

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,300

Open access books available

116,000

International authors and editors

125M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# A Survey on Soft Biometrics for Human Identification

---

Abdelgader Abdelwhab and Serestina Viriri

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.76021>

---

## Abstract

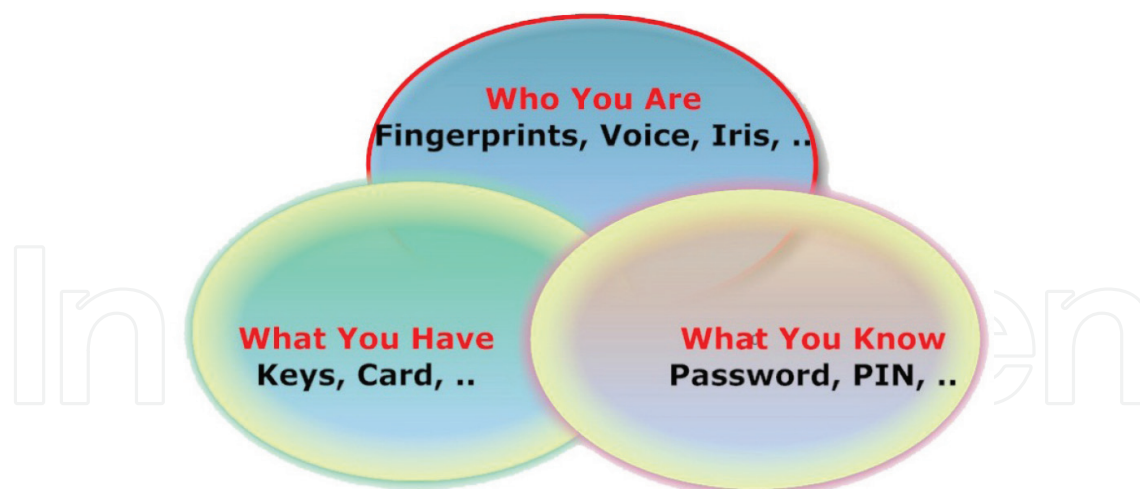
The focus has been changed to multi-biometrics due to the security demands. The ancillary information extracted from primary biometric (face and body) traits such as facial measurements, gender, color of the skin, ethnicity, and height is called soft biometrics and can be integrated to improve the speed and overall system performance of a primary biometric system (e.g., fuse face with facial marks) or to generate human semantic interpretation description (qualitative) of a person and limit the search in the whole dataset when using gender and ethnicity (e.g., old African male with blue eyes) in a fusion framework. This chapter provides a holistic survey on soft biometrics that show major works while focusing on facial soft biometrics and discusses some of the features of extraction and classification techniques that have been proposed and show their strengths and limitations.

**Keywords:** multi-biometrics, primary biometric, soft biometrics, gender, facial, ethnicity, fusion methods

---

## 1. Introduction

Along with the automation of our modern life, security issues become more critical and important. There are questions asked in our daily life such as “is this the right person to be allowed to access the system?”, “is this the authorized person to perform such action?”, and “does this person belong to this country?” [1]. There were two methods for answering these questions: first one based on “what you have” and called (knowledge factors), such as ID cards, and the second one based on “what you know” and called (ownership factors), such as passwords as shown in **Figure 1**. However both methods can be borrowed or copied or stolen, so users need to carry many IDs and memorize a lot of passwords. As reported banks, telecommunication companies, and governments are losing millions of dollars annually because of the violations of their password-based and card-based security policies [2]. To solve this person identification issue, biometrics is an opened field.



**Figure 1.** Information sets.

Biometrics rely on *what you are* called (inherence factors) so can natively differentiate between a permitted and illegal person [3, 4]. Biometric traits offer the following advantages [5]:

- They are unique for each individual.
- They cannot easily be forgotten, stolen, borrowed, shared, or observed.
- They always vary and are always available.
- They cannot easily be transferred to another individual.

A biometric-based security system is almost impossible to be fooled. The word biometric is a composite word bios, which refers to life, and metron, which refers to measure, coming from the Greek language. Biometric is sometimes defined as a research area focused on measuring and analyzing a person's unique characteristics [6] to identify or verify a person identity and is an essential daily task for a security system to make sure that the services are available for the permitted users only [7]. It can be divided into traditional, primary, and soft biometrics as shown: traditional biometric deals with physical, behavioral, and biological characteristics such as facial features, eye, signature, gait, voice, DNA, and fingerprints as shown in **Figure 2**. Soft biometrics are concerned with ancillary characteristics that provide some information not enough to identify a person clearly as gender, ethnicity, skin color, scars, and height [8, 9]. Behavioral or physiological human features must fulfill the following requirements to be recognized as can be used as a biometric characteristic [7, 10]:

1. Universal: each person has the trait.
2. Acceptable: available when needed.
3. Resistance to circumvention: not easy to cheat.
4. Distinctive: can be used to differentiate between persons.
5. Permanence: they don't change over a period of time.
6. Collectable: the characteristic can be easily collected and measured.

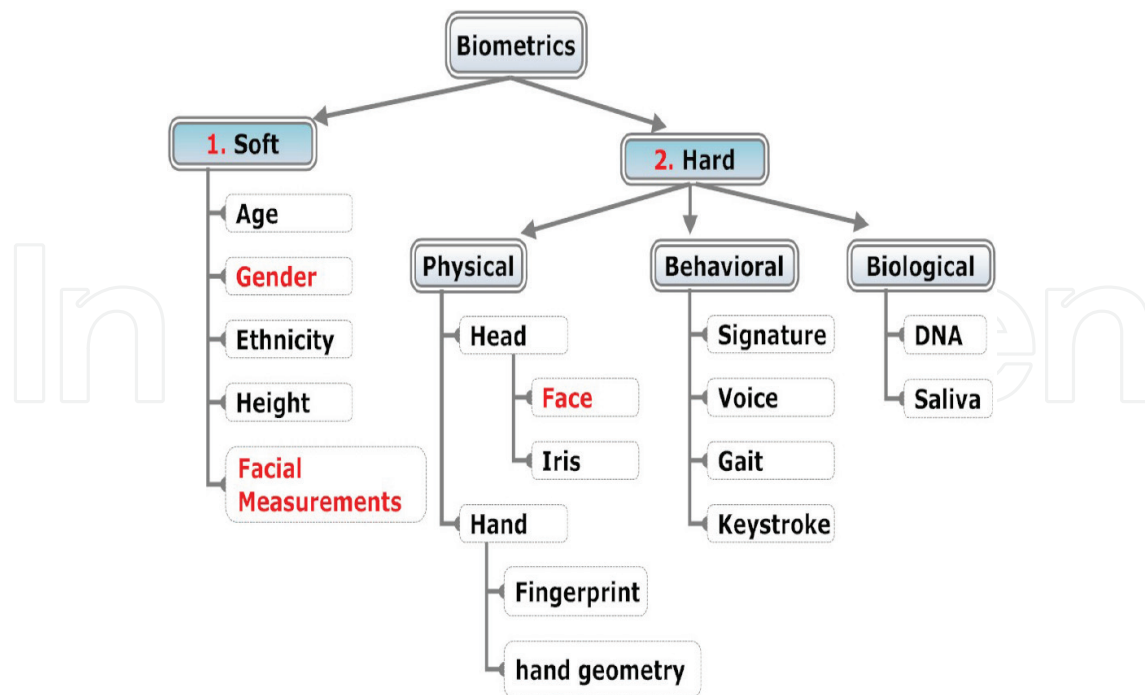


Figure 2. Biometrics type.

However, there is not a single biometric feature that satisfies all these characteristics identified above yet, so as a result, none of existing biometric system provides a precise foolproof recognition, so there is a gap for improving the recognition accuracy and speed of primary biometrics using soft biometrics.

This chapter is divided into five sections as follows: Section 2 shows the soft biometric benefits, unimodal biometric system limitation and how multimodal biometric system overcome this limitation, the need of biometric fusion for system performance, and the system performance measurement. A holistic survey on related works is presented in Section 3 while focusing on facial soft biometrics. In Section 4 we show the challenges and the limitations of the soft biometrics. Section 5 concludes the work.

## 2. Soft biometrics

Soft biometrics provide ancillary information but are not fully distinctive and permanent, so these features cannot provide a reliable person recognition. However, such ancillary information still can be used as a secondary information to complement the primary biometric traits (face, iris, etc.), and these features can be classified to physique (e.g., color skin, gender, ethnic origin), clothing (e.g., clothes' color), or accessories (e.g., glasses, hat) [11].

### 2.1. Benefits of soft biometrics

- Can be used to improve the recognition accuracy and speed of a primary biometric system [12].

- Can be used when there is a difficulty to collect a primary biometric trait or the collected data is not clear due to the sensor error or data collected from a distance with no cooperation with the user.
- Acceptable: collecting data for identification don't need cooperation between the person and the sensor and available.
- Soft facial biometrics are not expensive to compute since they can be acquired at the same time during primary face biometric collection.
- Enrolling person needs no cooperation and taken at distance even training of the system is done offline.
- Soft biometric bridges the gap between machine and human since they have a semantic meaning and can be understood by the human as old and short African male.
- Soft biometrics don't rise a privacy concern about collecting and saving data because they provide ancillary description and are not fully distinctive as old and short male.
- Filtering and indexing the large database to limit the number of searched data according to the connected person characteristics [13], for example, we can restrict the search for female gender.

## 2.2. Biometric system

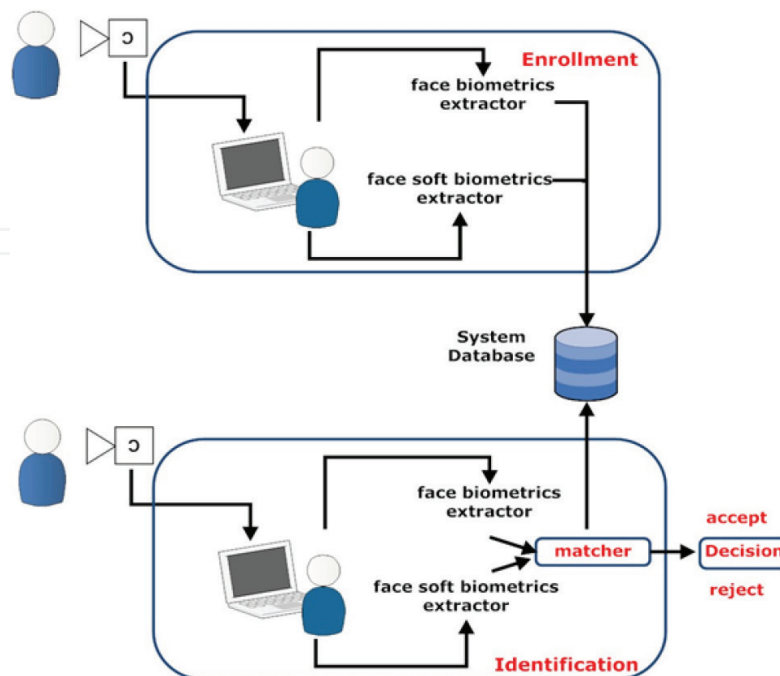
It is an essential pattern recognition system that uses the human characteristics in order to identify the person divided into unimodal system when using single trait and one that uses more than single traits called multi-biometric [14]; when developing a reliable biometric system, there are some concerns that need to be analyzed and balanced as needed [7]:

- Harmless to the users, as reported a research company put a SIM card under the skin for authentication.
- Performance, which means the highest recognition rate and system speed, while tolerance the environmental factors affecting the system, stable and time invariant.
- Acceptability, are the people ready to use their biometric trait?
- Circumvention means how easily your system can be overcome or bypassed using fake techniques.
- Accessible, easy to use.

Unimodal systems suffer from low-resolution data due to the person or the sensor, and this can lead to high failure to enroll rate, lacking people coverage area, and low recognition rate because cooperation with the user is needed to collect the data. So it is almost difficult to get very high recognition rates using unimodal system [14]; to improve the recognition rate, we need to acquire more than one trait from the same sensor or multiple sensors, but while increasing the recognition rate, the complexity and processing, which is time-consuming, increase.

Some problems associated with the unimodal biometric systems can be overcome by the use of the multi-biometric systems that combine the information obtained from multiple sources [15]. Still, such a system has two major limitations: first, the overall cost to construct the system can be prohibitive due to the need for more high-quality sensors, large storage capacity, and computational requirements. Second, the system requires a longer time for verification, hence causing inconvenience to the users [10]. However, soft biometrics are the solution to decrease the cost by using the same sensor [10]. The main steps for a biometric system are as follow [7, 16] as shown in **Figure 3**:

- Enrollment is the first step where biometric traits of the person are collected by the sensor and saved to the dataset as a template for verification purpose and later on used for identification. Successful biometric enrollment is necessary for the next steps.
- Enhancing the stored data to get high recognition rate by doing preprocessing as histogram equalization, clipping the area of interest dealing with the illumination.
- Extracting features vector from the individual for identification and match it with the stored template data.
- Template dataset: enrolling data means storing biometric data to the dataset as a template to be compared with the stored one. In the case of authentication, biometric data are matched against a reference template from the template database.
- Classification and matching: biometric feature data are validated against the template data in the dataset
- Decision can be rejected or accepted according to the matching similarity score or the threshold value.



**Figure 3.** Biometric system enrollment and identification diagram.

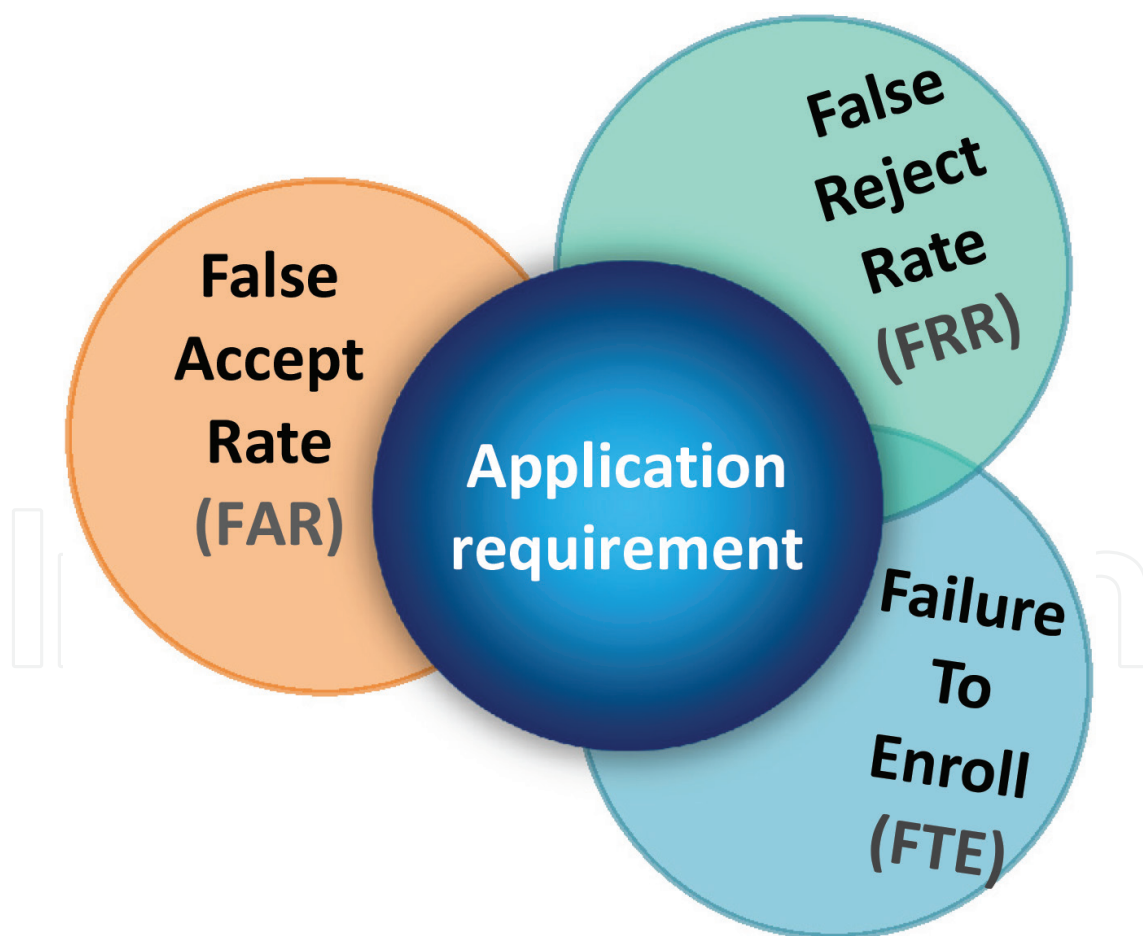


**Biometric system can work in two modes:**

Identification either in identification mode or verification mode. Identification mode works as one to many by comparing the individual with all the templates stored in the dataset, while a verification mode works as one to one by comparing the individual with his own template stored in the dataset.

**2.3. Biometric fusion**

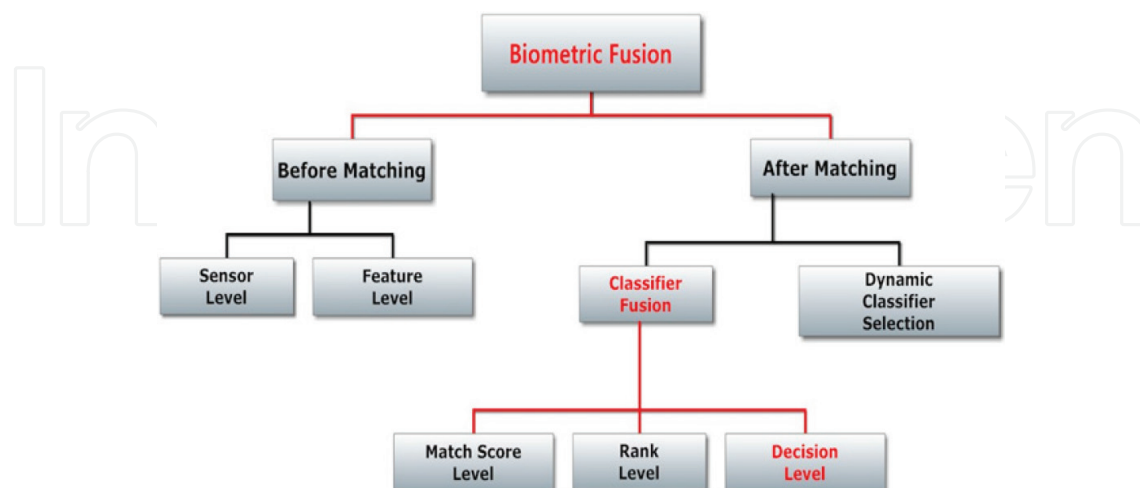
Biometric data may change over time or affected by environmental condition, so by fusing more than one trait or same trait from more than one source, we overcome the unimodal limitation and try to reduce one or more of the rejection and acceptance error rate based on the system requirements [17] as shown in **Figure 4**. Moreover, there is no one best biometrics since different applications require different policies such as distance learning, border control, and national identity card that require low false accept rate and failure to enroll. However, fusion is key to increase the recognition rate and can be taken at different stages (sensor, decision, feature extraction, classification stage).



**Figure 4.** Performance evaluation.

Sanderson and Paliwal et al. [18] divide the fusion into two categories: before classification called pre-matching and after classification called post-matching as shown in **Figure 5**:

- Pre-classification fusion [19–21]: before the classification level, the integration can be done in two ways as followed:
  1. Sensor level: integrating the raw data is difficult because it has a lot of unimportant features not only the region of interest and data collected from the sensors can be suffered from noisy as nonuniform illumination. Sensor-level fusion refers to raw data obtained using multiple sensors or multiple snapshots of a biometric using a single sensor. Face images collected from multiple sources with different resolutions may not be possible to integrate together.
  2. Feature level: in feature-level fusion, we get a lot of information by producing one feature set from fusing different features that are extracted from the captured images. So feature sets need to be tuned, normalized, transformed, and reduced. In practice, it is difficult to achieve feature-level fusion because concatenating different features may lead to dimensionality problem.
- Post-classification fusion [19–22]: the integration after the classification can be divided into three types:
  1. At score stage [23]: scores combined to generate one score value to and used for making decision according to the threshold value. Threshold making the system more reliable than using true and false since there is range can be tuned to increase or decrease the false acceptance rate and false rejected rate. However, a lower threshold decreases the rate of falsely rejected rate but also increases the rate of falsely accepted rate.



**Figure 5.** Fusion levels.



2. At rank stage [24]: the score values are arranged in descending order showing the possibility of the decision that at top list most preferred classes are placed and at down list least preferred classes.
3. Decision stage depends totally on the result value of the score stage, and final decision is taken whether the identified person is fake to reject or a unique to accept. Each classifier provides a hard decision. The decisions can be combined using:
  - Majority voting:
  - Decision is taken when a majority of the classifiers declare the same decision. To ensure a decision is taken, we must have classifiers more than the number of classes.
  - Logic operator (and, or):
  - And operator means all the classifiers give the same result whether reject or accept, and it is good when low false acceptance is required. Or operator is useful when low false reject is required.
  - Fuzzy logic [25]:
  - Instead of having reject or accept, we have a truth value between two values.

## 2.4. Performance evaluation

A biometric system needs to be evaluated and tested; there are some measurement concepts for evaluation as equal, false rejection and false acceptance rate [1, 26]:

- EER means both rate false accept and false reject are equal, and the more the EER, the more accurate the system is. The FRR refers to the rate of permitted users but are rejected by the system falsely.
- FAR means how many people don't have permission but the system accepts them as authorized person and falsely accepted.
- Failure to enroll (FTE) concerned with the rate of individuals not able to enroll in the system.
- FRR: the number of the authorized person but falsely rejected by the system.
- Failure to capture (FTC) concerned with the biometric traits are presented correctly, but the system was not able to capture them correctly.

$$\text{FRR} = (\text{number of false rejected}/\text{NAA}) \times 100\% \quad (1)$$

$$\text{FAR} = (\text{number of false accepted}/\text{NIA}) \times 100\% \quad (2)$$

NIA means number of impostor attempts and NAA means number of authorized attempts. The accuracy and recognition rate and performance measurements of a biometric system can be affected by some factors [26]:

- Environmental factors as high temperature, steam, and rain humidity lead to low accuracy. The features change over time as age and performance. The age, gender, ethnic, and face pose.
  - User willing and wishes: since users don't need to deal with the system intentionally, the system get affected and accuracy decrease.
  - The plastic surgery patients and people who don't have a hand cannot use a fingerprint.
- All the measurement rates are affected by the above factors, so any biometric system needs to calculate the error factors and tune and normalize them according to the system requirements and nature.

### 3. Literature review and related work

Alphonse Bertillon, who firstly introduced the idea of personal identification system based on biometric, morphological, and anthropometric using color of the eyes, hair, and skin in 1896. Face recognition is lower in uniqueness and more acceptable than iris but still is user-friendly, and people are willing to use it than other techniques [27]. The soft biometric is divided into three groups as follows [28]:

- Global traits are used for dataset indexing that remain fixed for the whole life as ethnicity and sex.
- Body features are used to describe an individual height and weight as tall or fat.
- Head features, this is where the research is heading now because of the rich feature in this body part as facial measurements and skin and hair color.

Soft biometric traits also can be classified according to permanence and distinctiveness as shown in **Table 1**. The permanence of a trait shows the strength of the trait over the period of time as gender and ethnicity don't change over time. Distinctiveness refers to the ability of a trait to differentiate between individuals.

Soft biometric traits	Face	Permanence	Distinctiveness
Facial measurements	Face	High	Medium
Gender	Face	High	Low
Skin color	Face	Medium	Low
Eye color	Face	Medium	Medium
Tattoo	Face	High	High
Age	Face	Low	Medium
Mustache	Face	Low	Low

**Table 1.** Facial soft biometric traits.

In this paper, we are focusing on the head soft biometric features. As shown in **Table 2**, humans can easily be identified by their faces because they don't change over a period of time and widely. According to Lin [47], face features provide different information when resized or clipped or shown from different sides.

The related works show some of the major works presented in timeline order starting from 2000 up to 2017 as shown in **Table 2**.

Jain and Dass et al. [13] the father of soft biometric who introduced it as ancillary information, but are not able to individually authenticate the person due to the lack of distinctiveness and permanence. They propose to use demographic information (gender, ethnicity, and height) as soft biometrics to improve the primary fingerprint system. Experiments show that recognition performance of fingerprint increased 5% by using soft biometrics.

Pedro and Julian et al. [28] experimental result shows that soft biometrics can be used as a secondary information to improve the primary biometrics and they can be acquired from distance; fusion is taken at score stage. Park and Jain et al. [34] use three feature extraction techniques:

- Active appearance model for extracting facial features as nose and eyes
- Laplacian of Gaussian
- Morphological operators

Two datasets are used to evaluate the system. They show that the use of soft biometrics (ethnicity, gender, facial marks) increases the recognition rate. Soft biometric traits can be considered as an alternative when face images are occluded or partially damaged. Gender and ethnicity of a person do not change over the lifetime, so they can be used to purge the database to narrow the search list. However, performance increased, but complexity also increased, and facial mark extraction depends on the image resolution and controlled environment needed.

Dantcheva and Velardo et al. [48] introduce two new soft biometric traits, called body weight and clothes color. Related promising results on the performance are provided. Dantcheva and Dugelay et al. [35] use eyes, skin, and hair color traits and cascade classifier; performance increased and balanced between complexity and performance. However, system suffers from illumination and poses, evaluated under one dataset and controlled environment. Soft biometric traits collected from a distance without user cooperation as shown by Denman and Fookes et al. [31] propose head and body traits and system evaluated using PETS 2006 small dataset and recognition rate decreased but the system can be used when primary data not available. Niinuma and Jain et al. [33] propose framework for continuous user authentication that uses clothing and skin colors fused with password. Soft biometric traits collected automatically every time user login with his password. Experiment results show the method effectiveness for continuous user authentication. However, system is evaluated with one dataset and suffering from illumination.

Ref.	Modalities	Database	Techniques	Results
[16] 2000	<ul style="list-style-type: none"> <li>• Face</li> <li>• Voice</li> <li>• Lip movement</li> </ul>	<ul style="list-style-type: none"> <li>• 150 persons for 3 months</li> <li>• A sample of audio and video frames</li> </ul>	<ul style="list-style-type: none"> <li>• Optical-flow technique</li> <li>• Fourier transformation</li> <li>• Hausdorff face location measurement</li> <li>• Synergetic</li> </ul>	<ul style="list-style-type: none"> <li>• Performance increased, and the false acceptance rate decreased</li> </ul>
[10] 2004	<ul style="list-style-type: none"> <li>• Fingerprint</li> <li>• Gender</li> <li>• Ethnicity</li> <li>• Height</li> </ul>	160 subjects	Bayes rule	The recognition rate increased by 6%
[29] 2008	Tattoo(human, plant, flag, symbol, object)	<ul style="list-style-type: none"> <li>• Image Database Web-based tattoo</li> <li>• Michigan State Police Tattoo Database (MI-DB)</li> </ul>	<ul style="list-style-type: none"> <li>• Scale-invariant feature transformer (SIFT)</li> <li>• Difference of Gaussian</li> <li>• Local extrema detection</li> <li>• Key point matching</li> </ul>	Performance increased to 77.2% on (MI-DB) and 98.6% on (web-based tattoo)
[30] 2008	<ul style="list-style-type: none"> <li>• Height</li> <li>• Body size</li> <li>• Gender</li> <li>• Stride/step</li> </ul>	<ul style="list-style-type: none"> <li>• USF outdoor dataset</li> <li>• SET HD indoor dataset</li> </ul>	A novel gait analysis in video surveillance	Height, gender, and body size show a better performance over stride/step lengths
[31] 2009	Legs, head color and size	PETS 2006	Active appearance model	The error rate is decreased
[11] 2009	<ul style="list-style-type: none"> <li>• Face</li> <li>• Facial marks</li> </ul>	<ul style="list-style-type: none"> <li>• FERET</li> <li>• Mugshot</li> </ul>	<ul style="list-style-type: none"> <li>• Morphological Operators</li> <li>• AAM</li> <li>• Laplacian of Gaussian</li> </ul>	System performance increased by 3 percentage to become 94.14%

Ref.	Modalities	Database	Techniques	Results
[32] 2009	<ul style="list-style-type: none"> <li>• Height and color of:               <ul style="list-style-type: none"> <li>○ Head</li> <li>○ Torso</li> <li>○ Legs</li> </ul> </li> </ul>	Subset of the PETS 2006	—	Equal error rate of 6.1% is achieved
[33] 2010	<ul style="list-style-type: none"> <li>• Face</li> <li>• Face color</li> <li>• Body location and size</li> <li>• Cloth color</li> </ul>	Video frames for 20 persons	<ul style="list-style-type: none"> <li>• PCA (Eigen)</li> <li>• Haar classifier</li> <li>• Bhattacharyya coefficient</li> </ul>	<ul style="list-style-type: none"> <li>• Real-time identification</li> <li>• Not affected by the person position</li> </ul>
[34] 2010	<ul style="list-style-type: none"> <li>• Face</li> <li>• Gender</li> <li>• Ethnicity</li> <li>• Facial marks (scars, moles, freckles)</li> </ul>	<ul style="list-style-type: none"> <li>• FERET</li> <li>• Mugshot</li> </ul>	<ul style="list-style-type: none"> <li>• AAM</li> <li>• Gaussian</li> <li>• PCA</li> <li>• Morphological operators</li> </ul>	<ul style="list-style-type: none"> <li>• The fusion of soft biometrics is able to improve the performance of face recognition</li> <li>• Facial marks can help in discriminating identical twins</li> </ul>
[35] 2010	<ul style="list-style-type: none"> <li>• Skin color</li> <li>• Hair color</li> <li>• Eye color</li> <li>• Beard</li> <li>• Mustache</li> <li>• Glasses</li> </ul>	Color FERET with 646 people	<ul style="list-style-type: none"> <li>• AdaBoost</li> <li>• Cascade</li> <li>• Histogram-based Bayes</li> </ul>	<ul style="list-style-type: none"> <li>• A proper balance between complexity and performance</li> <li>• Increase system reliability</li> </ul>
[36] 2011	<ul style="list-style-type: none"> <li>• Facial measurement of the lips and eyes</li> <li>• Face</li> </ul>	Yale	<ul style="list-style-type: none"> <li>• Biohashing</li> <li>• Hamming distances</li> <li>• Absolute differences</li> </ul>	The error rates have been decreased

Ref.	Modalities	Database	Techniques	Results
[37] 2011	<ul style="list-style-type: none"> <li>Sunglasses</li> <li>Scarf</li> </ul>	AR Face	<ul style="list-style-type: none"> <li>Gabor wavelets</li> <li>PCA</li> <li>SVM</li> </ul>	Recognition rate has increased
[38] 2012	Anthropometric body measures	<ul style="list-style-type: none"> <li>Medical chimera dataset</li> <li>NHANES dataset</li> <li>FERET dataset</li> </ul>	Eigenfaces	Anthropometric features improve performance in both accuracy and recognition speed
[39] 2012	<ul style="list-style-type: none"> <li>Gender</li> <li>Height</li> <li>Weight</li> <li>Blood group</li> </ul>	<ul style="list-style-type: none"> <li>IMS-BHU Indian hospital</li> </ul>	<ul style="list-style-type: none"> <li>Principal component analysis (PCA)</li> <li>Independent component analysis (ICA)</li> <li>Linear discriminant analysis (LDA) [27]</li> <li>Local binary pattern (LBP)</li> <li>Speeded up robust features (SURF)</li> </ul>	Soft biometrics improve primary face biometric performance by 6.5%
[40] 2013	<ul style="list-style-type: none"> <li>Facial measurement</li> <li>Skin color</li> <li>Hair color</li> </ul>	Face94	<ul style="list-style-type: none"> <li>Wavelet characterization</li> <li>SVM</li> </ul>	The recognition rate increased, equal error rate decreased, and skin color highly increases the performance
[41] 2013	Facial wrinkles	Well-known people from the Internet with high resolution	<ul style="list-style-type: none"> <li>Bipartite graph matching algorithm</li> <li>Curves to Line segments algorithm</li> <li>Modified Hausdorff distance (MHD)</li> <li>Curve proximity distance (CPD)</li> </ul>	<ul style="list-style-type: none"> <li>88% achieved by MHD</li> <li>87 achieved by CPD</li> <li>CPD performs better than MHD</li> <li>Fusing MHD and CPD increases the recognition rate to 93%</li> </ul>



Ref.	Modalities	Database	Techniques	Results
[42] 2014	<ul style="list-style-type: none"> <li>• Eyebrow size</li> <li>• Eye-to-eyebrow distance</li> <li>• Eyebrow length</li> </ul>	Southampton with video recordings from more than 200 subjects	Viola-Jones	The performance increased to 100% when using ten persons only
[43] 2014	<ul style="list-style-type: none"> <li>• Clothing attribute (head, upper body, lower body, foot, attached to body)</li> </ul>	<ul style="list-style-type: none"> <li>• Soton Gait database</li> </ul>	<ul style="list-style-type: none"> <li>• Soft-margin ranking SVM</li> <li>• Formulation of similarity constraints</li> </ul>	The identification rate increased from 78% to 95%
[44] 2015	Clothe color and type	Online shopping-labeled dataset	<ul style="list-style-type: none"> <li>• Deep learning based</li> <li>• RCNN detector</li> <li>• SV regression</li> </ul>	The performance increased but more dataset for training the module needed
[45] 2015	Body parts (clothe, hair, color)	VIPeR and GRID dataset	Multilevel CNN	Classification rate increased by 9% than support vector machine
[46] 2017	<ul style="list-style-type: none"> <li>• Eyebrow length</li> <li>• Eye size</li> <li>• Noise width</li> </ul>	Labeled faces in the wild Dataset	<ul style="list-style-type: none"> <li>• GIST descriptor</li> <li>• Deformable part model</li> </ul>	Accuracy rate increased using comparative features

**Table 2.** List of some of these works.

Asma and Souhir et al. [40] use facial measurements and skin and hair color as soft biometric traits. Support vector machine as a classifier is evaluated using one dataset. Results show equal error rate is decreased and recognition rate improved and requires no more cost since soft biometric traits are collected at the time of primary biometric collection by the same sensor. However, system needs to be tested with more difficult dataset and compared with another system. On the other hand, facial measurement features are very sensitive to pose and expression variation.

Nawaf and Nixon et al. [42] consider the eyebrow measurement distance and length from crowd sourcing. System is evaluated under one dataset with one classifier. Recognition rate increases but still needs to be tested with another dataset and compared with different classifiers. Jain and Park et al. [11] fuse face and facial marks. Their results show system performance increased up to 94.14%, but still facial mark extraction depends on the image resolution.

Min and Hadid et al. [37] propose facial occlusions as sunglasses, scarf, eye color, beard mustache, and glasses' traits. Experimental result shows that facial occlusions affect the system performance especially when user tries to use it to prevent himself from being recognized. However, they used one dataset for evaluation and did not compare it with other systems. Chen and Huang et al. [44] define new soft biometric traits to describe people based on their clothes' type, color, and pattern. RCNN body detector is used. However, they used their own dataset taken under controlled environment for training the RCNN, so the system cannot be compared with different systems and neural network needs more training data.

Jain, Dass, and Nandakumar et al. [10] combine gender, height, and ethnicity as soft biometric traits with fingerprint. The system performance increased by 6%. However, soft biometric traits did not extract automatically, and the system is evaluated by 160 subjects only. Lee, Jain, and Jin et al. [29] achieve a recognition rate of 98.6% on Web-DB with good quality taken under controlled environment and 77.2% on Michigan State Police Tattoo Database (MI-DB) using scale-invariant feature transform (SIFT) feature extractor. Experiment results show scars, marks, and tattoos (SMT) are more distinctive than other demographic biometrics such as ethnicity, gender, and weight to identify a person. However, tattoo dataset is collected under controlled environment at booking time.

Batool, Nazre, and Sima et al. [41] report a classification accuracy of 88% for facial wrinkles as a soft biometrics using modified Hausdorff distance (MHD) algorithm. There is no standard dataset to evaluate the system and compare with the other one. However, wrinkles are extracted manually by hand, and detecting wrinkles needs high-resolution image. Velardo, Carmelo, and Jean-Luc et al. [38] present a human body measurement (anthropometry) to prune primary biometric dataset. Their own medical dataset is collected from Indian hospital used for evaluating the body measurements and FERET data for face recognition. Results show system accuracy and recognition speed increased.

Saini and Sinha et al. [36] integrate the face and facial measurement of the lips and eyes as distance between two pupils, distance between the eyes and the lips, and length of the lips and the eyes to improve the recognition rate using hamming, absolute difference, and biohashing distance techniques. Experiment results on Yale dataset show error rate is decreased. However biohashing performances are poor when the tokenized random numbers are compromised; also only one dataset is used and results are not compared with another system.

Tiwari S, Singh A, and Singh SK et al. [39] propose an optimal framework for newborn recognition by fusing match scores from face and soft biometrics. Results on IMS-BHU Indian hospital dataset show that soft biometrics improve recognition rate by 5.6% over the primary biometric. However framework evaluated on one dataset has high-resolution image taken under controlled pose and illumination.

Jaha, Emad, and Mark et al. [43] show clothing traits can be used for identification of individual where clothing descriptions might be the only available feature. An, Chen, Kafai, Yang, and Bhanu et al. [49] aim to improve the re-identification performance by re-ranking the returned results based on soft biometric attributes. Experiments on challenging benchmark VIPeR dataset show that reranking improves the recognition accuracy.

## 4. Challenges and future work

Multimodal biometric systems are used to overcome the unimodal biometric system limitations by collecting multiple traits from multiple sensors. However, such a system will decrease the performance by increasing the processing duration and verification steps, and this causes users' troubles. So for developing reliable and user-friendly biometric system, we fuse soft and primary biometrics to improve the overall performance of the primary biometric system.

Soft biometrics inherit the nonintrusiveness and computational efficiency, which allow for fast, enrolment-free, and pose-invariant biometric analysis. However biometric system based on soft biometric trait only cannot provide accurate recognition because they change over time and lack distinctiveness, so there are still many challenges in this area. Parameter tuning as fusion rules and decision threshold otherwise error rate will increase and this can be improved using fuzzy logic.

Soft biometrics are very sensitive to illumination, expression variations, and pose variation, so we can use deep learning for preprocessing and feature extraction. New soft biometric traits can be also introduced as relative between the size of the head and body and facial distance measurement.

## 5. Conclusion

In a holistic survey on soft biometrics for user identification, we have seen that there is no one best biometric technology since it depends on the application requirement. A zero false acceptance rate is needed, for example, in security, and the false rejection rate needs to decrease, but in the civilian application, we need the opposite, so for any biometric system, we need to find a good balance between authentication reliability and complexity. As a result, traditional biometrics suffer from low recognition rate because they need cooperation with the user, operate in the controlled environment, and introduce privacy concern. So using multi-biometrics is the solution, but still, the system suffers from computation cost and long processing steps. However, another possible solution is to use soft biometrics to increase the population coverage and decrease the system cost and complexity.

## Author details

Abdelgader Abdelwhab<sup>1</sup> and Serestina Viriri<sup>2\*</sup>

\*Address all correspondence to: [viriris@ukzn.ac.za](mailto:viriris@ukzn.ac.za)

1 College of Computer Science and Information Technology, Sudan University of Science and Technology, Khartoum, Sudan

2 School of Maths, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

## References

- [1] Zhang DD. Automated Biometrics: Technologies and Systems. Vol. 7. Berlin: Springer Science & Business Media; 2013
- [2] Shoniregun CA. The future of internet security. *Ubiquity*. 2002;**2002**:8
- [3] Clarke R. Human identification in information systems: Management challenges and public policy issues. *Information Technology & People*. 1994;**7**(4):6-37
- [4] Srinivasa K, Gosukonda S. Continuous multimodal user authentication: Coupling hard and soft biometrics with support vector machines to attenuate noise. *CSI transactions on ICT*. 2014;**2**(2):129-140
- [5] Franke K, Ruiz-del-Solar J, Koppen M. Soft-biometrics: Soft-computing for biometric-applications. *International Journal of Fuzzy Systems*. 2002;**4**(2):665
- [6] Lu X, Jain AK. Ethnicity identification from face images. In: *Defense and Security*. International Society for Optics and Photonics; 2004. pp. 114-123
- [7] Jain AK, Ross A, Prabhakar S. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*. 2004;**14**(1):4-20
- [8] Jain A, Verma CK. A framework based on hybrid biometrics for personal verification systems. *International Journal of Applied*. 2012;**1**(1):55-58
- [9] Kim M-G, Moon H-M, Chung Y, Pan SB. A survey and proposed framework on the soft biometrics technique for human identification in intelligent video surveillance system. *BioMed Research International*. 2012;**2012**
- [10] Jain AK, Dass SC, Nandakumar K. Can soft biometric traits assist user recognition? In: *Defense and Security*. International Society for Optics and Photonics; 2004. pp. 561-572
- [11] Jain AK, Park U. Facial marks: Soft biometric for face recognition. In: *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE; 2009. pp. 37-40
- [12] Dantcheva A, Elia P, Ross A. What else does your biometric data reveal? A survey on soft biometrics. *IEEE Transactions on Information Forensics and Security*. 2016;**11**(3):441-467

- [13] Jain AK, Dass SC, Nandakumar K. Soft biometric traits for personal recognition systems. In: Zhang D, Jain AK, editors. *Biometric Authentication*. Berlin/Heidelberg: Springer; 2004. pp. 731-738
- [14] Zewail R, Elsafi A, Saeb M, Hamdy N. Soft and hard biometrics fusion for improved identity verification. In: *The 2004 47th Midwest Symposium on Circuits and Systems*, 2004. MWSCAS'04. Vol. 1. IEEE; 2004. pp. I-225
- [15] Khalifa AB, BenAmara NE. Contribution to the fusion of biometric modalities by the choquet integral. *International Journal of Image, Graphics and Signal Processing*. 2012;**4**(10):1
- [16] Frischholz RW, Dieckmann U. Biold: A multimodal biometric identification system. *Computer*. 2000;**33**(2):64-68
- [17] Sree SRS, Radha N. A survey on fusion techniques for multimodal biometric identification. *International Journal of Innovative Research in Computer*. 2014:7493-7497
- [18] Sanderson C, Paliwal KK. Information fusion and person verification using speech and face information. *Research Paper IDIAP-RR*; 2002. pp. 02-33
- [19] Ross A, Poh N. Multibiometric systems: Overview, case studies, and open issues. In: Tistarelli M, Li SZ, Chellappa R, editors. *Handbook of Remote Biometrics*. Berlin: Springer; 2009. pp. 273-292
- [20] Jain A, Nandakumar K, Ross A. Score normalization in multimodal biometric systems. *Pattern Recognition*. 2005;**38**(12):2270-2285
- [21] Hicklin A, Ulery B, Watson C. *A Brief Introduction to Biometric Fusion*. National Institute of Standards and Technology; 2006
- [22] Ho TK, Hull J, Srihari SN, Senior Member. Decision combination in multiple classifier systems. *Analysis*. 1994;**16**(I):66-75
- [23] Lam L, Suen SY. Application of majority voting to pattern recognition: An analysis of its behavior and performance. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*. 1997;**27**(5):553-568
- [24] Achermann B, Bunke H. Combination of Classifiers on the Decision Level for Face Recognition Combination of Classifiers on the Decision Level for Face Recognition, January; 1996
- [25] Zadeh LA. I. Introduction, and U. S. Navy, *Fuzzy Sets* \* -, vol. 353; 1965. pp. 338-353
- [26] Cavadini D, Fasel AMD, Cimasoni L. Introducing the biometrical electronic passport (epass); 2006
- [27] Shen W, Tan T. Automated biometrics-based personal identification. *Proceedings of the National Academy of Sciences*. 1999;**96**(20):11065-11066
- [28] Tome P, Fierrez J, Vera-Rodriguez R, Nixon MS. Soft biometrics and their application in person recognition at a distance. *IEEE Transactions on Information Forensics and Security*. 2014;**9**(3):464-475



- [29] Lee J-E, Jain AK, Jin R. Scars, marks, and tattoos (smt): Soft biometric for the suspect and victim identification. In: Biometrics Symposium, 2008. BSYM'08. IEEE; 2008. pp. 1-8
- [30] Ran Y, Rosenbush G, Zheng Q. Computational approaches for real-time extraction of soft biometrics. In: 19th International Conference on Pattern Recognition, ICPR 2008. IEEE; 2008. pp. 1-4 December
- [31] Denman S, Fookes C, Bialkowski A, Sridharan S. Soft biometrics: Unconstrained authentication in a surveillance environment. In: Digital Image Computing: Techniques and Applications, 2009. DICTA'09. IEEE; 2009. pp. 196-203
- [32] Denman S, Fookes C, Bialkowski A, Sridharan S. Soft-biometrics: Unconstrained authentication in a surveillance environment. In: Digital Image Computing: Techniques and Applications, IEEE 2009. DICTA'09. 2009. pp. 196-203
- [33] Niinuma K, Park U, Jain AK. Soft biometric traits for continuous user authentication. IEEE Transactions on Information Forensics and Security. 2010;5(4):771-780
- [34] Park U, Jain AK. Face matching and retrieval using soft biometrics. IEEE Transactions on Information Forensics and Security. 2010;5(3):406-415
- [35] Dantcheva A, Dugelay J-L, Elia P. Person recognition using a bag of facial soft biometrics (bofsb). In: 2010 IEEE International Workshop on Multimedia Signal Processing (MMSP). IEEE; 2010. pp. 511-516
- [36] Saini N, Sinha A. Soft biometrics in conjunction with optics based biohashing. Optics Communications. 2011;284(3):756-763
- [37] Min R, Hadid A, Dugelay J-L. Improving the recognition of faces occluded by facial accessories. In: 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011). IEEE; 2011. pp. 442-447
- [38] Velardo C, Dugelay J-L. Improving identification by pruning: A case study on face recognition and body soft biometric. In: 2012 13th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS). IEEE; 2012. pp. 1-4
- [39] Tiwari S, Singh A, Singh SK. Integrating faces and soft-biometrics for newborn recognition. International Journal of Advanced Computer Engineering and Architecture. 2012;2(2):201-209
- [40] Ghalleb AEK, Sghaier S, Amara NEB. Face recognition improvement using soft biometrics. In: 2013 10th International Multi-Conference on Systems, Signals & Devices (SSD). IEEE; 2013. pp. 1-6
- [41] Nazre B, Taheri S, Chellappa R. Assessment of facial wrinkles as a soft biometrics. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). 2013. pp. 1-7
- [42] Almudhahka N, Nixon M, Hare J. Human face identification via comparative soft biometrics. In: 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). IEEE; 2016. pp. 1-6



- [43] Jaha ES, Nixon MS. Soft biometrics for subject identification using clothing attributes. In: 2014 IEEE International Joint Conference on Biometrics (IJCB). 2014. pp. 1-6
- [44] Chen Q, Huang J, Feris R, Brown LM, Dong J, Yan S. Deep Domain Adaptation for Describing People Based on Fine-Grained Clothing Attributes; 2015. pp. 5315-5324
- [45] Zhu J, Liao S, Yi D, Lei Z, Li SZ. Multi-label CNN based pedestrian attribute learning for soft biometrics. In: Proceedings of 2015 International Conference on Biometrics, ICB 2015; 2015. pp. 535-540
- [46] Almudhahka NY, Nixon MS, Hare JS. Automatic semantic face recognition. In: 2017 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017); 2017. pp. 180-185
- [47] Lin D, Tang X. Recognize high resolution faces: From macrocosm to microcosm. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 2. IEEE; 2006. pp. 1355-1362
- [48] Dantcheva A, Velardo C, Dangelo A, Dugelay J-L. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*. 2011;51(2):739-777
- [49] An L, Chen X, Kafai M, Yang S, Bhanu B. Improving person re-identification by soft biometrics based reranking. In: 2013 7th International Conference on Distributed Smart Cameras, ICDSC 2013; 2013. pp. 4-9