is digitized after the writing process) and (ii) online or dynamic (the signature is digitized during the writing process). Signature biometric features are extracted by analyzing curves, edges, spatial coordinates, inclination, the center of gravity, pen pressure, and pen stroke of the signature samples in both off-line and online applications. However, dynamic information like writing speed and stroke order is available only in online signatures [19].

Mouse dynamics: It makes use of patterns in mouse or trackpad cursor movement including clicks, trajectories, direction changes, tracking speed, and the relationships between them. Mouse-generated movement features are relatively stable for the same individual and different compared to other users, as such can be used to authenticate individuals [20]. These methods are most often used to continuously verify the user's identity.

Keystrokes: Keystroke dynamics (also known as typing biometrics) include the tracking of the rhythm used to type on a keyboard. Two events constitute a keystroke event—key down and key up. The first one occurs when an individual presses a key, whilst the second one is associated with the event that occurs when the pressed key is

[1] STR is the DNA sequence of the short repeat region of the sequence in the noncoding region of the human genome [13].

released. Making use of these events, a set of inter-key and intra-key features known as delay times, hold times, and key down-key downtimes can be extracted. In general, keystroke recognition will work on the computer or virtual keyboards, mobile phones, smartwatches, and touchscreen panels, providing a low-cost authentication method that can be easily deployed in a variety of scenarios [21].

Voice: Voice recognition technology falls under both the physiological and behavioral biometric categories. Voice biometric recognition allows to distinguish among humans' voice for personal authentication as voice features include physical characteristics such as vocal tracts, nasal cavities, mouth, and larynx [22]. Behaviorally, the way a person speaks or says something, for example, tone, movement variations, accent, pace, and so on, is also considered unique to each individual. Using data from both physiological and behavioral biometrics creates, therefore, a precise vocal signature, though mismatches may occur due to illness or other factors.

2.3 Properties of biometric traits

The main requirements that should be satisfied before a trait can be characterized as suitable for its applicability in a biometric recognition system, are briefly discussed as follows [23].

- Universality: Every individual or at least most of them, accessing the biometric application should possess the characteristic.

- Distinctiveness (or uniqueness): The given trait should be sufficiently different across individuals comprising the user population. Otherwise, the proportion of times the biometric system grants access to unauthorized individuals would be unacceptably high.

- Permanence: The biometric trait of an individual should be sufficiently invariant (with respect to the matching criterion) over a period of time. This implies that the given trait should not change significantly over time otherwise the proportion of times the biometric system denies access to authorized individuals would be unacceptably high.

- Collectability: The biometric trait can be measured quantitatively with particular regard to the easiness of obtaining the biometric data using suitable devices that do not cause undue inconvenience to the user.

Even though any human characteristic can be used as a biometric trait as long as the previous requirements are satisfied, in real-world biometric recognition applications there are a number of other issues that should be considered, such as:

- Performance: This is a property aimed at assessing the verification or identification accuracy, the computational time required for a single recognition, as well as the operational and environmental factors that may affect or not the recognition accuracy and speed.

- Acceptability: It indicates the extent to which people are willing to accept the use of a specific biometric application as well as their willingness to provide their biometric data. Nowadays, this is a crucial aspect to be considered due to the current pandemic

Biometric trait	**Universality**	**Uniqueness**	**Permanence**	**Collectability**	**Performance**	**Acceptability**	**Circumvention**
Fingerprint	M	H	H	M	H	M	H
Face	H	L	M	H	L	H	H
Hand geometry	M	M	M	H	M	M	M
Iris	H	H	H	M	H	L	L
Ear	M	M	H	M	M	H	M
Vascular patterns	H	H	M	M	H	M	L
DNA	H	H	H	L	H	L	L
Gait	M	L	L	H	L	H	M
Signature	L	L	L	H	L	H	H
Keystroke dynamics	L	L	L	M	L	M	M
Voice	M	L	L	M	L	H	H

H = High; M = Medium; L = Low.

Table 1.
Comparison study of the most common traits based on the characteristics of biometric entities.

situation caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [24], raising questions about how safe using touch-based biometric systems really is as touching the sensors can potentially spread viruses. As a consequence, less-constrained biometrics will likely be the preferred modality, whilst there may be less demand for other solutions that rely on physical contact with a reader.

- Circumvention: This property reflects how easily the system can be deceived through potential spoofing attacks. It refers to the ways in which an attacker can endeavor to bypass a biometric system and finally attack the weak spot of such a system in order to gain unauthorized access.

Real-life biometric recognition systems ought to meet the requirements of accuracy, speed, and resource constraints, be harmless to the users, be accepted by the intended population as well as sufficiently robust to various fraudulent methods and attacks to the system [25].

Table 1 is reported a comparison study of the most popular traits based on the characteristics of biometric entities [26].

3. Biometric system operating modes

A biometric system can provide two kinds of operating modes (identity management functionalities), namely, *verification* and *identification*. Biometric systems can indeed automatically authenticate[2] or identify subjects in a reliable and fast way and

[2] Throughout this book chapter, the term authentication will be used as a synonym for verification.

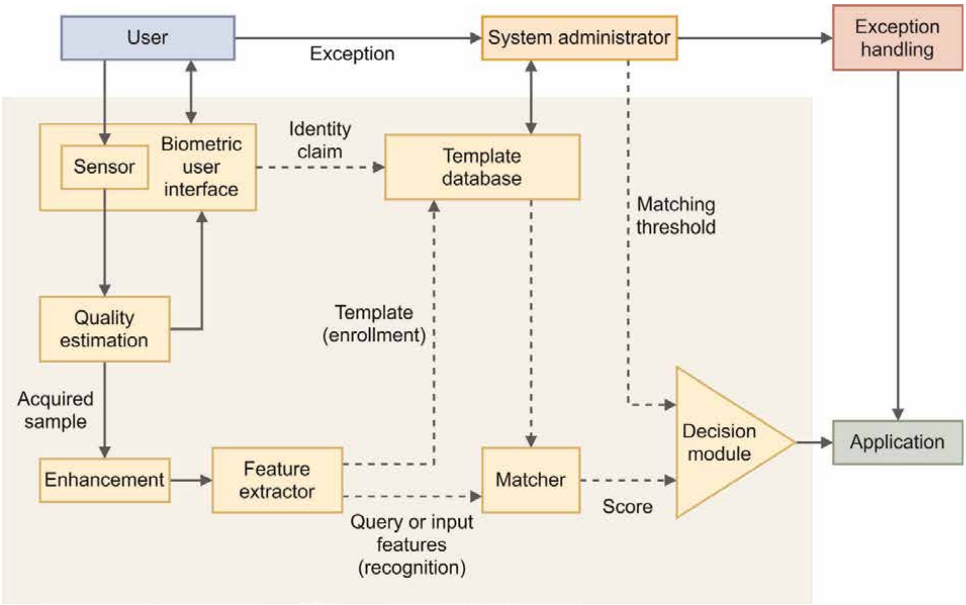

Figure 2.
Basic building blocks of a generic biometric recognition system.

are, therefore, suitable to be used in a wide range of applications to face the risks of unauthorized logical or physical access and identity theft, as well as new threats such as terrorism or cybercrime [27]. **Figure 2** provides a high-level view of a generic biometric recognition system as well as all its basic building blocks, whilst **Figure 3** depicts the enrollment and the biometric recognition schemes of the authentication and identification modalities.

3.1 Authentication

In the authentication mode, the purpose of the biometric system is to verify whether an individual's claimed identity is genuine or not (binary classification). Thus, the captured biometric data (query) is compared only with the biometric template(s) stored in the system database and corresponding to the claimed identity (one-to-one or one-to-few comparison). Given a claimed identity I and a query feature set $\boldsymbol{x}^Q$, the biometric system has to be categorized $(I, \boldsymbol{x}^Q)$ into "genuine" or "impostor" class. Let $\boldsymbol{x}_I^E$ be the stored biometric template corresponding to the identity I (i.e., the enrolled user with identity I). The similarity measure between $\boldsymbol{x}^Q$ and $\boldsymbol{x}_I^E$ gives, as a result, a matching score. Hence, the biometric system applies the decision rule given by

$$(I, \boldsymbol{x}^Q) \in \begin{cases} \text{genuine}, & \text{if } s(\boldsymbol{x}^Q, \boldsymbol{x}_I^E) \geq \xi, \\ \text{impostor}, & \text{otherwise}, \end{cases} \tag{1}$$

where s represents a similarity function and ξ represents a pre-defined threshold at which the system is intended to operate. The authentication mode is typically employed for positive recognition, where the aim is to prevent multiple people from using the same identity [28].

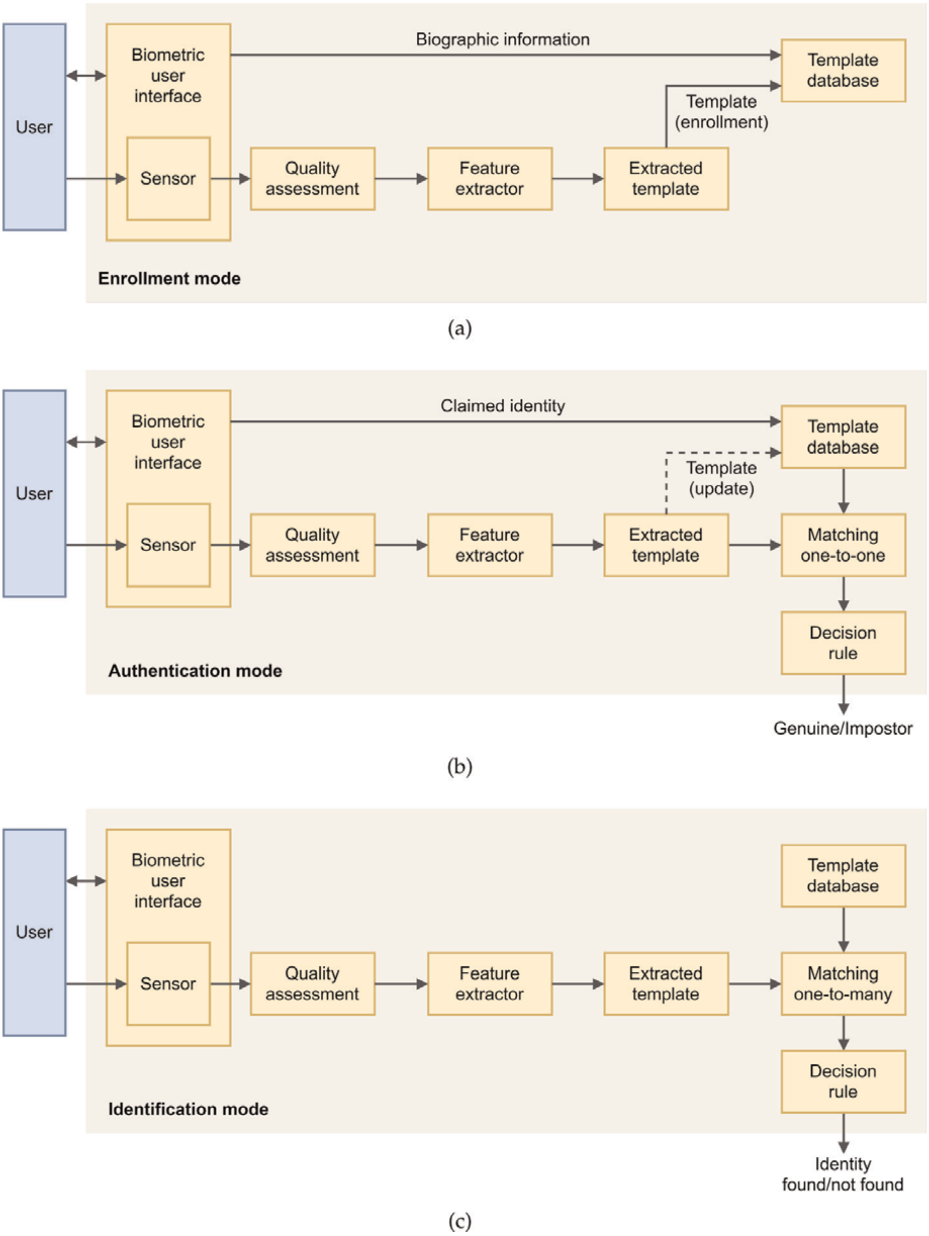

Figure 3.
Different operating modes of a biometric system—(a) enrollment mode, (b) authentication mode (the dashed line is an optional operation aimed at updating a specific user's template), and (c) identification mode.

3.2 Identification

In the identification mode, the purpose of the biometric system is to recognize an individual's identity by searching the templates of all the enrolled individuals in the system database for a match (one-to-many comparison) without the subject having to claim an identity.

This operating mode can be further split into negative and positive identification—in the negative identification (also known as *screening*), the user is considered to be hiding her/his true identity from the biometric system, whilst in the positive identification, the user tries to positively identify herself/himself to the system without explicitly claiming an identity. Given a query feature set $\boldsymbol{x}^Q$, the biometric system has to determine the identity $I_k \ \forall k \in \{1, 2, \dots, n, n+1\}$ where $\{I_1, I_2, \dots, I_n\}$ are identities of the enrolled users in the system, whilst I_{n+1} represents the failure case where no identity can be assigned for the given query (*open-set identification*). Hence, assuming that $\boldsymbol{x}^E_{I_k}$ is the stored template corresponding to the identity I_k, the biometric system applies the decision rule given by

$$\boldsymbol{x}^Q \in \begin{cases} I_k, & \text{if } \max\limits_k \left\{ s\left(\boldsymbol{x}^Q \boldsymbol{x}^E_{I_k}\right) \right\} \geq \xi, \\ I_{n+1}, & \text{otherwise}, \end{cases} \tag{2}$$

where s represents a similarity function and ξ represents a pre-defined threshold at which the system is intended to operate.

The identification mode is typically employed for screening[3], where the aim is to prevent a single person from using multiple identities [28].

4. Vulnerabilities

Biometric-based cybersecurity solutions ensuring tight access control are essential in preventing intrusions and unauthorized accesses. However, even though a biometric system enhances user convenience and security, does not necessarily mean that it is also exempt from security and privacy issues. Many security measures in biometric systems are designed to protect one or more facets of the CIA triad, which is a common framework that refers to confidentiality, integrity, and availability [31].

- Confidentiality is roughly equivalent to privacy. Measures undertaken to ensure confidentiality are designed to prevent sensitive information from reaching unauthorized people. It is perhaps the most obvious aspect of the CIA triad when it comes to security; but correspondingly, it is also the one which is attacked most often. Confidentiality covers a wide spectrum of access controls and measures that protect data from getting misused by any unauthorized access. Cryptography and encryption methods are an example of an attempt to prevent illegitimate access ensuring the confidentiality of (sensitive) data.

- Integrity of information refers to the ability to protect information from being modified or destroyed by unauthorized parties, thus ensuring nonrepudiation and authenticity of the information. Thus, integrity involves maintaining the consistency and trustworthiness of data. One type of security attack is to intercept some important data and make changes to it before sending it on to the intended receiver.

[3] In some real scenario, such as latent palmprint matching [29], it is preferable to use a semi-automated approach aimed at providing the top n identities that best match to the given template for further analysis by a human expert. Alternatively, it is possible to consider all the identities whose corresponding match scores exceed the threshold ξ that leads to a challenging task in a quite large database (e.g., FBI's next generation identification (NGI) system, which provides the world's largest repository of biometric and criminal history information [30]).

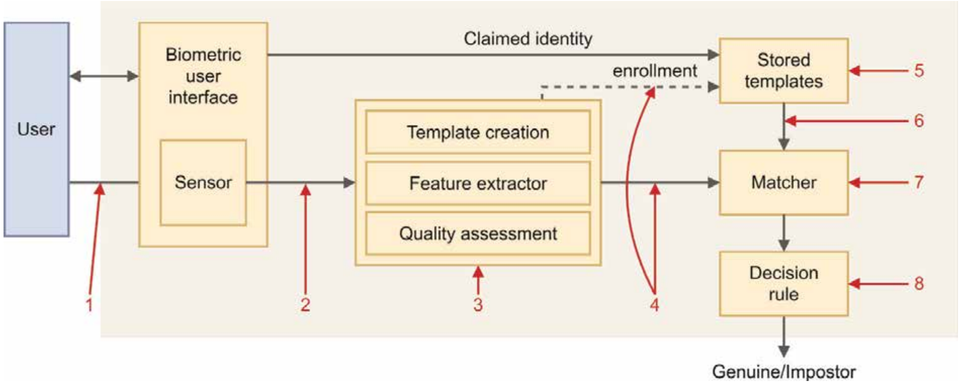

Figure 4.
Attack points of a general biometric system.

- Availability of information refers to ensuring that only legitimate and authorized parties are able to access the information when needed. Problems affecting the information system could make it impossible to access information, thereby making the information unavailable. Some types of security attacks attempt to deny access to the appropriate user, either for the sake of inconveniencing them, or because there is some secondary effect.

Biometric recognition systems implicitly (and effectively) address the authentication problem included in the last issue of the CIA triad, which consists in guaranteeing access to data only to authorized users. The reason for this is because biometric traits are (generally) not susceptible to loss, manipulation, or theft, and therefore overcome the security issues affecting the conventional methods for personal authentication, such as knowledge-based and token-based systems. However, it must be kept in mind that a biometric-based security solution is composed of several different components and the recognition module, which is only capable of addressing the authentication aspect, is just one of them. Thus, a logical structure-based approach of biometric systems is used to describe the eight points of attacks illustrated in **Figure 4**.

1. An attack on the biometric sensor consists of presenting a fake biometric trait (e.g., an artificial characteristic) to perform a spoofing attack aimed to either avoid detection (false negative) or masquerade as another (false positive). Methods used to prevent spoofing attacks include layered biometrics, liveness, and combining biometrics and conventional authentication methods such as passwords, tokens, or smart cards [32].

2. The connection between the biometric sensor and the subsequent modules of the system may be attacked to allow input of a stored digital biometric signal. This data can be obtained, for instance, by performing an eavesdropping (disclosure) attack [31].

3. Attacks on the feature extractor can be used either to create impostors or to evade detection. Hence, knowledge of the algorithms involved in this module[4]

[4] Since biometric recognition algorithms are likely susceptible to reverse engineering techniques, it is possible to conduct off-line experiments on a copy of the biometric software to be hacked in order to achieve the objective [32].

may be used to forge features in presented samples to cause computation of incorrect features. To achieve this, an attacker can replace the feature extractor with a Trojan horse program that produces the desired feature sets.

4. An attack on the output of the previous module consists of spoofing the legitimate biometric feature set to replace it with a synthetic one.

5. Vulnerabilities of template database concern modifying the storage (modifying, removing, or adding templates), copying stored data for future use (identity theft or directly using the acquired information to gain access), or modifying the identity to which the biometric is assigned.

6. The channel between the template database and the matching module is similarly vulnerable to the previous one, however, the attack against data transmission may be easier than against the template storage, especially in the case of an adversary able to intercept any information communicated by the system by observing the data (passive eavesdropping). Encryption is crucial in this case, but may still be vulnerable to key discovery [33].

7. The matcher module is responsible for computing a similarity score between two biometric templates in order to confer the likelihood that they are from the same subject. Even though it may not be possible to do it easily, an attack against the matcher can be possible in specific cases. For instance, it is possible to replace the matcher module with a Trojan horse program that always outputs high scores thereby defying system security [34].

8. An attack on the final decision module means that if the final decision can be inserted or blocked by the attacker then the authentication system function will be overridden. If it is instead reviewed by a human operator, a DoS (denial of service) attack may be performed to mislead it or to force it to mistrust the output of the system [35].

5. Criteria for performance evaluation

The reliability and validity of a biometric scheme as well as the selection of a certain biometric trait for an application are determined by specific measures that are used to evaluate the recognition accuracy and effectiveness as addressed in ISO/IEC Standards [36]. Accordingly, to evaluate the accuracy of the proposed method based on a single-sample approach for unimodal biometric systems, each sample in the database should undergo a one-to-one matching test against every single stored sample. Hence, a comparison between a subject with a real identity I_r and a subject with claimed identity I_c is aimed at testing the hypothesis:

$$H_0 : \{I_r = I_c\} \text{ versus } H_1 : \{I_r \neq I_c\} \tag{3}$$

where H_0 is the null hypothesis that the user is who s/he claims to be (genuine or intra-class matching), whilst H_1 is the alternative hypothesis that the user is not who s/he claims to be (impostor or inter-class matching). To test the hypothesis in (3), it is required to compute a similarity measure, $s(Q, T)$ where large (respectively, small) values of s indicate that the template T of the claimed identity I_c in the database and the biometric query Q of a real user I_r are close to (far from) each other. Formally, the verification problem consists of determining if a claimed identity I with biometric data Q belongs to the class H_0 or not:

$$(I, Q) = \begin{cases} H_0, & \text{if } s(Q, T) \geq \xi, \\ H_1, & \text{otherwise.} \end{cases} \tag{4}$$

Precisely, given a threshold ξ, all matching values s lower (respectively, greater) than ξ lead to the rejection (acceptance) of the null hypothesis [37]. Therefore, whether the hypothesis is accepted or not, the test is prone to two kinds of error:

- false acceptance rate (FAR), that is the probability of accepting the null hypothesis H_0 when input is not valid (type-I error),

- false rejection rate (FRR), that is the probability of rejecting the null hypothesis H_0 when input is valid (type-II error).

Let H_0 and H_1 be the labels that denote the genuine and impostor classes, respectively. Assume also that the $p(s|H_0)$ and $p(s|H_1)$ represent the probability density functions of the genuine and impostor scores, respectively. Then the FAR and FRR distributions are given by:

$$FAR(\xi) = p(s \geq \xi | H_1) = \int_{\xi}^{+\infty} p(s|H_1) ds, \tag{5}$$

$$FRR(\xi) = p(s < \xi | H_0) = \int_{-\infty}^{\xi} p(s|H_0) ds. \tag{6}$$

The false acceptance and false rejection rates are functions of the system threshold ξ and are closely related because the increase of one implies the decrease of the other. Hence, for a given biometric system, it is not possible to decrease both these errors at the same time by varying the threshold ξ [25]. The separation between the two distributions (or classes) indicates the ability of the system to distinguish the genuine user samples from those of the impostors. Indeed, the separation also provides a hint on the threshold point that maximizes the variance between the two classes in order to correctly mark a user sample image as authentic or impostor [23].

The genuine acceptance rate (GAR) is instead the probability of accepting the null hypothesis H_0 when input is valid, hence it can be used as an alternative to FRR:

$$GAR(\xi) = p(s \geq \xi | H_0) = 1 - FRR(\xi). \tag{7}$$

Depending on the security level required by the final application (i.e., forensics, surveillance and homeland security, civilian, or high-security applications), the same biometric system may operate at different threshold values (ξ), as illustrated in **Figure 5**.

Hence, in order to evaluate the biometric system performance as a function of the threshold ξ, the following curves can be considered:

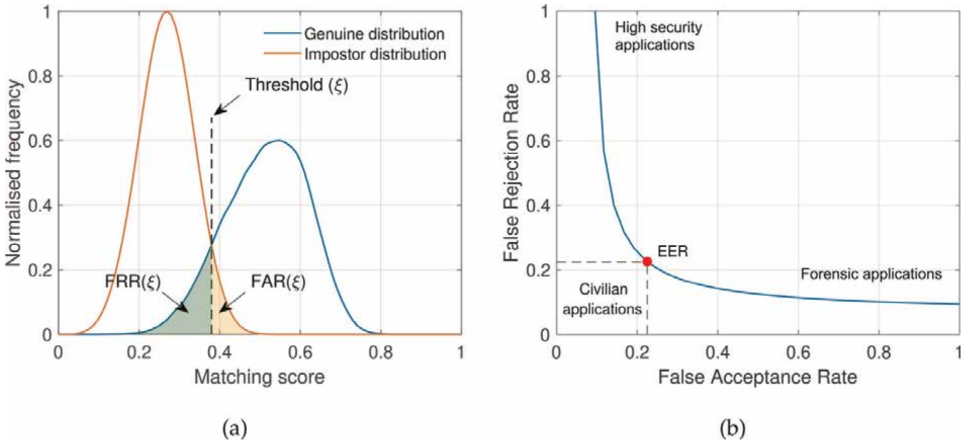

Figure 5.
Examples of biometric system error rates: (a) FAR and FRR for a given threshold ξ are displayed over the genuine and impostor score distributions and (b) typical operating points of different biometric applications are displayed on a DET curve aimed at relating FAR and FRR at different threshold values.

- The receiver operating characteristic (ROC) is a graphical plot that illustrates the trade-off between false acceptance and false rejection rates when the threshold varies, whilst the intersection point for which rejection and acceptance errors are equal is named equal error rate (EER). The curve is generated by plotting the genuine acceptance rate against the false acceptance rate at various threshold settings,

- The detection error trade-off (DET) is another graphical plot that illustrates the false rejection rate against the false acceptance rate at various threshold values. The two axes are scaled nonlinearly by their standard normal deviates[5] or just by logarithmic transformation.

Furthermore, the above-mentioned ROC and DET curves are threshold-independent, allowing performance comparison of different biometric systems under similar conditions [23], as illustrated in **Figure 6**. Given a set of thresholds $\{\xi_i\} \mid s_{\min} \leq \xi_i \leq s_{\max} \ \forall i \in \{1, 2, \dots, n\}$ where $s_{\min}$ and $s_{\max}$ are the minimum and maximum scores, respectively, in a given set of match scores $\{s_i\} \mid 0 \leq s_i \leq 1 \ \forall i \in \{1, 2, \dots, n\}$. Then, it is possible to generate a ROC curve computing the overall false acceptance and false rejection rates for each threshold value ξ as follows:

$$FAR = \frac{1}{N}\sum_{k=1}^{N} FAR(\xi), \tag{8}$$

$$FRR = \frac{1}{N}\sum_{k=1}^{N} FRR(\xi), \tag{9}$$

[5] In the normal deviate scale, the threshold values ξ correspond to linear multiples of standard deviation σ of a Gaussian distribution. Thus, if the FAR and FRR distributions are Gaussian, the corresponding DET curve would be linear [25].

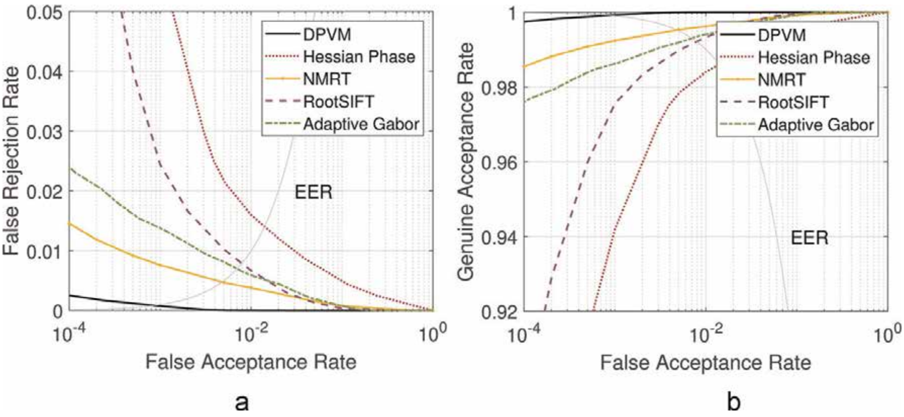

Figure 6.
Example of vascular-based biometric systems performance comparison [4]. Comparative graph of—(a) DET curves generated by plotting FRR against FAR and (b) ROC curves generated by plotting GAR against FAR.

where N represents all identities being evaluated by the system and

$$FAR(\xi) = \frac{\text{no.of FARs}}{\text{no.of impostor accesses}} \tag{10}$$

$$FRR(\xi) = \frac{\text{no.of FRRs}}{\text{no.of genuine accesses}}. \tag{11}$$

Since biometric systems cannot jointly provide a false acceptance rate equal to zero and a perfect verification/identification rate, the system threshold must be adjusted for the given application considering the trade-off between accuracy and false positives. Once the threshold has been set, the system can be evaluated by means of common measures that are used to assess the classification accuracy and effectiveness. In this context, we are interested in confirming or denying the identity of a subject leading thus to a dichotomous binary classification problem, where the labels are P (genuine) and N (impostor) and the predictions of the classifier are summarized in a 2×2 contingency table known as confusion matrix [38] (expanded in **Table 2**):

$$\mathbf{M} = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix} \tag{12}$$

		Predicted class		
		P	***N***	**Total**
Actual class	***P***	*TP*	*FN* (Type-II error)	*TP* + *FN*
	N	*FP* (Type-I error)	*TN*	*FP* + *TN*
	Total	*TP* + *FP*	*FN* + *TN*	

Table 2.
Example of confusion matrix for a dichotomous binary classification problem.

which completely describes the outcome of the classification task. This contingency table may be expressed using raw counts of the number of records from class times each predicted label is associated with each actual class. As illustrated in **Table 2**, the confusion matrix reports:

- true positive (TP), the probability of correctly accepting the null hypothesis;
- true negative (TN), the probability of correctly rejecting the null hypothesis;
- false positive (FP), the probability of falsely rejecting the null hypothesis;
- false negative (FN), the probability of falsely accepting the null hypothesis.

Based on the entries in the confusion matrix, the total number of correct predictions carried out by the model is $TP + TN$, whilst the number of incorrect predictions is $FP + FN$ [39]. Therefore, if.

$$\boldsymbol{M} = \begin{bmatrix} n^{+} & 0 \\ 0 & n^{-} \end{bmatrix} \tag{13}$$

where obviously $n^{+} = TP + FN$ and $n^{-} = FP + TN$, then the classification has been perfectly done. Conversely, if the confusion matrix is as follows

$$\boldsymbol{M} = \begin{bmatrix} 0 & n^{+} \\ n^{-} & 0 \end{bmatrix} \tag{14}$$

it represents the worst case (perfect misclassification).

Several measures have been defined to assess the quality of a prediction [40], aimed at conveying into a single figure the structure of $\boldsymbol{M}$. The most used functions are briefly described as follows.

Precision also known as positive predictive value (PPV) counts the true positives, how many samples are properly classified within the same cluster (closeness of the measurements to each other)

$$PPV = \frac{TP}{TP + FP}. \tag{15}$$

Sensitivity also known as recall or true positive rate (TPR) refers to the proportion of the samples properly classified as true positives out of the actual number of true positives

$$TPR = \frac{TP}{TP + FN}. \tag{16}$$

F-measure combines precision and recall in a single metric, indeed, it is the harmonic mean of precision and sensitivity and as a function of $\boldsymbol{M}$, has the following form:

$$F_1 = 2\frac{PPV \cdot TPR}{PPV + TPR} = \frac{TP}{TP + \frac{1}{2}(FN + FP)} \tag{17}$$

where the worst case ($F_1 = 0$) is achieved for $TP = 0$, whilst the best case ($F_1 = 1$) is reached for $FN = FP = 0$.

Accuracy represents the ratio between the correctly predicted instances and all the instances in the dataset, whose range is between 0 (worst case) and 1 (best case):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \tag{18}$$

Matthews correlation coefficient is the measure of the quality of binary (two-class) classifications:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{19}$$

it is a correlation coefficient between the actual and predicted binary classifications and it returns a value between −1 (worst case) and 1 (best case).

Accuracy and F-score computed on confusion matrices have been (and still are) among the most popular adopted metrics in binary classification tasks. However, these statistical measures can dangerously show overoptimistic inflated results, especially on imbalanced datasets [40]. Hence, among all the parameters described above, the Matthews correlation coefficient (MCC) is the only one that takes into account true and false positives and negatives and is generally regarded as a balanced measure that can be used even if the classes are of very different sizes [41].

6. Conclusions

Biometric-based technologies make use of unique behavioral (extrinsic) and/or physiological/biological (intrinsic) attributes to overcome the security issues affecting the conventional methods for identity authentication. Even though biometrics has been in use for decades, the advent of technology has expanded its application from primarily criminal identification to a wide range of everyday tasks, becoming a regular security process of our nowadays life. Accurate authentication or identification is fundamental to physical security, cyber security, military applications (e.g., biometric-driven lethal autonomous weapon systems), financial transactions, contracts and employment, public services, criminal justice, national security, and more. The approaches that have been proposed in literature depend on the type and the number of the underlying biometric traits, which, in general, cannot be easily transferred between people, and thereby represents a highly secure unique identifier. As a matter of fact, various biometric modalities have been developed over the years making the biometric technology landscape very vibrant. In this book chapter, we have provided an overview of the most commonly used biometric traits along with their properties, the various biometric system operating modalities as well as various limitations and weaknesses related to these systems. Indeed, biometric technologies have a number of vulnerabilities that underscore the concerns over their employment and may result in the failure of the technology to perform as anticipated. We have also discussed how the system threshold must be adjusted for the given application considering the trade-off between accuracy and false positives since biometric systems cannot jointly provide a FAR equal to zero and a perfect recognition rate. Finally, the criteria for performance evaluation have been defined to determine the system's accuracy and security which are related to the applicability in real-world deployments, even though the existing evaluation metrics are more related to the data quality than the security aspects of the overall system. However, despite the risks, biometrics provide very compelling security solutions remaining a growing way to verify identity offering tons of promise for the future of cybersecurity.

Conflict of interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Abbreviations

DET	Detection error trade-off
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
FAR	False acceptance related
FN	False negative
FRR	False rejection rate
FP	False positive
GAR	Genuine acceptance rate
MCC	Matthews correlation coefficient
NGI	Next-generation identification
PPV	Positive predictive value
ROC	Receiver operating characteristic
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
STR	Short tandem repeat
TN	True negative
TP	True positive
TPR	True positive rate

Author details

David Palma*† and Pier Luca Montessoro†
Polytechnic Department of Engineering and Architecture, University of Udine, Udine, Italy

*Address all correspondence to: david.palma@uniud.it

†D.P. and P.L.-M. designed the research; D.P. performed the research and wrote the paper. The results and the paper were analysed and reviewed by P.L.-M. All authors have read and agreed to the published version of the manuscript.

References

[1] Maltoni D, Maio D, Jain AK, Prabhakar S. Handbook of Fingerprint Recognition. London, UK: Springer Science & Business Media; 2009

[2] Zhao W, Rama Chellappa P, Phillips J, Rosenfeld A. Face recognition: A literature survey. ACM Computing Surveys. 2003;**35**(4):399-458

[3] Jain AK, Ross A, Prabhakar S. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology. 2004; **14**(1):4-20

[4] Palma D, Montessoro PL, Giordano G, Blanchini F. Biometric palmprint verification: A dynamical system approach. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2019;**49**(12):2676-2687

[5] Li SZ, Jain AK. Encyclopedia of Biometrics: I-Z. Boston, MA: Springer Science & Business Media; 2015

[6] Daugman J. How iris recognition works. In: The Essential Guide to Image Processing. Amsterdam, NL: Elsevier; 2009. pp. 715-739

[7] Akkermans AHM, Kevenaar TAM, Schobben DWE. Acoustic ear recognition for person identification. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID). New York, US: IEEE; 2005. pp. 219-223

[8] Kumar A, Hanmandlu M, Gupta HM. Online biometric authentication using hand vein patterns. In: IEEE Symposium on Computational Intelligence for Security and Defense Applications. New York, US: IEEE; 2009. pp. 1-7

[9] Palma D, Blanchini F, Giordano G, Montessoro PL. A dynamic biometric authentication algorithm for near-infrared palm vascular patterns. IEEE Access. 2020;**8**:118978-118988

[10] Zharov VP, Ferguson S, Eidt JF, Howard PC, Fink LM, Waner M. Infrared imaging of subcutaneous veins. Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery. 2004; **34**(1):56-61

[11] Uhl A. State of the art in vascular biometrics. In: Handbook of Vascular Biometrics. Cham: Springer; 2020. pp. 3-61

[12] Wübbeler G, Stavridis M, Kreiseler D, Bousseljot R-D, Elster C. Verification of humans using the electrocardiogram. Pattern Recognition Letters. 2007;**28**(10):1172-1175

[13] Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R. Evaluation of 13 short tandem repeat loci for use in personal identification applications. American Journal of Human Genetics. 1994;**55**(1):175

[14] Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature. 1985;**316**(6023): 76-79

[15] Tautz D. Hypervariability of simple sequences as a general source for polymorphic dna markers. Nucleic Acids Research. 1989;**17**(16):6463-6471

[16] Hu N, Tong H-L, Tan W-H, Yap TT-V, Chong P-F, Abdullah J. Human identification based on extracted gait features. International Journal on New Computer Architectures and Their Applications. 2011;**1**(2):358-370

[17] Mason JE, Traoré I, Woungang I. Machine Learning Techniques for Gait Biometric Recognition. New York, US: Springer; 2016

[18] Fierrez J, Pozo A, Martinez-Diaz M, Galbally J, Morales A. Benchmarking touchscreen biometrics for mobile authentication. IEEE Transactions on Information Forensics and Security. 2018;**13**(11):2720-2733

[19] Deore MR, Handore SM. A survey on offline signature recognition and verification schemes. In: International Conference on Industrial Instrumentation and Control (ICIC). New York, US: IEEE; 2015. pp. 165-169

[20] Sayed B, Traoré I, Woungang I, Obaidat MS. Biometric authentication using mouse gesture dynamics. IEEE Systems Journal. 2013;**7**(2):262-274

[21] Killourhy KS, Maxion RA. Comparing anomaly-detection algorithms for keystroke dynamics. In: IEEE/IFIP International Conference on Dependable Systems & Networks. New York, US: IEEE; 2009

[22] Delac K, Grgic M. A survey of biometric recognition methods. In: Proceedings Elmar-2004, 46th International Symposium on Electronics in Marine. New York, US: IEEE; 2004. pp. 184-193

[23] Palma D. A Dynamical System Approach for Pattern Recognition and Image Analysis in Biometrics and Phytopathology [PhD thesis]. Udine, IT: University of Udine; 2021

[24] Sarfraz M. Introductory chapter: On fingerprint recognition. In: Sarfraz M, editor. Biometric Systems. Rijeka: IntechOpen; 2021

[25] Jain AK, Ross A, Nandakumar K. Introduction to Biometrics. New York: Springer; 2011

[26] Dasgupta D, Roy A, Nag A, et al. Advances in User Authentication. New York, US: Springer; 2017

[27] Huang D, Tang Y, Wang Y, Chen L, Wang Y. Hand-dorsa vein recognition by matching local features of multisource keypoints. IEEE Transactions on Cybernetics. 2015;**45**(9):1823-1837

[28] Wayman JL. Fundamentals of biometric authentication technologies. International Journal of Image and Graphics. 2001;**1**(01):93-113

[29] Palma D, Montessoro PL, Giordano G, Blanchini F. A dynamic algorithm for palmprint recognition. In: 2015 IEEE Conference on Communications and Network Security (CNS). New York, US: IEEE; 2015. pp. 659-662

[30] Federal Bureau of Investigation (FBI). Next Generation Identification (NGI). Washington DC, US: 2021. Available from: https://www.fbi.gov/

[31] Palma D. Detection of Stealthy False-data Injection Attacks on Safety-Critical Cyber-Physical Systems. London, UK: Technical report, Imperial College of Science, Technology and Medicine; 2019

[32] Adler A, Schuckers SAC. Biometric Vulnerabilities: Overview. US, Boston, MA: Springer; 2009. pp. 1-11

[33] Sheldon FT, Weber JM, Yoo S-M, Pan WD. The insecurity of wireless networks. IEEE Security Privacy. 2012; **10**(4):54-61

[34] Prasad PS. Vulnerabilities of biometric system. International Journal

of Scientific & Engineering Research. 2013;**4**(6):1126-1129

[35] Ferguson N, Schneier B. Practical Cryptography. Vol. 141. New York: Wiley; 2003

[36] ISO/IEC JTC 1/SC 37 Biometrics. Information technology – biometric performance testing and reporting – part 1: Principles and framework. ISO/IEC. 2006;**1**:19795-19791

[37] Dass SC, Zhu Y, Jain AK. Validating a biometric authentication system: Sample size requirements. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;**28**(12): 1902-1319

[38] David MW Powers. Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. Journal of Machine Learning Technologies. 2011;**2**(1):37-63

[39] Gan G, Ma C, Jianhong W. Data Clustering: Theory, Algorithms, and Applications. Pennsylvania, US: SIAM; 2020

[40] Chicco D, Jurman G. The advantages of the matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020;**21**(1):6

[41] Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using matthews correlation coefficient metric. PLoS One. 2017;**12**(6):e0177678