APSIPA Transactions on Signal and Information Processing, 2023, 12, e38

Overview Paper

Exploring Human Biometrics: A Focus on Security Concerns and Deep Neural Networks

Waleed H. Abdulla[1]*, Felix Marattukalam[1] and Vedrana Krivokuća Hahn[2]

[1] *The University of Auckand, New Zealand*
[2] *Idiap Research Institute, Switzerland*

ABSTRACT

Biometric technology is rapidly growing due to the urgent need to secure people's properties, from goods to information, in the overwhelming digital technology proliferation in all aspects of society. In this paper, biometric recognition is defined as the automated recognition of individuals based on biological or behavioral characteristics, such as fingerprints, facial recognition, and speech patterns. The authors emphasize that a robust biometric system consists of a combination of physiological and behavioral features. However, using biometrics for identification raises privacy concerns and the paper addresses the need to balance privacy and security. A comprehensive section on biometric template protection is introduced to address biometrics privacy and different attack protections. It discusses deep neural network-based models to segment real-world features and match them for authentication. It presents a case study of a new model based on the Siamese neural network. It explains how the Siamese neural network can be used for biometric recognition and how it compares to other deep learning models commonly used in the field. Lastly, the paper discusses state-of-the-art methods to secure information and provides a futuristic view of the technology. This paper provides a comprehensive overview of biometric technology, its advantages, and the associated privacy concerns.

*Corresponding author: Waleed H. Abdulla, w.abdulla@auckland.ac.nz.

Received 02 April 2023; Revised 12 July 2023
ISSN 2048-7703; DOI 10.1561/116.00000021

Keywords: Biometrics, computer vision, deep learning, template protection and recognition.

1 Introduction

The 2001 MIT Technology Review indicated that biometrics is one of the ten emerging technologies that will profoundly impact the economy and how we live and work. Biometric technology is initially treated as an exotic topic. At the same time, it is a fast-growing industry due to the urgent need to secure people's properties, from goods to information, in the overwhelming digital technology proliferation in all aspects of society.

We need to know what biometrics means as a first step to get into this exciting topic. Biometric recognition, or biometrics, is the automated recognition of individuals based on biological or behavioral characteristics. With factors like the rapid digitalization of data, the need to secure digital information, repeated privacy attacks, and the impact of Covid-19 to work in a contact-free environment, this technology has become increasingly important. Biological features include the face, fingerprints, palm prints, iris patterns, palm veins, and many more. Some behavioral features include the way one walks, the way one speaks even the way one types! Often a robust biometric system consists of a combination of biological and behavioral features. Most biometrics combine physiological and behavioral features (traits) and should not be exclusively classified into physiological or behavioral characteristics. For example, speech is partially determined by the biological structure of the speaker's vocal tract and partially by how a person speaks. Fingerprints may be physiological in nature, but the input device's usage (e.g., how a user touches the fingerprint scanner and the pressure on the sensor) depends on the person's behavior. A car mechanic has a different touch from a computer geek! Thus, the input to the recognition engine is a combination of physiological and behavioral characteristics. Behaviors can help distinguish the confusion when identifying parents, children, and siblings in their voice, gait, and signature. The same argument applies to facial recognition. Faces of identical twins may completely match at birth, but facial features change during growth based on the person's behavior developed from the profession, way of living, environment, and more.

Biometric systems have many advantages over classical ones; one is that everyone owns biometrics, and it is always available to the person. However, despite all the advantages biometrics recognition facilitates, not everyone supports using biometrics. Biometrics proliferation for recognizing people raised concerns from civil rights advocates. The privacy issue is a big concern and where privacy and security should meet. More security might reach the point of privacy breaching, which is unacceptable to most people. Also, the compromise of the biometrics data is one of the main concerns; what if someone unlawfully attains personal data?

This paper will systematically introduce the field of biometrics and discuss the common methods used in different biometric recognition systems. It will also focus on the effectiveness of recently developed deep neural network-based models to accurately segment real-world features and match them for authentication. A new model based on the Siamese neural network will be explained. Further, the privacy risks associated with biometrics will be discussed along with the state-of-the-art methods to secure information giving a futuristic view of the technology.

2 Human Biometrics Definitions

The word "biometrics" is derived from the Greek words 'bios' and 'metric'; which means life and measurement, respectively. This directly translates into: "life measurement." Human Biometrics is the automated recognition of a person using adherent distinctive physiological and/or involuntary behavioural features. Physiological features include facial characteristics, fingerprints, palm prints, iris patterns, and many more. Examples of behavioural features are signature, gait, voice, and keyboard typing dynamics [5, 35, 78, 119]

3 Biometrics Recognition Systems Operative Modes OR Biometric Features and Types of Biometrics

Recognition systems work in two modes: authentication and identification. It is essential to distinguish between the two operational modes. Biometrics recognition comprises authentication or verification and identification. However, the primary need in civilian applications is for authentication, while identifications are more toward governmental and law enforcement applications. Authentication serves the right person with the right privileges and access at the right time. The subject wants the identity to be verified and, accordingly, be very cooperative; it is a one-to-one mapping.

For example, a subject introduces the identity to an ATM machine to withdraw money. The machine will verify whether the claimed identity belongs to the right person who claimed it to approve the request or not decline it. Authentication can be pursued by: Something You Own, Something You Memorize, and Something You Carry, as shown in Figure 1. On the other hand, in identification, the subject doesn't want to prove their identity, but it is required by an investigation party, usually a government authority. Identification mode is a one-to-many search problem, where the attributes of a subject are compared with all subject's characteristics stored in a database.

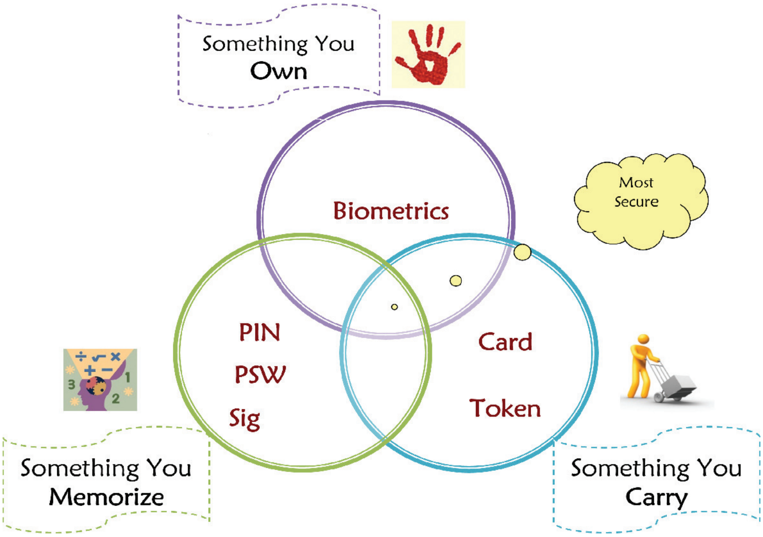
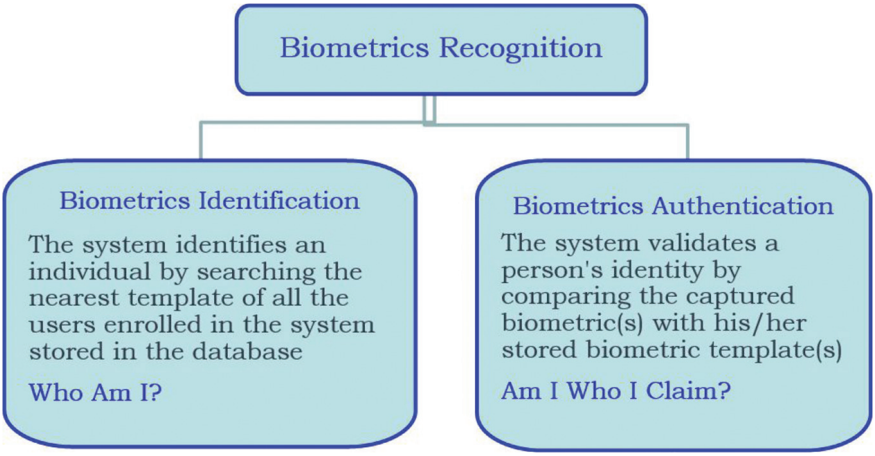


Figure 1: Authentication modes.

For example, fingerprints acquired from a crime scene are compared to all population fingerprints to match the suspect's identity. The architectures of the authentication and identification systems will be discussed in Section 15.

Figure 2 shows the recognition system's definitions of authentication and identification modes.



The system identifies an individual by searching the nearest template of all the users enrolled in the system stored in the database

Who Am I?

The system validates a person's identity by comparing the captured biometric(s) with his/her stored biometric template(s)

Am I Who I Claim?

Figure 2: Biometric systems recognition modes and definitions.

4 Biometrics versus Classical Recognition

The trivial question is how biometrics-based systems differ from classical recognition techniques. They have many differences, although they aim for the same objectives. However, the main differences can be summarized in Table 1. There are many other differences, such as in biometrics systems, you don't need to memorize or carry your identity as it is always with you. Biometrics systems are more secure, but they could be less accurate. Identity fraud with biometrics is far more challenging than classical techniques. These are just a few different examples from many others.

Table 1: Fundamental Differences between the classical and biometrics recognition techniques.

Classical	Biometrics
The password and PIN must be secret, and the token has to be well-kept in a secure place.	Biometric features are well revealed to all and are better to be high quality!
To be authenticated, the claimant must submit the exact password, pin, and source token.	Biometric features cannot be the same for the same claimant in multiple submissions.
The claimant is authorized for the exact information submissions.	The claimant must be fraudulent for the identical biometric submissions!

5 Biometrics Favored Attributes

People adhere to many biometrics that can refer to their identities. Some biometrics are commonly used, others are rarely used, and others are not. Each biometric has its pros and cons. Therefore, the choice of a biometric for a particular application depends on various issues besides its matching performance. Favoring a biometric over the other depends on the attributes of that biometric. Seven essential attributes can be used to evaluate the suitability of a biometric. These attributes are:

1. **Universality:** Every individual accessing the application should possess the trait. This is one reason to make fingerprints the most common biometric. Each finger can uniquely identify its corresponding individual. Usually, people are asked to submit their ten fingers in case any finger is lost or suffers from losing the patterns the other fingers can be used for recognition.

2. **Distinctiveness (Uniqueness):** The biometric trait should be as unique as possible to individuals. Some biometrics, such as hand

geometry, can easily mix up people, while others can accurately identify them, such as fingerprint patterns. The favorite biometric is the one that carries the most distinctive traits.

3. **Permanence:** The biometric trait of an individual should be minimally varied over time. It refers to the stability of the biometric feature or trait over time, health conditions, environmental change, and other effects. The feature should be as stable as possible under adversary conditions.

4. **Collect-ability (Ability to measure):** It should be possible to accurately acquire and digitize the biometric trait using suitable devices that cause no or minimal inconvenience to the individual. The acquired features should be sensed and processed within a reasonable time; otherwise, those features will be limited to fewer applications. The sensors and processing equipment should not be costly to make the technology affordable.

5. **Performance:** The recognition accuracy and the resources required to achieve that accuracy should meet the constraints imposed by the application.

6. **Acceptability:** Individuals in the target population that will utilize the application should be willing to present their biometric traits to the system.

7. **Circumvention:** This refers to the ease with which the trait of an individual can be imitated using artifacts (e.g., fake fingers) in the case of physical traits and mimicry in the case of behavioral traits.

A comparison of various biometrics concerning these attributes is shown in Table 2.

The favored biometric is the one that has most of the attributes labeled H, and circumvention is L. Those labels can also change with the development and advancements of the technologies used to acquire them and people's convenience and acceptability. For example, the iris acceptability in Table 2 is labeled L. Nowadays; it can be labeled H as the technology of acquiring the iris from a comfortable distance and the convenience of using this biometric is well advanced. Iris biometrics is adopted in most airports for recognition purposes. However, the iris circumvention is labeled L *(desired attribute)*, while it currently can be labeled H *(easily faked)*. This is because of the introduction of very thin colored lenses and special eye drops that can foul the biometric system. Not a single biometric is expected to accurately meet all the seven requirements (e.g., accuracy, practicality, cost) desired by all applications (e.g., Digital Rights Management (DRM), access control, welfare

Table 2: Comparison of biometric technologies. High, Medium, and Low are denoted by H, M, and L, respectively.

Attributes	**Universality**	**Distinctiveness**	**Permanence**	**Collect-ability**	**Performance**	**Acceptability**	**Circumvention**
Fingerprints	H	H	H	H	H	M	M
Iris	H	H	H	M	H	L	L
Face	H	M	M	H	M	H	H
Gait	M	L	L	H	L	H	M
Ear	H	M	H	M	M	M	M
Palmprint	H	H	H	M	H	M	L
Palm Geometry	M	M	M	H	M	M	M
Palm Vein	H	H	H	M	H	H	L
Retina	H	H	H	L	M	L	L
Voice	M	M	L	H	M	H	M
Signature	L	L	L	H	M	H	H
Keystrokes	L	L	L	M	L	M	M

distribution). In other words, no biometric is ideal, but many of them are admissible. Thus, fusing multibiometrics (multimodal systems) becomes a norm in highly secured recognition systems. The relevance of a specific biometric to an application is established depending upon the application's nature and requirements and the biometric characteristic's properties [36]. Now, having specified how to quantify each attribute, finding the best biometric system looks easy! Unfortunately not! Setting weighting for each attribute to value the system for a particular application is governed by the biometric system administration, which is application dependent. System administrators prioritize the attributes differently based on the application. The focus may be on robustness more than distinctiveness or vice versa. The same thing can be said about the other attributes. Thus, the interaction with the biometric systems is highly changeable, and what is suitable for a certain application might not be so for another one. Accordingly, the impact of the attributes of any biometric cannot be unified among all applications. However, for the users, it is simply to decide based on the following: Is the biometric system easier and friendlier to use, faster in response, accurate in results than its rivals, and cheaper?

6 Applications of Biometrics

The authentication of identity is something that everyone is aware of these days due to its importance in our modern society. No one would like to see their identity stolen, which may incur devastating consequences. The typical questions that biometric-enabled systems might ask:

1. Is the person a genuine claimant?

2. Is this person authorized to use/access this facility?

3. Is the person on the wanted list by the authorities?

These questions are posed in various scenarios ranging from issuing a driver's license to gaining entry into a country. The applications can be categorized into three categories: Commercial, Government, and Civilian. Some of the desired applications based on these categories are:

1. Commercial applications include computer network login, electronic data security, e-commerce, Internet access, ATM or credit card use, physical access control, mobile phone, PDA, medical records management, distance learning, etc.

2. Government applications include national ID cards, managing inmates in a correctional facility, driver's licenses, social security, welfare disbursement, border control, passport control, etc.

3. Forensic applications include corpse identification, criminal investigation, parenthood determination, etc.

Figure 3 summarizes some typical applications according to their categories.

7 Common Questions on the Design of Biometric Systems

When a biometric system is designed or even purchased, there are some questions the design engineer should consider. Based on these questions, the system's suitability can be conformed to the application requirements.

1. Which biometric is the most suitable for an application?

2. What accuracy can a particular biometric offer?

3. What are the overall capital and running costs?

4. What happens when the system fails?

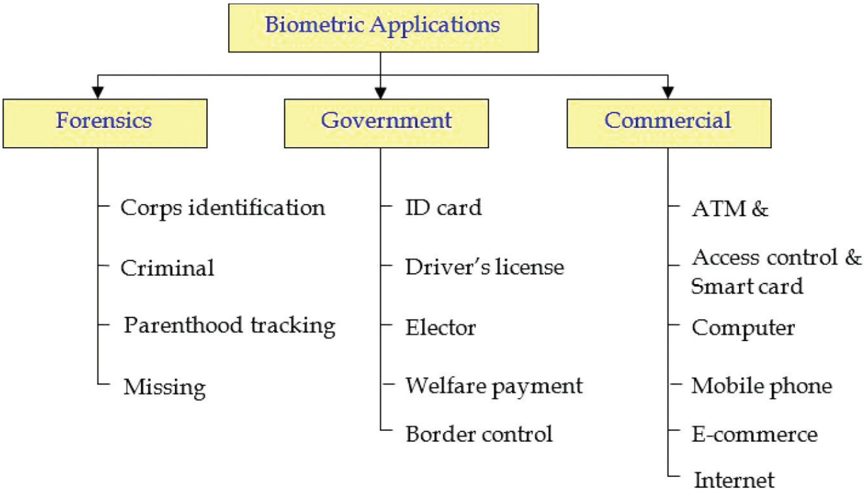


Figure 3: Common biometrics applications with respect to their categories.

5. What is the system's response time, and how is it affected by increasing the number of enrollments?
6. How secure is the system against different attacks?
7. How are the privacy issues dealt with?
8. How are possible new security holes created due to using a biometric?
9. How acceptable is the selected biometrics to the users?
10. How does the selected biometric system comply with the favored attributes in Table 2?

Companies market their products very professionally and do not cover the responses to the above questions in their advertisements. Customers should seek answers to all the questions to make the right decision and have the optimum compensation for the non-achieved requirement.

8 Biometric Technology

An abundance of biometrics can help distinguish people from each other. However, not all of them are suitable to use according to their attributes. Some of the common biometrics are shown in Figure 4. The physical biometrics are fingerprints, ear, palmprint, palm geometry, palm vein, iris, retinal veins,

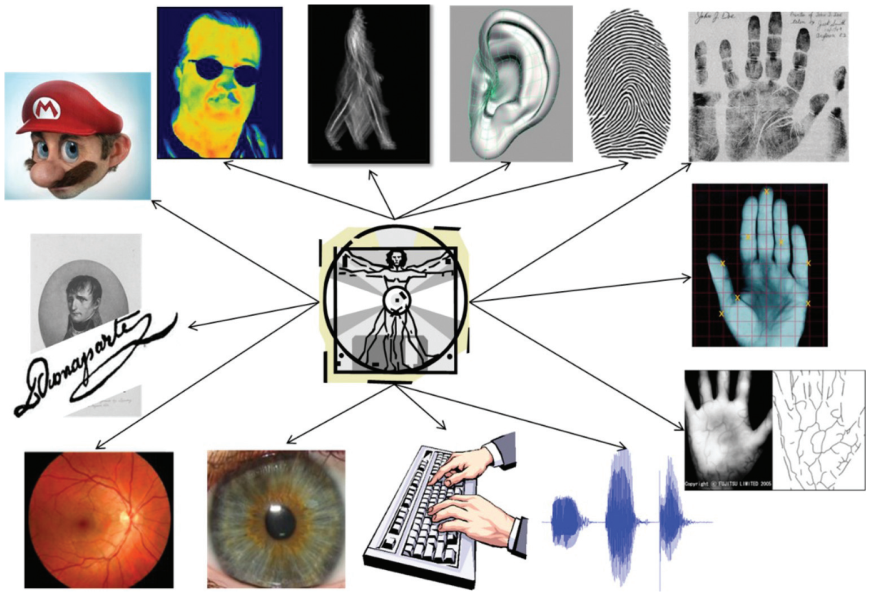


Figure 4: Some of the common biometrics.

face, and infrared thermography. On the hand, behavioral biometrics are gait, voice, keyboard typing, and signature. There are more types of biometrics, but they are less common due to their lack of compliance with one or more of the favored features depicted in Table 2. This section discusses some common biometrics and specifies their advantages and disadvantages.

8.1 *Fingerprints Biometric*

Fingerprints biometric [67] is based on the oldest unique biometric, which replaces ink paper with a digital technique. A user submits their finger on a small flat scanner or swipes it over a line scanner, and special algorithms are programmed to recognize the subject. It is basically the automated version of the original manual methods. It is highly distinctive and robust, yet problems arise from poor finger scanning, such as partial or noisy images acquired, the dryness of the fingers, and the resolution of the scanners and compatibility. Also, producing a thin membrane carrying the fingerprints of any person is easy and cheap. New scanners can detect these membranes, though. Fingerprint-based biometric systems are currently the leading technology in the biometric market share; they occupy more than 50% of the biometric market. The following are the main contribution to this domination.

1. Universality - Almost everyone owns fingerprints of at least one finger.

2. Uniqueness - Every single finger has a unique fingerprint pattern. This uniqueness is among all humanity, dead or alive.

3. Permanence and invariance with aging - Fingerprints form in the 7th month of fetal development and remain unchanged during an individual's lifetime. Only the finger's size changes, but the pattern remains the same.

4. Suitable collectability; ease and low cost of acquisition - Fingerprint scanners have become very cheap and easy to integrate with any computational system. It is also easy to use by non-trained users.

The question now is how fingerprints are described. The epidermis of a fingertip constitutes a pattern of parallel lines called ridges and valleys. The dark lines represent the ridges, and the valleys are represented by the light areas between ridges, as shown in Figure 5. These patterns have many common characteristics which the recognition systems can identify.

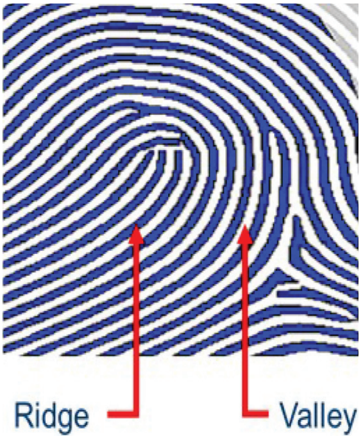


Figure 5: Ridges and valleys of the fingerprints.

Local and global features can be extracted from these patterns. Local features are referred to as minutiae (small details) and are defined by the ridges themselves at a local level. Minutiae are extensively used for the matching process as minutiae matching is the traditional way of matching fingerprint used by experts and provide good accuracy and low error rates. Around 157 different types of minutiae have been identified, most of which are rarely identified in most fingerprints and are greatly dependent on impression conditions and fingerprint quality. The most prominent minutiae used for fingerprint matching are terminations, where a ridge line abruptly ends, and bifurcations, where a ridge line divides into two different lines. Figure 6 shows examples of the termination and bifurcation of local features.

On the other hand, the global features include singularities, frequency of ridges, and their overall orientation. The singularities are the points where there is maximum orientation, called core and the points where ridges diverge, called delta. One interesting property of the singularity points is that they can group the possible patterns into 5 to 8 classes. They combine into different orientations in fingerprints to distinguish the various classes. Figure 7 shows five main classes: arch, left loop, right loop, tented arch, and whorl.

Figure 8 shows the local and global features extracted from a fingerprint.

From the quick details we presented, we can realize how simple the features can be used to recognize millions of people. The challenge here is finding efficient techniques to detect the local and global features accurately. Recently many efficient methods have been implemented to detect these biomarkers

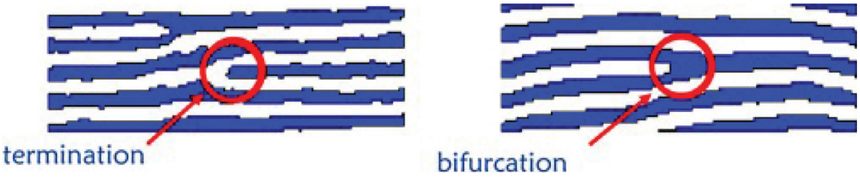


Figure 6: Termination and bifurcation of local features.

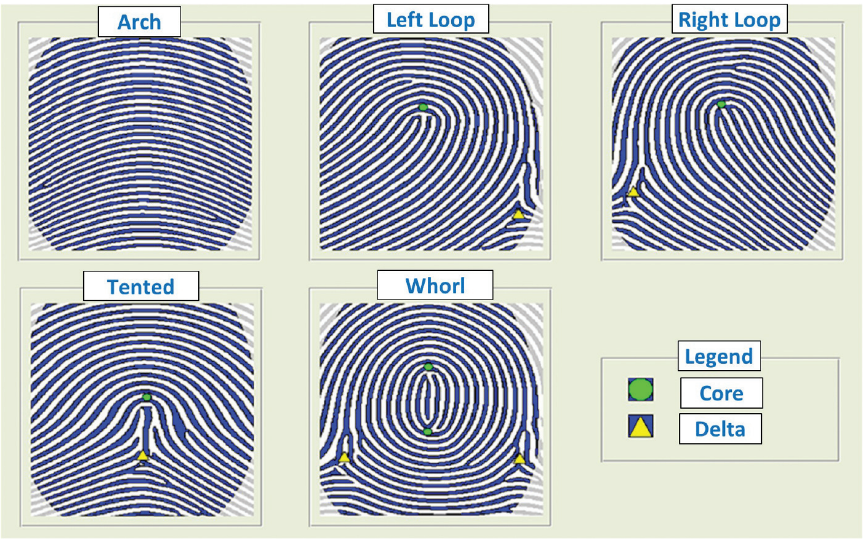
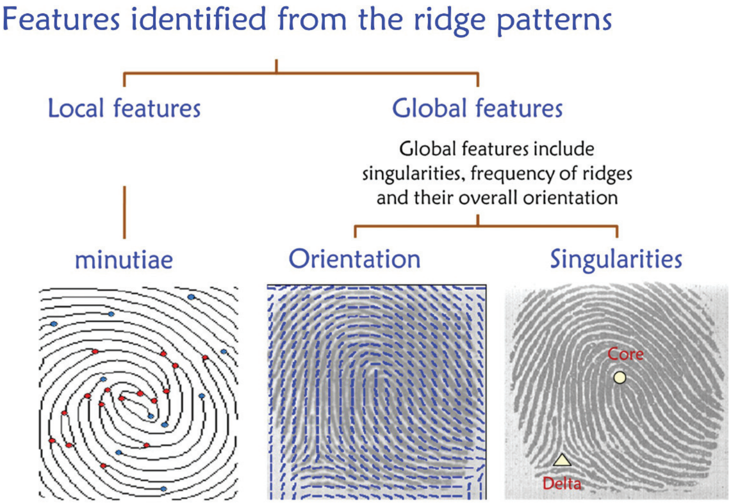


Figure 7: Global features-based classification.


Features identified from the ridge patterns
Local features
Global features
Global features include singularities, frequency of ridges and their overall orientation
minutiae
Orientation
Singularities
Core
Delta

Minutiae [mi-noo-shee-ee] is the plural of minutia [mi-noo-shee-uh, -shuh]

Figure 8: Ridge patterns features.

correctly. Interestingly, know that fingerprints are not solely for humans. Some animals have fingerprints quite similar to human fingerprints. The same algorithms can be applied to recognize the animal! Figure 9 shows only two examples of chimp and koala. Caw muzzles and dog nose prints are also unique, and it is an exciting line of research.

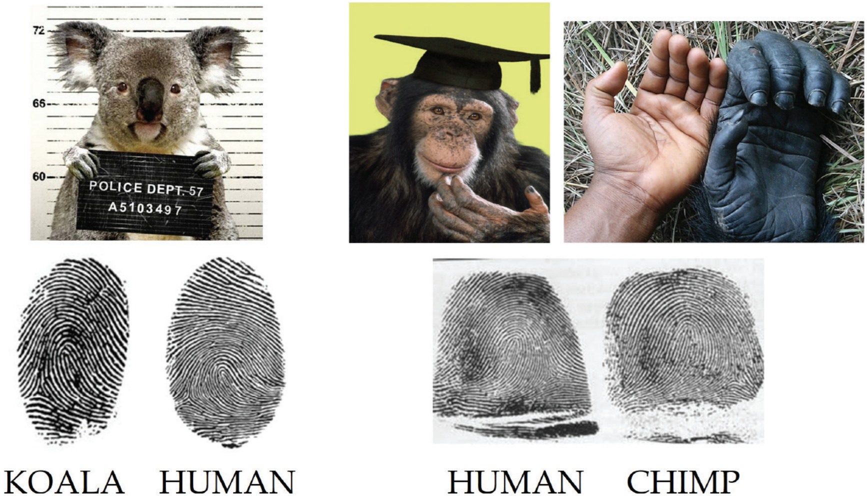


Figure 9: Fingerprints of koalas and chimps and similarity to humans.

8.2 Iris Biometric

This method extracts distinctive features from the iris pattern in the colored part of each eye. The iris reveals rich random, interwoven patterns in the visible band of light. The iris patterns of the left and right eyes are different. Genetically identical-twin eyes have iris patterns that are different and uncorrelated in detail. One of the vital advantages is contactless, which is an advantage over fingerprint biometrics. The problems with the Iris biometric:

1. It can be disturbed by colored contact lenses and some eye drops.
2. Some eye sicknesses have severe effects on the iris patterns.

Figure 10 shows an interesting incidental tracking by National Geography magazine. Shrbat Gulat's photo was taken over several periods of her life. Her face and identity changed during her struggling life when she migrated to Pakistan and returned to her home country in 2017. She was identified through her iris biometric only.

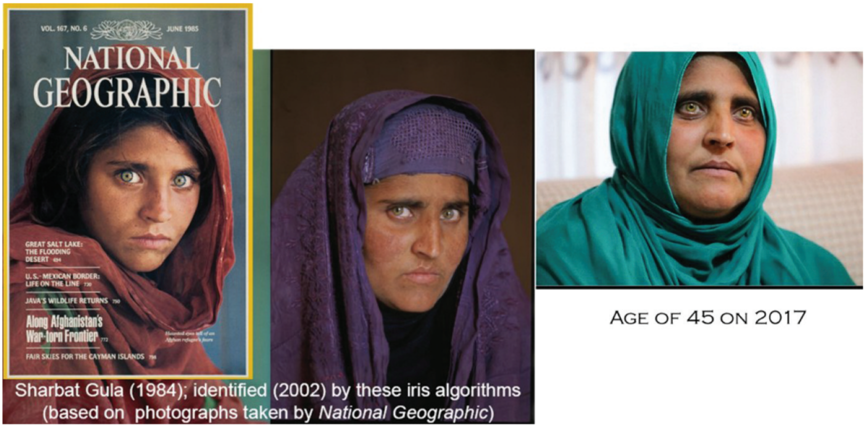

VOL. 167, NO. 6
JUNE 1985
NATIONAL GEOGRAPHIC
GREAT SALT LAKE: THE FLOODING DESERT
U.S.-MEXICAN BORDER: LIFE ON THE LINE
JAVA'S WILDLIFE RETURNS
Along Afghanistan's War-torn Frontier 772
FAIR SKIES FOR THE CAYMAN ISLANDS
Sharbat Gula (1984); identified (2002) by these iris algorithms
(based on photographs taken by National Geographic)
AGE OF 45 ON 2017

Figure 10: Sharbat Gulat's photos over time and the persistence of her iris patterns.

Interestingly, animals can also be identified through their iris patterns.

8.3 Retina Vein Patterns Biometric

Physicians with different imaging modalities usually observe retinal blood vessels (veins), which deliver nutrients to various tissues in the eye. The retinal vessel system will not change in its lifetime except for pathological changes. Retinal vascular characteristics, such as vascular thickness, reflectivity, and curvature, can be used as essential biomarkers for many retinal and haematological-related diseases [121]. Also, the patterns of these vessels are unique among people, so they can be used as biometrics to recognize people in a big population. This method scans the patterns of the blood vessels in the back of the eye. This requires the subject to stand away by a few inches from a light focused on their eyes to scan the pattern of the blood vessels. However, this biometric is not welcomed by people due to the invasive way of taking retinal images, the long time to acquire them, and the expensive imaging units. This biometric is not suitable for border security for the reasons mentioned above despite its uniqueness. Figure 11 shows an example of a retinal image and what the vein patterns look like.

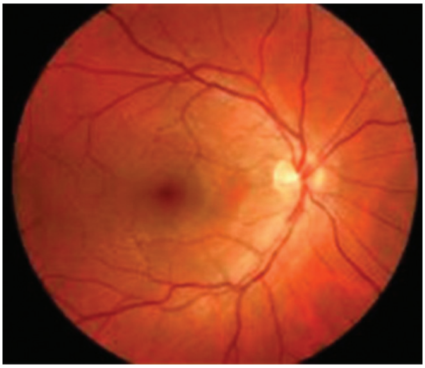


Figure 11: Retinal veins.

8.4 Face Recognition Biometric

One of the early common biometrics used is face features. Many features can be extracted from faces; accordingly, different methods are available to define this biometric. The face image can be acquired from a distance by suitable cameras, an excellent tracking method without being noticed by the tract subject. It is used for border control, law enforcement, and fingerprint and iris biometrics to constitute a robust multimodal recognition system. Initial research in face recognition depends on extracting distinctive facial features, such as the distance between the eyes, nose, ears, and lips, shapes, and locations. Then use machine learning techniques to recognize people. Recently deep neural networks have been used without the necessity of extracting the features specified by the feature engineers. The recognition will be a straightforward task if a suitable dataset with an abundance of images is available. However, despite all the favored characteristics of facial recognition, many problems may deter the system from the correct recognition. Some adversary situations are variations in illumination, pose, occlusions, facial expression, and aging.

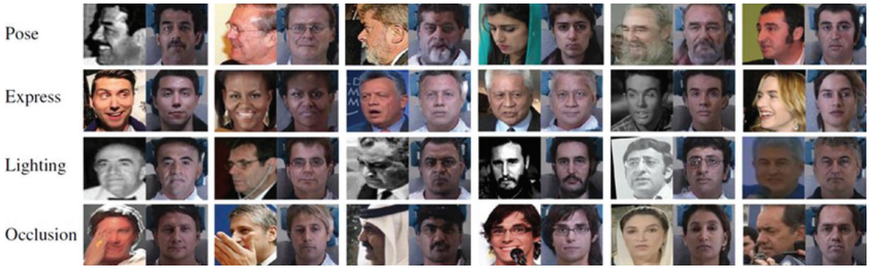


Figure 12: Situations where face recognition performance degradation is expected [89].

Figure 12 shows some of the problems with face recognition. However, there is intensive research in rectifying these problems, and the recognition accuracy can be well improved in many cases.

The most challenging problem is to distinguish genetically identical people! In this case, we have to use other recognition modalities, such as iris and fingerprints.

Figure 13 shows an example where face recognition fails with a genetically identical triplet. This biometric form cannot be used with genetically identical people as no technique can rectify this problem.

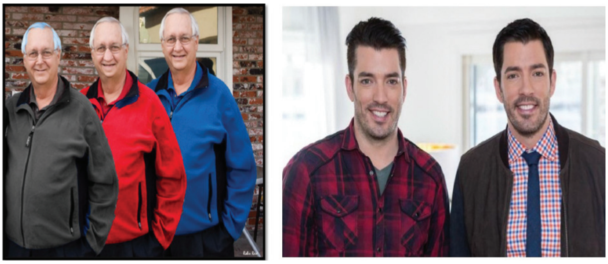


Figure 13: Identical twin and triplet where face recognition biometric fails.

8.5 Voice Biometric

Voice is the most natural way of communication, and it can be acquired by telephone or mobile microphones [1, 2, 17, 18]. It is an involuntary behavior biometric. It does not need the overhead for special sensors or equipment. The disadvantages of voice biometrics are its vulnerability to noise, respiratory sicknesses, aging, and coding techniques. Yet, it is used widely in telephone banking, verification over phones, and ATM machines.

8.6 Gait Biometric

People can be recognized by the way they walk [102]. It is a non-intrusive method and can be measured from a distance. It can be efficiently augmented with other biometrics, such as facial biometrics. It is not very accurate, yet it can provide helpful information about the subject under surveillance. This biometric is a spatio-temporal phenomenon that characterizes the motion characteristics of an individual, and it is extracted from a sequence of the subject frames. One of the most straightforward features to be extracted is from the subject silhouettes. A set of feature vectors is derived from the width of the outer contour of the temporally ordered subject silhouettes. The width is simply the difference in the location of the rightmost and leftmost boundary pixel in that row, as shown in Figure 14. Each subject's frame is represented by one vector, with each coefficient representing one width from a specific location of the frame.

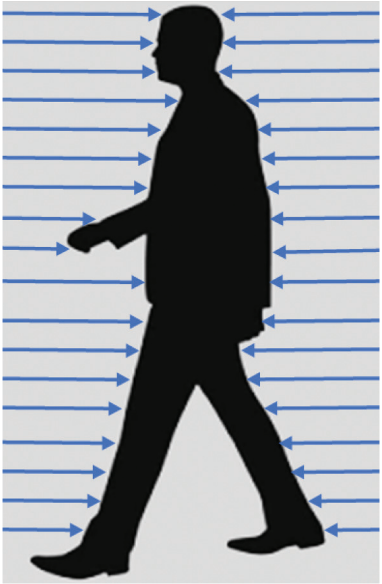


Figure 14: Gait Feature vector.

This biometric is non-invasive and doesn't need any cooperation from the subject. This is very useful for tracking and recognizing people from a distance. However, it has many disadvantages, as the accuracy can be affected by the walking speed and phase, posture, and dress. Also, it is view-dependent [84, 85, 126]. There are intensive research activities on overcoming such adversary conditions, especially using deep learning techniques.

8.7 Ear Pattern Biometric

The Pinna pattern of the ear has different patterns and can be used to recognize people from close proximity [3, 13, 23, 34]. It is claimed to be unique among all people. It is a bit intrusive as the hair must not cover the ears, which may offend some people. It requires well-illuminated ears to avoid the shadow effect. The interest in this biometric started from a tragic story. Mark Dallagher was convicted of murdering an older woman in Huddersfield, a market town in West Yorkshire, England. He was convicted after the prosecution showed that he could have only left ear prints on a newly washed window as he listened for signs of movement inside the house. The ear print of the murderer extracted is shown in Figure 15.

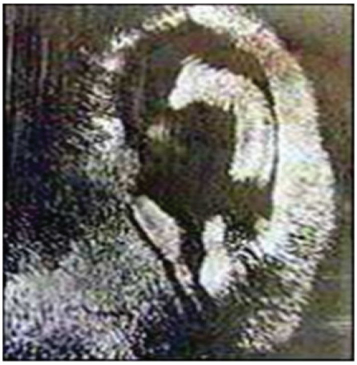


Figure 15: Criminal Ear Print extracted.

Airport security could soon be looking at the shape of ears when deciding whether to allow a person into the country. Figure 16 shows the different ear images of David Cameron, King Charles, and Daniel Craig, which are very different.

8.8 Palm Geometry

This biometric depends on the geometry of the hand and fingers without taking the fingerprint or the palm print. Hand geometry features are extracted from the dimensions of fingers, the location of joints, shape, and size of the palm, Figure 17. The first commercial system was introduced in 1974. The hand is placed over a suitable sensor using specific guides to place the hand correctly. These days there are 3D palm scanners that can provide more features about the hand, such as the thickness and nail shapes.

Palm geometry is fast to process and a more compatible image acquisition device. Any scanner or low-resolution cheap webcam can be used to acquire the palm image. However, it is not unique among a large population and

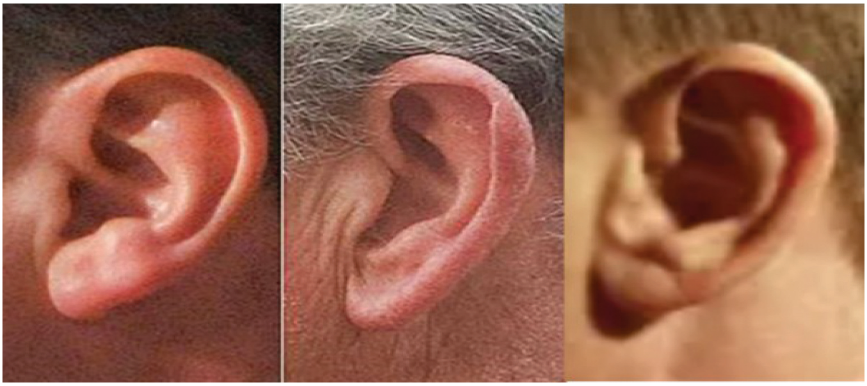


Figure 16: Ear images from the left, David Cameron, King Charles, and Daniel Craig.

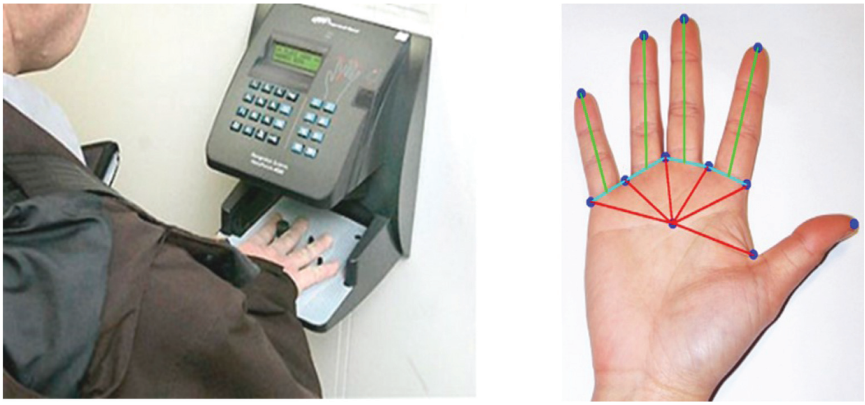


Figure 17: Palm geometry device and some palm features.

needs a wide contact space for imaging the hand. Also, it is changing with aging. Recently, some contactless techniques need further investigations to be commercially adopted [4, 22, 62]. This biometric has been successfully implemented for physical access control, time and attendance, and personnel identification applications in small-scale companies.

8.9 Palmprints Biometric

Another biometric that can be invested from palms is palmprints [128, 129]. Palm biometrics, like fingerprints, uses ridge and valley patterns on the palm surface. Palmprint is also considered a useful biometric and can be measured simultaneously by scanning the hand geometry and fingerprints. Palmprint carries more accurate information about the person than palm geometry,

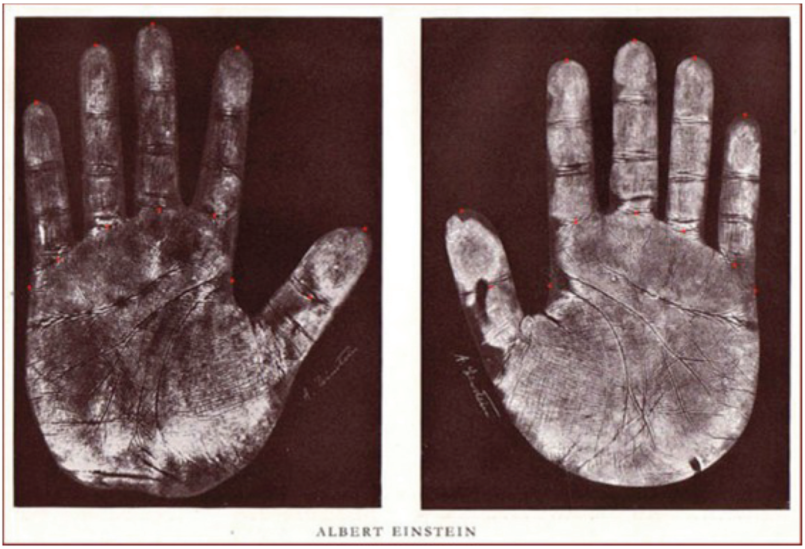


Figure 18: Einstein palmprints.

which is considered unique. Yet it required bigger, with at least big handful sizes to acquire the images than fingerprint scanners. Palmprint sensors are consequently bulkier and more expensive than fingerprint sensors. The acquired images carry the lifelines and wrinkles information in addition to the palm ridges and valleys. Figure 18 shows the palmprints of Einstein. Albert Einstein (1879–1955) was a German theoretical physicist. Recently, deep learning techniques have been used to implement successful palmprint recognition systems [9].

8.10 Palm Vein Pattern Biometric

This biometric is a fairly new growing fast contactless biometric [71–73, 106, 122]. Vein pattern recognition uses an infrared light source to scan for hemoglobin in the blood. De-oxygenated hemoglobin absorbs the light creating an image of the vein pattern that is reflected and captured by the scanner. The backs of hands and palms have more complex vascular patterns than fingers and provide more distinct features for pattern matching and authentication. Fujitsu company pioneered the palm vein scanners shown in Figure 19(a), and now many labs and companies are developing such scanners. A scanner developed at the University of Auckland, New Zealand is shown in Figure 19(b). This is a palm vein scanner that is capable of capturing high resolution infrared vein images and is developed using a cost efficient Rasberry Pi computer with Rasberry Pi NoIR camera. This scanner is currently under

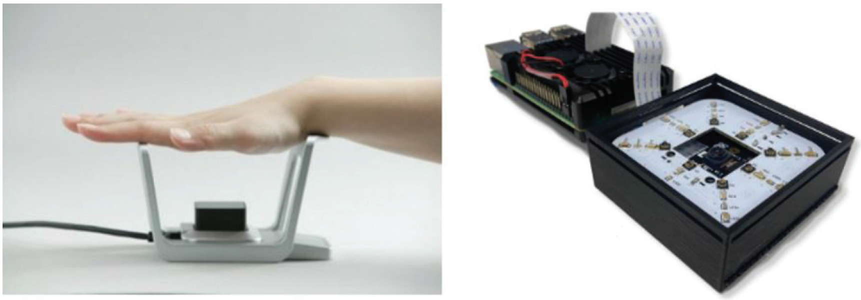


Figure 19: Palmprint scanners (**a**) Fujitsu Palmprint scanner, (**b**) UoA palm vein scanner.

patent review. Vein recognition technology is secure because the authentication data exists beneath the skin and is very difficult to forge. It is also highly accurate - according to the Fujitsu report, in testing using 140,000 palm profiles of 70,000 individuals, it had a false acceptance rate of less than 0.00008% and a false rejection rate of 0.01%. As indicated, veins reside hidden beneath the skin; thus, it is more difficult to attack because the enrolled users do not leave vein biometric data behind, as with fingerprints. Imitation of the body parts cannot be used to fool the palm vein recognition system because the palm vein sensors depend on the blood flow, thus the aliveness of the subject.

Due to the virtues of the hidden veins, finger veins have also been used for recognition. Hitachi company developed a suitable scanner for that, Figure 20. They are used independently or augmented with the palm vein for a multimodal system [100, 101]. Recently that interest has been oriented toward the wrist vein, which also showed great potential in using them for people recognition. This paper's separate section will focus on implementing a system based on the wrist scanner. However, the veins-based systems are not immune from pitfalls. Body temperature, ambient temperature, humidity, uneven distribution of heat, heat radiation, the proximity of the vein to the surface, camera calibration, and focus affects the image quality. The scanners are still relatively expensive. They are still untested widely, like the fingerprints, because they haven't yet been deployed globally.

8.11 Less Commonly Used Biometrics

There are other biometrics that we have not elaborated on, yet they have the potential to be used more commonly in the future. These biometrics include facial infrared thermogram, keystroke, odor, and signature. They haven't been proliferated because they currently do not fulfill the desirable attributes listed in Table 2. DNA is another biomarker that can be considered a biometric. DNA is still not fully automated and needs human interference, and it cannot

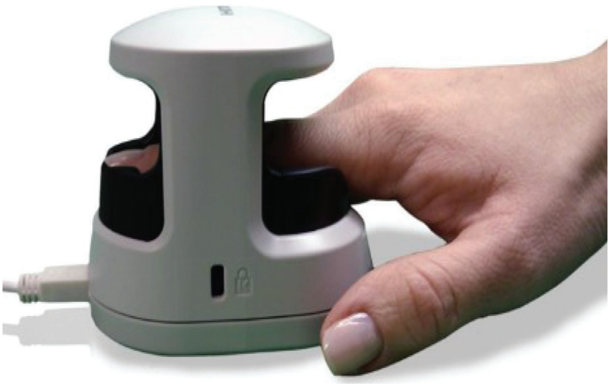


Figure 20: Hitachi finger vein scanner.

be used in real-time due to the extensive processing time. Yet, it is vital in forensics, proof of kinships, and genealogy.

9 Structures of the Human Biometric Recognition Systems

Biometric recognition is a pattern recognition system and comprises biometric verification and identification, as indicated in Figure 2. In both cases, a subject (person in this case) must enroll his/her biometrics in the system through a training or enrolment phase. The next stage is the identification or verification phase, where the association of the subject to a targeted objective is confirmed or denied.

9.1 Biometric Enrolment

In this phase, the user introduces his/her biometrics to the system with biographic information. The biometrics and the related information are stored in a secured database to be used for identification or verification phase later. The enrolment system is shown in Figure 21.

In this phase, the biometrics features (minutia fingerprint here) are extracted and stored in a database. The quality assessment module determines if the feature extractor can effectively use the sensed data. The subject must enter the bio information such as name, date of birth, bank account, and more or less info depending on the application the recognition system will be used in.

9.2 Biometric Verification

The system validates a person's identity by comparing the captured biometric(s) with their stored biometric template(s). In such a system, an individual who

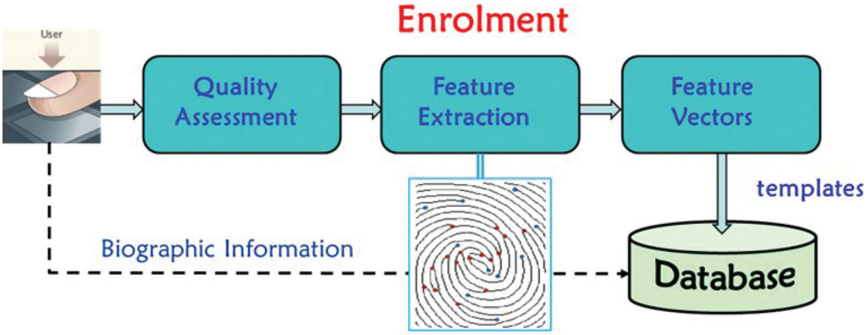


Figure 21: Enrolment System structure.

desires to be recognized claims an identity, and the system conducts a one-to-one comparison to determine whether the claim is a true claimant or not. The system structure is similar to the enrolment system added to comparison processing, as shown in Figure 22.

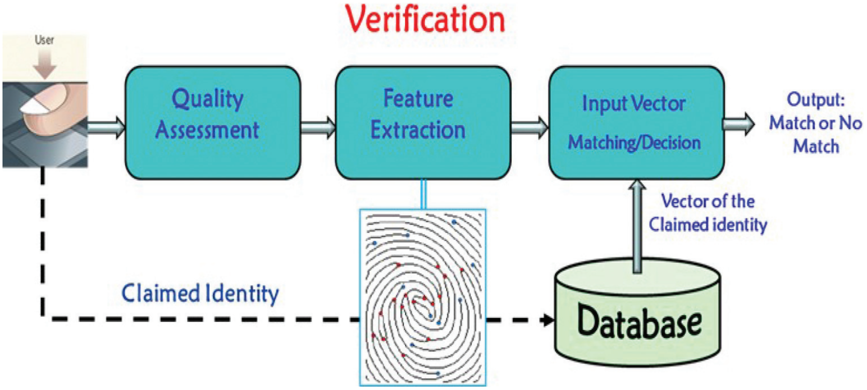


Figure 22: Verification system architecture.

A token card, memory stick, pin number (PIN), or any other identity source can be used to claim the identity. The claimed identity will act as a pointer to pull the corresponding biometric features stored in the database during the enrolment phase to camper it with the current submitted biometric features. The output is either "accept" if the matching happens or "decline" the other way. The system response time does not depend on the number of enrolled subjects in the database, as it is a one-to-one matching.

9.3 Biometric Identification

The system identifies an individual by searching the nearest template of all the users enrolled in the system stored in the database. It is a one-to-many matching. In this case, the processing time depends on the number of populations registered in the system. The reason is that it must compare the input biometric with all the people features in the database. The response time for 100 enrolments is faster than when they are one million. Figure 23 shows the structure of the identification system.

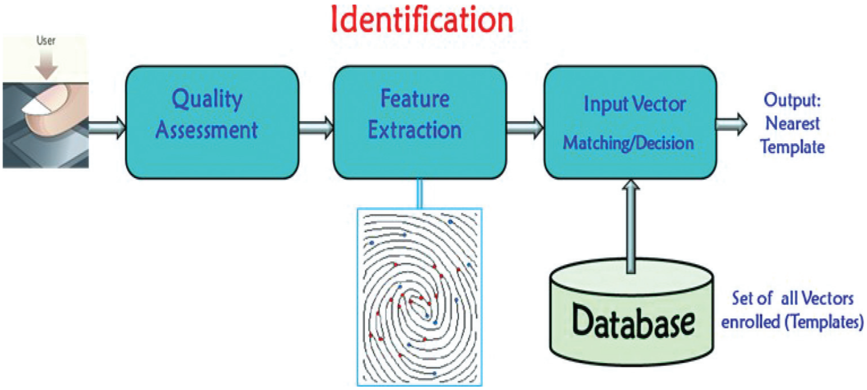


Figure 23: Identification system architecture.

10 Positive and Negative Recognition

Positive recognition is to prevent multiple people from using the same identity. The owner can only access a laptop using his fingerprint as a biometric. Contrary to the password or token (which can be shared), no one can use the laptop, but the owner as identity is associated with the owner's fingerprint, which is unique! On the other hand, Negative recognition is to prevent a single person from using multiple identities. Negative recognition is more challenging. Suppose you want a new passport. Your biometrics allow the authority to know that you do not have a passport with a different name. A biometric database is used to find duplicate requests by matching biometric traits (features) (fingerprint, face, etc.) with all the individuals in the database who have already been issued passports. Figure 24 clarifies the notions of positive and negative recognition.

Biometrics can be used for positive and negative recognition. Verification is typically used for positive recognition to prevent multiple people from abusing the same identity. Identification is used for negative recognition and can also be used for positive recognition (the user is not required to claim

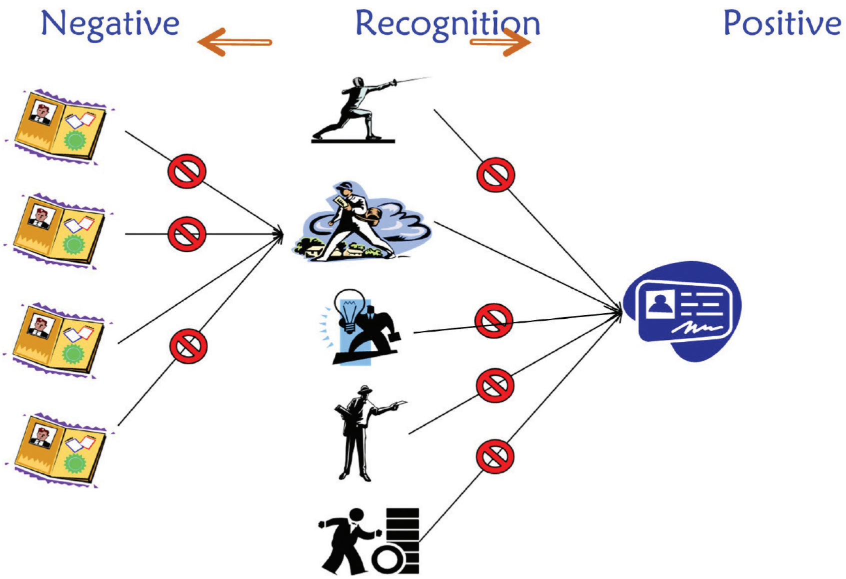


Figure 24: Positive and negative recognition interpretation.

an identity here). Here we may notice one more advantage of biometric systems. Negative recognition can only be established through biometrics and is impossible through traditional methods. Yet, the conventional techniques of people recognition, such as passwords, PINs, keys, and tokens, may weakly work for positive recognition because one can lend the personal token to any other person to use or share the password or pin numbers. To make it strong, they must be kept well secured.

11 Biometric System Performance Measures

In the traditional password, token, PIN-based systems, a perfect match between two alphanumeric strings is a must to validate a user's identity. However, getting two biometric samples with exactly matched feature sets in the biometric systems is impossible. This is due to imperfect sensing conditions (e.g., noisy fingerprint), alterations in the user's biometric characteristic (e.g., speech for speaker recognition), changes in ambient conditions (e.g., noisy environment for speaker recognition), and variations in the user's interaction with the sensor (e.g., occluded iris or partial fingerprints). Having said that, it is impossible for two feature sets originating from the same biometric of a user to look exactly the same. A perfect match between two feature sets might indicate the possibility of a replay attack being launched against the system. The variability observed in the biometric feature set of an individual is referred to as intra-class variation. While the variability between feature sets originating from two different individuals is known as inter-class variation. The desired feature set exhibits slight intra-class variation and significant inter-class variation.

A similarity score indicates the degree of similarity between two biometric feature sets. A similarity match score is a genuine or authentic score if it results from matching two samples of the same biometric trait of a user. A similarity match score is known as an impostor score if it involves comparing two biometric samples originating from different users. A similarity impostor score that exceeds a threshold η results in a false accept (or a false match). A genuine similarity score that falls below a threshold η results in a false reject (or a false non-match). The False Accept Rate (FAR) (or the False Match Rate (FMR)) of a biometric system can therefore be defined as the fraction of similarity impostor scores exceeding a threshold η. However, the False Reject Rate (FRR) (or the False Non-match Rate (FNMR)) of a biometric system may be defined as the fraction of genuine scores falling below a threshold η. Optionally used, the Genuine Accept Rate (GAR) is the fraction of similarity genuine scores exceeding a threshold η and can be related to FRR by:

$$GAR = 1 - FRR \tag{1}$$

Regulating the value of η changes the FRR and the FAR values. A given biometric system can't decrease FAR and FRR errors simultaneously. However, we can do that only if we fuse two or more biometrics. Figure 25 demonstrates these notions clearly.

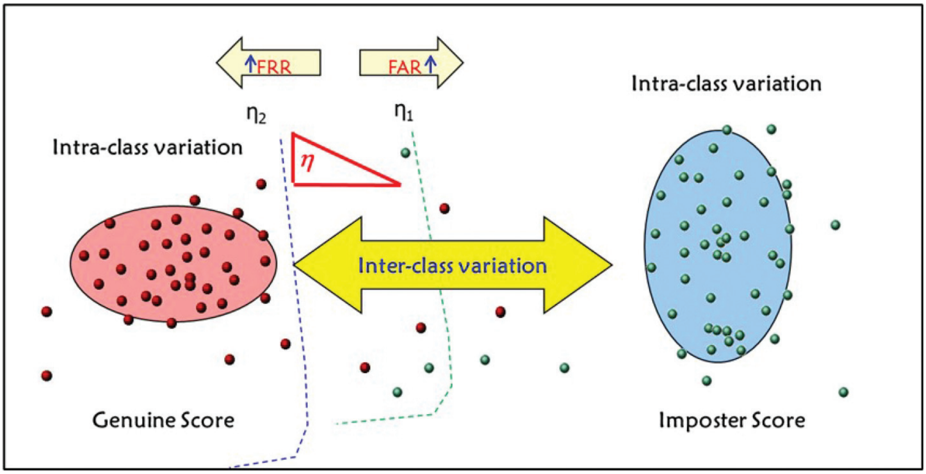
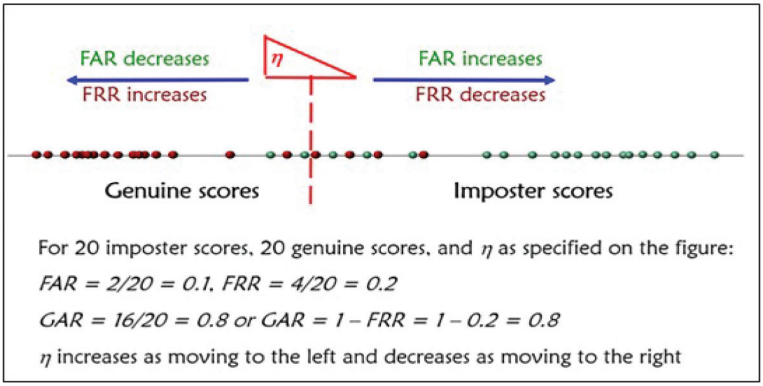


Figure 25: Inter-class and intra-class notions.



For 20 imposter scores, 20 genuine scores, and η as specified on the figure:

FAR = 2/20 = 0.1, FRR = 4/20 = 0.2

GAR = 16/20 = 0.8 or GAR = 1 – FRR = 1 – 0.2 = 0.8

η increases as moving to the left and decreases as moving to the right

Figure 26: FAR, FRR, and GAR Calculations.

From Figure 25, we may notice that when we increase the desired similarity threshold i.e., $\eta_2 > \eta_1$, we increase the FRR as we only include the genuine samples, and we reduce the FAR as we discard many genuine samples. The vice versa situation happens when we move toward $\eta_2 < \eta_1$.

Example. *Calculate the FAR, FRR, and GAR for one-dimensional biometric features that have 20 genuine samples and 20 imposter samples. Figure 26 shows the solution of such a scenario with the threshold value η as indicated. Try to shift η to the left or right, then calculate the performances and verify that when FRR increases, FAR decreases and vice versa.*

Other measures may also assess the performance of a biometric system. Equal Error Rate (EER) is an essential measure for assessment. The EER refers to where the FAR equals the FRR, as shown in Figure 27. A lower EER value indicates better performance.

If we want to compare two or more biometric algorithms or systems, we can look at the EER level. Whichever delivers the lowest EER level is the best-performing one. Furthermore, the FAR and FRR scores at various values of η can be summarized using a Detection Error Tradeoff (DET) curve that plots the FRR against the FAR at multiple thresholds. It is important to note that the decision to find a suitable value of the threshold η to set the balance between false accepts and false rejects is not unique across all users or applications of biometric systems. Figure 28 shows the DET of the biometric systems, i.e., BS1, BS2, and BS3, with the EER line. The system with the lowest DET curve corresponds to the lowest EER and thus is the best-performing one, in this case, BS1.

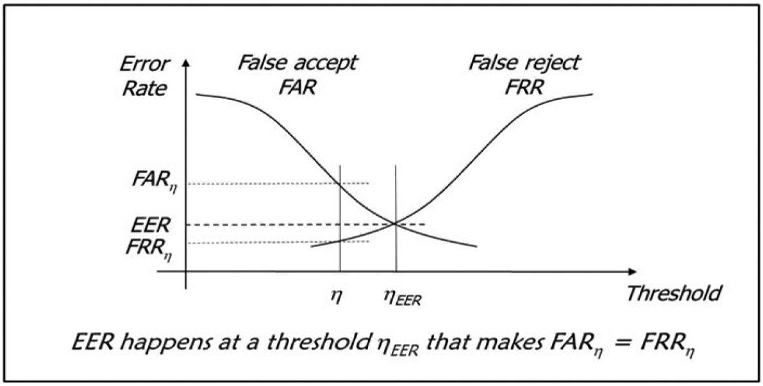


Figure 27: The relation between FAR, FRR, and EER.

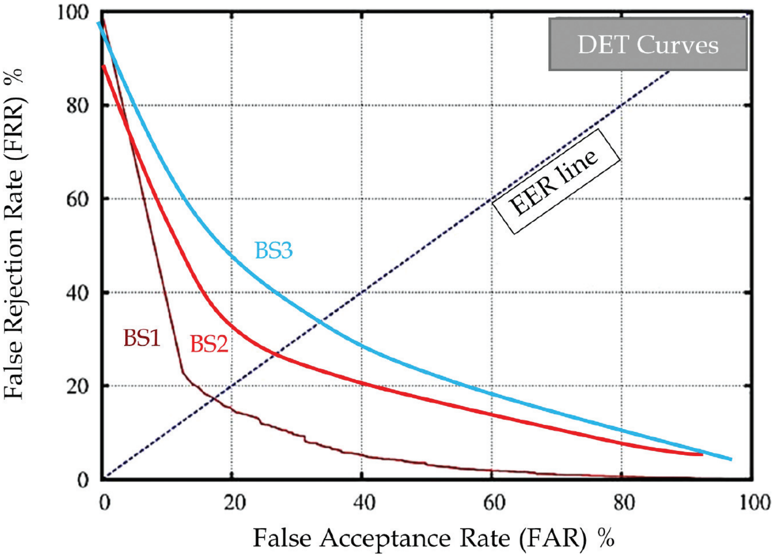


Figure 28: Detection Error Tradeoff of three biometric systems.

12 Biometrics Fusion

As we have mentioned in Section 5, we highlighted the limitations of relying on a single biometric for achieving comprehensive security. In other words, no single biometric can fulfill all the desired attributes and address all potential security vulnerabilities. Therefore, combining different biometrics to enhance recognition systems and achieve optimal results is necessary. We can integrate various biometrics such as fingerprints, facial features and iris patterns, palm prints, geometric characteristics, and vein patterns. The fusion of biometrics depends on the specific application at hand. Combining biometrics is based on human experience, as individuals recognize each other by considering evidence from multiple biometric traits (both physical and behavioral) and contextual details related to the environment. Each biometric, referred to as a modality, cannot always be relied upon solely for recognition. However, when the information provided by these multiple experts is consolidated, it enables accurate verification or identification. Additionally, biometric systems can be designed to recognize individuals by leveraging information obtained from various biometric sources. Such systems, commonly known as multibiometric or multimodal systems, are expected to achieve higher accuracy levels due to the availability of multiple pieces of evidence. Numerous types of research have been conducted on biometric fusion. We highlight here some of the efforts in this direction.

Ross *et al.* [97], explored multibiometric systems and their advantages, discussing integrating different biometric information sources. It covers acquisition and processing schemes and examines various fusion architectures. The fusion levels—sensor-level, feature-level, rank-level, and decision-level—are explained in detail, along with integration strategies and examples. The book also focuses on score-level fusion, discussing integration strategies and improving system performance with user-specific parameters. Additionally, the book explores incorporating ancillary information, such as data quality and soft biometric traits, in a fusion framework. It provides an information fusion framework for including soft biometric traits in the authentication. Finally, the book lists databases used for evaluating multibiometric algorithm performance. Overall, it offers a comprehensive understanding of multibiometrics, fusion schemes, integration strategies, and the incorporation of ancillary information.

Yang *et al.* [123] proposed a multimodal biometric system that combines fingerprint, palm print, and hand geometry for personal identity verification. The system was tested on a database of 98 persons, and the test performance results indicate the feasibility of the combination. The three biometrics (fingerprint, palm print, and hand geometry) can be taken from the same image.

Wang *et al.* [118] proposed a novel biometric recognition system that fuses a human hand's palmprint and hand geometry. It uses image morphology and the Voronoi diagram concept to cut the palm's image into several irregular blocks. Statistic characteristics of the gray level in the blocks are employed as characteristic values, resulting in an encouraging performance with a FAR of 0.0035% and FRR of 5.7692%.

Ramachandra *et al.* [90] proposed a new algorithm for bimodal biometric authentication. The algorithm uses fingerprint and face images to identify a person. The proposed algorithm outperforms other transformation domain techniques in terms of equal error rate (EER) and true success rate (TSR). Minchev [75] proposed a “scenario method” approach, in combination with experts-based decision support and users’ biometric “validation-in-advance” as a framework for multiple human biometrics fusion in support of cyber threats identification. Practical examples are given to illustrate the proposed ideas.

Charfi *et al.* [14] proposed a bimodal biometric system based on hand shape and palmprint modalities for person identification. The fusion step is carried out at the rank level after the classification step using SVM (Support Vector Machines) classifier. The experimentation is performed on the IITD hand database, and results demonstrate encouraging performances achieving IR = 99.34%. SIFT descriptors were extracted and represented sparsely using the sparse representation method. Walia *et al.* [114] proposed a multimodal biometric system based on an optimal score-level fusion model. The system integrates three complementary biometric traits: iris, finger vein, and fingerprint. They optimized individual classifier performance using the evolutionary Backtracking Search Optimization Algorithm (BSA). They also resolved conflicting beliefs from individual classifiers using proportional conflict redistribution rules (PCR-6). On average, the system achieves an accuracy of 98.43% and an EER of 1.57%. Singh *et al.* [103] discussed the development of numerous biometric fusion schemes over the past two decades. The paper focuses on three key questions: what to fuse, when to fuse, and how to fuse. It comprehensively reviews techniques that incorporate ancillary information in the biometric recognition pipeline. The discussed topics include incorporating data quality, combining soft biometric attributes with primary identifiers, utilizing contextual information, and continuous authentication using ancillary information. The paper also mentioned applying information fusion principles in presentation attack detection and multibiometric cryptosystems.

Chang *et al.* [11] proposed a multibiometric fusion framework- BIOFUSE, that combines fuzzy commitment and fuzzy vault using the format-preserving encryption scheme. BIOFUSE makes it improbable for an attacker to get unauthorized access to the system without impersonating the genuine user’s biometric inputs simultaneously. The results show a 0.98 true match rate at 0.01 false match rate on a virtual IITD-DB1 database, indicating that the proposed work achieves significantly good recognition performance while providing high security.

Stahlschmidt *et al.* [105] paper focuses on analyzing multimodal biomedical data using deep learning-based data fusion strategies. It highlights the need to capture complex relationships among biological processes and reviews the current state-of-the-art methods in this field. The paper proposes a detailed taxonomy to aid in selecting fusion strategies for biomedical applications and research novel methods. The proposed taxonomy categorizes fusion strategies into subcategories, each with its own advantages and drawbacks. It suggests that joint representation learning, particularly for intermediate fusion strategies, is the preferred approach as it effectively captures the intricate interactions among different levels of biological organization. The paper also highlighted transfer learning as a technique to overcome sample size limitations in multimodal datasets. With the momentum of intensive multi-biometric research, what can the fusing of biometrics offer? We can summarise some of the main advantages over single biometric systems in the following points:

1. Improved matching accuracy: Multi-biometric systems can simultaneously reduce both the False Acceptance Rate (FAR) and the False Rejection Rate (FRR) of the verification process, leading to enhanced accuracy.

2. Increased feature space: Incorporating multiple sources of biometric information expands the feature space available for identification, thereby increasing the system's capacity to accommodate more individuals.

3. Compensation for non-universality: Multibiometric systems address the issue of non-universality in biometric traits. For instance, if a person has dry fingers that prevent successful enrollment into a fingerprint system or if their fingerprints are lost, the availability of other biometric traits can still facilitate verification or identification.

4. Flexibility: Multibiometric systems offer flexibility in enrollment and authentication processes. Users can enroll using several different traits (e.g., face, voice, fingerprint, iris, hand), but only a subset of these traits may be required during authentication. This makes it challenging for impostors to spoof multiple randomly requested biometric systems.

5. Verification of live users: Multibiometric systems can verify user interaction by randomly requesting the presentation of a subset of traits. For example, the system may ask the user to say a specific sequence of digits and then submit a fingerprint, ensuring the presence of a live user.

6. Noise resilience: Multibiometric systems effectively handle noisy data, which is particularly important in adverse conditions where certain biometric traits cannot be reliably extracted. For instance, if an individual's voice characteristics are difficult to extract due to ambient noise, the multibiometric system can use facial authentication features.

7. Enhanced tracking capability: Multibiometric systems can facilitate the monitoring or tracking individuals in situations where a single trait is insufficient or temporarily unavailable. For example, in a crowded environment, a person's face and gait cues can be used for recognition, depending on the distance and pose to the camera.

8. Fault tolerance: Multibiometric systems are considered fault-tolerant, as they can continue to operate even when specific biometric sources become unreliable due to sensor or software malfunctions or deliberate user manipulation. This fault tolerance is particularly beneficial in large-scale authentication systems, such as border control systems, which handle numerous users. These advantages highlight the significant potential of multibiometric systems in achieving more robust and accurate recognition capabilities, which motivates further research to be conducted. In short, by fusing multiple biometric modalities, we can overcome individual biometric limitations and enhance the recognition system's overall performance and reliability.

13 Multibiometric System Design Considerations

Several issues must be considered by the design engineer when looking at fusing biometrics [97]. Some of these issues are:

1. Cost benefits: What is the tradeoff between the added cost and the improvement in matching performance? The cost depends on the number and type of sensors deployed, the time taken to acquire the biometric data, the storage requirements, the processing time of the algorithm, and more.

2. Determining sources of biometric information: What are the various sources of biometric information that can be used in a multibiometric system? Which of these sources are relevant to the application at hand?

3. Acquisition and processing sequence: Should the data corresponding to multiple information sources (e.g., modalities) be acquired simultaneously or at different instances? Similarly, should the information obtained be processed sequentially or simultaneously?

4. Types of information:

   (a) What types of information or attributes (i.e., features, match scores, decisions, etc.) are to be fused?

   (b) What is the impact of correlation among the sources of information on the performance of the fusion system?

5. Fusion methodology:

   (a) What should fusion scheme be employed to combine the information presented by multiple biometric sources?

   (b) Is it possible to predict the performance gain obtained using different fusion methodologies to determine the optimal one?
   (c) Is the fusion better in the feature space or decision stage?

6. Levels of fusion: Multimodal biometric systems are designed to take input from single or multiple sensors that measure two or more different modalities of biometric characteristics. The essence of multimodal biometrics lies in the fusion of various biometric modes. A typical multimodal biometric system comprises four essential modules:

   (a) Sensor level fusion: This fusion strategy involves acquiring raw data from multiple sensors, which are then processed and integrated to generate new data. This integrated data allows the extraction of features. Sensor level fusion can be implemented when multiple cues of the same biometric are obtained from compatible sensors.

   (b) Feature level fusion: The feature set is extracted from multiple sources of information and combined into a joint feature vector. This high-dimensional feature vector represents an individual. In feature level fusion, it is often necessary to apply reduction techniques to select only the most relevant and informative features.

   (c) Match score level fusion: Match score level fusion involves comparing the similarity between the input biometric and template biometric feature vectors. Each subsystem calculates its own match score value based on the similarity of the feature vectors and the templates. These individual scores are then combined to obtain a total score, which is passed to the decision module for further processing and recognition.

   (d) Rank level fusion: Rank level fusion is typically employed for person identification rather than verification. It involves consolidating the multiple ranks associated with an identity and determining a new rank that aids in establishing the final decision regarding the individual's identity.

   (e) Decision level fusion: At the decision level, fusion takes place when only the decisions outputted by individual biometric matchers are available. Each biometric trait produces a separate authentication decision, which is then combined to arrive at a final vote or decision. Various strategies can be employed to combine the decisions of individual modalities into a final authentication decision. However, fusion at this level is considered more rigid compared to other fusion schemes due to the limited availability of information.

By incorporating these fusion techniques at different levels, multimodal biometric systems aim to enhance the accuracy, robustness, and reliability of biometric recognition. The combination of multiple modalities provides a more comprehensive and holistic representation of individuals, resulting in improved performance in various applications.

14 Biometric Template Protection

Due to the pervasiveness of biometric technologies in our everyday lives, our biometric data is being collected by an increasingly larger number of applications. Consequently, there are rising concerns about the potential misuse of our biometric data, which would present a threat to both our security and privacy when this data is used to make decisions about us. This threat is especially serious in light of the fact that biometric data is irreplaceable, meaning that any compromise of this highly personal data would lead to a lifelong compromise of our security and privacy in the context of our biometric identities. For this reason, it is of paramount importance that biometric data be protected when it is used in biometric recognition systems, especially during storage in a system's database (when the data is most vulnerable to attacks, such as database hacking). This issue is the focus of the Biometric Template Protection (BTP) research field.

Biometric Template Protection (BTP) aims to secure biometric "templates" when they are stored and processed in biometric recognition systems. A "template" usually refers to the set of features extracted from a raw biometric signal, which is used to represent the underlying biometric characteristic. For example, a fingerprint image is traditionally represented in terms of a minutiae template, which specifies the locations and orientations of ridge terminations and bifurcations, and a face image is today most commonly represented in terms of a template or "embedding" learned from a face image using a deep neural network. The type of BTP method used to protect a biometric template depends on many factors, including the nature or format of the template, the level of protection required in terms of the perceived threats to the biometric system, the system's computing resources, the amount of convenience required for the users of the biometric system, etc. In general, however, it is agreed-upon that an ideal BTP method should satisfy three main criteria:

1. **Recognition accuracy:** The incorporation of the BTP method into a biometric recognition system should not (significantly) degrade the system's recognition accuracy.

2. **Irreversibility:** It should be impossible (or at least computationally infeasible) to recover the original biometric template from the protected template.

3. **Renewability/Unlinkability:**[1] It should be possible to generate multiple *diverse* protected templates from the same person's original template(s), such that the protected templates *cannot be linked* to the same identity. This would allow for the *cancellation/revocation* and subsequent *renewal* (replacement) of compromised templates, as well as the use of the same biometric characteristic across multiple applications, without the risk of cross-matching the protected templates.

Keeping in mind these criteria, we are now ready to dive into some examples of the types of BTP methods that have been proposed in the literature. Existing BTP approaches may be categorised into two main types, as described in [54]: Handcrafted (designed by humans) and Learned (learned by a neural network).[2] Sections 14.1 and 14.2, respectively, present examples of Handcrafted and Learned BTP methods from the literature, then Section 14.3 compares the two types of approaches in terms of their advantages and disadvantages.

14.1 Handcrafted BTP Methods

Handcrafted BTP methods refer to algorithms that have been explicitly formulated by humans. These include the traditional methods that, until recently, were the only BTP methods in the literature. Depending on the nature of the BTP algorithm, Handcrafted BTP methods are most commonly categorised into *Feature Transformations* and *Biometric Cryptosystems*, which are discussed in Sections 14.1.1 and 14.1.2.

14.1.1 Feature Transformations

A Feature Transformation uses a specific function to transform a biometric template into a protected version of its former self. This transformation is usually *user-specific*, meaning that it is different for each user of the biometric recognition system. In particular, the unprotected biometric template, T, of a user to be enrolled in the biometric system is transformed into a protected template, T', via a transformation function, F. The transformation function is characterized by a set of user-specific parameters, which are normally derived from a randomly-generated key, K. Thereafter, only the protected template,

[1] This criterion is often split into two or more separate (but related) criteria, such as *renewability*, *cancellability*, *revocability*, *unlinkability*, *diversity*, etc. However, in this paper we will adopt the single definition presented in [54], since it makes sense to combine the underlying, related concepts.

[2] [54] used the terms "Non-NN" and "NN-learned" to refer to BTP methods that are *not* formulated by a neural network (NN) versus those that *are*. In this paper, we use the more general terms "Handcrafted" and "Learned" to refer to the essential ideas behind the "Non-NN" and "NN-learned" categories, respectively, since we do not assume the existence of a neural network in the biometric recognition system (unlike [54]).

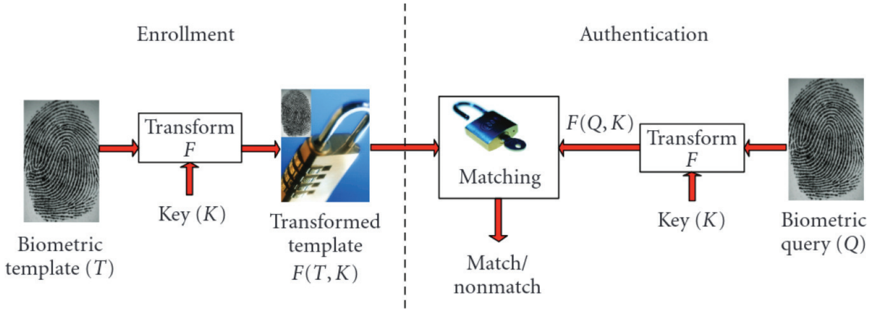


Figure 29: The enrollment and authentication stages in a fingerprint recognition system employing Feature Transformation to secure its fingerprint templates. Image from Jain *et al.* [37].

$F(T, K)$, is stored in the system's database. The enrollment process in a Feature Transformation approach is illustrated on the left side of Figure 29. During authentication, which is depicted in the right half of Figure 29, the user-specific transformation function, F, and its governing parameters, K, are applied to the unprotected query/probe[3] biometric data, Q, such that comparison between the enrolled and query templates occurs in the transformed space, i.e., $F(T, K)$ is compared against $F(Q, K)$.

Feature Transformations were pioneered by Ratha *et al.* [91], who introduced the concept of *cancellable biometrics* in order to alleviate the issue of the permanence of biometric data in the event of compromise. The transformation can be applied to the biometric data either in the signal domain (e.g., to the acquired fingerprint image) or the feature domain (where the biometric signal's extracted features are transformed). One example of a signal-level transform from [91], suggested for the face image, is *grid morphing*, as illustrated in Figure 30. The basic idea is to overlay a grid onto the face image, then distort the grid lines to *warp* the face image. More recently, face image warping was studied in [48]. Here, a user-specific warping function was defined in terms of a key, K, which specifies the size of the blocks into which the face image should be divided, the maximum amount by which the edges of the blocks should be randomly offset, and a seed which initialises the random edge offset. This process is illustrated in Figure 31, where we see that, the larger the image blocks and maximum edge offsets, the more distorted or warped the image becomes.

In [48], the warped images were input into a neural-network-based face feature extractor, and the resulting features were used for face recognition purposes. This study actually recommended that warping *not* be used as a

[3] The terms "query" and "probe" will be used interchangeably to refer to the biometric sample acquired during the authentication/recognition stage.

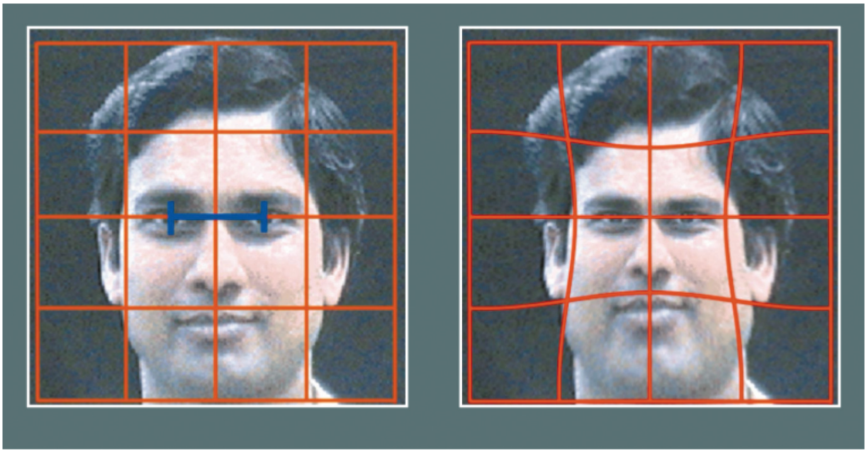


Figure 30: Grid morphing applied to a face image. Image from Ratha *et al.* [91].

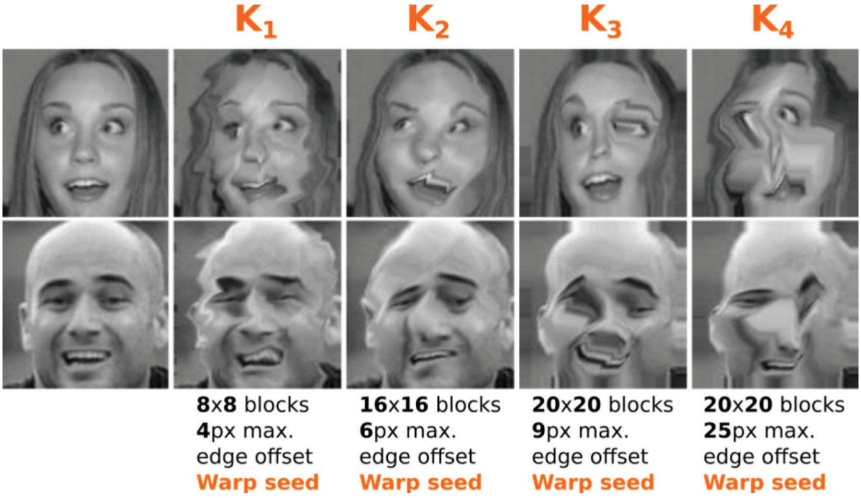


Figure 31: Warping a face image as per the method in [48]. The warping function is defined by a user-specific key, K, which specifies the size of the blocks into which the image should be divided, the maximum offset by which the block edges should be offset, and a seed which initialises the random edge offset. Image adapted from Krivokuća Hahn [53].

BTP method in practice, because for certain warping parameters, the warping function was found to be effectively ignored by the neural network as a form of intra-class variability [54]. More specifically, it was shown that, the greater the image distortion, the worse the recognition accuracy, whereas for lower distortions the warped images can actually be matched to their original counterparts. So, there is a trade-off between the "recognition accuracy" and "irreversibility" BTP criteria, which is typical of Feature Transformations in general. Furthermore, [48] also found that, for lower amounts of distortion, face images warped using different keys (warping parameters) can be *linked*, meaning that it may be difficult to satisfy the "renewability/unlinkability" criterion in practice.

Grid morphing, or warping, is an example of a signal-level (or image-level) Feature Transformation. Most Feature Transformations in the literature, however, are feature-level methods, meaning that they are applied to features extracted from the biometric signal instead of to the raw signal itself. One of the most well-known and widely-studied feature-level Feature Transformations, is *BioHashing*, which was originally proposed for protecting fingerprint features in [39]. In this approach, the first step is to generate a fingerprint feature vector from the acquired fingerprint image. The original BioHashing publication [39] proposes using a feature vector that is invariant to translation, rotation, and scaling (e.g., generated by applying the Wavelet Fourier-Mellin Transform (WFMT) to the fingerprint image[4]). The second step is to project the feature vector onto a randomly-generated, user-specific matrix, then binarise the projected vector to generate the protected template, which is referred to as a "BioHash".

More specifically, during enrollment, each user of the biometric system is presented with a secret seed, K, which is stored on an external device such as a USB token or a smart-card. This seed is used to generate a set of m pseudorandom vectors, $r_1, \dots, r_m$, which are orthonormalised (e.g., using the Gram-Schmidt orthonormalisation method). Then, the dot product between the orthonormal set of vectors, $\hat{r_1}, \dots, \hat{r_m}$, and the invariant biometric feature vector, x, is computed. The resulting vector is binarised to generate the protected biometric template, an m-bit code referred to as a "BioHash". The binarisation is performed using a pre-set threshold, τ, where 0 corresponds to a dot product that is less than or equal to τ, while 1 represents a dot product greater than τ. The threshold, τ, is selected based on the criterion that the expected number of zeros in the resulting BioHash is equal to the expected number of ones, to maximize the entropy of the protected template. Figure 32 illustrates the creation of a BioHash from an invariant biometric feature vector.

[4] Take the FFT of an image – the resulting spectral magnitude is translation invariant. Then define rotation and scale in terms of translation. Do this by first defining the spectral magnitude in terms of polar coordinates, to decouple rotation and scaling – rotation is now expressed in terms of translation. Reduce scaling to a translation by expressing the radial coordinate in terms of a logarithmic scale. The resulting image is now translation, rotation, and scale invariant. Flatten this image to produce the corresponding invariant feature vector.

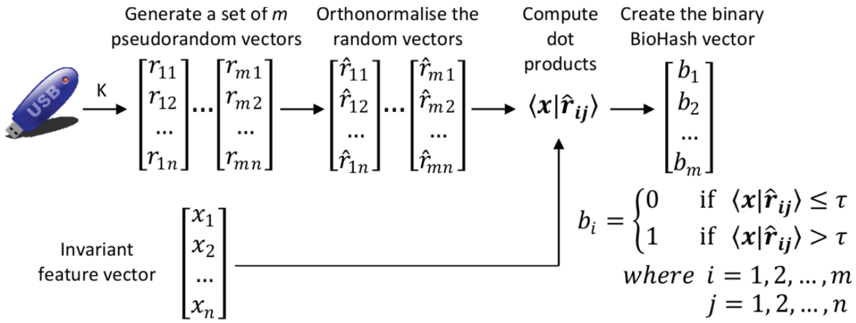

Generate a set of *m* pseudorandom vectors
Orthonormalise the random vectors
Compute dot products
Create the binary BioHash vector
USB
K
$\begin{bmatrix} r_{11} \\ r_{12} \\ \dots \\ r_{1n} \end{bmatrix} \dots \begin{bmatrix} r_{m1} \\ r_{m2} \\ \dots \\ r_{mn} \end{bmatrix}$
$\begin{bmatrix} \hat{r}_{11} \\ \hat{r}_{12} \\ \dots \\ \hat{r}_{1n} \end{bmatrix} \dots \begin{bmatrix} \hat{r}_{m1} \\ \hat{r}_{m2} \\ \dots \\ \hat{r}_{mn} \end{bmatrix}$
$\langle \boldsymbol{x} | \hat{\boldsymbol{r}}_{\boldsymbol{ij}} \rangle$
$\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$
Invariant feature vector
$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$
$b_i = \begin{cases} 0 & \text{if } \langle \boldsymbol{x} | \hat{\boldsymbol{r}}_{\boldsymbol{ij}} \rangle \leq \tau \\ 1 & \text{if } \langle \boldsymbol{x} | \hat{\boldsymbol{r}}_{\boldsymbol{ij}} \rangle > \tau \end{cases}$
$where\ i = 1, 2, \dots, m$
$j = 1, 2, \dots, n$

Figure 32: The creation of a user's BioHash (protected template) from an invariant (to translation, rotation and scaling) biometric feature vector. Image from Krivokuca [50].

During authentication, the invariant query biometric feature vector is transformed in the same fashion, and the resulting BioHashes are compared using Hamming distance. Variations of the BioHashing approach have been applied to several biometric modalities, including fingerprints (e.g., [7, 39]), face (e.g., [6, 82, 109–111], palm prints (e.g., [20, 88]), iris (e.g., [16]), and finger veins (e.g., [51, 83]).

The most important advantage of BioHashing is that, when each user of the biometric system uses their own K, it is possible to obtain an Equal Error Rate (EER) of zero when comparing biometric templates in the protected domain – so, BioHashing satisfies the "recognition accuracy" BTP criterion. However, it has been shown (e.g., [49]) that, in the scenario where an adversary gains access to a genuine user's K and applies it to their own biometric features to generate the corresponding BioHash (referred to as the "stolen-token" scenario), the resulting recognition accuracy is worse than that obtained when comparing the unprotected biometric templates. This means that a user's K must be kept secret. This key secrecy is also important considering that it has been shown (e.g., [15]) that a BioHash is relatively easy to invert to recover a close approximation of the original biometric template, when a user's K is known to an adversary – so, the ability of BioHashing to satisfy the "irreversibility" criterion usually depends on the secrecy of K. It has also been suggested [131] that the combination of different BioHashes of the same user can leak important information about the original biometric template – so, although "renewability" is technically possible by changing the user-specific K, the resulting BioHashes may not be entirely "unlinkable".

A more recent example of a feature-level Feature Transformation is the *PolyProtect* method proposed in [52]. PolyProtect was applied to the protection of face "embeddings", which are face features learned from face images using pre-trained deep neural networks. PolyProtect involves transforming face embeddings into protected templates via multivariate polynomials, which are parameterised by user-specific coefficients and exponents. More specifically, sets of m consecutive elements from the face embedding are passed to the user-specific polynomial in turn, where each embedding element constitutes one variable (hence the term “multivariate”), to generate the different elements of the protected face template. The amount of overlap between the sets of elements being passed to the polynomial can be varied, and the larger the overlap the larger the dimensionality of the protected template. Figure 33 illustrates the PolyProtect transformation from a 128-dimensional face embedding, V, to a protected template, P, when the polynomial consists of 5 variables (i.e., $m = 5$) and when the overlap is set to 4. Figure 34 depicts the PolyProtect transformation for the same face embedding, but for different amounts of overlap.

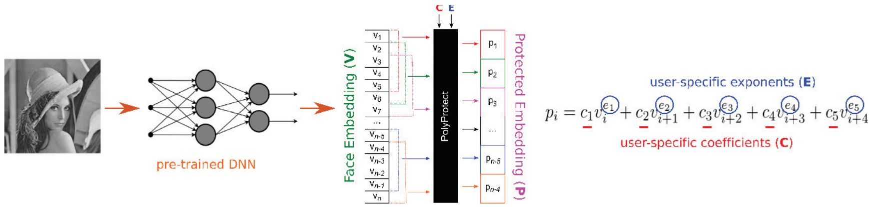

pre-trained DNN
Face Embedding (V)
v_1 v_2 v_3 v_4 v_5 v_6 v_7 … v_{n-5} v_{n-4} v_{n-3} v_{n-2} v_{n-1} v_n
C E
PolyProtect
p_1 p_2 p_3 … p_{n-5} p_{n-4}
Protected Embedding (P)
user-specific exponents (E)
$p_i = c_1 v_i^{e_1} + c_2 v_{i+1}^{e_2} + c_3 v_{i+2}^{e_3} + c_4 v_{i+3}^{e_4} + c_5 v_{i+4}^{e_5}$
user-specific coefficients (C)

Figure 33: Using PolyProtect to transform a 128-dimensional face embedding, V (learned from a face image by a deep neural network) to a protected template, P. The polynomials used in the transformation are parameterised by user-specific coefficients, C, and exponents, E, and they consist of 5 variables extracted from consecutive sets of 5 embedding elements.

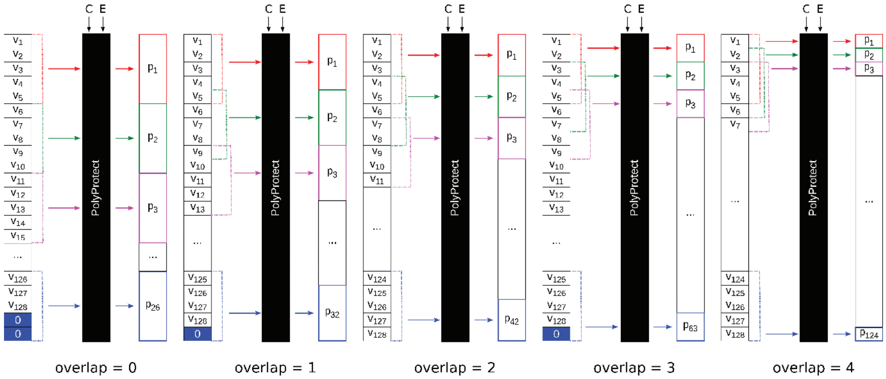

C E
PolyProtect
v_1 … v_{128}
0
p_1 p_2 p_3 … p_{26}
overlap = 0
p_1 p_2 p_3 … p_{32}
overlap = 1
p_1 p_2 p_3 … p_{42}
overlap = 2
p_1 p_2 p_3 … p_{63}
overlap = 3
p_1 p_2 p_3 … p_{124}
overlap = 4

Figure 34: The PolyProtect transformation from Figure 33, when the amount of overlap between the sets of embedding elements that are passed in turn to the polynomial, increases from 0 to 4. Image from Krivokuća Hahn and Marcel [52].

In a biometric recognition system that employs PolyProtect to protect its biometric template, PolyProtect would be applied to both the reference (generated during enrollment) and probe (generated during authentication) templates (e.g., embeddings). Then, the authentication stage would take place in the protected domain, whereby the protected reference and probe templates would be compared using the cosine distance. Note that, although [52] applies PolyProtect to face embeddings, the nature of this BTP algorithm suggests that it could be suitable for protecting any real-number, ordered biometric template.

As is typical of Feature Transformation BTP methods, PolyProtect was shown [52] to exhibit a trade-off between the "recognition accuracy" and "irreversibility" BTP criteria; however, [52] demonstrated that this trade-off can be balanced in practice by tuning the "overlap" parameter adopted for the PolyProtect transformation. The "irreversibility" of PolyProtect was evaluated based on the assumption of a fully-informed attacker, who has access to the entire PolyProtect algorithm, including the user-specific polynomial coefficients and exponents. Even in this worst-case scenario, it was shown [52] that PolyProtected templates generated using overlaps of 0-2 are practically irreversible and those generated using an overlap of 3 are only partially reversible, whereas protected templates generated using the maximum overlap of 4 were found to be almost fully reversible (so it was recommended that the highest overlap be avoided practice). Finally, [52] showed that it is possible to achieve effectively full "unlinkability" between multiple PolyProtected templates from the same person's face embeddings by changing the user-specific coefficients and exponents used to parameterise the transformation polynomials. This enables the renewal of compromised protected templates and the use of different protected templates from the same person in different applications without the risk of cross-matching.

14.1.2 Biometric Cryptosystems

A Biometric Cryptosystem incorporates ideas from traditional cryptographic protection schemes with biometrics. In fact, the initial motivation was to either use biometric features to secure a cryptographic key or to directly generate a cryptographic key from the biometric features themselves. This led to the two sub-categories of Biometric Cryptosystems, namely *Key-binding* and *Key-generating* systems. Key-binding techniques involve "binding" an external, randomly-generated key with the biometric template, while Key-generating methods try to extract a unique key from the biometric template itself. In the BTP literature, Key-binding Biometric Cryptosystems are the more common approach. The two most well-known Key-binding Biometric Cryptosystems are the Fuzzy Commitment scheme [44] and the Fuzzy Vault scheme [43], both of which have been applied to a number of different biometric modalities (e.g., fingerprints [77, 79, 80, 108, 112, 124], face [29, 76, 92], iris [58, 94], signature [26, 66]).

In the Fuzzy Commitment scheme, a randomly-generated, user-specific codeword is "bound" with the user's biometric template, by calculating the difference between the two entities (e.g., in terms of subtraction or the XOR operation). The binding calculated from the reference template and chosen codeword during enrollment constitutes the protected reference template, which is stored in the template database along with a cryptographic hash of the codeword. During authentication, the aim is to "unbind" the protected reference template to recover this codeword, which should be possible provided that the probe template is sufficiently similar to the reference template (to which the codeword was bound). If this is the case, then an appropriate error-correction mechanism should be able to correct any errors in the recovered codeword, such that the cryptographic hash of the recovered codeword matches the hash of the reference codeword stored in the biometric system's database. A match between the two hashes would indicate a successful authentication attempt. Figure 35 illustrates the enrollment (referred to as "commitment") and recognition (referred to as "decommitment") stages in a typical Fuzzy Commitment scheme.

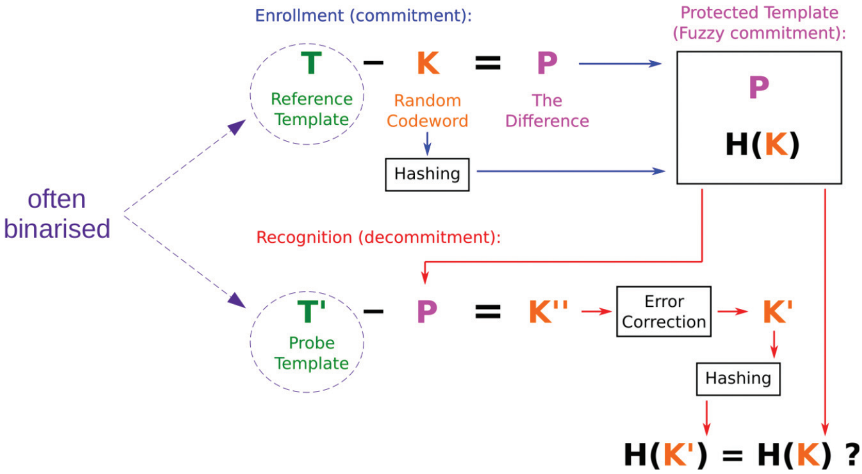


Figure 35: The enrollment ("commitment") and recognition ("decommitment") stages in a biometric system whose biometric templates are protected by the Fuzzy Commitment scheme. Image from Krivokuća Hahn [53].

One of the main advantages of the Fuzzy Commitment scheme is that the cryptographic hash function, which is used to secure the user-specific codeword, should be non-invertible, meaning that it should be practically impossible to recover this codeword from its hash (i.e., K from $H(K)$ in Figure 35). Without access to the codeword, it should be very difficult for an attacker to recover the reference template that is bound with this codeword. In other words, the binding between the biometric template and the codeword is meant to protect *both* of these entities, since it should be impossible to recover one without access to the other. From this standpoint, the Fuzzy Commitment scheme would be considered to satisfy the "irreversibility" BTP criterion to a high degree. However, it has recently been shown [45, 46] that the binding between a biometric template and a random codeword could be undone in a way that is easier than by cracking the cryptographic hash. Specifically, [45, 46] demonstrated that it may be possible for an attacker to guess a close enough approximation of the reference biometric template, which would allow them to undo the binding to recover a close enough approximation of the bound codeword. If the hashed version of this codeword matches the reference hash stored in the biometric system's database, then the attacker knows that the unbinding attempt has succeeded.

In this case, binding the same biometric template with a different codeword would not help, since the attacker already has access to a close approximation of the biometric template and can thus use it to unbind any new binding – for this reason, it is usually difficult for the Fuzzy Commitment scheme to satisfy the "renewability/unlinkability" BTP criterion. Finally, it should be noted that the recognition accuracy in the protected domain generally relies on the error-correction step, which is limited in its capability to correct errors in the recovered codeword. Since cryptographic hash functions are extremely sensitive to small changes in the input, such that even a tiny difference would result in a completely different hash, this means that the comparison between the hash of the recovered codeword and the hash of the reference codeword would fail unless the two codewords are completely identical. For this reason, biometric systems protected via the Fuzzy Commitment scheme may suffer from a high False Reject Rate (FRR), meaning that the ability of the Fuzzy Commitment scheme to satisfy the "recognition accuracy" criterion may be limited in practice.

The Fuzzy Vault scheme is similar to the Fuzzy Commitment scheme in that it also involves binding an external key with a biometric template, except this binding is performed in a different way, by projecting the template onto a polynomial. More specifically, during enrollment a randomly-generated, user-specific (secret) key is used to define the coefficients of a polynomial, P. This polynomial is then evaluated at each element of the biometric template, T (i.e., the template elements are treated as distinct x-coordinate values), to generate a set of "true points" (i.e., points that lie on the polynomial). Finally, some noise is added in the form of "chaff points", which are random points that do not lie on the polynomial, in order to hide the polynomial from an attacker. The final set of points (i.e., the true points plus the chaff points) constitutes the fuzzy vault, V, for this particular user. The resulting fuzzy vault is stored in the database as the protected template (or "binding"), along with a cryptographic hash of the user-specific key (which defines the polynomial's coefficients). During the recognition stage, the probe biometric template, T', is presented to the biometric system. Note that pre-alignment of the enrolled biometric template, T, and the probe biometric template, T', is assumed. If the elements of T' are sufficiently similar to the elements of T (regardless of their ordering), then T' can be used to reconstruct the polynomial, P, because we would be able to use the elements of T' to identify enough true points that lie on the polynomial. In this case, the recovered key, K' (i.e., the coefficients of the reconstructed polynomial) should be close enough to the secret key, K, such that an appropriate error-correction mechanism applied to K' should ensure that the cryptographic hash of K' matches the hash of K stored in the system's database. A correct match would indicate a successful authentication attempt. Figure 36 illustrates the enrollment (referred to as "locking the vault") and recognition (referred to as "unlocking the vault") stages in a typical Fuzzy Vault scheme.

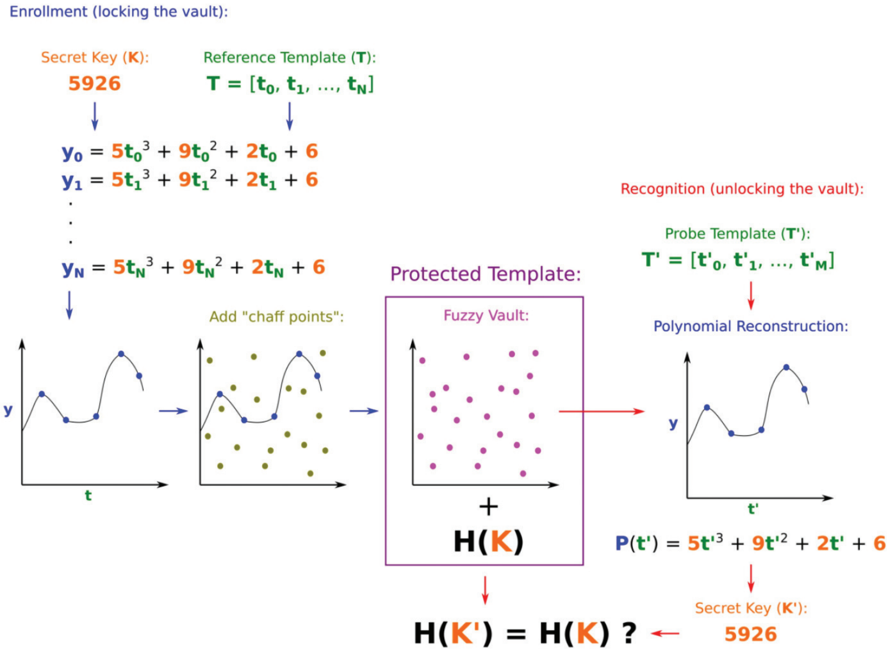

Enrollment (locking the vault):
Secret Key (K):
5926
Reference Template (T):
T = [t0, t1, ..., tN]
y0 = 5t0^3 + 9t0^2 + 2t0 + 6
y1 = 5t1^3 + 9t1^2 + 2t1 + 6
yN = 5tN^3 + 9tN^2 + 2tN + 6
y
t
Add "chaff points":
Protected Template:
Fuzzy Vault:
+
H(K)
Recognition (unlocking the vault):
Probe Template (T'):
T' = [t'0, t'1, ..., t'M]
Polynomial Reconstruction:
y
t'
P(t') = 5t'^3 + 9t'^2 + 2t' + 6
Secret Key (K'):
5926
H(K') = H(K) ?

Figure 36: The enrollment ("locking the vault") and recognition ("unlocking the vault") stages in a biometric system whose biometric templates are protected by the Fuzzy Vault scheme. To simplify the diagram, the error-correction process for K' is not explicitly shown. Image from Krivokuća Hahn [53].

As for the Fuzzy Commitment scheme, one of the main advantages of the Fuzzy Vault scheme is that it should be impossible to recover the secret key, K, from its cryptographic hash, $H(K)$. Without access to K, it should be difficult to separate the “true points” from the “chaff points” in the stored fuzzy vault, and consequently it should be very difficult to reconstruct the secret polynomial. Furthermore, in general the greater the number of chaff points, the more concealed the polynomial will be, since more “spurious” polynomials will appear to exist. So, we can effectively strengthen the binding between the key and the biometric template by adding more chaff points. This means that, if the polynomial is considered well-hidden, then the Fuzzy Vault scheme would be assumed to satisfy the “irreversibility” BTP criterion. However, as for the Fuzzy Commitment scheme, the “renewability/unlinkability” criterion would generally be difficult to satisfy for the Fuzzy Vault scheme. This is because, even if the same biometric template is bound with different polynomials, the x-coordinates of the “true points” in each of the resulting fuzzy vaults would be the same (since they come from the elements of the same biometric template) – so, different fuzzy vaults from the same person could be correlated to reveal the underlying biometric template (e.g., [47, 99]). Finally, similarly to the Fuzzy Commitment scheme, the recognition accuracy of the Fuzzy Vault scheme would depend on the ability to correct errors in the recovered key, K' during authentication. In particular, recognition accuracy may be adversely affected if the number of true points identified during recognition is too low for polynomial reconstruction, or if too many chaff points are added in close proximity to the true points (which may result in incorrect polynomial reconstruction). So, the ability of the Fuzzy Vault scheme to satisfy the “recognition accuracy” criterion would depend on its implementation.

Although Key-binding techniques, such as Fuzzy Commitment and Fuzzy Vault, appear to be the more common type of Biometric Cryptosystem considered in the literature, Key-generating methods have also been studied. In a Key-generating system, the idea is to generate a key directly from the biometric data, rather than binding an existing key with the biometric template as in Key-binding systems. In this case, the generated key represents the protected biometric template. One of the most popular Key-generating approaches involves the use of quantisation boundaries to help generate a stable key from the biometric template (e.g., see the survey in [93]). Figure 37 and Figure 38 show a high-level illustration of the enrollment and recognition stages, respectively, for a Key-generating method based on quantisation boundaries.

During enrollment (Figure 37) a biometric template would be quantised using user-specific quantisation boundaries, and the quantised template would be binarised to generate the biometric key, P, which would then be cryptographically hashed and stored in the biometric system’s protected template database along with the employed quantisation boundaries. Then, during the recognition stage (Figure 38), the user-specific quantisation boundaries would be retrieved from the database and used to quantise the probe biometric template, which would subsequently be binarised and corrected using an appropriate error-correction mechanism to generate the final probe biometric key, P'.

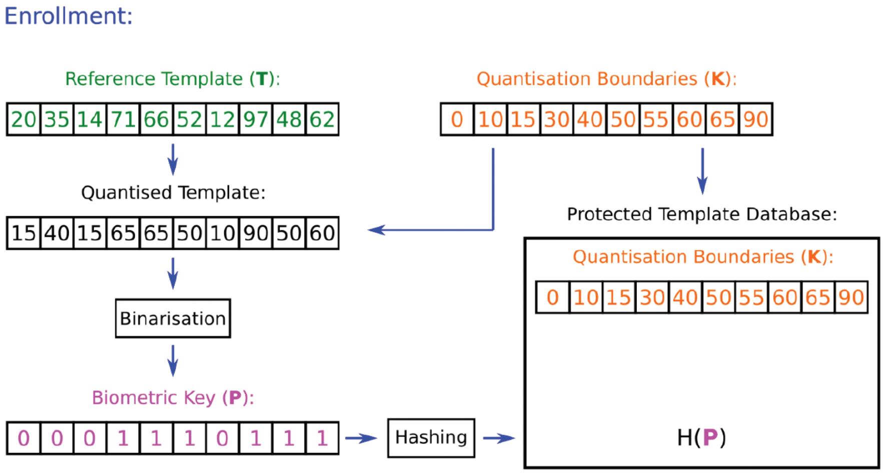


Figure 37: The enrollment stage of a Key-generating Biometric Cryptosystem based on the use of quantisation boundaries.

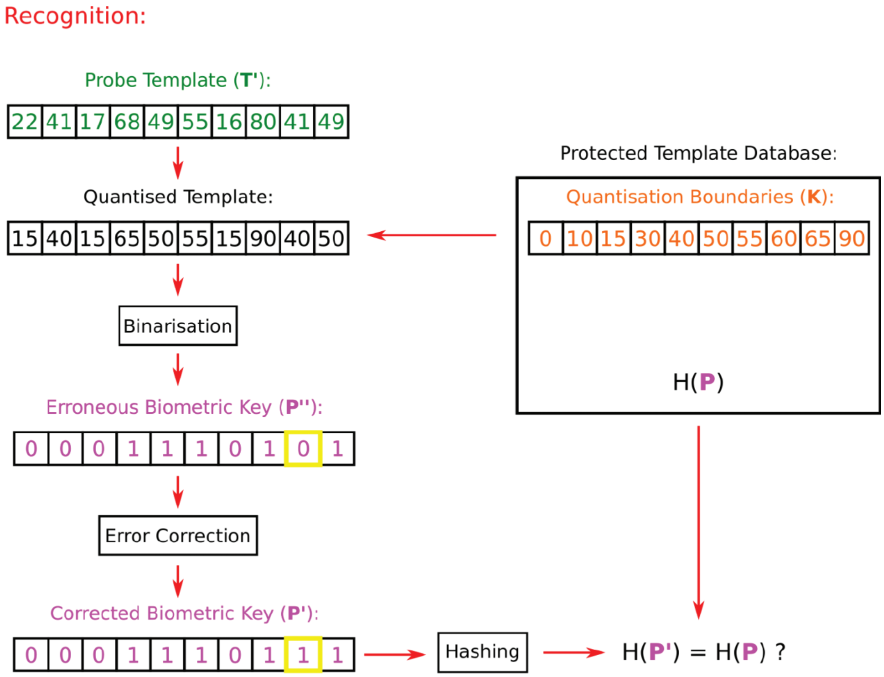


Figure 38: The recognition stage of a Key-generating Biometric Cryptosystem based on the use of quantisation boundaries.

The cryptographic hash of P' would then be compared to the cryptographic hash of P stored in the database, and a match between the two hashes would indicate a successful authentication attempt.

One of the biggest issues with Key-generating methods is the challenge of generating a biometric key that simultaneously possesses high stability and high entropy. A highly stable key with zero entropy would mean that the same key is generated regardless of the input biometric template, which would likely lead to a high False Accept Rate (FAR), whereas a key possessing high entropy but no stability would result from a scheme that generates a different key for each variation of the same user's biometric template, which would likely lead to a high False Reject Rate [50]. For this reason, it may be difficult for Key-generating Biometric Cryptosystems to satisfy the "recognition accuracy" BTP criterion in practice. Furthermore, similar to Key-binding schemes, the ability of Key-generating methods to fulfil the "renewability/unlinkability" criterion is not evident in practice. For example, using different quantisation boundaries for the same person may result in the generation of different biometric keys, but the extent of these differences may be insufficient for the different keys to be considered "unlinkable". As for the "irreversibility" criterion, Key-generating systems may be considered highly irreversible from the point of view of the difficulty of breaking the cryptographic hash that typically protects the extracted key. However, if the helper data (e.g., quantisation boundaries) is stolen from the system's database, it may reveal information that could allow an attacker to guess a close approximation of the underlying biometric template, which may then enable them to guess a close enough approximation of the generated biometric key. For these reasons, it is generally a good idea to combine Key-generating Biometric Cryptosystems with other BTP methods, such as Feature Transformations [50, 93] – in fact, such hybrid approaches may be recommended for Key-binding Biometric Cryptosystems as well. Several Feature Transformation plus Biometric Cryptosystem hybrid BTP methods have already been proposed, for example: hardening a fingerprint-based fuzzy vault with a user-specific password [81]; applying a cryptographic one-way hash function to a face template protected using a BioHashing-like approach [24]; generating irrevocable keys from cancellable fingerprint templates [56].

14.2 Learned BTP Methods

Learned BTP methods refer to algorithms that are learned by a (deep) neural network.[5] This is in contrast to Handcrafted BTP algorithms, which are explicitly formulated by humans. As specified in [54], existing Learned BTP

[5] Although BTP algorithms could also be learned using non-neural-network machine learning techniques, in this paper we refer only to neural-network-based learning as this seems to be the current trend in the literature.

methods focus on two main approaches.[6] The first approach involves training a neural network to learn the mapping from a person's biometric template to a *pre-defined*, randomly-generated code, which represents the person's protected template. The second approach involves training a neural network to learn its *own representation* of a protected template, without forcing it to conform to a pre-defined representation. Sections 14.2.1 and 14.2.2 present examples of methods that fall into these two categories, respectively. Note that, thus far, it appears that Learned BTP methods have been mainly investigated for the protection of face templates, which may be attributed to the current popularity of neural-network-based face recognition systems. A comprehensive survey of these techniques is presented in [54], so Sections 14.2.1 and 14.2.2 present an overview of a few selected methods to explain the two main categories of approaches. Although these methods will be explained in the context of protecting face templates, the concepts could be extended to other biometric modalities as well.

14.2.1 Learning Mapping to Pre-defined Protected Template

The best way to explain this type of BTP approach is in terms of the most well-known method in this category: Maximum Entropy Binary (MEB) codes [86, 87]. In fact, other methods that learn a mapping to a pre-defined protected template tend to build upon the MEB codes method. This method was proposed for the face modality (although it could be generalised to other biometric modalities as well), and it starts by assigning a random, maximum-entropy, binary code to every user that is to be enrolled in a particular face recognition system. Then, a Convolutional Neural Network (CNN) is trained to map each user's face image to their corresponding code. During training, the input to the CNN is a face image, and the output is a set of n floating-point numbers produced by n sigmoid activation functions. Each of these outputs is compared to the corresponding bit in the pre-defined n-bit code in terms of a binary cross-entropy loss, and the training continues until the n-valued output is as close as possible to the n-bit MEB code (i.e., the loss is reduced as much as possible) for each user of the face recognition system. This would indicate that the system has learned how to map each user's face image to their corresponding MEB code. Once training is finished, the MEB codes are cryptographically hashed, and the resulting hashes are stored in the system's database as the protected face templates of their users. Then, during the recognition stage, the input face image is passed through the trained CNN to once again produce the n sigmoid outputs, which are now binarised (by setting values above 0.5 to 1 and the rest to 0) to generate the n-bit MEB

[6] Although [54] focuses specifically on *face* template protection, the BTP method categorisations seem sufficiently generalisable to all biometric modalities.

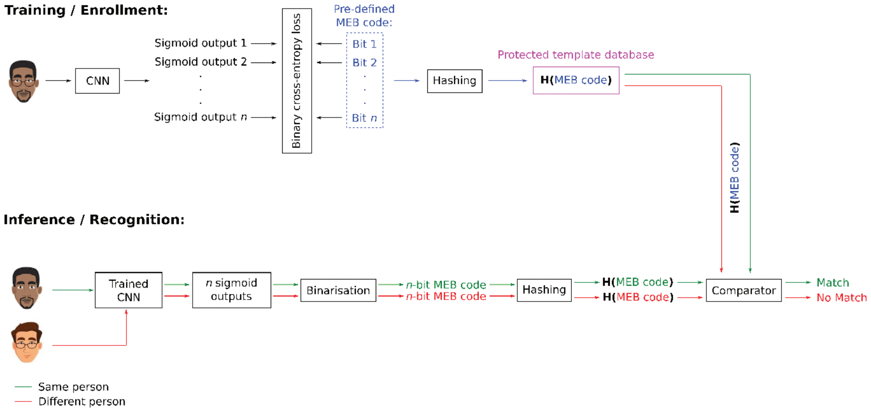


Figure 39: The training/enrollment and inference/recognition stages in a face recognition system that learns protected face templates using the MEB codes BTP method. Image from Krivokuća Hahn [53].

code. This code is then cryptographically hashed and compared to the same user's hash stored in the database. If the generated MEB code is the same as the code assigned to this user during enrollment, then the hashes should match, but if the codes are different or they come from different users then the hashes will not match. Figure 39 illustrates the training/enrollment and inference/recognition stages in a face recognition system protected by the MEB codes BTP method.

One of the most important advantages of the MEB codes method is that the cryptographic hash function, which is used to secure the pre-assigned MEB codes, is non-invertible, which means that it should be practically impossible to recover an MEB code from its hash. Even if an MEB code is somehow recovered, it does not reveal details about the face to which it was assigned since the code was randomly generated (so it is fundamentally unrelated to the assigned face identity). We may conclude, therefore, that the MEB codes method is able to satisfy the “irreversibility” property.[7] Furthermore, this method could technically satisfy the “renewability/unlinkability” criterion by assigning a new MEB code to the same user. However, the main issue with this approach is that each new MEB code would necessitate the re-training (either full or partial) of the neural network (e.g., each time a new user wishes to enroll into the face recognition system or when a compromised user needs

[7] However, as pointed out in [54], we may imagine the scenario where an attacker with full access to the trained CNN could use this information to gain some insight into how the neural network learns a mapping between a given face image and a pre-assigned binary code. This could potentially be exploited to recover information about the faces of the enrolled subjects in different layers of the trained neural network.

to be re-enrolled with a new protected template). So, the scalability of such methods in practice is questionable [54]. Finally, since the trained CNN must be able to generate exactly the same MEB code for the same user during each recognition attempt (otherwise the hash comparison will fail), this type of BTP approach may result in quite high False Reject Rates (FRR) in practice – so, the ability of the MEB codes method to satisfy the "recognition accuracy" BTP criterion may be difficult.

The issue of improving the recognition accuracy of face recognition systems protected by the MEB codes method, has been a point of interest in the literature, with several improvements having being proposed. One example of a suggested improvement [41] is to first use a pre-trained deep neural network to extract face templates from the face images, then start the training/mapping from these templates instead of from the images as in [86, 87]. The idea here is to take advantage of robust feature extractors, which could help to improve the ability of the neural network to map a given face image to its correct MEB code. Another example of a proposed improvement [42], which actually builds upon the approach in [41], is to additionally incorporate some user-specific randomness into the extracted face templates before mapping them to the users' pre-defined MEB codes. Specifically, [42] suggests projecting a user's face template onto a random subspace defined by a user-specific matrix, which would help separate different users even further during the mapping stage, thereby increasing the recognition accuracy of the protected face recognition system.

14.2.2 Learning Own Representation of Protected Template

As discussed in Section 14.2.1, the main issue with approaches that learn a mapping from a face image/template to a *pre-defined* code (e.g., MEB codes) is that the neural network would need to be re-trained for each new enrollment (either for a new user of the biometric recognition system, or for existing users whose codes have been compromised and need to be replaced). So, to avoid network re-training and thus improve the scalability of the BTP method in practice, an alternative approach to Learned BTP methods is to train a neural network to learn its *own representation* of a protected template (instead of forcing it to conform to a pre-defined representation). However, since the neural network would be trained to learn the *same* (or similar) representation of a protected template for each instance of the same person's face, renewability of compromised protected templates would only be possible with the incorporation of some user-specific randomness into the learning process. A couple of examples of Learned BTP methods that fall into this category, are Deep IoM Hashing [21] and SecureFace [65].

The Deep IoM Hashing method [21] emulates the Handcrafted Index-of-Max (IoM) hashing method [40]. The Handcrafted approach, illustrated in

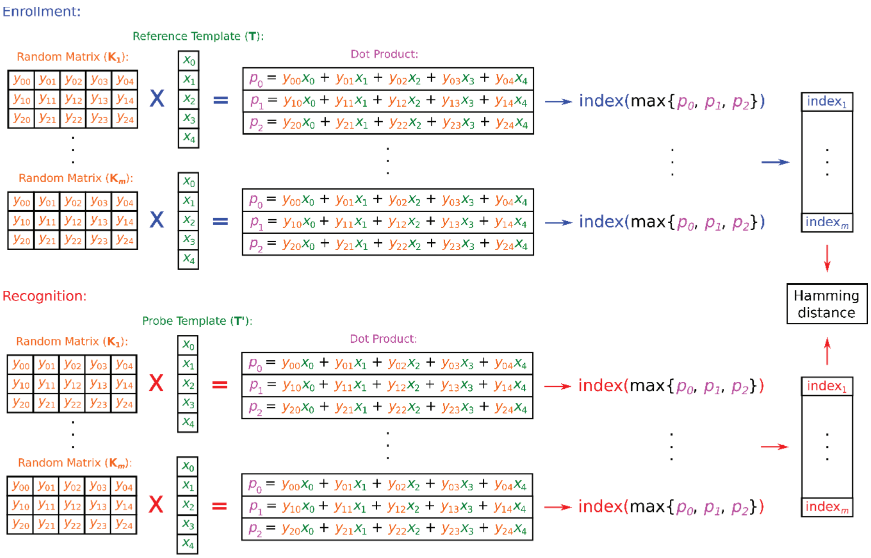


Figure 40: The enrollment and recognition stages for biometric templates protected using the Handcrafted IoM Hashing method. Image from Krivokuća Hahn [53].

Figure 40, is similar to the BioHashing method discussed in Section 14.1, except that the user-specific projection of the biometric template (or feature vector) is based on multiple projection matrices, instead of only one as in BioHashing. In particular, during enrollment a set of m random matrices is generated for a user of the biometric system, and their biometric template is then projected onto each of these matrices in turn via the dot product operation. The maximum value in the vector resulting from each projection operation is then located, and the index of this value is recorded. The set of all m maximum value indices constitutes the user's "hash", which corresponds to the protected reference template. During the recognition stage, the process is repeated to generate a hash from the probe biometric template. Finally, the probe IoM hash is compared to the reference IoM hash (stored in the system's database) in terms of Hamming distance: the smaller the distance, the more likely it is that the two hashes correspond to the same biometric identity.

In the Handcrafted IoM Hashing method, the user-specific matrices are randomly generated during the enrollment stage. In contrast, for the Deep IoM Hashing method, these matrices are *learned* by a neural network, and the "user-specific randomness" comes from a permutation step. Figure 41 illustrates the process by which a Deep IoM hash is learned for a particular face image. In particular, a face template is first learned from a face image, using a pre-trained deep neural network (DNN). This template is next permuted via a

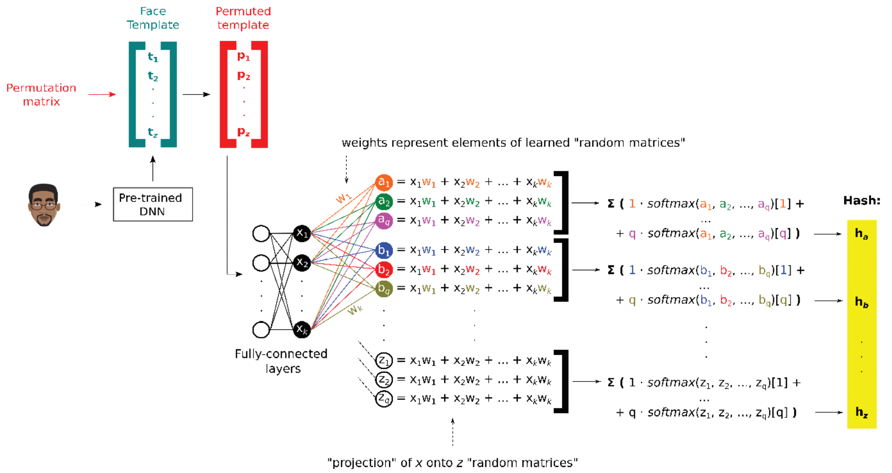

Face Template
Permuted template
Permutation matrix
t₁
t₂
t_z
p₁
p₂
p_z
Pre-trained DNN
weights represent elements of learned "random matrices"
w₁
w_k
x₁
x₂
x_k
Fully-connected layers
a₁ = x₁w₁ + x₂w₂ + ... + x_kw_k
a₂ = x₁w₁ + x₂w₂ + ... + x_kw_k
a_q = x₁w₁ + x₂w₂ + ... + x_kw_k
b₁ = x₁w₁ + x₂w₂ + ... + x_kw_k
b₂ = x₁w₁ + x₂w₂ + ... + x_kw_k
b_q = x₁w₁ + x₂w₂ + ... + x_kw_k
z₁ = x₁w₁ + x₂w₂ + ... + x_kw_k
z₂ = x₁w₁ + x₂w₂ + ... + x_kw_k
z_q = x₁w₁ + x₂w₂ + ... + x_kw_k
Σ (1 · softmax(a₁, a₂, ..., a_q)[1] + ... + q · softmax(a₁, a₂, ..., a_q)[q])
Σ (1 · softmax(b₁, b₂, ..., b_q)[1] + ... + q · softmax(b₁, b₂, ..., b_q)[q])
Σ (1 · softmax(z₁, z₂, ..., z_q)[1] + ... + q · softmax(z₁, z₂, ..., z_q)[q])
Hash:
h_a
h_b
h_z
"projection" of x onto z "random matrices"

Figure 41: The generation of a hash from a person's face image, using the Learned Deep IoM Hashing method. Image from Krivokuća Hahn [53].

randomly-generated, user-specific permutation matrix. The permuted template is then passed on to two fully-connected neuron layers, each consisting of k neurons. Each neuron in the second layer, $x_1, \ldots, x_k$, is then connected to z sets of q neurons each. We may thus visualise $x_1, \ldots, x_k$ as the elements of a "template" that is "projected" onto z "random matrices". For example, the weights connecting $x_1, \ldots, x_k$ to $a_1, \ldots, a_q$ represent the elements of matrix a. So, applying these weights to the neurons in $x_1, \ldots, x_k$ is equivalent to performing the dot product between $x_1, \ldots, x_k$ and these "matrix elements", where $a_1, \ldots, a_q$ corresponds to the output of this operation. Consequently, this operation may be seen as the "projection" of $x_1, \ldots, x_k$ onto matrix a. Similarly, the dot product between $x_1, \ldots, x_k$ and the weights connecting these neurons to $b_1, \ldots, b_q$ can be seen as the projection onto matrix b, and so on up to matrix z. Finally, a softmax operation is applied to the set of q outputs of each "matrix projection" to find the index of the maximum value. This is done by applying softmax to each of the q rows separately and multiplying by the row index, then summing the outputs, which will return the approximate maximum-value index. The set of maximum-value indices from all z "matrices" is then concatenated to generate the user's hash, which corresponds to the protected face template.

Unlike for the Handcrafted IoM Hashing method, where the projection matrices are randomly-generated and user-specific, the projection matrices in the Deep IoM Hashing method are *learned* by a neural network and the learned weights, which represent the elements of the random matrices, are the *same* for all users of the biometric system. For this reason, the user-specific permutation step prior to the matrix projection is necessary, in order to incorporate some user-specific randomness into the matrix projections. In fact, the Deep IoM Hashing method should be able to satisfy the "renewability" criterion by changing the user-specific permutation matrix; however, there does not seem to exist any analysis on how "unlinkable" the resulting hashes would be. Considering the "recognition accuracy" criterion, it was demonstrated [21] that Deep IoM Hashing may be able to achieve slightly better accuracy than the Handcrafted IoM Hashing method, which was attributed to the fact that the Deep method was trained on data as opposed to being defined by a human in a data-agnostic manner. Furthermore, one important advantage of the Deep IoM Hashing method is that, unlike the MEB codes method discussed in Section 14.2.1, it seems generalisable to users that were not seen during the training process; in fact, in [21] this method was trained and tested on different subjects, which suggests good generalisability. This generalisability comes, in large part, from the fact that training is based on a *pairwise* loss function, which aims to maximise the similarity of hashes from the same person and minimise the similarity of hashes from different subjects – so, training is not linked to specific users. Finally, the ability of Deep IoM Hashing to satisfy the "irreversibility" criterion in practice is currently not clear, since, to the best of our knowledge, there is no analysis evaluating the irreversibility of the hash codes under the assumption that the learned "projection matrix" weights and the user-specific permutation matrices are known to an adversary. This may be considered analogous to knowing the projection matrices in Handcrafted IoM Hashing, which could indicate that at least a partial inverse of the generated hashes may be possible [53].

Another example of a Learned BTP method that learns its own representation of a protected template, is SecureFace [65]. This method trains a "randomized CNN" in an end-to-end fashion (starting from the face image and ending up with a protected face template), incorporating user-specific randomness in several different stages of the process [54]. The most important component, which sets this method apart from other Learned BTP methods, is a neural network referred to as RandNet, in which randomly-selected neurons are activated or deactivated to change the network's architecture. In particular, we start with a "standard" neural network, whose architecture is modified in a user-specific way (via a user-specific randomisation key), such that the resulting network becomes "user-specific" – this is illustrated in Figure 42. Consequently, different users' face templates will be passed through different RandNet networks to generate their corresponding protected templates.

The training of RandNet in SecureFace is based on using a pairwise loss function,[8] which aims to generate *similar* protected templates from the *same*

[8] Another loss function is used to train the feature extraction network, before the RandNet stage.

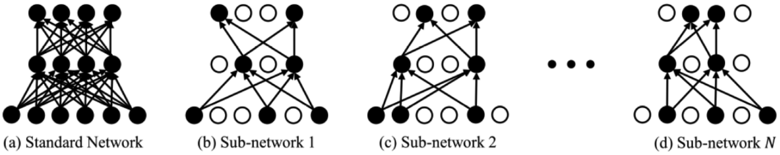


Figure 42: In the SecureFace BTP method, the architecture of a standard neural network is modified by randomly activating and deactivating certain neurons (black and white circles, respectively), to create user-specific sub-networks. Image from Mai *et al.* [65].

subject with the *same* randomisation key, and *dissimilar* protected templates from *different* subjects or the *same* subject with *different* randomisation keys. Furthermore, the method was trained and tested on different datasets [65], which suggests that this method has good generalisability to subjects unseen during training. So, if the method is well-trained, it should be able to satisfy both the "recognition accuracy" and "renewability/unlinkability" BTP criteria in practice (however, the scalability of this effort until the "standard" neural network needs to be changed, has not been evaluated). The "irreversibility" of this method was largely evaluated under the assumption that the user-specific randomisation key would not be leaked to an attacker [65]; however, to the best of our knowledge, there is no analysis for the worst-case scenario where this key might be leaked during the recognition stage, especially in the case of a fully-informed attacker with access to the trained neural network (including its learned parameters) and the various user-specific information that is employed to generate the protected template.

14.3 Handcrafted versus Learned BTP Methods

As noted in [54], thus far Handcrafted BTP methods have been the more popular approach in the literature. This makes sense, since Handcrafted methods include the traditional BTP techniques that have been studied for much longer than Learned BTP methods. Furthermore, compared to Learned BTP methods, Handcrafted methods tend to be more flexible for integration into existing biometric recognition systems, since they can often simply be added after an existing feature extractor. On the contrary, for Learned BTP methods, the neural network that is responsible for learning the BTP algorithm must usually be trained in the context of a specific biometric system, especially if it is designed to be trained in an end-to-end manner. Another advantage of Handcrafted BTP techniques is that they tend to be easier to understand and, therefore, evaluate, since they are explicitly formulated by humans. This is in contrast to Learned BTP methods, for which we generally do not have a complete understanding of how they work, since the BTP algorithms are learned by neural networks.

Although Handcrafted methods are currently more popular and have a number of advantages over Learned BTP methods (as discussed above), interest in Learned techniques is steadily growing. This is mainly due to the potential of these methods to incorporate a higher level of complexity (non-linearity) into the learned BTP algorithm, which may allow us to generate more robust (e.g., more irreversible) protected templates. In comparison, the complexity of Handcrafted BTP methods is limited by the capabilities of the human designer. However, the higher complexity of Learned BTP algorithms may come at the cost of too much ignorance about how the method works – indeed, this is the trade-off for relinquishing the algorithm design to a neural network.

Overall, therefore, whether a Handcrafted or Learned BTP method would be more suitable in practice, would depend on many factors, such as the type of biometric system into which the BTP method must be integrated, resource constraints, the required level of template protection, etc. In general, however, as suggested in [54], we might consider combining the two types of approaches to build on their individual strengths and reduce their respective weaknesses. This may allow us to balance the potentially higher complexity of Learned BTP methods with the more precise algorithm definition of Handcrafted methods, which in turn may help us generate more robust protected templates using a method that we are able to understand and whose efficacy can, therefore, be clearly evaluated.

15 Deep Learning Approach to Developing Human Biometric Systems

Deep learning has been increasingly applied to biometric systems to improve accuracy and reliability, especially after its success recent in pattern recognition problems. This has been supported by the availability of large databases with biometric data and computing systems with higher processing power.

A biometric system comprises of three important stages - acquisition, segmentation and matching. Deep learning has found its place in each of these stages. In the acquisition stage the biometric modality is captured and deep learning can be used. For example, if the acquired modality is an image, there are multiple deep learning algorithms that can help improve the quality of the acquired image, rectify the orientation of the image and highlight essential features, thereby eliminating noise. In segmentation stage, deep learning methods have found their place in tasks such as edge enhancement, line tracking and feature enhancement. Lastly, in the matching stage, deep learning - based networks are capable of examining the patterns in the biometric modality and make relevant decisions for the biometric system.

This has resulted in deep learning becoming a preferred choice in biometric systems. Some examples include using deep learning methods for palm vein segmentation, wrist vein segmentation, palm vein matching [70, 73], facial recognition, voice recognition, iris recognition and other behavioral biometrics.

Overall, deep learning has the potential to significantly improve the accuracy and reliability of biometric systems, which can have important applications in security, law enforcement, and other areas where identification is critical.

In the next section, we will discuss in detail the application of specific deep learning algorithms namely, U-net for segmentation and Siamese neural network for matching. As an example, we consider hand veins as the biometric modality of choice, although these concepts can accordingly be extended to other biometric modalities.

15.1 Deep Learning for Feature Extraction - Biometric Image Segmentation with U-Net

Feature extraction in biometric systems includes extracting relevant features from the captured biometric modality. To train deep learning models to capture features effectively, large databases where the features are marked (manually or automatically) are essential. These large databases with features marked are often referred to as gold-standard images. Considering blood vessel segmentation, it has been a topic of great interest in medical image segmentation problems as it helps triage diseases and identify the appropriate prognosis. In literature, vein segmentation systems are mostly applied to retinal vessel segmentation, brain vessels, and other medical scan images [33]. The purpose of this section in this article is to highlight one powerful deep learning-based biometric feature segmentation method that can be extended to other biometric modalities effectively. Before going into further details of the deep learning method U-Net, a few conventional methods are investigated here for the sake of completeness.

As already seen, there are many state-of-the-art vein recognition methods that involve steps of vein image recognition, feature extraction, and matching. In the world of biometrics, the most used methods are based on shape or texture. Shape-based methods focus on the structures of the vessels, extracting patterns like bifurcations and endpoints. These patterns can be picked using algorithms like local binary patterns (LBP) [132], biometric graph matching (BGM) [55], and width skeleton model (WSM) [61]. In [132], global and local shape representations are combined using cross-sectional profiles of the veins followed by Gaussian matched filter and skeletonization. Partition LBP (PLBP) [60] is a method where images are divided into sub-regions and partial LBPs are computed to extract desired features. This is then followed by the comparison of graph-like templates. The complete process encompasses of template registration and the graph pattern comparison with the distance computation between the probe and target image.

Texture descriptors such as Gabor filters [98], Scale Invariant Fourier Transform [12], Contrast Limited Adaptive Histogram Equalization [113] have proved to be successful in feature segmentation in cases where the biometric modality is within the textural information. In vein recognition, extra textural information is undesired and is often considered noise. In [32], Oriented Gradient Maps (OGM) have been used to extract vein features. This was done with the help of SIFT matching. The most common classifiers used for this purpose are k-NN [55] or SVM, and in recent years Convolutional Neural Networks (CNN) became of choice. An analysis of this was done and comparative information was drawn in our previous work in [68]. CNNs are known in the hand vein recognition world for authentication and verification but not as commonly for vein segmentation. The reason for this is the unavailability of annotated vein database necessary for segmentation. Even for authentication, the proposed methods from researchers make use of proprietary databases with minimal information disclosure making the evaluation of such studies difficult. In [116], a four-layered recurrent CNN is presented and is tested on a self-made database. AlexNet, VGG-16 have been applied for personal identification in [61]. [115] proposed a VGG-based ensembled CNN constructed with Squeeze-Net layers. In [10], Deep Hashing Networks (DHN) were introduced for hand-vein feature extraction and matching. This was followed by simplified CNN introduced in [130] again for matching.

CNNs have been extremely successful in medical image segmentation as compared to traditional methods. Fully convolutional networks were improvised versions of AlexNet, VGG, and ResNet with adaptations that have an encoder path of down convolution and a decoder path of up convolution. These networks could capture information i.e. features in every stage where the resolution of the original image was decreasing. Therefore, making it possible for performing pixel-wise segmentation of the whole input image. Although the drawback in this was the loss of information in the convolution and pooling process resulting in uneven segmentation. The most powerful pixel-wise segmentation method was introduced in [96] where skip connections were introduced between the upsampling and downsampling path of the CNN. This refined the features in the upsampling path by concatenating the encoder feature map with its corresponding decoder pair. UNets have been popular mostly in the medical world where they have been applied to segment organs or tumours. It has also found its application in other engineering domains for fluid dynamics [57], concrete crack detection [107], pattern denoising [31], and detection of manufacturing defects [120]. Other CNN structures inspired by UNet like VNet, 3D UNet [19, 74] were also successful for prostrate segmentation and kidney segmentation.

UNet and its modified variations have also found themselves to be a prominent method in retinal vessel segmentation [30]. Vessel segmentation is computationally complex and UNet is a supervised learning method that demands the original and mask image, also commonly referred to as gold-standard image. Retinal databases are usually accompanied by gold-standard images that are annotated by experts in the medical field. This makes the optimization and evaluation of UNet based segmentation method easier as compared to the biometric world. Subcutaneous veins in the finger, palm, and wrist are obtained from near-infrared images. The databases for these biometric modalities have images that have low contrast and low resolution. They are not expertly annotated, therefore, segmentation is only possible with the help of manual segmentation and other traditional methods. In most cases, the databases do not have mask images for evaluation. Now, a simple UNet is first explained followed by the modified UNet architecture that has been introduced and tested on palm, and wrist vein images for the generation of masks which in turn have been successfully used for the development of a complete vascular biometric system [71, 73]. The specifics of the UNet structure with its modifications and targeted use in vascular biometric image segmentation are described. This can be extended to other image modalities used in biometric recognition systems.

15.1.1 UNet Architecture

Figure 43 shows the original UNet architecture proposed by [96]. It consists of a contractive path and an expansive path. The contractive path resembles the typical architecture of a CNN where it has repeated application of a pair of 3 × 3 convolutions that are unpadded followed by a rectified linear unit (ReLU) and a 2 × 2 max pooling operation with a downsampling stride of 2. At each downsampling step, the feature channel numbers are doubled. In the expansive path i.e. the upsampling path, the up convolution of the feature map happens by a 2 × 2 convolution which reduces the number of feature channels by half followed by the corresponding concatenation of the cropped feature map and in turn, has 3 × 3 convolutions followed by a single ReLU. The cropping is necessary due to the loss of border pixels in every convolution. The final layer has a 1 × 1 convolution that is used to map each 64-component feature vector to its corresponding class. The total network has 23 convolutional layers. The authors have indicated that an even x and y size is necessary to allow seamless tiling of the segmentation map generated. The complete training information of this UNet architecture can be found in [96]. The general architecture was covered so that the readers can appreciate the effectiveness of this architecture for semantic segmentation and also easily distinguish the modifications carried out for biometric image segmentation in turn witnessing its strong ability to generate mask images that could be used as ground truth images for biometrics.

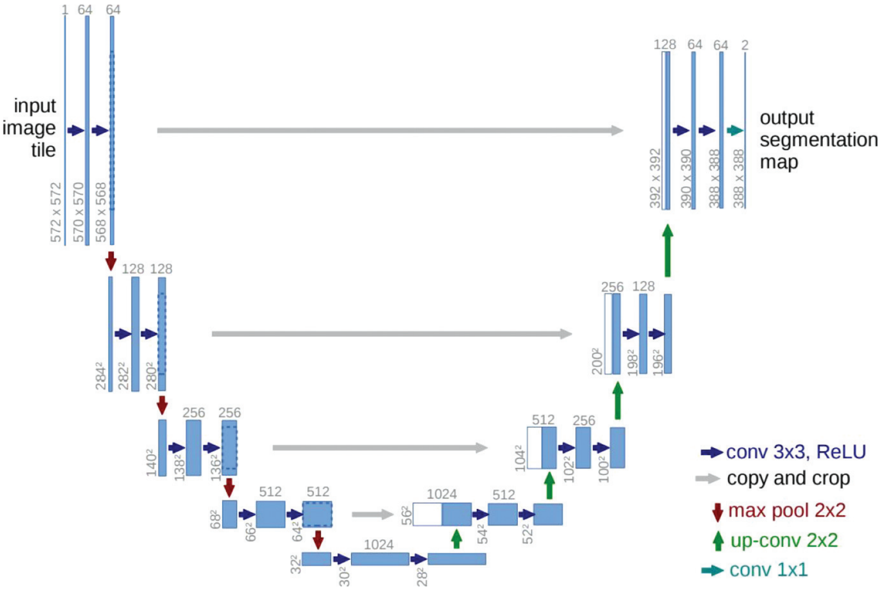


Figure 43: Original UNet. Image from Ronneberger *et al.* [96].

15.1.2 UNet for Hand Vein Segmentation

In [71], a modified UNet was proposed for the segmentation of palm vein images. This showed the effectiveness of UNet on vein image segmentation. The precise differences of these modifications are mentioned here. The input image from the palm vein scanner in the case of a palm vein image is greyscale near-infrared (NIR) image. The greyscale NIR image shows the vein networks and other undesired features like palmprint, geometry, and blemishes due to the uneven illumination of the infrared source combined with other image quality issues occurring during the acquisition phase. One example original image acquired from the HK PolyU palm vein database [127] is shown in Figure 44.

This is a palm vein NIR image and it can be seen that the vein network is not explicitly visible so as to be used in a recognition system. Often the entire palm vein image is of a bigger size and needs to be cropped down to the region of interest (ROI). ROI images are smaller in size, faster to process, and contain all the essential information needed from the biometric modality for the recognition engine. There are multiple algorithms proposed in the literature for ROI images [117]. One such method applied to obtain ROI images of 128×128 resolution for palm vein images is described in [71]. Once the ROI image is obtained, the next step was the use of morphological operations to extract

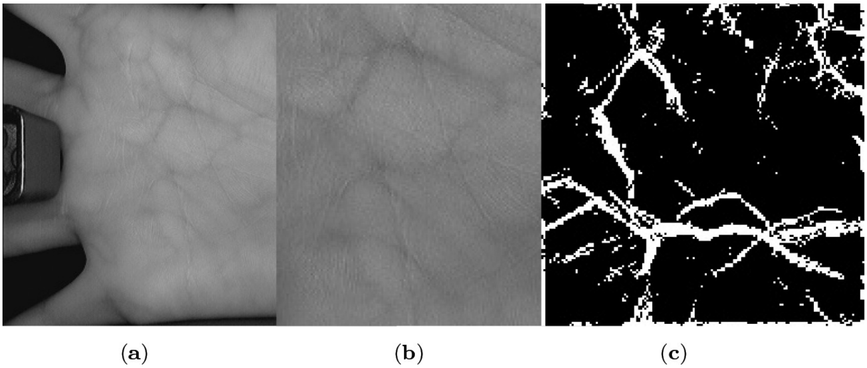


Figure 44: Mask image generation (**a**) Original image from HK PolyU, (**b**) ROI image, (**c**) Mask Generated with U-Net.

vein features so that they could be used as input into the UNet structure for feature extraction. Pixel thresholding was used on the original image shown in Figure 44 where all pixels above a particular value would be forced to 255 and below a particular value would be forced to 0. This would generate a binary image of the input image. Unlike the standard global thresholding method, adaptive thresholding was also implemented to improve the quality of the binary image. The binary image obtained using such methods have a lot of imperfections in them. Morphological methods such as erosion, dilation, and skeletonization can help further improve the quality of the binary image. This binary image is then provided as input to the UNet.

Figure 45 shows the UNet structure that can segment hand vein images that has been tested for palm vein images [71] and wrist vein images [69].

The input and output have the same dimensions with the input image having a resolution of 128 × 128 pixels for palm vein images (ROI image) and 256 × 256 for wrist vein images. Since the UNet does not contain fully connected layers, the number of trainable parameters do not increase. The input and output sizes can also change without re-training as the network weights do not change. The primary differences include the modification of the first contracting block which has the number of filters decreasing from 64 to 26, and the replacement of the 2D convolutional layer with a dropout layer in each expansive block to reduce overfitting. In the output layer, a custom Gabor filter kernel is introduced in the first contracting block. The input resolution of the UNet was changed according to the input image or the input ROI image. This method was used to generate the mask images for all the images in the database before using it in the recognition engine for matching. The specifics of the Gabor filter used and the parameters for the kernel are described in [71] for the reproducibility of the research.

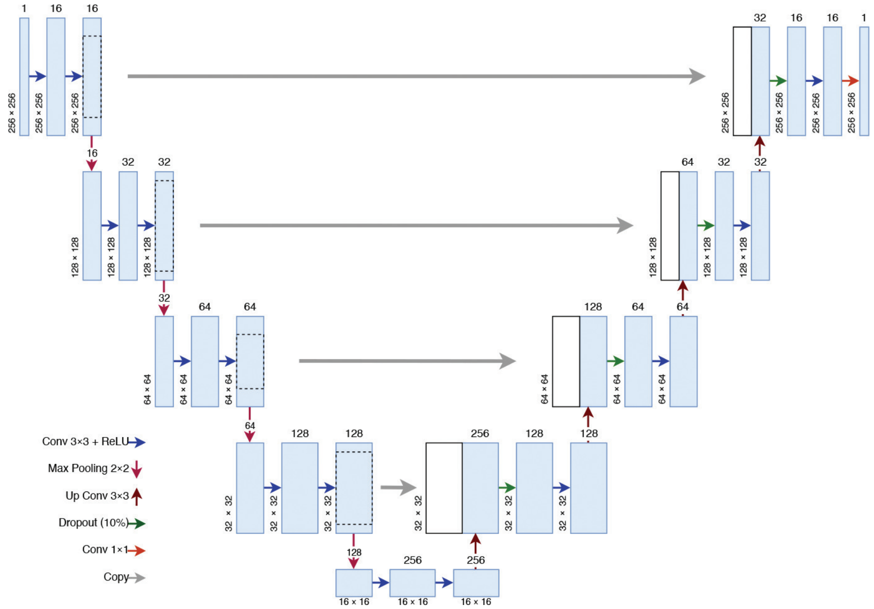


Figure 45: UNet architecture for vein image segmentation.

After the introduction of vein segmentation with the help of UNet structure, other unsupervised vein segmentation models using UNet variance became more popular in literature [59]. Automated labelling is of interest among researchers and variations of UNet structure have proven to be successful. Automatic labelling for hand vein images has been attempted with the help of the modified versions of several different types of UNets which include the simple UNet with modified filter structure, VGG-UNet structure, ResNet-UNet structure, and the UNet++ structure.

15.2 Deep Learning for Feature Matching: Biometric Vein Image Matching using Siamese Neural Networks

In the previous section, it was seen how a deep learning method, CNN based UNet was effectively used to segment vascular biometric vein image which then can be used as template for a biometric matching system. Here, in this subsection, an effective matching engine using a deep learning method known as Siamese Neural Networks is described. As in any biometric system, the typical process of vascular biometrics consists of image acquisition, preprocessing, feature extraction, and matching. Deep learning techniques have emerged to be an effective tool for pattern recognition in many computer vision tasks. It has already outperformed traditional algorithms. However, most deep learning methods proposed for matching purposes have a significant application disadvantage. One major issue is it requires large amount of training samples and its corresponding labels. Obtaining labelled images itself is a challenging task and to ease this out, novel UNet structure was used. Collecting data is laborious and labelling them accurately is cost and time intensive. Also, because of the biometric template involved, there are privacy considerations that need to be accounted for the template being used. Even if there is sufficient data, training on large dataset would be expensive with difficulty in generalization of the methods used. Therefore, investigating into a recognition system that requires less training data and is still accurate was imperative. This issue of scarcity of images in a particular sample has been common in biometrics and is referred to as small sample size problem [64]. This situation led to the advent of few shot recognition where a dataset containing N sample images for each category that have k samples labelled. Then the task would be to recognize the rest of the images in each category with the few remaining labelled data. Relating this to a real life biometric recognition system scenario where a set of images are registered into the database and the query image, also known as probe image, is to be matched with one of registered image to establish that it matches with one of the probe images i.e. identified and authenticated or rejected.

15.2.1 Few Shot Learning for Biometric Image Recognition

Spatial filters compute the weighted sum in the image when CNN is applied. This is computationally expensive and is difficult to generalize. Few shot learning was originally proposed to solve the problem of overfitting when the number of samples to train the network was limited [38]. The three common types of few shot learning methods are based on recurring neural networks (RNN), metric based, and initialization based models. RNN relies on memory of the previous iteration to optimize the output of the current iteration. When facing tasks with different distribution, it again relies on memory. Metric based model learn embeddings in image i.e. subspace features to help towards classification. Initialization based methods rely on parametric updation based on the observations of parameters after a few gradient steps. In [104], prototypical networks for few shot learning is proposed and distances between the features of every category is mapped to perform classification. Ren *et al.* [95] proposed a novel few shot learning algorithm with unlabelled examples. The iterations learned to leverage information from unlabelled examples. Finn *et al.* [25] introduced a model-agnostic meta training algorithm where explicit training was possible to adjust parameters using very few data for a new task. Recent literature also shows researchers propose generative neural networks (GNNs) and graph algorithms for few shot learning [28, 125]. [63] introduced transductive propagation network for few shot classification. It proposed to construct graphical model to exploit the mainfold data structure and classify the entire test set at once. These recent studies show the prominence of few shot learning which works using a special CNN based neural network structure known as Siamese Neural Networks. This was applied on palm vein images and wrist vein images successfully as part of our previous works [69, 72].

N-shot recognition can be termed as training the classifier to recognize images using a few labelled images as reference from each category. If the classifier is trained traditionally using optimizers like Softmax, the classifier would easily suffer overfitting. Describing this in a broader way, few-shot learning is a subset of machine learning algorithm where only a few samples are available for supervised learning. When referring to few shot learning, it usually means n-way k-shot classification where n represents the number of classes and k is the number of samples in each class for training. Few-shot, one-shot, and zero-shot learning are subfields of n-shot learning. Zero-shot learning aims to classify unseen classes without seeing any training examples. One-shot learning has one sample of each class in the training phase and few-shot has two,three or maybe even five samples per class. For vascular biometrics, the subset of n-shot learning, few-shot method has been most successful where there are n-class labels and k-labelled images for each class, and a query image also referred to as probe image. If the query image is to be classified among n-classes, then n × k samples in the training set are available. In biometrics, for a verification scenario, where similarity is the key aspect for matching engine decision, the underlying concept of Siamese neural networks is key. The application of Siamese neural networks in the biometric world is now reviewed and an overview of Siamese neural networks is being presented followed by the key details of its application for palm and wrist vein recognition systems. Similarity evaluation has been a key in verification setting within biometrics. There are different ways of comparison, for example, Eucledian distance, other correlation co-efficients, Spearman's rank etc. Siamese neural networks consists of two artificial neural networks (ANN), each capable of learning features from the input. The two networks are feed forward networks and employ back propagation during training. They share weights and work parallely to eventually compare their outputs through a distance function. This output can be a semantic similarity function and can behave as a matching engine. Siamese neural networks were introduced by Bromley *et al.* [8] to detect forged signatures by comparing two handwritten signatures and determine if the probe signature was forged or not. Siamese neural network is a feed forward network based on the perceptron model. The first layer reads the input value, multiplies it by a weight, and forwards it to the neurons in the following layer. The neurons of each layer beyond the input layer do exactly the same and keep passing the task to the subsequent neurons in the following layer.

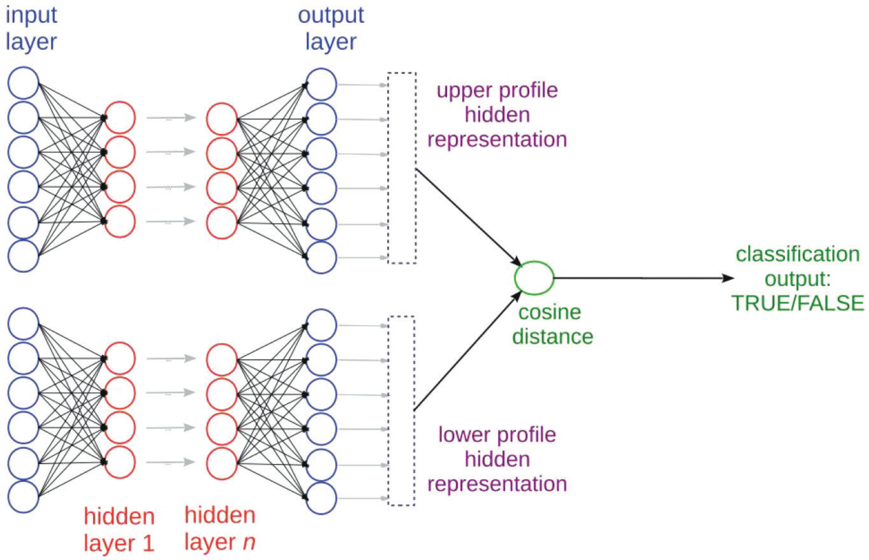

input layer
output layer
upper profile hidden representation
classification output: TRUE/FALSE
cosine distance
lower profile hidden representation
hidden layer 1
hidden layer n

Figure 46: General structure of siamese neural network with cosine distance as measure of similarity.

During training, the values produced by the neural network with its corresponding ground truth is compared and statistical error is computed. Examples of statistical error are Mean Square Error (MSE) or cross entropy error. Then this error is propagated backwards to update the neuron weights, often referred to as back propagation. The training stops when maximum number of iterations initially set is reached. At this point, the network is said to be trained and ready to be subjected to a test dataset. Once the network has generated predicted value for the test set, a confusion matrix can be drawn. Feed forward networks with back propagation is used in Siamese neural networks. As shown in Figure 46, there are two identical networks, each having the perceptron model [27]. In the training phase, each input is processed separately and the weights are updated on each of the networks through back propagation, finally generating a lower dimension output vector that can be compared easily. The algorithm compares the output of the upper neural network with the output of the lower neural network as seen in Figure 46. The similarity score is generated using cosine distance thus indicating if the inputs are similar or different.

15.2.2 Siamese Neural Network for Verification in Hand Vein Biometrics

The siamese neural network architecture for verification setting in hand vein biometrics is briefly discussed in this subsection. In a verification setting where the identity of the probe image is already known and is to be verified before access is granted, siamese neural network structure is beneficial. Figure 47 shows the application of siamese neural network for this purpose [69]. This has been effectively applied for palm vein and wrist vein images in our previous work. Two binary mask images that are generated with the help of algorithms like UNet discussed in Section 15.1.2, one represents the input image and the other represents the enrolled image. Both the binary images are simultaneously processed using the identical neural network that share weights and produce feature vectors. The Eucledian distance between the two feature vectors are then calculated and fed into a Sigmoid activation function. A pre-determined threshold is set into this function and depending on the distance output, the decision is made as genuine or imposter.

Figure 48 shows the feature extractor sub network used within the siamese neural network architecture. The same sub network processes each of the input that is shown in Figure 47. The input is converted into 1 dimensional feature vector by the sub network. The network has 3 convolutional blocks and 1 fully

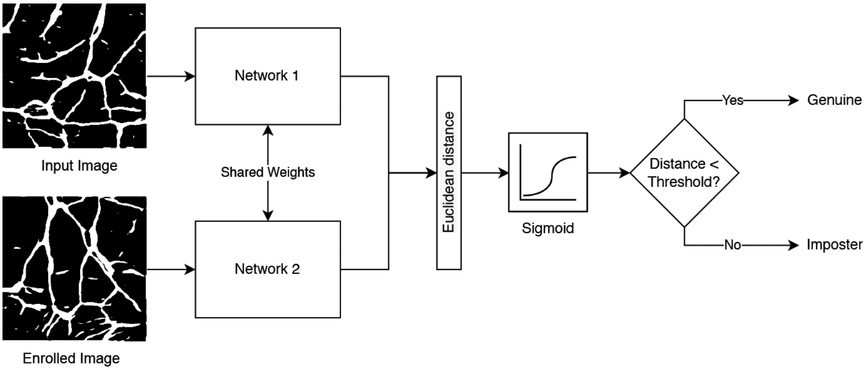


Figure 47: Siamese neural network for vascular biometric verification.

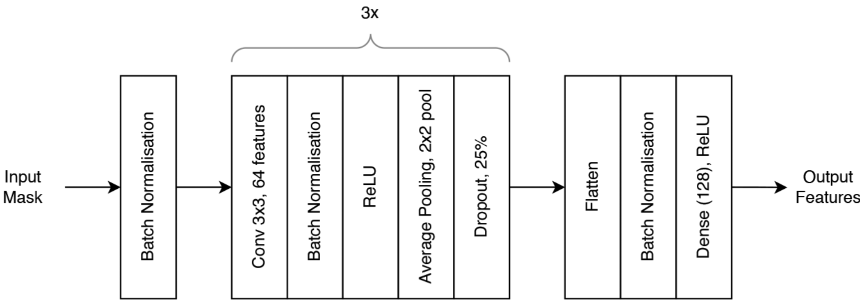


Figure 48: Sub network for feature extraction.

connected block. Each convolutional block has 3 × 3 kernel 2D convolutional layer producing 64 filter layers that helps with the extraction of 1 dimensional feature. Batch normalization improves the training process along with its speed and stability. ReLU activation function succeeds normalization layer with pooling layer of 2×2 to downsample the feature maps followed by dropout to prevent overfitting. The output of convolutional block is flattened through batch norm. In case of wrist vein images, into fully connected layer of 128 neurons representing the feature set. The model designed consists of 11 tunable layers with 768, 602 trainable parameters [69]. Contrastive loss is the usual preferred loss function for Siamese neural networks. It aims at maximization of the distance between non-matching feature set and minimization of the distance between matching feature sets. Mathematically it is represented by:

$$L = mean((1 - Y)(P^2) + Y(max(M - P, 0)^2)) \tag{2}$$

Here L represents the calculated loss, the known values is represented by Y, the predicted values are represented by P, and M is the distance reference to consider the images dissimilar. The margin is set to 1. This network for palm vein and wrist vein images was implemented using TensorFlow and Keras 2.1. The UNet discussed in the previous section was helpful to mass generate the mask images in both scenarios.

In a verification setting, the matching process is of comparison between the two images that are subjected to the Siamese neural network. In the training phase, same pairs were assigned true label and dissimilar pairs were assigned false labels. The paired dataset was used in the training phase to estimate the distance function and train the network. To compensate for the variations that occur during the image acquisition process, data augmentation was also considered to introduce variation in the input image. Adam Optimizer with the learning rate 0.003 was used and the feature extractor used binary cross entropy loss function with batch size of 32 and finally the siamese neural network used contrastive loss with batch size of 16. These parameters are customized based on our application in our previous work in [69] and can further be tailored depending on the input biometric modality, capture setting, feature extraction method, and matching engine type.

Section 15 showed how deep learning methods could effectively be used in vascular biometric systems, specifically in hand based recognition system. This is a recent topic and further research is being carried out in this area with more variations of UNet for segmentation and siamese neural networks for matching, thus promising the effectiveness of deep learning methods in all the stages of biometric recognition.

16 Conclusion and Future Perspective

In conclusion, this paper has underscored the increasing importance of securing people's properties and information in the digital age, as traditional security measures are vulnerable to hacking and theft. Biometric recognition systems are a more secure and convenient form of authentication as they utilize individuals' unique biological or behavioral characteristics. However, using biometric data raises privacy concerns as it is sensitive and cannot be changed, and there is a risk of it being stolen or misused. To address these concerns, researchers are exploring Biometric Template Protection (BTP) methods that ensure secure storage and processing of biometric templates. BTP methods should maintain recognition accuracy, be irreversible, and be renewable/unlinkable. This paper has explained two main types of BTP methods: handcrafted and learned. Handcrafted methods rely on expert knowledge to design secure systems, while learned techniques use machine learning algorithms to learn how to protect biometric templates. Moreover, this paper has emphasized the effectiveness of deep neural network-based models in accurately segmenting real-world features and matching them for authentication. Deep learning algorithms have been used to improve the quality of acquired images, enhance features, and make relevant decisions for the biometric system. The paper has explored specific deep learning algorithms, such as U-net for segmentation and Siamese neural network for matching, using hand veins as the biometric modality of choice. The results have shown that deep learning-based biometric systems have the potential to outperform traditional methods, indicating a promising future for biometrics research.

In summary, the importance of biometric technology in securing people's properties and information cannot be overstated. However, using biometric data raises privacy concerns that need to be addressed by developing effective and secure biometric systems, including Biometric Template Protection methods. Additionally, deep learning algorithms have shown great potential in enhancing the accuracy and effectiveness of biometric systems, indicating a promising future for biometric research.

References

[1] W. H. Abdulla, "Robust speaker modeling using perceptually motivated feature," *Pattern recognition letters*, 28(11), 2007, 1333–42.

[2] W. H. Abdulla and Y. Zhang, "Voice biometric feature using gammatone filterbank and ica," *International Journal of Biometrics*, 2(4), 2010, 330–49.

[3] O. Aiadi, B. Khaldi, and C. Saadeddine, “MDFNet: an unsupervised lightweight network for ear print recognition,” *Journal of Ambient Intelligence and Humanized Computing*, 2022, 1–14.

[4] D. W. Alausa, E. Adetiba, J. A. Badejo, I. E. Davidson, O. Obiyemi, E. Buraimoh, A. Abayomi, and O. Oshin, “Contactless palmprint recognition system: a survey,” *IEEE Access*, 10, 2022, 132483–505.

[5] J. Ashbourn, *Biometrics: advanced identity verification: the complete guide*, Springer, 2014.

[6] M. Azadmanesh, “Siamese Neural Networks for Biometric Hashing,” *PhD thesis*, 2014.

[7] R. Belguechi, C. Rosenberger, and S. Ait-Aoudia, “Biohashing for securing minutiae template,” in *2010 20th International Conference on Pattern Recognition*, IEEE, 2010, 1168–71.

[8] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. Lecun, C. Moore, E. Säckinger, and R. Shah, “Signature verification using a “Siamese” time delay neural network. Series in Machine Perception and Artificial Intelligence. 1994; 25–44,” 1994.

[9] D. Brown and K. Bradshaw, “Deep palmprint recognition with alignment and augmentation of limited training samples,” *SN Computer Science*, 3, 2022, 1–17.

[10] Z. Cao, M. Long, J. Wang, and P. S. Yu, “Hashnet: Deep learning to hash by continuation,” in *Proceedings of the IEEE international conference on computer vision*, 2017, 5608–17.

[11] D. Chang, S. Garg, M. Ghosh, and M. Hasan, “BIOFUSE: A framework for multi-biometric fusion on biocryptosystem level,” *Information Sciences*, 546, 2021, 481–511.

[12] S. Chanthamongkol, B. Purahong, and A. Lasakul, “Dorsal hand vein image enhancement for improve recognition rate based on SIFT keypoint matching,” in *2nd International Symposium on Computer, Communication, Control and Automation*, Atlantis Press, 2013, 174–7.

[13] P. S. Chanukya and T. Thivakaran, “Multimodal biometric cryptosystem for human authentication using fingerprint and ear,” *Multimedia Tools and Applications*, 79, 2020, 659–73.

[14] N. Charfi, H. Trichili, and B. Solaiman, “Bimodal Biometric Method Fusing Hand Shape and Palmprint Modalities at Rank Level,” in *Computational Collective Intelligence: 9th International Conference, ICCCI 2017, Nicosia, Cyprus, September 27-29, 2017, Proceedings, Part I 9*, Springer, 2017, 538–47.

[15] K. H. Cheung, A. W.-K. Kong, J. You, D. Zhang, *et al.*, “An Analysis on Invertibility of Cancelable Biometrics based on BioHashing.,” in *CISST*, Vol. 2005, Citeseer, 2005, 40–5.

[16] C. S. Chin, A. T. B. Jin, and D. N. C. Ling, “High security iris verification system based on random secret integration,” *Computer Vision and Image Understanding*, 102(2), 2006, 169–77.

[17] D. Chow and W. Abdulla, “Robust speaker identification based on perceptual log area ratio and Gaussian mixture models,” in *Eighth International Conference on Spoken Language Processing*, 2004.

[18] D. Chow and W. H. Abdulla, “Speaker identification based on log area ratio and Gaussian mixture models in narrow-band speech: speech understanding/interaction,” in *PRICAI 2004: Trends in Artificial Intelligence: 8th Pacific Rim International Conference on Artificial Intelligence, Auckland, New Zealand, August 9-13, 2004. Proceedings 8*, Springer, 2004, 901–8.

[19] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: learning dense volumetric segmentation from sparse annotation,” in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, 424–32.

[20] T. Connie, A. Teoh, M. Goh, and D. Ngo, “Palmhashing: a novel approach for cancelable biometrics,” *Information processing letters*, 93(1), 2005, 1–5.

[21] J. Cui and A. B. J. Teoh, “Deep Index-of-Maximum Hashing for Face Template Protection,” in *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, 2020, 413–8, DOI: 10.1109/ICCCS49078.2020.9118594.

[22] M. Dvořák, M. Drahanský, and W. H. Abdulla, “On the fly biometric identification system using hand-geometry,” *IET Biometrics*, 10(3), 2021, 315–25.

[23] K. Faez, S. Motamed, and M. Yaqubi, “Personal verification using ear and palm-print biometrics,” in *2008 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2008, 3727–31.

[24] Y. Feng, P. C. Yuen, and A. K. Jain, “A hybrid approach for face template protection,” in *Biometric Technology for Human Identification V*, Vol. 6944, SPIE, 2008, 58–68.

[25] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” in *International conference on machine learning*, PMLR, 2017, 1126–35.

[26] M. Freire-Santos, J. Fierrez-Aguilar, and J. Ortega-Garcia, “Cryptographic key generation using handwritten signature,” in *Biometric technology for human identification III*, Vol. 6202, SPIE, 2006, 225–31.

[27] Y. Freund and R. E. Schapire, “Large margin classification using the perceptron algorithm,” in *Proceedings of the eleventh annual conference on Computational learning theory*, 1998, 209–17.

[28] V. Garcia and J. Bruna, "Few-shot learning with graph neural networks," *arXiv preprint arXiv:1711.04043*, 2017.

[29] B. P. Gilkalaye, A. Rattani, and R. Derakhshani, "Euclidean-Distance Based Fuzzy Commitment Scheme for Biometric Template Security," in *2019 7th International Workshop on Biometrics and Forensics (IWBF)*, 2019, 1–6, DOI: 10.1109/IWBF.2019.8739177.

[30] C. Guo, M. Szemenyei, Y. Pei, Y. Yi, and W. Zhou, "SD-UNet: A structured dropout U-Net for retinal vessel segmentation," in *2019 IEEE 19th international conference on bioinformatics and bioengineering (BIBE)*, IEEE, 2019, 439–44.

[31] J. Gurrola-Ramos, O. Dalmau, and T. Alarcón, "U-Net based neural network for fringe pattern denoising," *Optics and Lasers in Engineering*, 149, 2022, 106829.

[32] D. Huang, Y. Tang, Y. Wang, L. Chen, and Y. Wang, "Hand vein recognition based on oriented gradient maps and local feature matching," in *Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part IV 11*, Springer, 2013, 430–44.

[33] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "Unet 3+: A full-scale connected unet for medical image segmentation," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020, 1055–9.

[34] S. M. Islam, M. Bennamoun, R. Owens, and R. Davies, "Biometric approaches of 2D-3D ear and face: A survey," in *Advances in computer and information sciences and engineering*, Springer, 2008, 509–14.

[35] A. Jain, R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*, Vol. 479, Springer Science & Business Media, 1999.

[36] A. K. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*, Springer Science & Business Media, 2007.

[37] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," *EURASIP Journal on advances in signal processing*, 2008, 2008, 1–17.

[38] C. Jiang, H. Xu, X. Liang, and L. Lin, "Hybrid knowledge routed modules for large-scale object detection," *Advances in Neural Information Processing Systems*, 31, 2018.

[39] A. T. B. Jin, D. N. C. Ling, and A. Goh, "Biohashing: two factor authentication featuring fingerprint data and tokenised random number," *Pattern recognition*, 37(11), 2004, 2245–55.

[40] Z. Jin, J. Y. Hwang, Y.-L. Lai, S. Kim, and A. B. J. Teoh, "Ranking-Based Locality Sensitive Hashing-Enabled Cancelable Biometrics: Index-of-Max Hashing," *IEEE Transactions on Information Forensics and Security*, 13(2), 2018, 393–407, DOI: 10.1109/TIFS.2017.2753172.

[41] A. K. Jindal, S. Chalamala, and S. K. Jami, "Face Template Protection Using Deep Convolutional Neural Network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, 575–5758, DOI: 10.1109/CVPRW.2018.00087.

[42] A. K. Jindal, S. R. Chalamala, and S. K. Jami, "Securing Face Templates using Deep Convolutional Neural Network and Random Projection," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, 2019, 1–6, DOI: 10.1109/ICCE.2019.8662094.

[43] A. Juels and M. Sudan, "A Fuzzy Vault Scheme," *Designs, Codes and Cryptography*, 38(2), 2006, 237–57.

[44] A. Juels and M. Wattenberg, "A Fuzzy Commitment Scheme," in *Proceedings of the 6th ACM Conference on Computer and Communications Security*, *CCS '99*, Kent Ridge Digital Labs, Singapore: Association for Computing Machinery, 1999, 28–36, ISBN: 1581131488, DOI: 10.1145/319709.319714.

[45] D. Keller, M. Osadchy, and O. Dunkelman, "Fuzzy Commitments Offer Insufficient Protection to Biometric Templates Produced by Deep Learning," *arXiv preprint arXiv:2012.13293*, 2020.

[46] D. Keller, M. Osadchy, and O. Dunkelman, "Inverting Binarizations of Facial Templates Produced by Deep Learning (and Its Implications)," *IEEE Transactions on Information Forensics and Security*, 16, 2021, 4184–96.

[47] A. Kholmatov and B. Yanikoglu, "Realization of correlation attack against the fuzzy vault scheme," in *Security, forensics, steganography, and watermarking of multimedia contents X*, Vol. 6819, SPIE, 2008, 263–9.

[48] S. Kirchgasser, A. Uhl, Y. Martinez-Diaz, and H. Mendez-Vazquez, "Is Warping-based Cancellable Biometrics (still) Sensible for Face Recognition?" In *2020 IEEE International Joint Conference on Biometrics (IJCB)*, 2020, 1–9, DOI: 10.1109/IJCB48548.2020.9304870.

[49] A. Kong, K.-H. Cheung, D. Zhang, M. Kamel, and J. You, "An analysis of biohashing and its variants," *Pattern recognition*, 39(7), 2006, 1359–68.

[50] V. Krivokuca, "Fingerprint template protection using compact minutiae patterns," *PhD thesis*, The University of Auckland, 2015.

[51] V. Krivokuca and S. Marcel, "On the recognition performance of biohash-protected finger vein templates," *Handbook of Vascular Biometrics*, 2020, 465–80.

[52] V. Krivokuća Hahn and S. Marcel, "Towards Protecting Face Embeddings in Mobile Face Verification Scenarios," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(1), 2022, 117–34, DOI: 10.1109/TBIOM.2022.3140472.

[53] V. Krivokuća Hahn, "Biometric Template Protection for Face Recognition Systems: A behind-the-scenes look at the Motivation, Methods, and Metrics," Tutorial presented at IJCB2022, 2022, https://drive.switch.ch/index.php/s/wNlOfS1LX00pXkl.

[54] V. Krivokuća Hahn and S. Marcel, "Biometric Template Protection for Neural-Network-Based Face Recognition Systems: A Survey of Methods and Evaluation Techniques," *IEEE Transactions on Information Forensics and Security*, 18, 2023, 639–66, DOI: 10.1109/TIFS.2022.3228494.

[55] S. M. Lajevardi, A. Arakala, S. Davis, and K. J. Horadam, "Hand vein authentication using biometric graph matching," *IET Biometrics*, 3(4), 2014, 302–13.

[56] N. Lalithamani and K. Soman, "An effective scheme for generating irrevocable cryptographic key from cancelable fingerprint templates," *International Journal of Computer Science and Network Security*, 9(3), 2009, 183–93.

[57] Q. T. Le and C. Ooi, "Surrogate modeling of fluid dynamics with a multigrid inspired neural network architecture," *Machine Learning with Applications*, 6, 2021, 100176.

[58] Y. J. Lee, K. Bae, S. J. Lee, K. R. Park, and J. Kim, "Biometric key binding: Fuzzy vault based on iris images," in *Advances in Biometrics: International Conference, ICB 2007, Seoul, Korea, August 27-29, 2007. Proceedings*, Springer, 2007, 800–8.

[59] S. Lefkovits, S. Emerich, and L. Lefkovits, "Boosting Unsupervised Dorsal Hand Vein Segmentation with U-Net Variants," *Mathematics*, 10(15), 2022, 2620.

[60] K. Li, G. Zhang, Y. Wang, P. Wang, and C. Ni, "Hand-dorsa vein recognition based on improved partition local binary patterns," in *Biometric Recognition: 10th Chinese Conference, CCBR 2015, Tianjin, China, November 13-15, 2015, Proceedings 10*, Springer, 2015, 312–20.

[61] X. Li, D. Huang, R. Zhang, Y. Wang, and X. Xie, "Hand dorsal vein recognition by matching width skeleton models," in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, 3146–50.

[62] X. Liang, Z. Li, D. Fan, B. Zhang, G. Lu, and D. Zhang, "Innovative contactless palmprint recognition system based on dual-camera alignment," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(10), 2022, 6464–76.

[63] Y. Liu, J. Lee, M. Park, S. Kim, E. Yang, S. J. Hwang, and Y. Yang, "Learning to propagate labels: Transductive propagation network for few-shot learning," *arXiv preprint arXiv:1805.10002*, 2018.

[64] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized discriminant analysis for the small sample size problem in face recognition," *Pattern recognition letters*, 24(16), 2003, 3079–87.

[65] G. Mai, K. Cao, X. Lan, and P. C. Yuen, “SecureFace: Face Template Protection,” *IEEE Transactions on Information Forensics and Security*, 16, 2021, 262–77, DOI: 10.1109/TIFS.2020.3009590.

[66] E. Maiorana, P. Campisi, and A. Neri, “User adaptive fuzzy commitment for signature template protection and renewability,” *Journal of Electronic Imaging*, 17(1), 2008, 11011–11.

[67] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar, *et al.*, *Handbook of fingerprint recognition*, Vol. 2, Springer, 2009.

[68] F. Marattukalam, W. Abdulla, A. Swain, C. R Wanigasekara Mudiyanse Ralahamillage, and J. James, “Palm Vein Recognition using SVM and CNN: A Comparative Performance Investigation,” *Transactions on Computational Science & Computational Intelligence*, 2021.

[69] F. Marattukalam, W. Abdulla, D. Cole, and P. Gulati, “Deep Learning-Based Wrist Vascular Biometric Recognition,” *Sensors*, 23(6), 2023, 3132.

[70] F. Marattukalam and W. H. Abdulla, “On palm vein as a contactless identification technology,” in *2019 Australian & New Zealand Control Conference (ANZCC)*, IEEE, 2019, 270–5.

[71] F. Marattukalam and W. H. Abdulla, “Segmentation of palm vein images using u-net,” in *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, IEEE, 2020, 64–70.

[72] F. Marattukalam, W. H. Abdulla, and A. Swain, “N-shot palm vein verification using siamese networks,” in *2021 International Conference of the Biometrics Special Interest Group (BIOSIG)*, IEEE, 2021, 1–5.

[73] F. Marattukalam, D. Cole, P. Gulati, and W. H. Abdulla, “On Wrist Vein Recognition For Human Biometrics,” in *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, IEEE, 2022, 66–73.

[74] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 fourth international conference on 3D vision (3DV)*, Ieee, 2016, 565–71.

[75] Z. Minchev, “Multiple human biometrics fusion in support of cyberthreats identification,” *Cybernetics and Information Technologies*, 15(7), 2015, 67–76.

[76] D. D. Mohan, N. Sankaran, S. Tulyakov, S. Setlur, and V. Govindaraju, “Significant Feature Based Representation for Template Protection,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019, 2389–96, DOI: 10.1109/CVPRW.2019.00293.

[77] A. Nagar, K. Nandakumar, and A. K. Jain, “Securing fingerprint template: Fuzzy vault with minutiae descriptors,” in *2008 19th International conference on pattern recognition*, IEEE, 2008, 1–4.

[78] S. Nanavati, M. Thieme, and R. Nanavati, “Biometrics, Identity Verification in a Networked World, Wiley Computer Publishing, 2002.”

[79] K. Nandakumar, “A fingerprint cryptosystem based on minutiae phase spectrum,” in *2010 IEEE International Workshop on Information Forensics and Security*, IEEE, 2010, 1–6.

[80] K. Nandakumar, A. K. Jain, and S. Pankanti, “Fingerprint-based fuzzy vault: Implementation and performance,” *IEEE transactions on information forensics and security*, 2(4), 2007, 744–57.

[81] K. Nandakumar, A. Nagar, and A. K. Jain, “Hardening fingerprint fuzzy vault using password,” in *Advances in Biometrics: International Conference, ICB 2007, Seoul, Korea, August 27-29, 2007. Proceedings*, Springer, 2007, 927–37.

[82] H. Otroshi Shahreza, V. Krivokuća Hahn, and S. Marcel, “On the Recognition Performance of BioHashing on state-of-the-art Face Recognition models,” in *2021 IEEE International Workshop on Information Forensics and Security (WIFS)*, IEEE, 2021, 1–6.

[83] H. Otroshi Shahreza and S. Marcel, “Towards protecting and enhancing vascular biometric recognition methods via biohashing and deep neural networks,” *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(3), 2021, 394–404.

[84] N. Pandey, W. Abdulla, and Z. Salcic, “Gait-based person identification using multi-view sub-vector quantisation technique,” in *2007 9th International Symposium on Signal Processing and Its Applications*, IEEE, 2007, 1–4.

[85] N. Pandey, W. Abdulla, and Z. Salcic, “Multi-view Gait recognition using sparse representation,” in *2019 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, IEEE, 2019, 1–6.

[86] R. K. Pandey, Y. Zhou, B. U. Kota, and V. Govindaraju, “Deep Secure Encoding for Face Template Protection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016, 77–83, DOI: 10.1109/CVPRW.2016.17.

[87] R. K. Pandey, Y. Zhou, B. U. Kota, and V. Govindaraju, “Maximum Entropy Binary Encoding for Face Template Protection,” *arXiv preprint arXiv:1512.01691*, 2015.

[88] Y.-H. Pang, A. Teoh, and D. Ngo, “Palmprint based cancelable biometric authentication system,” *International Journal of Signal Processing*, 1(2), 2004, 98–104.

[89] Y. Qian, W. Deng, and J. Hu, “Unsupervised face normalization with extreme pose and expression in the wild,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, 9851–8.

[90] A. Ramachandra, S. Abhilash, K. Raja, K. Venugopal, and L. Patnaik, "Feature level fusion based bimodal biometric using transformation domine techniques," *IOSR Journal of Computer Engineering (IOSR-JCE)*, 3(3), 2012, 39–46.

[91] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM systems Journal*, 40(3), 2001, 614–34.

[92] C. Rathgeb, J. Merkle, J. Scholz, B. Tams, and V. Nesterowicz, "Deep face fuzzy vault: Implementation and performance," *Computers & Security*, 113, 2022, 102539.

[93] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP journal on information security*, 2011(1), 2011, 1–25.

[94] C. Rathgeb and A. Uhl, "Adaptive fuzzy commitment scheme based on iris-code error analysis," in *2010 2nd European Workshop on Visual Information Processing (EUVIP)*, IEEE, 2010, 41–4.

[95] M. Ren, E. Triantafillou, S. Ravi, J. Snell, K. Swersky, J. B. Tenenbaum, H. Larochelle, and R. S. Zemel, "Meta-learning for semi-supervised few-shot classification," *arXiv preprint arXiv:1803.00676*, 2018.

[96] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, Springer, 2015, 234–41.

[97] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of multibiometrics*, Vol. 6, Springer Science & Business Media, 2006.

[98] O. Russakovsky, Y. Lin, K. Yu, and L. Fei-Fei, "Object-centric spatial pooling for image classification," in *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II 12*, Springer, 2012, 1–15.

[99] W. J. Scheirer and T. E. Boult, "Cracking fuzzy vaults and biometric encryption," in *2007 Biometrics Symposium*, IEEE, 2007, 1–6.

[100] J. Shen, N. Liu, C. Xu, H. Sun, Y. Xiao, D. Li, and Y. Zhang, "Finger vein recognition algorithm based on lightweight deep convolutional neural network," *IEEE Transactions on Instrumentation and Measurement*, 71, 2021, 1–13.

[101] A. Sierro, P. Ferrez, and P. Roduit, "Contact-less palm/finger vein biometrics," in *2015 International Conference of the Biometrics Special Interest Group (BIOSIG)*, IEEE, 2015, 1–12.

[102] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "A survey of behavioral biometric gait recognition: Current success and future perspectives," *Archives of Computational Methods in Engineering*, 28, 2021, 107–48.

[103] M. Singh, R. Singh, and A. Ross, "A comprehensive overview of biometric fusion," *Information Fusion*, 52, 2019, 187–205.

[104] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," *Advances in neural information processing systems*, 30, 2017.

[105] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren, "Multimodal deep learning for biomedical data fusion: a review," *Briefings in Bioinformatics*, 23(2), 2022, bbab569.

[106] J. Sun and W. Abdulla, "Palm vein recognition using curvelet transform," in *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, 2012, 435–9.

[107] Y. Tang, Z. Chen, Z. Huang, Y. Nong, and L. Li, "Visual measurement of dam concrete cracks based on U-net and improved thinning algorithm," *J. Exp. Mech*, 37, 2022, 209–20.

[108] A. B. J. Teoh and J. Kim, "Secure biometric template protection in fuzzy commitment scheme," *IEICE Electronics Express*, 4(23), 2007, 724–30.

[109] A. B. Teoh, A. Goh, and D. C. Ngo, "Random multispace quantization as an analytic mechanism for biohashing of biometric and random identity inputs," *IEEE transactions on pattern analysis and machine intelligence*, 28(12), 2006, 1892–901.

[110] A. B. Teoh and D. C. Ngo, "Cancellable biometerics featuring with tokenised random number," *Pattern recognition letters*, 26(10), 2005, 1454–60.

[111] A. B. Teoh, D. C. Ngo, and A. Goh, "An integrated dual factor authenticator based on the face data and tokenised random number," in *Biometric Authentication: First International Conference, ICBA 2004, Hong Kong, China, July 15-17, 2004. Proceedings*, Springer, 2004, 117–23.

[112] U. Uludag and A. K. Jain, "Fuzzy fingerprint vault," in *Proc. Workshop: Biometrics: Challenges arising from theory to practice*, 2004, 13–6.

[113] V. Vijayan, K. W. Bowyer, P. J. Flynn, D. Huang, L. Chen, M. Hansen, O. Ocegueda, S. K. Shah, and I. A. Kakadiaris, "Twins 3D face recognition challenge," in *2011 international joint conference on biometrics (IJCB)*, IEEE, 2011, 1–7.

[114] G. S. Walia, T. Singh, K. Singh, and N. Verma, "Robust multimodal biometric system based on optimal score level fusion model," *Expert Systems with Applications*, 116, 2019, 364–76.

[115] H. Wan, L. Chen, H. Song, and J. Yang, "Dorsal hand vein recognition based on convolutional neural networks," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, 2017, 1215–21.

[116] J. Wang and G. Wang, "Hand-dorsa vein recognition with structure growing guided CNN," *Optik*, 149, 2017, 469–77.

[117] P. Wang and D. Sun, "A research on palm vein recognition," in *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, IEEE, 2016, 1347–51.

[118] W.-C. Wang, W.-S. Chen, and S.-W. Shih, "Biometric recognition by fusing palmprint and hand-geometry based on morphology," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2009, 893–6.

[119] J. L. Wayman, "National Biometric Test Center: Collected Works 1997-2000," *Biometric Consortium of the US Government interest group on biometric authentication) San Jose State University, CA*, 2000.

[120] V. W. H. Wong, M. Ferguson, K. H. Law, Y.-T. T. Lee, and P. Witherell, "Segmentation of Additive Manufacturing Defects Using U-Net," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 85376, American Society of Mechanical Engineers, 2021, V002T02A029.

[121] H. Wu, W. Wang, J. Zhong, B. Lei, Z. Wen, and J. Qin, "Scs-net: A scale and context sensitive network for retinal vessel segmentation," *Medical Image Analysis*, 70, 2021, 102025.

[122] W. Wu, S. J. Elliott, S. Lin, S. Sun, and Y. Tang, "Review of palm vein recognition," *IET Biometrics*, 9(1), 2020, 1–10.

[123] F. Yang, B. Ma, Q. xia Wang, D. Yao, C. Fang, S. Zhao, *et al.*, "Information fusion of biometrics based-on fingerprint, hand-geometry and palm-print," in *2007 IEEE Workshop on Automatic Identification Advanced Technologies*, IEEE, 2007, 247–52.

[124] S. Yang and I. M. Verbauwhede, "Secure fuzzy vault based fingerprint verification system," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.* Vol. 1, IEEE, 2004, 577–81.

[125] H. Yao, C. Zhang, Y. Wei, M. Jiang, S. Wang, J. Huang, N. Chawla, and Z. Li, "Graph few-shot learning via knowledge transfer," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34, No. 04, 2020, 6656–63.

[126] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *18th international conference on pattern recognition (ICPR'06)*, Vol. 4, IEEE, 2006, 441–4.

[127] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo, "An online system of multispectral palmprint verification," *IEEE transactions on instrumentation and measurement*, 59(2), 2009, 480–90.

[128] D. D. Zhang, "Palmprint segmentation by key point features," *Palmprint Authentication*, 2004, 73–83.

[129] Y. Zhang, Q. Duan, C. Shao, and Y. Shi, “Parallel Population-Based Simulated Annealing for High-Dimensional Black-Box Optimization,” in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, 2021, 1–7.

[130] D. Zhong, H. Shao, and Y. Liu, “Hand dorsal vein recognition based on deep hash network,” in *Pattern Recognition and Computer Vision: First Chinese Conference, PRCV 2018, Guangzhou, China, November 23-26, 2018, Proceedings, Part I 1*, Springer, 2018, 26–37.

[131] X. Zhou and T. Kalker, “On the security of biohashing,” in *Media forensics and security II*, Vol. 7541, SPIE, 2010, 266–73.

[132] X. Zhu, D. Huang, and Y. Wang, “Hand dorsal vein recognition based on shape representation of the venous network,” in *Advances in Multimedia Information Processing–PCM 2013: 14th Pacific-Rim Conference on Multimedia, Nanjing, China, December 13-16, 2013. Proceedings 14*, Springer, 2013, 158–69.