Received: 17 March 2020 | Revised: 2 September 2020 | Accepted: 15 September 2020

DOI: 10.1049/bme2.12001

ORIGINAL RESEARCH PAPER

The Institution of Engineering and Technology WILEY

A multimodal retina-iris biometric system using the Levenshtein distance for spatial feature comparison

Vincenzo Conti[1] | **Leonardo Rundo**[2,3] | **Carmelo Militello**[4] | **Valerio Mario Salerno**[1] | **Salvatore Vitabile**[5] | **Sabato Marco Siniscalchi**[1]

[1]Faculty of Engineering and Architecture, University of Enna KORE, Enna, Italy

[2]Department of Radiology, University of Cambridge, Cambridge, UK

[3]Cancer Research UK Cambridge Institute, Cambridge, UK

[4]Institute of Molecular Bioimaging and Physiology, Italian National Research Council (IBFM-CNR), Cefalù, Italy

[5]Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University of Palermo, Palermo, Italy

Correspondence

Vincenzo Conti, Faculty of Engineering and Architecture, University of Enna KORE, Enna 94100, Italy.
Email: vincenzo.conti@unikore.it

Abstract

The recent developments of information technologies, and the consequent need for access to distributed services and resources, require robust and reliable authentication systems. Biometric systems can guarantee high levels of security and multimodal techniques, which combine two or more biometric traits, warranting constraints that are more stringent during the access phases. This work proposes a novel multimodal biometric system based on iris and retina combination in the spatial domain. The proposed solution follows the alignment and recognition approach commonly adopted in computational linguistics and bioinformatics; in particular, features are extracted separately for iris and retina, and the fusion is obtained relying upon the comparison score via the Levenshtein distance. We evaluated our approach by testing several combinations of publicly available biometric databases, namely one for retina images and three for iris images. To provide comprehensive results, detection error trade-off-based metrics, as well as statistical analyses for assessing the authentication performance, were considered. The best achieved False Acceptation Rate and False Rejection Rate indices were and 3.33%, respectively, for the multimodal retina-iris biometric approach that overall outperformed the unimodal systems. These results draw the potential of the proposed approach as a multimodal authentication framework using multiple static biometric traits.

1 | INTRODUCTION

The evolution of technology and the increasing presence of distributed services and resources have boosted the demand for secure access systems. The current frontier and the state of advancement of recognition and authentication systems coincide with the use of multimodal biometric approaches [1]. The impossibility of obtaining a correct acquisition of one biometric trait in certain conditions, such as face, fingerprint, palm of the hand, voice, makes unimodal systems useless in environments, where the security needs are high. A combination of two or more biometric traits, called "multimodal systems", is preferred to the classic unimodal approaches due to their higher reliability in all critical environments, where stringent security conditions have to be met [2–5].

There exist numerous unimodal and multimodal biometric systems, offering both software and hardware implementations [6,7]. In the literature, some of those systems—combining different biometric traits—are implemented using features in the spatial/structural domain [8,9], whilst the majority in the frequency domain [10–12]. Even though the terminology used for fusion in multimodal biometric systems is standardised in [13], it is difficult to define a proper guideline in terms of both the various processing stages and specific algorithms that are generally applicable to multiple biometric traits/features [1]. However, the use of multiple biometric measurements from independent biometric sensors, traits, or algorithms can achieve accurate and reliable performance [13]: multimodal systems improve the robustness and reduce the risk, by making decisions that rely upon the combination of distinct biometric traits under different circumstances. To some extent, frequency-based and space-based approaches might be seen as complementary and there is no well-established clue for choosing one out of the two strategies. Frequency-based techniques certainly involve a higher computational load, mainly due to the involved bidirectional mapping into the transformed domain (e.g. Fourier transform, and wavelets), with respect to the spatial approaches: this can represent a limitation in the case of system deployment onto hardware platforms with limited resources [3,14]. This work has the objective of devising an innovative methodology for the realisation of biometric, unimodal, and multimodal authentication systems based on the spatial domain, by analysing the same type of micro-features. More specifically, the novel part focuses on the possibility of creating a single system for the feature extraction, processing, and comparison algorithm based on the Levenshtein Distance (LD) [15,16]. The final phase leverages a fusion at the level of comparison score [2,17]. To the best of our knowledge, the authors are the first to propose an authentication system based on iris and retina entirely implemented on spatial domain and using the LD for comparison algorithm. Furthermore, another reason for choosing the iris and retina as biometric traits is the high level of uniqueness, performance, universality, and circumvention [12]. In the literature, only a few publications propose a multimodal authentication system using iris and retina; the most representative ones are [12,18,19], even though they relied upon the frequency domain alone. Furthermore, the aforementioned work involved vessel-based matching by using feature points, that is minutiae points. It is well-known that vessel segmentation and minutiae point extraction are time-consuming tasks. This is another challenge directly tackled by our contribution, aiming at realising a system that can also be used in real-time by following a novel approach with respect to the methods proposed in the state-of-the-art so far [20]. For the biometric terminology, we relied upon standardised vocabularies [13,21].

The main contributions of the work are:

- Most of the literature work based on iris leveraged transforms in the frequency domain (e.g. Daugman approach [22]). We aimed to investigate a system entirely developed in the spatial domain
- Pair of biometric traits (i.e. iris and retina) not quite commonly used in the literature on multimodal systems
- The features extracted from the retina and iris images are homogeneous and analogous to the well-known fingerprint features [23]: end-points and bifurcations for retina, end-points and local peaks (similar to bifurcations) for iris
- The unimodal iris and retina biometric systems are implemented using the same algorithms and the same processing pipeline
- A general and universal approach that can be extended to any system operating in the spatial domain and based on static biometric features
- All databases exploited in the experiments are public, freely available and used in papers already published in the literature. This denotes a necessary condition in order to guarantee result repeatability and benchmarking

The remainder of this manuscript is structured as follows. Section 2 outlines the related work. Section 3 describes the used recognition methodology. Section 4 describes the proposed multimodal retina-iris biometric system, whilst Section 5 shows the experimental results. Section 6 analyses and discusses comparisons against the state-of-the-art. Finally, conclusive remarks and future directions are provided in Section 7.

2 | RELATED WORK

Traditionally, unimodal biometric systems might have many limitations [1,2], which can be addressed and reduced by combining multiple biometric traits [3,24]. To this purpose, multimodal biometric systems are based on different biometric traits and/or introduce various fusion schemes on the extracted features. The fusion can be performed at: (*i*) feature-level [25]; (*ii*) comparison score level [12]; (*iii*) decision level, also by exploiting Soft Computing techniques [24]. Many researchers have experimentally demonstrated that the fusion process is effective, because fused scores provide a notably higher discrimination than individual scores [2,25]. Regarding the extracted features, biometric recognition systems can be implemented using features in the frequency domain (e.g. Fourier/wavelet transforms [12], and log-Gabor [11]) or to the spatial domain (e.g. position, orientation, and structural properties [8,9]). Such results have been achieved by means of a variety of fusion techniques. In what follows, some relevant studies about biometric recognition and authentication systems, based on iris or/and retina, are briefly described. Waheed et al. in [9] proposed a retina recognition matching without using minutiae points, resulting in a simple and fast non-vascular based retina recognition system. A similarity measure was computed via features based on illuminance, contrast, and structural features from a colour retina image The extracted attributes were combined by means of an empirically optimised function to generate a similarity score between two candidate images. The matching decision was obtained according to the highest score value. This approach was tested on two retinal image databases collected from local sources, namely, Retina Identification database (RIDB) and Armed Forces Institute of Ophthalmology (AFIO). The proposed system was tested by calculating the False Rejection Rate (FRR) and False Acceptance Rate (FAR) and achieved an average identification rate of 92.50% and 85.75% on the RIDB and AFIO databases, respectively.

Lajevardi et al. in [10] presented an automatic retina verification framework based on the Biometric Graph Matching (BGM) algorithm. The retinal vasculature was extracted using a family of matched filters in the frequency domain and morphological operators. Then, the retinal templates were defined as formal spatial graphs derived from the retinal vasculature. The BGM algorithm—which is a noisy graph matching algorithm, robust to translation, non-linear distortion, and small rotations—was used to compare retinal templates. The BGM algorithm exploited the graph topology to define three pairwise distance measures between graphs. A Support Vector Machine (SVM) classifier was used to distinguish genuine and imposter comparisons. By using single as well as multiple graph measures, the classifier achieved complete separation on a training set of images from the VARPA Retinal Images for Authentication (VARIA) database (60% of the data), achieving comparable performance with the state-of-the-art for retina verification. Because the available dataset was small, Kernel Density Estimation (KDE) of the genuine and imposter score distributions of the training set were used to assess the performance of the BGM algorithm. In the one-dimensional case, the KDE model was validated on the test set: a 0 Equal Error Rate (EER) on testing showed that the KDE model is a good fit for the empirical distribution. For the multiple graph measures, a novel combination of the SVM boundary and the KDE model was devised to obtain a fair comparison with the KDE model for the single measure. A clear benefit in using multiple graph measures over a single measure to distinguish genuine and impostor comparisons was revealed by boosting the theoretical error between 60% and more than two orders of magnitude.

Saha et al. in [12] proposed a multimodal user authentication system developed by feature-level fusion of iris and retina recognition. The 'curse-of-dimensionality' problem arising in feature-level fusion was minimised to some extent by applying a Principal Component Analysis (PCA) on the augmented feature template in the frequency domain. To validate this approach, iris and retina images obtained from the Indian Institute of Technology Delhi (IITD) and Digital Retinal Images for Vessel Extraction (DRIVE) databases, respectively, were used. The achieved Recognition Rate (RR) was 98.37% for the multimodal approach, whereas it was 96.74% and 94.56% for separate iris and retina recognition, respectively.

In Choras [18], the iris and retina features were combined for biometric recognition. In this multimodal biometric system, the Gabor transform was used to extract the features from the iris and retina to perform a fusion at feature-level. The extraction of the features was not homogeneous, because different filters and features were used for the iris and retina. The fused template was obtained by the concatenation of iris and retina descriptors. No detail was provided by the authors on the databases used, number of images and comparison methods, hampering the quantitative evaluation of this approach. Also Modarresi et al. [26] used a fusion at feature-level. In fact, retina and iris features (mean energy, standard deviation, entropy, contrast and homogeneity) were extracted, by means of the Contour let transform, and subsequently fused. The Hamming Distance (HD) was used for the matching to provide a higher accuracy than unimodal systems, by obtaining an EER = 0.0413%. Sen and Islam [27] developed a multimodal biometric system that uses fusion of iris and retina recognition. Iris feature extraction and template creation were performed by using 1D log-Gabor filter. The retina descriptor was obtained by extracting retinal blood vessels and, successively, the discrete wavelet transform was used for feature extraction. The fused multimodal template was obtained by means of concatenation of two (iris and retina) feature vectors. The achieved results showed that the RR of the multimodal system outperformed the iris and retina unimodal recognition systems. The achieved authentication results for this multimodal system are: FAR = 2.041%, FRR = 0% and RR = 97.959%. Considering that FAR > FRR, this represent a represents an anomalous 'operating point' with respect to biometric good practices. Both Modarresi et al. [26] and Sen and Islam [27] used the Chinese Academy of Sciences, Institute of Automation (CASIA) and DRIVE databases for iris and retina.

Latha et al. [19] proposed a multimodal hybrid approach based on images of both (left and right) irises and retinal images. The iris characteristics were encoded by means of a 1D log-Gabor filter to extract the underlying information and generate the binary iris template that is used in the HD matching. The retina template was composed of the blood vessel intersection points. From the processing point of view, this system is hybrid because the iris and retina templates contained frequency and spatial domain information, respectively. Iris and retina images were taken from CASIA v3 and VARIA databases, respectively. Finally, the irises and retina recognition scores were combined together using a simple weighted-sum fusion rule. This system obtained ERR = 0.01% and RR = 99.3%.

Kihal et al. in [28] proposed a multimodal biometric authentication with a multimodal ocular biometric system based on the iris pattern and the three-dimensional shape of the cornea. In this work, the authors showed that the cornea can be used as a biometric trait, by then proposing an intraocular fusion with iris features to improve the overall performance of the system. Feature extraction was performed by modelling the shape of the cornea with a Zernike polynomial expansion, whilst the iris texture was analysed with a typical methodology via Gabor filtering and phase encoding. Finally, the fusion was performed at the matching score level using min, max, sum and weighted-sum rules. The experimental results on a new database constructed for this bimodal study showed an EER decreasing to 0% with the weighted-sum rule.

Un-like the multimodal approaches described so far, Meenakshi et al. [29] proposed a fuzzy vault framework to encrypt both the retina and iris templates to enhance the security in biometric systems. The proposed multimodal fuzzy vault was tested with spatial features extracted from retina and iris images, obtained from the DRIVE and Chinese University of Hong Kong (CUHK) databases, respectively. Since the purpose of this study is to improve safety, no recognition result was reported.

In this work, the main objective is to devise a novel fusion approach based on a comparison algorithm that can improve the state-of-the-art on the retina-iris traits. With the goal of devising a flexible and extensible biometric framework, we decided to exploit the spatial domain, considering that most literature approaches—based on iris and/or retina biometry—proposed solutions in the frequency domain. Interestingly, our approach leverages a fusion at a comparison score-level [17], and the proposed minutiae-based comparison algorithm, using on features extracted in the spatial domain, is invariant to roto-translation problems.

3 | REMARKS ON IRIS BIOMETRY, RETINA BIOMETRY AND LEVENSHTEIN DISTANCE

The biometric systems developed in the domain of space and focused on the comparison of physiological traits are typically based on the extraction, processing, and comparison of peculiar micro-features, commonly called minutiae, obtained from the scanned digital images [8,10,30].

Once the retrieval phase is accomplished, either manually or through optical readers, the digital image is processed using different filters, with the aim of cleaning it from negligible details and clearly identifying the fundamental features (i.e. lines, crests, minutiae) useful for comparison purposes. Typically, in all existing systems, the first step includes the segmentation and binarisation of the image, of which the goal is to transform the digital image, initially in grey-scale, into a binary image. Subsequently, a thinning operation is performed, which reduces the size of the reference lines contained in the binary image and represents them as lines composed of a single pixel of thickness. Next, the identification of the characteristic points via search algorithms and selection of the minutiae are executed. Finally, the obtained data are processed and matched for the biometric task. These approaches usually include a phase of comparison at the level of Cartesian positioning amongst the minutiae detected, measuring the Euclidean Distance (ED) or similar metrics, between the points and verifying if this measurement is lower than a pre-set threshold. However, roto-translation problems could perceptibly affect the final result, since two equal images could have undergone a roto-translation with a measurement higher than the fixed threshold value. In such a case, the high ED between the minutiae would not allow the system to make the appropriate assessments, yielding an incorrect negative outcome (i.e. false negative).

The proposed method, on the other hand, aims to overcome this problem by making use of distance vectors. The measured numerical references remain, indeed, the same even in the case of a strong roto-translation, since the reference is set to the mutual EDs of the minutiae rather than to their individual positions. Thus, the problem of roto-translation would be eliminated. Our work presents a comparison phase, which is not based on the classic ℓ_2 ED, but exploits the LD [15,16]. The LD is typically used to measure the difference of two strings, defined on an alphabet Ω of symbols, and to provide the information of the related additions, deletions, and settings required to identify the two sequences as the same sequence. The goal of our methodology is to standardise the extraction phase of the characteristics from different types of images, the features themselves in the spatial domain and, eventually, the comparison algorithm. The objective is to devise a multimodal recognition system in the spatial domain using a novel and universal approach for all types of static biometric features. For this reason, we choose to avoid the use of a hybrid approach for each unimodal authentication sub-system, combining spatial and frequency biometric characteristics. This strategy allows for a high degree of homogeneity in the treatment of the most common biometric features.

The basic concepts on retina-based and iris-based biometry, as well as the formalism of the LD, are introduced in the following sections. Finally, the basic concepts about the LD are introduced.

3.1 | Retina-based biometry

The retina acquisition is based on the uniqueness of its vascular scheme, and, for this reason, the main features extracted from the retinal vascular tree are usually employed as a reference for comparisons between retinas, namely:

- optic disc: yellow-whitish area of the retina, which is the point of convergence of the nerve fibres for the optic nerve formation. Starting from this branch, the vessels form the retinal vascular tree;
- bifurcations of the vascular tree: points where a branch divides into two sub-capillary branches;
- terminations of the vascular tree: end-points of the individual capillaries.

The proposed system focuses on bifurcations and terminations because they are sufficient for a recognition system. The reliable detection of the optic disc and the certain positioning inside the retinal vascular tree is currently the subject of numerous approaches [31]. Indeed, the optical disk is not included because, due to its considerable size that occupies a larger area than the aforementioned characteristic points, it is immediately detectable but does not provide a precise positioning in terms of (x, y) coordinates. Figure 1 shows an example of retina processing, from retinal vascular tree extraction to minutiae detection.

3.2 | Iris-based biometry

In this work, we process iris images in the spatial domain [32] leveraging peculiar features extracted from the scanned image segmentation process. We therefore do not follow the Daugman method [22], such as in [33], which identifies an iris-code relying upon the frequency domain. Our approach takes care of detecting and isolating the ring part of the iris in a meticulous manner (to the best of one pixel), excluding negligible details and possible concealments due to eyelids, reflections, shadows, and eyelashes. Furthermore, we perform an image normalisation procedure, based on linearised representation of iris image with fixed dimensions allowing for quantitative comparisons (Figure 2). This is necessary because the iris area for a given eye is not always constant but varies depending on the dilation of the pupil. For this reason, two images of the same iris might have distinctive characteristics whilst having the same spatial location, under different conditions [34].

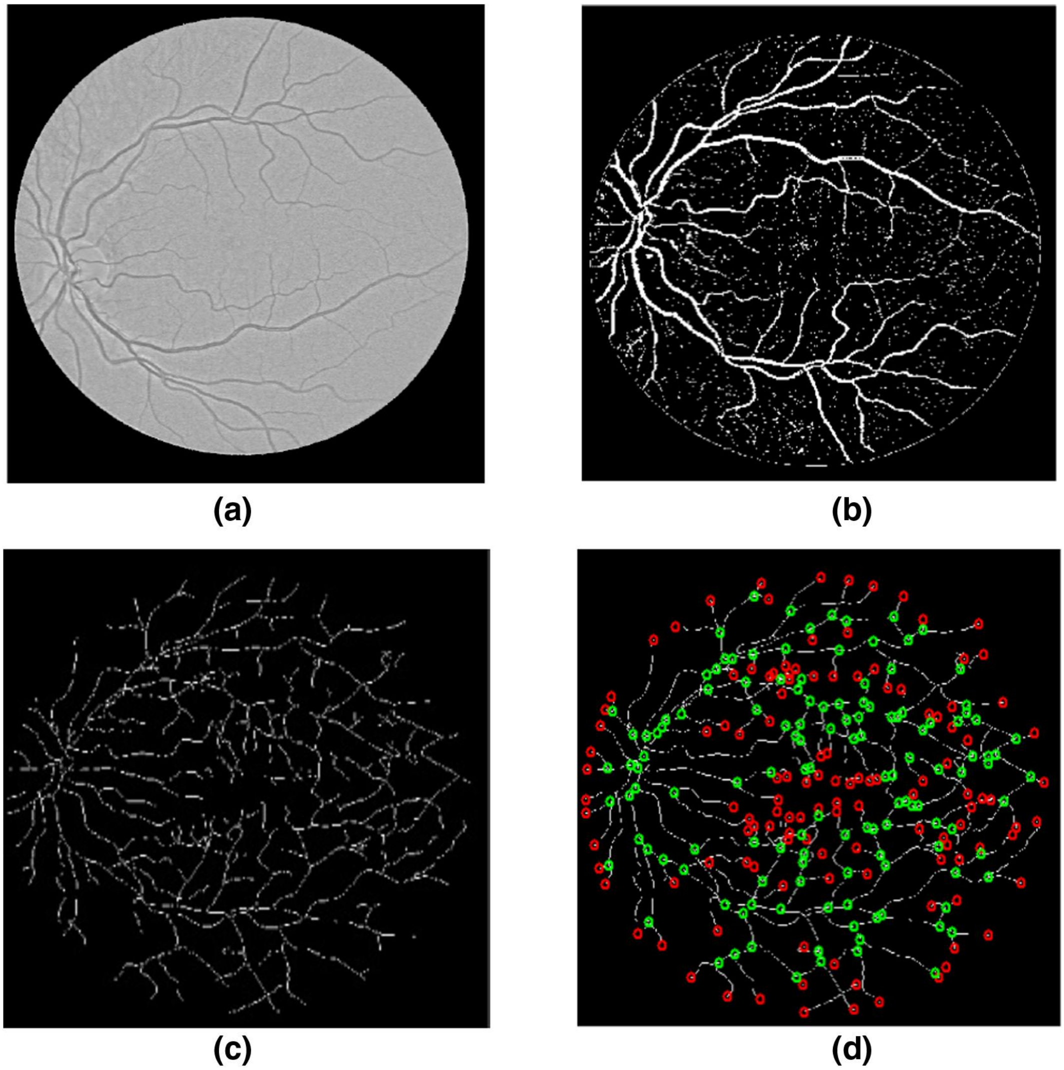

FIGURE 1 (a) Retina scan; (b) image after binarization; (c) image after thinning and (d) detected minutiae

Accordingly, the iris is segmented into the following micro-features:

- core: variation of the circular symmetry of the pupil contour;
- valleys: regions that are characterised by a higher pigmentation within the iris, visually corresponding to areas with a darker shade;
- collarettes: ideal jagged outline that divides the iris into two zones, the former is near the pupil with a darker shade, and the latter is next to the sclera having a lighter shade. The collarettes are usually characterised by the presence of a thick ray of fibres, which have a darker colour than the rest of the rainbow crown;
- radiant channels: minutiae present outside the boundary defined by the collarettes. They have the shape of arcs of circumference and are characterised by a darker shade than the pigmentation that surrounds them.
- Accurate iris image analyses revealed a rich textural content. In particular, the collarette region conveys the highest texture information amongst the iris-derived micro-features [35]. The proposed system employs only collarettes thanks to its discriminant characteristics in iris recognition, following the well-known literature approach proposed in [36]. A function that detects the contours is then applied to the collarette image obtained by the segmentation. Two types of characteristic points are investigated on the resulting image:
- terminal points, wherein the detected contour stops;
- local peak points, wherein the contour of the collarette changes in height.

The output coordinates of the detected characteristic points are used for the comparison between images.

3.3 | Levenshtein distance

The LD [15,16], also called edit distance, quantifies the difference between two strings: this metric determines how

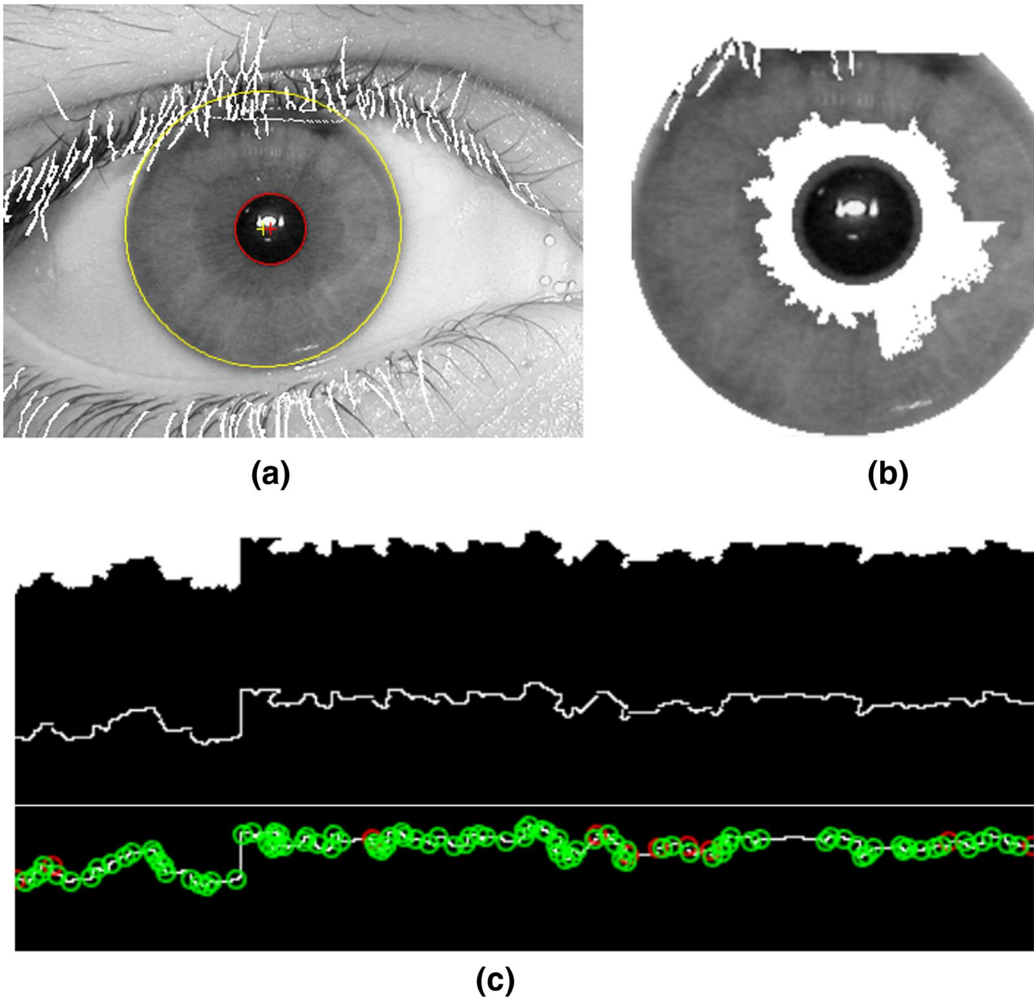

FIGURE 2 (a) Iris scan; (b) segmented iris with highlighted collarette (in white) and (c) collarette extracted from the iris

much two strings are dissimilar. This concept has been effectively applied to multiple areas of interest, such as in spelling check algorithms or in similarities search between images, sounds, texts, sequences. With particular interest to biometry, the LD was only applied to unimodal methods, such as in the case of binary iris-code matching (efficiently outperforming the HD) [33].

Formally, the LD between two strings $\mathbf{s}_1$ and $\mathbf{s}_2$ is defined as the minimum number of elementary modifications that allow us to transform the string $\mathbf{s}_1$ into the string $\mathbf{s}_2$. Elementary modifications denote the following operations:

- deletion of a character;
- insertion of a character;
- replacement of a character with a different one.

This metric, therefore, represents the cost of the optimal alignment of two strings; the total cost of alignment is calculated by summing-up the costs of each position. The algorithm for calculating the LD between two strings is based on the following observation. Let $\mathbf{s}_1 = a_1 a_2 \cdots a_m$ and $\mathbf{s}_2 = b_1 b_2 \cdots b_n$ be two strings. Let $e\ (i, j)$ denote the LD between the two prefixes $\mathbf{p}_1 = a_1 a_2 \cdots a_i$ and $\mathbf{p}_2 = b_1 b_2 \cdots b_j$ of $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively. If $e\ (i-1, j-1)$, $e\ (i-1, j)$ and $e\ (i, j-1)$ are known, we can compute $e\ (i, j)$.

Each of the three values corresponds to an excellent alignment:

- $e\ (i-1, j-1)$ is the cost of the optimal alignment of $a_1 \cdots a_{i-1}$ and $b_1 \cdots b_{j-1}$;
- $e\ (i-1, j)$ is the cost of the optimal alignment of $a_1 \cdots a_{i-1}$ and $b_1 \cdots b_1$;
- $e\ (i, j-1)$ is the cost of the optimal alignment of $a_1 \cdots a_i$ and $b_1 \cdots b_{j-1}$.

Each of the three alignments can be extended to an alignment between $a_1 \cdots a_i$ and $b_1 \cdots b_j$:

- aligning a_i with b_j;
- aligning a_i with a gap;
- aligning a gap with b_j.

To achieve the optimality, it is necessary to select the operation that obtains the minimum cost alignment. Note that the generic value $e\ (i, j)$ depends only on $e\ (i-1, j-1)$, $e\ (i, j-1)$ ed $e\ (i-1, j)$, and can be calculated by means of Equation (1):

$$e(i,j) = \min \begin{cases} e(i-1, j-1) + c(a_i, b_j) \\ e(i, j-1) + 1 \\ e(i-1, j) + 1 \end{cases} \tag{1}$$

We can conveniently store these values in an $(m + 1)\times(n + 1)$ matrix, such as in Table 1, where the row and column indices start from zero and ε denotes the empty string.

It is clear that when the alignment of an empty string with a string $x_1\cdots x_i$ costs i, i characters must be inserted into the empty string ε to obtain $x_1\cdots x_i$: $e\ (0, i) = e\ (i, 0) = i$. The alignment between two empty strings has no cost: $e\ (0, 0) = 0$. The LD matrix in Table 1 can therefore be rewritten as in Table 2, and it can be incrementally built row-by-row, that is from left to right, and from top to bottom. The value $e\ (m, n)$ represents the LD between $\mathbf{s}_1$ and $\mathbf{s}_2$. Algorithm 1defines the procedure for computing the LD matrix $\mathbf{M}_L$ between $\mathbf{s}_1 = a_1\cdots a_m$ and $\mathbf{s}_2 = b_1\cdots b_n$.

Algorithm 1 Pseudo-code of the iterative procedure for constructing the LD matrix $\mathbf{M}_L$

Require Strings $\mathbf{s}_1 = a_1\cdots a_m$ and $\mathbf{s}_2 = b_1\cdots b_n$
Ensure Matrix $\mathbf{D}_E$

Initialise an $(m + 1)\times(n + 1)$ matrix called $\mathbf{M}_L$
$\mathbf{M}_{L_{i,0}} \leftarrow i$, for $0 \le i \le m$
$\mathbf{M}_{L_{0,j}} \leftarrow j$, for $0 \le j \le n$
for $i = 1, \ldots, m$ **do**
  for $j = 1, \ldots, n$ **for**
    if then
      $c = 0$
    else
      $c = 1$
    end if
    $\mathbf{M}_{L_{i,j}} \leftarrow \ldots$
    $\min\{\mathbf{M}_{L_{i-1,j-1}} + c, \mathbf{M}_{L_{i,j-1}} + 1, \mathbf{M}_{L_{i-1,j}} + 1\}$
  end for
end for

TABLE 1 LDs between the two prefix strings $\mathbf{p}_1 = a_1a_2\cdots a_i$ and $\mathbf{p}_2 = b_1b_2\cdots b_j$

	ε	b_1	b_2	…	b_n
ε	e (0, 0)	e (0, 1)	e (0, 2)	…	e (0, n)
a_1	e (1, 0)	e (1, 1)	e (1, 2)	…	e (1, n)
a_2	e (2, 0)	e (2, 1)	e (2, 2)	…	e (2, n)
…	…	…	…	…	…
a_m	e (m, 0)	e (m, 1)	e (m, 2)	…	e (m, n)

Abbreviations: LDs, Levenshtein distance; P, prefixes.

TABLE 2 LD matrix $\mathbf{M}_L$ (in Table 1) rewritten considering that $e\ (0, i) = e\ (i, 0) = i$, when the alignment of an empty string ε with a string $x_1\cdots x_i$ costs i, exactly i characters must be inserted into the empty string to obtain $x_1\cdots x_i$

		b_1	b_2	…	b_n
ε	0	1	2	…	e (0, n)
a_1	1	e (1, 1)	e (1, 2)	…	e (1, n)
a_2	2	e (2, 1)	e (2, 2)	…	e (2, n)
…	…	…	…	…	…
a_m	m	e (m, 1)	e (m, 2)	…	e (m, n)

Abbreviations: ε, empty string; LDs, Levenshtein distance; ML, LD matrix.

4 | PROPOSED MULTIMODAL RETINA-IRIS BIOMETRIC SYSTEM

Recognition systems based on physiological traits typically operate on images and Cartesian coordinates. The minutiae detected by the process described in Section 3 are stored in $M\times2$ matrices as Cartesian coordinates (x, y), where M denotes the number of the extracted minutiae that might be different for each matrix/image. Images are usually compared by measuring the ED between the detected minutiae. If this distance is lower than a pre-set threshold value, the comparison provides a positive outcome; otherwise, it yields a negative outcome.

The proposed solution aims to apply the LD algorithm on the minutiae detected by the retina- and iris-based recognition systems. The expected result is to obtain the minimum EDs whether the two scans belong to the same individual, or high ED values if the two scans belong to distinct individuals.

Figure 3 shows the overall block scheme of the proposed system. Each component is described in the following sections.

4.1 | Biometric spatial feature extraction

The following description is valid for unimodal recognition systems based on either retina or iris scanning, analysing the Cartesian coordinates of the previously extracted minutiae. Therefore, we obtain two-dimensional matrices $M\times2$, corresponding to the Cartesian coordinates of the minutiae computed from the processing of the two biometric images to be compared. The minutiae under consideration are: (i) terminations/bifurcations, with regard to the retina-based recognition; (ii) terminations/local peaks, when considering iris-based recognition. Thus, these operations are independently performed on both retina-based and iris-based minutiae matrices.

We use a 5×2 coordinate matrix, which is suitable to represent the preliminary operations to the subsequent comparison step on a reduced scale (Table 3), to exemplify our procedure. Each row represents a minutia, identified by its coordinates (x, y), within the image. The minutiae are sorted according to the value of the x coordinate.

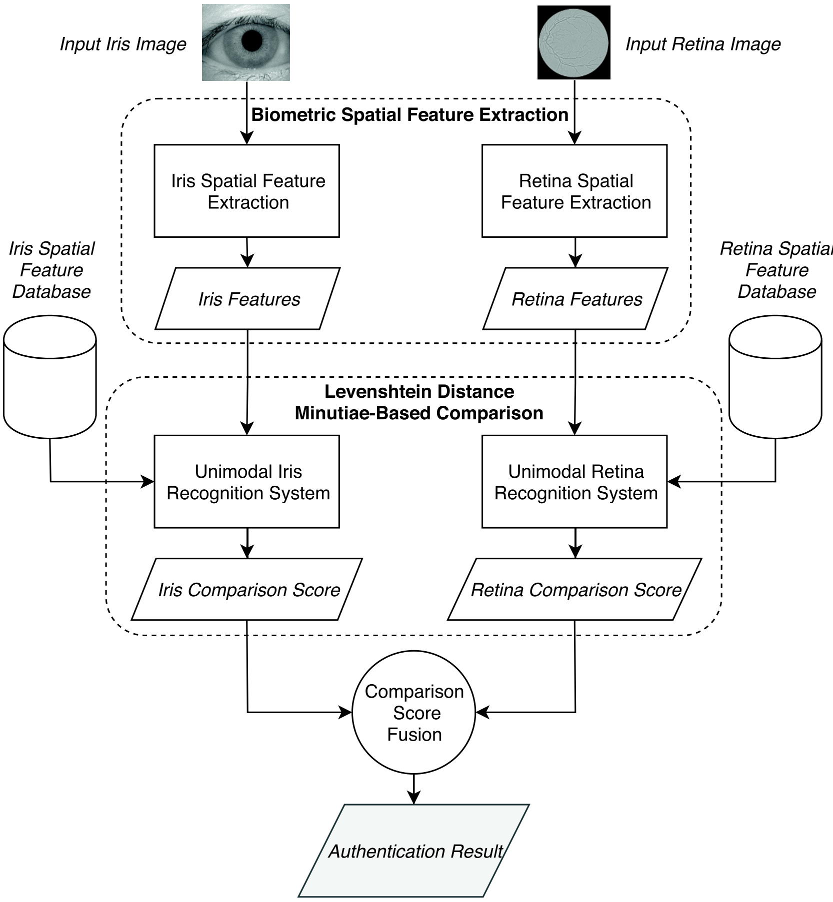

FIGURE 3 Overall block scheme of the proposed multimodal retina-iris system that performs a comparison score fusion based on the Levenshtein distance

TABLE 3 Example of coordinate matrix

x	y
11	130
17	310
39	291
96	122
155	97

TABLE 4 ED matrix obtained from the coordinate matrix in Table 4

0	180.10	163.42	85.38	147.73
180.10	0	29.07	203.92	253.80
163.42	29.07	0	178.35	226.04
85.38	203.92	178.35	0	64.08
147.73	253.80	226.04	64.08	0

Abbreviation: ED, Euclidean distance.

In the first step, the ED between the minutiae contained in the two-dimensional coordinate matrix is obtained. This operation allows us to quantitatively measure the distance between all the minutiae of the same image, obtaining informative values that can be properly assessed in the following phases. The result of this operation consists of an $M \times M$ matrix, in which each row stores the EDs between a given minutia and the others: the first row contains the ED values measured between the first minutia of the coordinate matrix and the others, the second row considers the second detail as its reference, and so on. This direct correspondence between the lines of the coordinate matrix and the rows of the $\mathbf{D}_E$ matrix is extremely important and is maintained for all the subsequent processing steps. Table 4 shows the $\mathbf{D}_E$ matrix obtained starting from the coordinate matrix used as an example (Table 3).

In the second step, the ED matrix is sorted in a row-wise ascending order: the values of each row are neatly arranged from the minimum to the maximum values. Table 5 shows the resulting ED matrix corresponding to the example in Table 3. As expected, the first column is composed of only 0 values, since it corresponds to the ED between the detail taken into consideration and itself. This column can therefore be removed because it is ineffective for comparison purposes. Table 6 shows the final $\mathbf{D}_E$ matrix.

The operations executed so far result in a $\mathbf{D}_E$ matrix $M \times N$, with $N = M-1$, which consists of M rows with the ED values arranged in ascending order and are obtained starting from M minutiae contained in the coordinate matrix used in the example. Each row can be interpreted as a real number string that contains unique information on the associated reference minutiae, as well as on the corresponding EDs measured from them. Therefore, not only information on the Cartesian positioning of the single minutia is obtained, but also a complete mapping of all the other characteristic points detected by the system using the ED as a metric.

4.2 | Levenshtein distance for minutiae-based comparison

The LD overcomes the problem of descriptors with variable lengths coming from different images that represent either the same biometric trait or different biometric traits [15,16]. The points characterising the various biometric traits in the spatial domain cannot be always the same for reasons related to noise, roto-translation problems, extraction algorithms and the types of points themselves. The LD, designed to evaluate if parts of text sequences belong to the same sequence [33], allows us to manage and perform the matching with biometric descriptors of variable lengths.

Here, we show how to compare the $M \times 2$ coordinate matrices obtained in the previous step. Our fusion scheme proposes an algorithm that consists of a series of row-by-row tests between the coordinate matrices, aiming at verifying the presence of similarities between them. Once the direct correspondence between each minutia detail of the $M \times 2$ coordinate matrix and each row of the corresponding $\mathbf{D}_E$ matrix is set—i.e. the first row contains the ED values measured between the first minutia of the coordinate matrix and the others, the second row has the second detail as its reference, and so on—a measurement of the ED between the minutiae to compare is performed. If the value is lower than a threshold value, which is just useful to avoid comparisons between too different minutiae (thus assuming the inequality) and the consequent system overload, the corresponding rows are compared using the LD algorithm that interprets them as two sequences.

The function associated with this algorithm requires six input parameters:

TABLE 5 ED matrix sorted in a row-wise ascending ordered

0	85.38	147.73	163.42	180.10
0	29.07	180.10	203.92	253.80
0	29.07	163.42	178.35	226.04
0	64.08	85.38	178.35	203.92
0	64.08	147.73	226.04	253.80

Abbreviation: ED, Euclidean distance.

TABLE 6 $\mathbf{D}_E$ matrix after removing the 0-valued column

85.38	147.73	163.42	180.10
29.07	180.10	203.92	253.80
29.07	163.42	178.35	226.04
64.08	85.38	178.35	203.92
64.08	147.73	226.04	253.80

- the first sequence;
- the second sequence;
- the length of the first sequence;
- the length of the second sequence;
- two distance-based threshold values to increase the reliability in minutiae comparison score computation and fusion, respectively.

This procedure is sequentially executed for each minutia and the $\mathbf{d}_L$ vector is updated accordingly. Each row of the first $\mathbf{D}_E$ matrix is compared with each row of the second $\mathbf{D}_E$ matrix, after verifying the ED within the threshold. The result of the developed algorithm consists of a vector of length M that contains in each cell the minimum value computed by the LD algorithm for each minutia. Then, the first cell of the output vector $\mathbf{d}_L$ denotes the minimum LD value obtained between the first row of the first $\mathbf{D}_{E1}$ matrix and all the rows of the second $\mathbf{D}_{E2}$ matrix; the second cell denotes the smallest value computed between the second row of the first $\mathbf{D}_{E1}$ matrix and all the rows of the second $\mathbf{D}_{E1}$ matrix, and so on. The reference is set to the smallest value, since the aim of this algorithm is to find the similarity between the rows of the matrices $\mathbf{D}_{E1}$ and $\mathbf{D}_{E2}$. An LD value close to zero, indicates that the two compared sequences are similar.

If one of the tests between sequences returns the value 0, it means that the algorithm stopping condition (i.e. the exact equality according to the LD) is reached and the search ends. In all the other cases, the tests continue aiming at determining the calculated minimum value. If a match is not found (i.e. the LD is higher than the threshold value), an out-of-range numerical value is assigned to identify the failed test. At the end of the algorithm execution, the yielded result consists of the $\mathbf{d}_L$ vector of length M (i.e. the dimension of the largest matrix) that contains either the minimum values for each tested minutiae detail, or the best correspondences between the sequences analysed by the LD algorithm. Figure 4 depicts the flow diagram of the proposed algorithm.

Variables and constants used in the flow chart are defined as follows:

- $\mathbf{C}_1$ and $\mathbf{C}_2$ are the coordinate matrices of all minutiae of the first and second biometric features, respectively;
- $\mathbf{D}_{E1}$ and $\mathbf{D}_{E2}$ are the matrices of the EDs amongst all minutiae of the first and second biometric features, respectively, lower than a set threshold θ_E;
- θ_E is the maximum ED threshold for considering minutiae pairs;
- θ_L is the LD threshold able to determine whether the comparison was successful;
- $\mathbf{d}_L$ is the distance vector obtained by the comparison scheme based on the LD.

To verify the correspondence between the $\mathbf{D}_E$ matrices compared, the average μ of the values contained in the $\mathbf{d}_L$ is calculated. In this phase, however, this average value cannot be considered as a precise and reliable indicator, because the

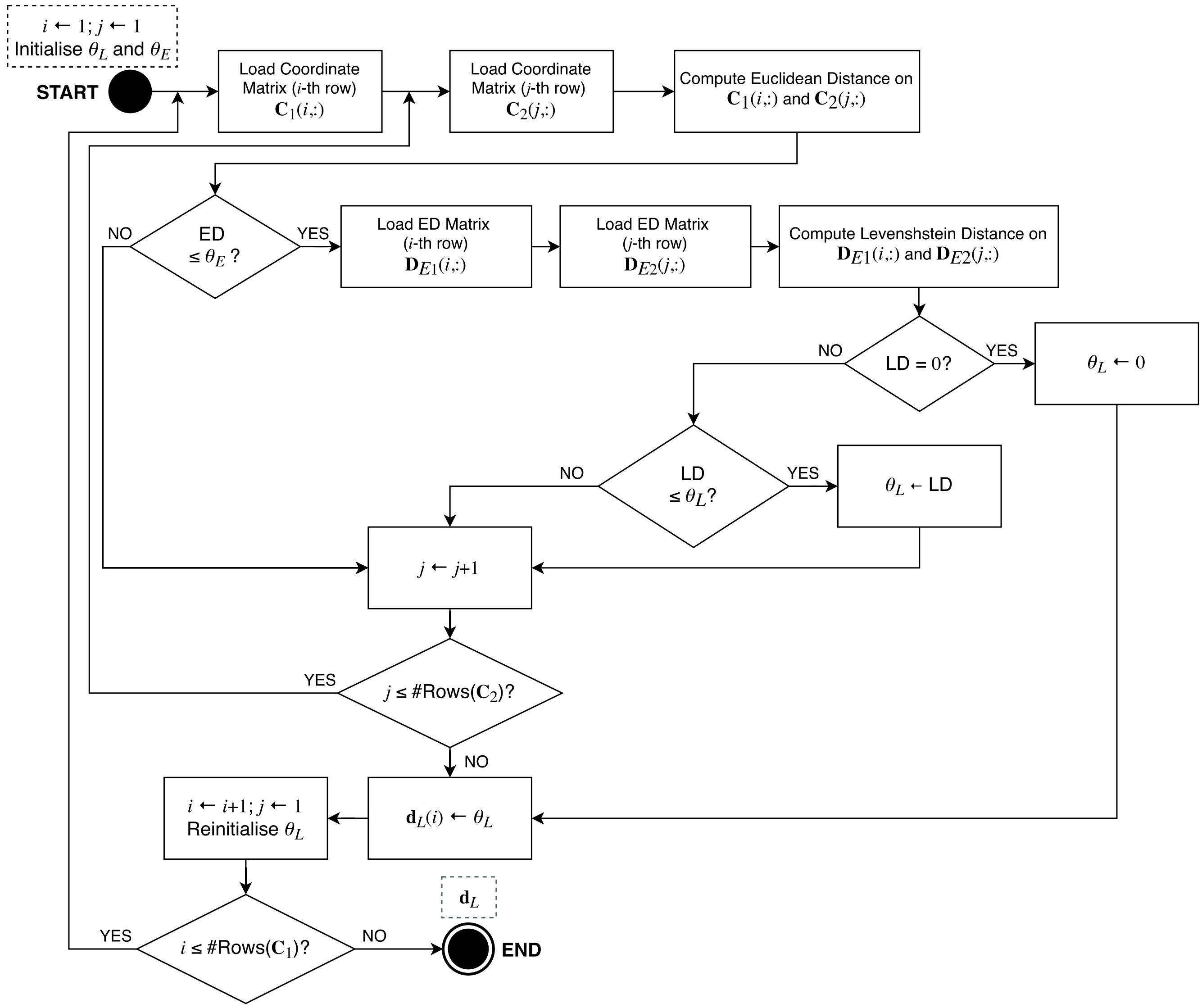

FIGURE 4 Overall flow diagram of the iterative minutiae-based comparison algorithm based on the Levenshtein distance. The i-th element of the output vector $\mathbf{d}_L$ denotes the minimum LD value obtained between the i-th row of the first $\mathbf{D}_{E1}$ matrix and all the rows of the second $\mathbf{D}_{E2}$ matrix

$\mathbf{d}_L$ vector could also include some out-of-range numerical values, possibly added during the procedure described above. For this reason, the standard deviation σ amongst these values is calculated: this statistical dispersion measure, together with the previously calculated average value, allows for defining an interval that includes only the useful values, excluding those that might alter the final result. The described operation is carried out by a function, which takes care of checking for each element whether its value in included in the interval $[\mu - \sigma, \mu + \sigma]$: if the current element is included within this range, it is taken into consideration for the subsequent calculation of the average; otherwise, it is set to zero. The aforementioned function is effective regarding the removal of the out-of-range values inserted by the algorithm in the test phase, as well as the exclusion of any spurious low value under the condition of diversity between the tested $\mathbf{D}_E$ matrices. Once only the effective values are selected via this function, the average of the elements in $\mathbf{d}_L$ is more reliably recomputed and stored in V_L. By so doing, the whole procedure yields a single numerical value V_L, which can be evaluated by defining a simple, yet precise, classification based on numerical ranges according to each biometric recognition system currently used. Considering the output mean LD value V_L and setting two distinct thresholds S_1 and S_2 (with $S_1 < S_2$), the final result is calculated by comparing the value V_L and a result classification based on four distinct intervals that are associated with the following four different outcomes:

- $V_L = 0$, exact equality between the examined $\mathbf{D}_E$ matrices;
- $0 < V_L < S_1$, high degree of similarity between the examined $\mathbf{D}_E$ matrices;
- $S_1 \leq V_L < S_2$, high degree of uncertainty on the current comparison, owing to just few correspondences between the matrices;
- $V_L \geq S_2$, two distinct $\mathbf{D}_E$ matrices.

With more details, in terms of FAR, a comparison is correct if the result is higher than S_2; considering FRR, a comparison is correct if the result is lower than S_1. Typically, results between S_1 and S_2 are labelled as doubtful cases (i.e. uncertain correspondence). In this work, in order to obtain a reliable authentication system in terms of safety (FAR = 0), all uncertain cases were considered as a non-matches. From an algorithmic point of view, this choice is equivalent to having a single threshold S_1.

We aimed at developing a fusion technique at the level of comparison score: the individual scores are combined into a single total score representing the decision outcome. More specifically, the similarity scores for each unimodal system are fused using the weighted-sum rule via the coefficients $\kappa_r, \kappa_i \in [0, 1]$, which weight the scores achieved by the individual retina and iris authentication systems, respectively. From now on, the unimodal systems are denoted as sub-systems. Let V_r and V_i be the results of the retina and iris sub-systems, respectively, the weighting coefficients κ_r and κ_i are applied to achieve the final result, V_{final}, according to Equation (2):

$$V_{\text{final}} = \kappa_r \times V_r + \kappa_i \times V_i. \tag{2}$$

The computational complexity of the LD is $\mathcal{O}(l \cdot l)$ where l_1 and l_2 denote the lengths of the two vectors (in our case, the lengths of the currently analysed rows $\mathbf{D}_{E1}(i)$ and $\mathbf{D}_{E2}(j)$ of the matrices of the EDs amongst all the minutiae of the first and second biometric features) that are currently involved in the pairwise distance calculation. Since the LD calculation is usually performed recursively, it can be halted if it does not quickly converge towards zero or a value smaller than a threshold. This condition corresponds to the equality of the two descriptors. If the distance does not decrease and stabilises at a value above a tolerance threshold, then the calculation can be halted and the two descriptors defined are not equal. It is worth noting that the computational cost $\mathcal{O}(l \cdot l)$ represents the worst case scenario, whilst the halting conditions are more likely to occur and lead to a lower number of iterations. Overall, this calculation is applied by the pairwise comparison between all the rows #Rows ($\mathbf{C}_1$) and #Rows ($\mathbf{C}_1$) of the two coordinate matrices.

4.3 | Software implementation details

The proposed framework was designed and developed using the MatLab® R2016b (64-bit version) environment (The MathWorks, Natick, MA, USA). The Graphical User Interface (GUI) for the proposed multimodal system was developed using the GUI Development Environment (GUIDE), integrated in MatLab. The main phases of the multimodal system are the following:

1. Loading the scanned images;
2. Segmentation process;
3. Loading the database profile: the system requires to load the second profile to be compared, labelled as 'database Profile". The user has to look for the image in the database, using the name or the identification code of a certain person;
4. Starting the image processing: the step-by-step execution of the individual processing functions is indicated by the colour change in the box containing the operation being performed;
5. At the end of the image processing, an evaluation is performed to choose the right reference minutiae, both for the retina and iris. The system automatically decides which characteristic points to use and relies upon user interaction to start the comparison test.

At the end of the test, the system provides in the "Image Comparison" box all the results of the comparison test, for both retina and for iris, which include: the total of the expected tests, the actually carried out tests (within the threshold), the average of the minimum values, and the standard deviation. The final result of the comparison V_L is displayed in the "Resulting final value", next to a label that is coloured according to the location of the resulting value within the ranges. Furthermore, the result of the comparison is reported in textual form. Figure 5 shows the final phase of the GUI application in the case of negative matching.

5 | EXPERIMENTAL RESULTS

The performance of a biometric system has to be evaluated according to specific metrics, such as the dimensions of the realised model, speed of response, and accuracy. In the computer security field, accuracy, to be intended as the capability of the system in making the correct decision, is a rather critical parameter.

The proposed solution involves the development of a retina-iris multimodal system, which combines the advantages offered by the two unimodal approaches, whilst overcoming limits of the single unimodal solution. The proposed solution is favoured by the fact that these two biometric technologies can be considered quite similar to some extent and make use of the same human anatomical organ: the eye. The key concept for the development of a valid and efficient multimodal biometric system is the analysis and evaluation on heterogeneous data, with the aim of providing a single final result that represents an answer as accurate as possible to a certain authentication problem. In this case, the results obtained by the unimodal recognition systems—based on retina and iris scans alone—must be properly fused into a single comprehensive result. The values of the coefficients κ_r and κ_i, which weight the scores achieved by the individual sub-systems, are experimentally determined. These values were chosen with the goal of optimising the recognition capabilities of the multimodal authentication system (in terms of the achieved FAR and FRR). Therefore, exhaustive tests were performed, by varying κ_r and κ_i, and evaluating the corresponding FAR and FRR values.

After a description of the used datasets, the experimental results concerning the developed unimodal and multimodal

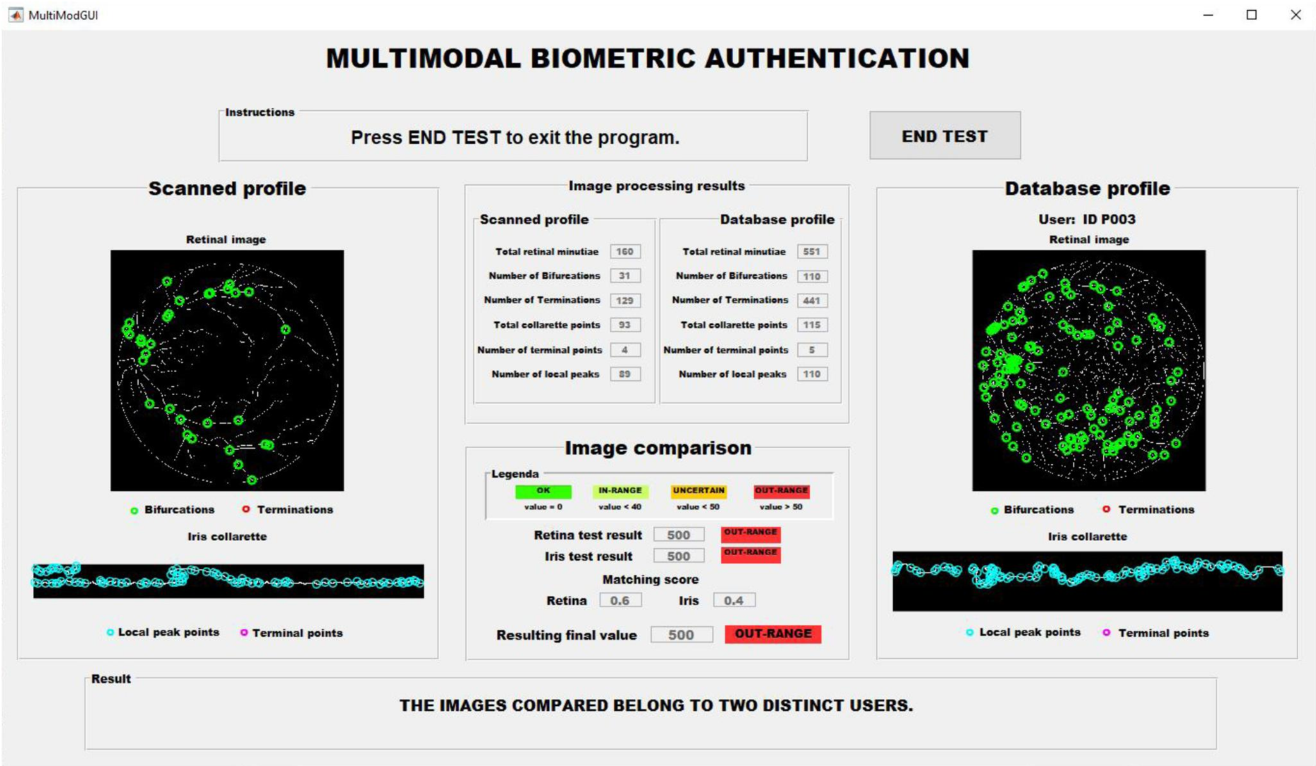

FIGURE 5 The GUI of the proposed multimodal biometric authentication system combining retina and iris. The screenshot shows the comparison, as well as the corresponding result, between two biometric descriptors belonging to different individuals. GUI, graphical user interface

systems are presented in what follows. Moreover, the FAR and FRR values by varying the coefficients κ_r and κ_i are also assessed for all the three tests. Finally, to confirm that the proposed multimodal system achieves statistically significant better performance than the unimodal sub-systems, McNemar test on paired authentication results [37] was performed. This non-parametric test on paired nominal data assesses whether the accuracy achieved by two classification models is statistically different. This test uses the null hypothesis: *the predicted class labels from the two compared models have equal accuracy for predicting the ground-truth class labels*. Since this statistical analysis involves multiple comparison tests, the authors adjusted their p-values using the Bonferroni–Holm method [38,39]. In all the tests, a significance level of $\alpha = 0.05$ was used.

5.1 | Datasets

To test the effectiveness of the proposed multimodal retina-iris system, three configurations, exploiting a retina database [40] and three different iris databases [41], were considered:

1. DRIVE retina database + University of Bath iris database;
2. DRIVE retina database + University of Beira Interior Iris (UBIRIS) [42] iris database;
3. DRIVE retina database + CASIA [43] iris database.

Unfortunately, there are only very few databases that aim at recognition purposes. Moreover, in the case of the retina, the database should provide multiple acquisitions for each retina image of the same eye to allow for comparisons. In several literature approaches, the authors built their own databases starting from a small set of images and applying geometric transformations (e.g. rotation and scaling) to increase the number of images to analyse, such as in [44]. Another solution could be to collect self-built databases. This second alternative undermines the reproducibility of the recognition performance, and the evaluation with respect to the state-of-the-art—which uses publicly available databases—is impossible. To exploit publicly available biometric databases, the authors combined three (big) databases containing iris images, built for developing recognition systems, but only one (small, even though well-prepared and commonly used) of the retina with a very limited number of users. Instead of using data augmentation procedures, sometimes exploited in biometric system experiments [26], the authors preferred to use small yet publicly available dataset accessible by the scientific community allowing for result reproducibility and comparability. Furthermore, since in the literature there are no databases combining together retina and iris images of the same person, three datasets were composed by combining single-feature belonging to different databases as follows: an image from the retina database was coupled with an image from the iris database, thus creating a new 'virtual person' with both biometric traits. Thus, considering that only a limited number of large-scale retina databases is publicly available, in this initial phase the authors built a 'virtual' multimodal retina-iris database with a satisfactory number of images. With more detail, three different acquisitions of the same subject were selected for each unimodal database: in each composed database, each 'virtual person' had three scans of the same retina and three scans of the same iris, in order to constitute three distinct retina-iris pairs. This procedure was repeated for 20 individuals according to Figure 6.

5.2 | Authentication performance

For all the three composed datasets, 1770 tests were performed, distributed as follows:

- number of tests amongst different images: 1710;
- number of tests amongst similar images: 60. For the sake of clarity, the term similar images refers to images belonging to the same individual, but which are not identical.

The tests considered useful for the calculation of the FAR and FRR indices are those between different images (for the FAR index) and those between similar images (for the FRR index). In all the experiments, the thresholds θ_E were set to 4 and 18 for the recognition sub-systems based on retina and iris, respectively, whilst the θ_L was initialised to 500 for the multimodal fusion. These settings relied upon a calibration set consisting of a random selection of images from the four available databases. Furthermore, as a common standard for

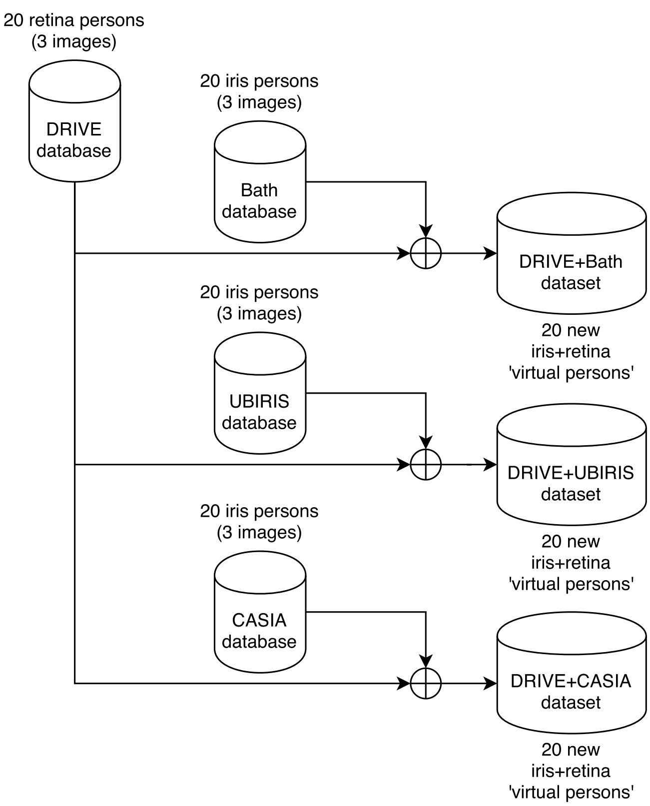

FIGURE 6 Block scheme representing the methodology used to obtain the new datasets combining retina and iris images

the three testing configurations, the comparison ranges (described in Section 4.2) are set as follows. S_1 and are two thresholds used to determine whether the result of a comparison is: (*i*) correct match, (*ii*) uncertain, or (*iii*) non-match. With more details, in terms of FAR, a comparison is correct if the result is higher than S_2; considering FRR, a comparison is correct if the result is lower than S_1. Typically, results between S_1 and S_2 are labelled as doubtful cases (i.e. uncertain correspondence). Since an authentication system must prioritise cautious and safe decisions, in this work, in order to obtain a reliable authentication system in terms of safety (FAR = 0), all uncertain cases were considered as a non-matches. From an algorithmic point of view, this choice is equivalent to having a single threshold S_1. The best values of S_1 and S_2, used in the final implementation, were experimentally determined after a set of tests with variable thresholds (in the range 30-60 using 10 as a step) aiming at optimising the system recognition performance. In particular, S_1 must first minimise the FAR and then also the FRR values. The value of S_2 was set as a margin of 10, according to the selected S_1 value, for uncertain cases. Since an authentication system must prioritise cautious and safe decisions, all the uncertain cases were classified as non-matches. The selected values were applied to all the three multimodal tests. The best values of S_1 and S_2 were 40 and 50, respectively:

- $0 \leq V_{\text{final}} \leq 40$: matching;
- $40 < V_{\text{final}} < 50$: uncertain correspondence (i.e. 'undetermined');
- $V_{\text{final}} \geq 50$: non-matching.

The reliability tests carried out on the unimodal recognition systems described so far offer good performance in terms of safety and robustness, although the two systems are not yet adequate to be considered truly reliable (Tables 7 and 8). In order to select the best decision threshold S_1, the FAR and FRR values were calculated by varying S_1 in the range [20, 70], with a step of 5.

As stipulated in the ISO/IEC 19,795-1 standard [45], the Detection Error Trade-off (DET) curves for the unimodal retina and iris recognition sub-systems are shown in Figure 7. The Area Under the Receiver Operating Characteristic Curve (AUC), EER and FNMR1000 (i.e. False Non-Match Rate (FNMR) at False Match Rate (FMR) of 0.1%). It is worth to note that the FAR and FRR indices can be used interchangeably with FMR and FNMR, respectively, when no repeated access attempts are considered during a transaction.

5.2.1 | Test 1—DRIVE retina database + bath iris database

The FAR and FRR values achieved by the proposed multimodal retina-iris system on Test 1, by varying the weighting coefficients κ_r and κ_i, are shown in Figures 8a,b, respectively. This system achieves an excellent 0% in terms of FAR index, an improvement when compared to both unimodal systems.

TABLE 7 FAR values achieved by the unimodal sub-systems, by varying the decision threshold S_1

S_1 value	DRIVE (%)	Bath (%)	UBIRIS (%)	CASIA (%)
20	0.64	0.53	0.58	2.16
25	0.64	0.58	0.64	4.09
30	0.70	0.64	0.64	6.02
35	0.82	0.64	0.64	9.01
40	**0.82**	**0.64**	**0.64**	**12.16**
45	0.82	0.64	0.64	15.91
50	0.82	0.64	0.64	20.58
55	0.88	0.64	0.64	24.44
60	0.88	0.64	0.64	27.84
65	0.88	0.64	0.64	31.20
70	0.88	0.64	0.64	33.86

Note: Total number of comparisons: 1710. Boldface indicates the results obtained with the chosen S_1 value.

Abbreviations: CASIA, Chinese academy of sciences institute of automation; DRIVE, digital retinal images for vessel extraction; FAR, false acceptance rate; UBIRIS, University Of Beira Interior Iris.

TABLE 8 FRR values achieved by the unimodal sub-systems, by varying the decision threshold S_1

S_1 Value	DRIVE (%)	Bath (%)	UBIRIS (%)	CASIA (%)
20	8.33	36.67	33.33	85.00
25	8.33	23.33	16.67	68.33
30	6.67	13.33	8.33	55.00
35	0.00	8.33	3.33	41.67
40	**0.00**	**6.67**	**3.33**	**33.33**
45	0.00	5.00	3.33	33.33
50	0.00	5.00	3.33	25.00
55	0.00	5.00	3.33	21.67
60	0.00	5.00	3.33	20.00
65	0.00	5.00	3.33	20.00
70	0.00	5.00	3.33	18.33

Notes: Total number of comparisons: 60. Boldface indicates the results obtained with the chosen S_1 value.

Abbreviations: CASIA, Chinese academy of sciences institute of automation; DRIVE, digital retinal images for vessel extraction; FAR, false acceptance rate; UBIRIS, University of Beira Interior Iris.

Considering the FRR level, there is a slight improvement with respect to the unimodal iris system, whilst a considerable loss is observed when compared to the 0% obtained by the unimodal retina-based system.

5.2.2 | Test 2—DRIVE retina database + UBIRIS iris database

The FAR and FRR values achieved by the proposed multimodal retina-iris system on Test 2, by varying the weighting coefficients κ_r and κ_i, are shown in Figures 9a,b, respectively. The fusion of the two most reliable systems, amongst the tested configurations, allows the FAR index to reach an excellent value of 0%, whilst the FRR index remains the same compared to the unimodal iris system, even though it is worse than the unimodal retina system which achieved one FRR = 0%.

5.2.3 | Test 3—DRIVE retina database + CASIA iris database

Figures 10a,b depict the FAR and FRR indices in the case of Test 3, by varying the weighting coefficients κ_r and κ_i. Comparing the FAR and FRR indices of the multimodal system with those of the unimodal iris system CASIA (FAR = 12.16%, FRR = 33.3%), the remarkable improvements obtained are evident. More specifically, the FAR index is even acceptable, whilst the FRR index, although considerably improved, it might be not yet acceptable.

5.2.4 | Overall experimental findings

Tables 9 and 10 summarise the authentication accuracy, in terms of FAR and FRR, for each combined dataset used in the three investigated multimodal retina-iris tests. For an ideal authentication system, both the FAR and FRR indices should be 0. The aforementioned result may be reached by online biometric authentication systems, because they have the freedom to reject the low-quality acquired items. On the contrary, official ready-to-use databases (e.g. DRIVE, CASIA, Bath and UBIRIS) contain images with highly different quality, including low-, medium-, and high-quality biometric acquisitions. For this reason, these databases cannot achieve the ideal result. To increase the related security level, the system parameters are usually fixed in order to achieve the FAR = 0% point and a corresponding FRR point. Considering the trend of the FAR and FRR revealed in Figures 8, 9, and 10, the comparison score coefficients, κ_r and κ_i, were chosen to optimise the multimodal system performance. Indeed, Tables 7 and 8 show the best trade-off of FAR and FRR obtained with the following values:

- retina weighting coefficient $\kappa_r = 0.6$;
- iris weighting coefficient $\kappa_i = 0.4$.

Considering that the retina is generally a more reliable biometric descriptor than the iris, it was expected to experimentally achieve κ_r higher than κ_i. To conclude, the authors can argue that the fusion technique at the comparison score level enabled a remarkable improvement of the FAR and FRR indices with respect to the unimodal biometric sub-systems. Figure 11 shows the DET curves for the three investigated retina-iris multimodal tests. The AUC, EER and FMR1000 metrics confirm the excellent authentication performance and

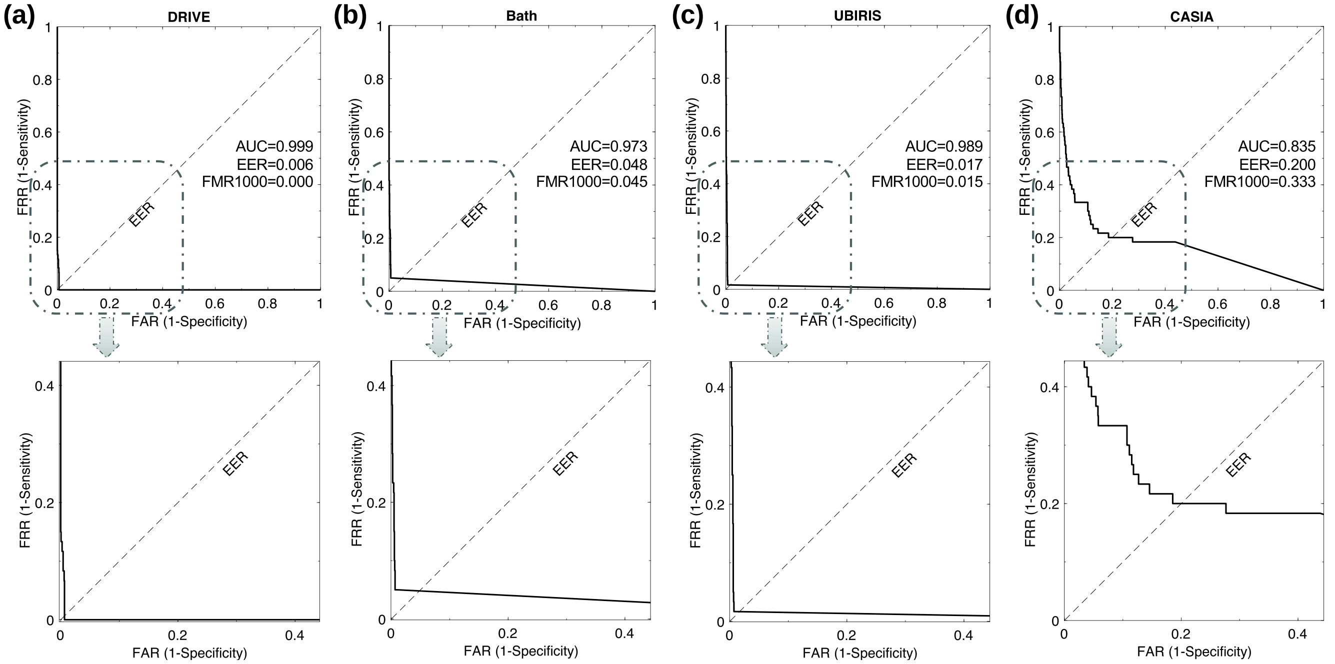

FIGURE 7 DET curves for the unimodal sub-systems on the datasets: (a) DRIVE; (b) Bath; (c) UBIRIS; (d) CASIA. In each figure the AUC, the EER and the FMR1000 are reported. The practically relevant area of the DET curves (dash-dotted grey box at the bottom-left corner) is zoomed in the plots in the bottom panel. AUC, Area Under the Receiver Operating Characteristic Curve; CASIA, chinese academy of sciences institute of automation; DET, detection error trade-off; DRIVE, digital retinal images for vessel extraction; EER, equal error rate; FMR, false match rate; UBIRIS, university of beira interior iris

FIGURE 8 Trend of the FAR and FRR indices by varying the coefficients κ_r and κ_i for Test 1. The two cells in light blue denote the values obtained by the two unimodal systems. The green cell represents the value chosen to optimise the accuracy of the final multimodal system. FAR, false acceptance rate; FRR, false rejection rate

(a)

Retina DRIVE (κ_r) \ Iris Bath (κ_i)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											0.64%
0.10										0.53%	
0.20									0.47%		
0.30								0.29%			
0.40							0.18%				
0.50						0.12%					
0.60					0.00%						
0.70				0.18%							
0.80			0.47%								
0.90		0.64%									
1.00	0.82%										

FAR

(b)

Retina DRIVE (κ_r) \ Iris Bath (κ_i)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											6.67%
0.10										6.67%	
0.20									6.67%		
0.30								6.67%			
0.40							5.00%				
0.50						5.00%					
0.60					5.00%						
0.70				3.33%							
0.80			1.67%								
0.90		0.00%									
1.00	0.00%										

the improvement provided by the multimodal approach, especially in the case of Test 3.

Finally, relying upon the results shown in Tables 5–8, the authors selected the single threshold $S_1 = 40$ for all the tests (the corresponding rows are denoted in boldface). This choice was motivated by the best compromise between the FAR and FRR indices to obtain a system with a higher security level, as also confirmed by the DET curves in Figures 7 and 11. It is important to point out that the only considered threshold S_1 was not fine-tuned for each experimental condition, but it was selected according to a correct functioning for all the implemented unimodal and multimodal systems. This characteristic

Iris UBIRIS (κ_i)

Retina DRIVE (κ_r)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											0.64%
0.10										0.53%	
0.20									0.47%		
0.30								0.29%			
0.40							0.18%				
0.50						0.12%					
0.60					0.00%						
0.70				0.18%							
0.80			0.47%								
0.90		0.64%									
1.00	0.82%										

(a) FAR

Iris UBIRIS (κ_i)

Retina DRIVE (κ_r)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											3.33%
0.10										3.33%	
0.20									3.33%		
0.30								3.33%			
0.40							3.33%				
0.50						3.33%					
0.60					3.33%						
0.70				1.67%							
0.80			1.67%								
0.90		0.00%									
1.00	0.00%										

(b) FRR

FIGURE 9 Trend of the FAR and FRR indices by varying the coefficients κ_r and κ_i for Test 2. The two cells in light blue denote the values obtained by the two unimodal systems. The green cell represents the value chosen to optimise the accuracy of the final multimodal system. FAR, false acceptance rate; FRR, false rejection rate

Iris CASIA (κ_i)

Retina DRIVE (κ_r)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											12.16%
0.10										10.18%	
0.20									8.19%		
0.30								7.13%			
0.40							5.32%				
0.50						2.57%					
0.60					0.35%						
0.70				0.41%							
0.80			0.53%								
0.90		0.70%									
1.00	0.82%										

(a) FAR

Iris CASIA (κ_i)

Retina DRIVE (κ_r)	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00											33.33%
0.10										31.67%	
0.20									30.00%		
0.30								25.00%			
0.40							25.00%				
0.50						20.00%					
0.60					18.33%						
0.70				16.67%							
0.80			13.33%								
0.90		10.00%									
1.00	0.00%										

(b) FRR

FIGURE 10 Trend of the FAR and FRR indices by varying the coefficients κ_r and κ_i for Test 3. The two cells in light blue denote the values obtained by the two unimodal systems. The green cell represents the value chosen to optimise the accuracy of the final multimodal system. L FAR, false acceptance rate; FRR, false rejection rate

is very important because it allows us to keep the ultimate goal of realising a method that can process and combine mm static biometric features in a homogeneous manner.

Table 11 shows the p-values of the McNemar test on the pairwise comparisons of the authentication results achieved by the multimodal retina-iris system and the unimodal sub-systems on the three tests described in Section 5.1. The results confirm that the proposed multimodal retina-iris system significantly outperforms the unimodal sub-systems for both Tests 1 and 2. As arguable, in Test 3, the retina classification significantly improves the unimodal iris sub-system, mostly affected by the relatively high FRR obtained on the CASIA iris

TABLE 9 FAR values achieved by the multimodal retina-iris system (using $\kappa_r = 0.6$ and $\kappa_r = 0.4$ for the fusion at the comparison score level) on the three tests described in Section 5.1

S_1 value	Test 1 (%)	Test 2 (%)	Test 3 (%)
20	0.00	0.00	0.00
25	0.00	0.00	0.23
30	0.00	0.00	0.23
35	0.00	0.00	0.29
40	**0.00**	**0.00**	**0.35**
45	0.00	0.00	0.35
50	0.00	0.00	0.35
55	0.00	0.00	0.35
60	0.00	0.00	0.41
65	0.00	0.00	0.41
70	0.00	0.00	0.41

Notes: The results were calculated by varying the decision threshold S_1. Total number of comparisons: 1710. Boldface indicates the results obtained with the chosen S_1 value.

TABLE 10 FRR values achieved by the multimodal retina-iris system (using $\kappa_r = 0.6$ and $\kappa_r = 0.4$ for the fusion at the comparison score level) on the three tests described in Section 5.1

S_1 value	Test 1 (%)	Test 2 (%)	Test 3 (%)
20	13.33	11.67	28.33
25	13.33	10	26.67
30	6.67	5	26.67
35	5.00	3.33	18.33
40	**5.00**	**3.33**	**18.33**
45	5.00	3.33	18.33
50	5.00	3.33	18.33
55	5.00	3.33	18.33
60	5.00	3.33	18.33
65	5.00	3.33	18.33
70	5.00	3.33	18.33

Notes: The results were calculated by varying the decision threshold S_1. Total number of comparisons: 60. Boldface indicates the results obtained with the chosen S_1 value.
Abbreviation: FAR, false acceptance rate.

dataset. These findings corroborate the significant improvement provided by our multimodal fusion scheme.

6 | DISCUSSION AND COMPARISONS

With reference to the latest advances in computer science, deep learning approaches have been showing cutting-edge performance in supervised classification tasks [46] and, more recently, also in biometry [47]. However, these feature learning methods need large-scale datasets with huge amounts of accurately labelled samples for training. Importantly, multimodal biometric authentication systems aim to effectively combine two or more physical or behavioural biometric traits to obtain adequate FAR and FRR indices that are beneficial for system dependability. We propose here a general framework based on fully unsupervised techniques that do not require any training and can be applied to datasets with limited amounts of unlabelled samples. Our multimodal method combines the minutiae from two different biometric traits and it is not just a classification technique.

Currently, no joint retina-iris database is available and this represents an inherent limitation for those researchers involved in the development and testing of multimodal biometric systems. Indeed, in the literature, only a few multimodal biometric systems based on retina and iris have been proposed so far (Table 11).

Latha et al. [19] were the only one proposing a score level fusion with a hybrid approach using frequency and spatial characteristics for iris and for retina, respectively. Choras [18], Modarresi et al. [26], Sen and Islam [27] and Saha et al. [12] proposed frequency-based approaches and feature-level fusion, nevertheless some important aspects have to be highlighted (Table 12). Choras [18] did not report any information on either the databases and the number of images used or the scores obtained. Modarresi et al. [26] used 40 iris images from the CASIA database that became 400 after a data augmentation process by means of 10 random rotations. Sen and Islam [27] reported an unusual 'operating point' in their work. Considering that the purpose of a biometric system is to increase the access security, such a system with FAR ¿ FRR (FAR = 2.041% and FRR = 0%) represents an anomaly with respect to biometric good practice, according to which it is preferable have a system with false rejections rather than with false acceptances. Lastly, Saha et al. [12] presented a multimodal user authentication system based on a feature-level fusion of iris and retina recognition, by considering features in the frequency domain. The RR for the proposed multimodal biometric system was 98.37%, whereas the distinct unimodal iris and retina recognition systems achieved 96.74% and 94.56%, respectively.

There are substantial differences between the proposed method and the approach in Saha et al. [12]. Primarily, the features in [12] were extracted in the frequency domain, by using the Discrete Wavelet Trans-form (DWT) and the 1D log-Gabor filters for retina and iris feature extraction, respectively. Especially, the feature-level fusion required a normalisation step for both unimodal sub-systems exploiting a DWT-based encoding and the Daugman rubber sheet model [22] for retina and iris feature representation, respectively. Afterwards, the concatenation of the normalised iris and retinal templates was not clearly described, especially how the iris and retinal feature vectors have the same number of components despite the normalisation step. This feature-level fusion was not suitable for our spatial feature vectors: the variable number of minutiae does not allow us to define any fixed template. Regarding the dimensionality reduction based on the PCA, after the augmented feature template, neither the number of principal

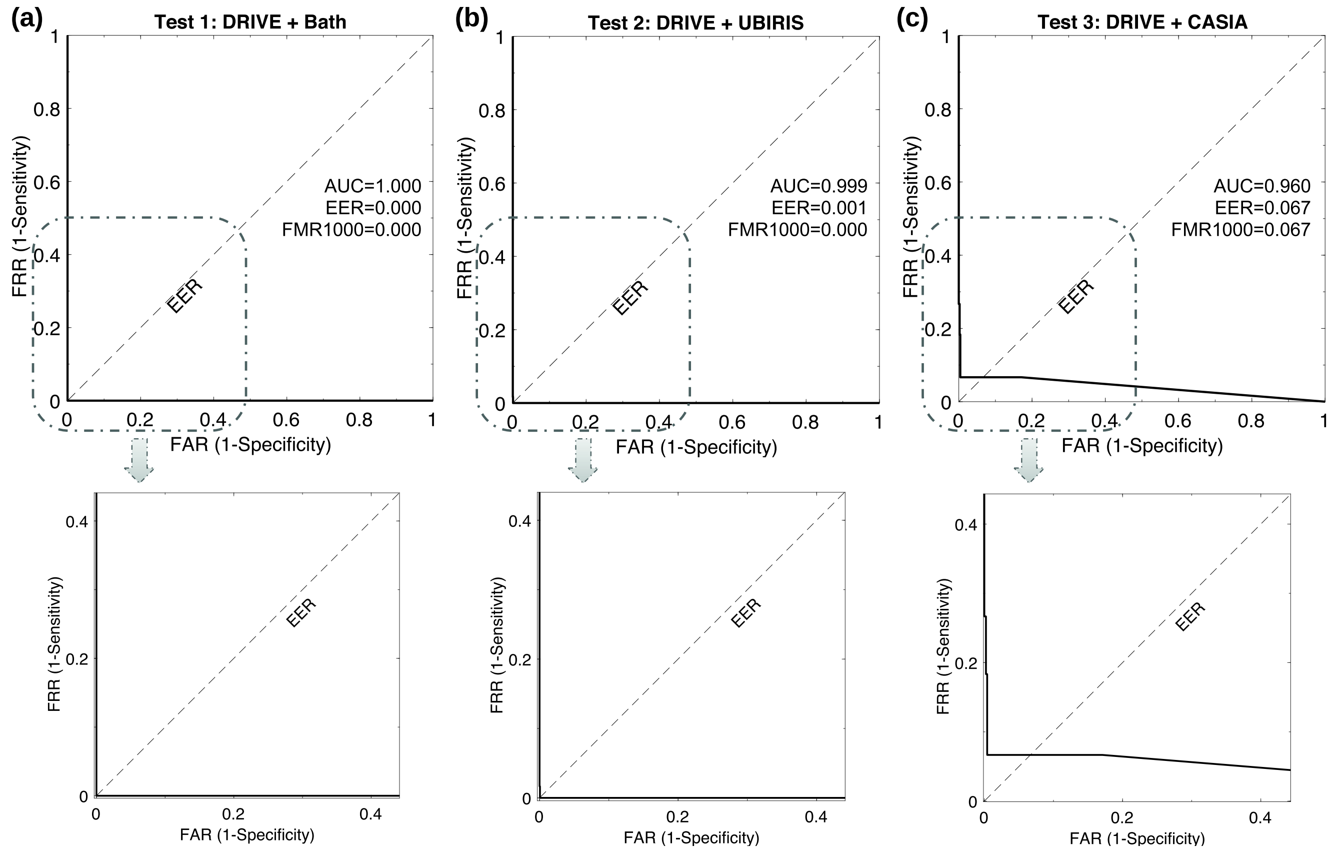

FIGURE 11 DET curves for the multimodal sub-systems on the multimodal retina-iris datasets: (a) Test 1; (b) Test 2 and (c) Test 3. In each figure, the AUC, the EER and the FMR1000 are reported. The practically relevant area of the DET curves (dash-dotted grey box at the bottom-left corner) is zoomed in the plots in the bottom panel. AUC, Area Under the Receiver Operating Characteristic Curve; DET, detection error trade-off; EER, equal error rate; FMR, false match rate

TABLE 11 McNemar test *p*-values for the pairwise comparison of the authentication results achieved by the multimodal retina-iris system (using $\kappa_r = 0.6$ and $\kappa_r = 0.4$ for the fusion at the comparison score level)

	Test 1	Test 2	Test 3
Multimodal versus unimodal iris	**0.0007**	**0.0015**	4.378×10^{-31}
Multimodal versus unimodal retina	**0.0425**	**0.0148**	2.379×10^{-1}
Unimodal iris versus unimodal retina	0.5716	0.8450	1.440×10^{-28}

Notes: The unimodal sub-systems on the three tests described in Section 5.1. (total number of comparisons: 1710). A significance level of $\alpha = 0.05$ with the Bonferroni-Holm correction method for multiple comparisons was used [39]. Boldface indicates that the null hypothesis can be rejected.

components nor the selection criteria (e.g. percentage of variance) are reported. The lack of these important implementation details hampers the replicability of the method for a quantitative comparison.

The multimodal retina-iris authentication system proposed in this paper presents a novel comparison score level fusion. More specifically, a large amount of meaningful points for retina and iris images are extracted in the spatial domain and, then, a comparison score is obtained for each unimodal sub-system. Interestingly, for the first time, the fusion at the level of comparison score relies upon the LD. Finally, a weighted-sum was used to obtain the final score for the multimodal authentication system. The best pair of FAR and FRR obtained was of 0% and 3.33%, respectively. With respect to the feature-level early fusion (i.e. early phase feature concatenation) used in [12], our comparison score-level fusion allows for more flexibility in terms of both used features and integration of additional biometric traits. More specifically, the LD-based comparison can be applied to any spatial feature based on minutiae, whilst the late fusion scheme can be adapted to a different set of biometric measurements.

7 | CONCLUSION AND FUTURE DIRECTIONS

This work aimed to investigate a system that leverages the best-performing biometric features, namely, retina and iris. The

TABLE 12 Comparison of the existing multimodal iris-retina approaches and the main characteristics of the systems described in Section 2

Reference	Approach type	Fusion type	Databases	Used images	Obtained Scores
Latha *et al.* [19]	Hybrid (Frequency for iris; Spatial for retina)	Score level	Iris: CASIA v3 Retina: VARIA	Iris: 500 Retina: 500	EER = 0.01% RR = 99.3%
Choras [18]	Frequency (Gabor transform)	Feature-level	N.A.	N.A.	N.A.
Modarresi *et al.* [26]	Frequency (Contourlet transform)	Feature-level	Iris: CASIA Retina: DRIVE	Iris: 40 Retina: 400	EER = 0.0413%
Sen and Islam [27]	Frequency	Feature-level	Iris: CASIA v1 Retina: DRIVE	Iris: 147 Retina: 147	FAR = 2.041% FRR = 0% RR = 97.959%
Saha *et al.* [12]	Frequency	Feature-level	Iris: IITD Retina: DRIVE	Iris: 184 Retina: 184	FAR = 0% FRR = 1.63% RR = 98.37%

Abbreviations: CASIA, Chinese academy of sciences institute of automation; EER, equal error rate; DRIVE, digital retinal images for vessel extraction; FAR, false acceptance rate; FRR, false rejection rate; IITD, Indian Institute Of Technology Delhi; RR, recognition rate; UBIRIS, University of Beira Interior Iris; VARIA, VARPA Retinal Images for Authentication.

proposed multimodal system exploited iris and retina, as well as the LD [15,16], in an innovative way, allowing us to overcome the typical issues in spatial approaches, often due to misalignment of the templates to be compared [48]. The tests aimed at evaluating the performance of the multimodal retina-iris system on multiple retina and iris database configurations.

The authors used publicly available databases accessible by the scientific community allowing for result reproducibility and comparability. In order to provide comprehensive results, the authors plotted the DET curves, as well as calculated the AUC, EER and FMR1000 metrics. The best FAR and FRR values achieved by our multimodal biometric approach were 0% and 3.33%, respectively. The multimodal retina-iris approach outperformed the corresponding unimodal systems, so drawing out its potential in authentication systems. Therefore, these experimental findings showed that our multimodal solution can guarantee a high level of reliability and be beneficial to computer security applications. Adaptive weights for the comparison score-level fusion might be employed to cope with the variability of the environmental conditions that could affect the quality of the traits acquired by the biometric sensors. For this reason, it might be useful to consider variable weights in order to dynamically manage this variability, as proposed in [49]. Since the authors analyse biometric images acquired in 'controlled' environments, the use of dynamic weights is not mandatory.

The authors are currently attempting to increase the size of the tested 'virtual' multimodal retina-iris database to validate our approach on a large-scale database [50]. With the goal of keeping result reproducibility and comparability, more public available databases might be combined to achieve larger datasets, such as in the particular case of retina databases. However, it is worth noting that the majority of retina databases were collected for the research and analysis in clinical scenarios tailored to anomaly or disease detection (e.g. diabetic retinopathy, glaucoma) and are not suitable for biometric purposes. In the near future, the authors aim to extend the same multimodal approach with other static biometric features that allow for identifying and extracting the minutiae in the spatial domain. As a matter of fact, The authors plan to develop a multimodal framework with a fusion scheme at the template-level to combine and standardise multiple biometric approaches into one system [3], in order to obtain a novel and universal approach for any type of static biometric features.

Lastly, the authors would like to deploy the proposed multimodal authentication system onto embedded devices [6,7], by leveraging the efficiency of the developed unimodal LD-based comparison algorithms in the spatial domain.

ORCID

Vincenzo Conti https://orcid.org/0000-0002-8718-111X

REFERENCES

1. Ross, A, Nandakumar, K, Jain, AK.: Handbook of multibiometrics. (1st ed.) International Series on Biometrics. Springer-Verlag, Berlin (2006)
2. Ross, A, Jain, AK.: Information fusion in biometrics. Pattern. Recogn. Lett. 24, 2115–2125 (2003)
3. Conti, V, Militello, C, Sorbello, F, Vitabile, S.: A frequency-based approach for features fusion in fingerprint and iris multimodal biometric identification systems. IEEE Trans. Syst. Man Cybern. Syst. 40, (4), 384–395 (2010)
4. Bhartiya, N, Jangid, N, Jannu, S.: Biometric authentication systems: security concerns and solutions. In: 2018 3rd international conference for convergence in technology (I2CT). pp. 1–6. IEEE, Piscataway, NJ (2018)
5. Barni, M, Droandi, G, Lazzeretti, R, Pignata, T.: SEMBA: secure multibiometric authentication. IET Biom. 8, (6), 411–421 (2019)

6. Yoo, JH, et al.: Design of embedded multimodal biometric systems. In: Proceedings of 3rd international IEEE conference on signal-image technologies and internet-based system (SITIS), pp. 1058–1062. IEEE, Los Alamitos, CA 2007
7. Conti, V, Militello, C, Sorbello, F.: Vitabile: Biometric sensors rapid prototyping on FPGA. Knowl. Eng. Rev. 30, (2), 201–219 (2015)
8. Militello, C, Conti, V, Sorbello, F, Vitabile, S.: A fast fusion technique for fingerprint and iris spatial descriptors in multimodal biometric systems. Int. J. Comput. Syst. Sci. Eng. 29, (3), 205–217 (2014)
9. Waheed, Z, Waheed, A, Akram, MU.: A robust non-vascular retina recognition system using structural features of retinal image. In: Proceedings of 13th IEEE international bhurban conference on applied science and technology (IBCAST), pp. 101–105. IEEE, Piscataway, NJ (2016)
10. Lajevardi, SM, Arakala, A, Davis, SA, Horadam, KJ.: Retina verification system based on biometric graph matching. IEEE T. Image Process. 22, (9), 3625–3635 (2013)
11. Elhoseny, M, Essa, E, Elkhateb, A, Hassanien, AE, Hamad, A.: Cascade multimodal biometric system using fingerprint and iris patterns. In: Proceedings of International conference on advanced intelligent systems and informatics (AISI), pp. 590–599. Springer, Cham, Swiss (2017)
12. Saha, A, Saha, J, Sen, B.: An expert multi-modal person authentication system based on feature level fusion of iris and retina recognition. In: Proceedings of IEEE international conference on electrical, computer and communication engineering (ECCE), pp. 1–5, IEEE, Piscataway, NJ (2019)
13. ISO/IEC TR 24722: Information technology—Biometrics—Multimodal and other multibiometric fusion. International Organization for Standardization, Geneva (2015)
14. Militello, C, Conti, V, Sorbello, F, Vitabile, S.: A novel embedded fingerprints authentication system based on singularity points. In: Proceedings of International IEEE conference on complex, intelligent and software intensive systems (CISIS), pp. 72–78. IEEE, Los Alamitos, CA (2008)
15. Levenshtein, VI.: Binary codes capable of correcting deletions, insertions, and reversals. Sov. Phys. Dokl. 10, (8), 707–710 (1966)
16. Yujian, L, Bo, L.: A normalised Levenshtein distance metric. IEEE Trans Pattern Anal Mach Intell. 29, (6), 1091–1095 (2007)
17. Kabir, W, Ahmad, MO, Swamy, M.: A multi-biometric system based on feature and score level fusions. IEEE Access. 7, 59437–59450 (2019)
18. Choraś, RS.: Multimodal biometric personal authentication integrating iris and retina images. In: Image processing and communications challenges 2, Vol. 84 of Advances in Intelligent and Soft computing, Springer. pp. 121–131. Springer Berlin, Heidelberg, Germany (2010)
19. Latha, L, Thangasamy, S.: A robust person authentication system based on score level fusion of left and right irises and retinal features. Procedia. Comput. Sci. 2, pp. 111–120 (2010)
20. Drozdowski, P, Rathgeb, C, Busch, C.: Computational workload in biometric identification systems: an overview. IET Biom. 8, (6), 351–368 (2019)
21. ISO/IEC 2382-37: Information technology—Vocabulary—Part 37: Biometrics. International Organization for Standardization, Geneva (2017)
22. Daugman, J.: The importance of being random: statistical principles of iris recognition. Pattern. Recogn. 36, (2), 279–291 (2003)
23. Conti, V, Rundo, L, Militello, C, Mauri, G, Vitabile, S.: Resource-efficient hardware implementation of a neural-based node for automatic fingerprint classification. J. Wirel. Mob. Netw. Ubiquit. Comput. Depend. Appl. 8, 19–36 (2017)
24. Walia, GS, Singh, T, Singh, K, Verma, N.: Robust multimodal biometric system based on optimal score level fusion model. Expert. Syst. Appl. 116, 364–376 (2019)
25. Hezil, N, Boukrouche, A.: Multimodal biometric recognition using human ear and palmprint. IET Biom. 6, (5), 351–359 (2017)
26. Modarresi, M, Oveisi, IS.: A contourlet transform based for features fusion in retina and iris multimodal biometric system. In: Proceedings of International workshop on biometric authentication (BIOMET), Vol. 8897 of LNCS, pp. 75–90. Springer, Cham, Switzerland (2014)
27. Sen, B, Islam, MR.: Iris and retina recognition based multimodal person identification system. Curr. Trends. Inf. Technol. 5, (1), 22–28 (2015)
28. Kihal, N, et al.: Efficient multimodal ocular biometric system for person authentication based on iris texture and corneal shape. IET Biom. 6, (6), 379–386 (2017)
29. Meenakshi, V, Padmavathi, G.: Retina and iris based multimodal biometric fuzzy vault. Int. J. Comput. Appl. 1, (29), 67–73 (2010)
30. Daniel, DM, Mihaela, C, Romulus, T.: Combining feature extraction level and score level fusion in a multimodal biometric system. In: Proceedings of 11th IEEE international symposium on electronics and telecommunications (ISETC), pp. 1–4. IEEE, Piscataway, NJ (2014)
31. Mittal, G, Sivaswamy, J.: Optic disk and macula detection from retinal images using generalised motion pattern. In: Proceedings of 5th national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG), pp. 1–4. IEEE, Piscataway, NJ (2015)
32. Conti, V, Milici, G, Vitabile, S, Sorbello, F.: A novel iris recognition system based on micro-features. In: Proceedings 5th IEEE workshop on automatic identification advanced technologies, pp. 253–258. IEEE, Piscataway, NJ (2007)
33. Uhl, A, Wild, P.: Enhancing iris matching using Levenshtein distance with alignment constraints. In: Proceedings of International Symposium on Visual Computing (ISVC), vol. 6453 of LNCS, pp. 469–478. Springer Berlin, Heidelberg, Germany (2010)
34. Muron, A, Pospísil, J.: The human iris structure and its usages. Physica. 39, 87–95 (2000)
35. Barpanda, SS, Sa, PK, Marques, O, Majhi, B, Bakshi, S.: Iris recognition with tunable filter bank based feature, Multimed. Tool. Appl. 77, (6), 7637–7674 (2018)
36. Sung, H, Lim, J, Park, J, Lee, Y.: Iris recognition using collarette boundary localization. In: Proceedings of 17th International Conference on Pattern Recognition (ICPR), vol. 4, pp. 857–860. IEEE, Los Alamitos, CA (2004)
37. McNemar, Q.: Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 12, (2), 153–7 (1947)
38. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, (2), 65–70 (1979)
39. Westfall, PH, Troendle, JF, Pennello, G.: Multiple McNemar tests. Biometrics. 66, (4), 1185–1191 (2010)
40. Image Sciences Institute, Department of Radiology, University Medical Center Utrecht, The Netherlands. DRIVE: digital retinal images for vessel extraction. http://www.isi.uu.nl/Research/Databases/DRIVE/. Accessed 16 March 2020
41. The International Association for Pattern Recognition. Iris database: CASIA, Bath, UBIRIS. http://www.cbsr.ia.ac.cn:8080/iapr_database.jsp. Accessed 16 March 2020
42. Proença, H, Filipe, S, Santos, R, Oliveira, J, Alexandre, LA.: The UBIRIS. v2: a database of visible wavelength iris images captured on-the-move and at-a-distance. IEEE Trans. Pattern. Anal. Mach. Intell. 32, (8), 1529–1535 (2009)
43. Tan, T, He, Z, Sun, Z.: Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition, Image. Vis. Comput. 28, (2), 223–230 (2010)
44. Meng, X, Yin, Y, Yang, G, Xi, X.: Retinal identification based on an improved circular Gabor filter and scale invariant feature transform. Sensors 13, (7), 9248–9266 (2013)
45. ISO/IEC 19795-1. Information technology—Biometric performance testing and reporting—Part 1: Principles and framework. International Organization for Standardization, Geneva (2006)
46. Ravì, D, Wong, C, Deligianni, F, et al.: Deep learning for health informatics. IEEE J. Biomed. Health. Inform. 21, (1), pp. 4–21 (2017)
47. Sundararajan, K, Woodard, DL.: Deep learning for biometrics: a survey. ACM Comput. Surv. 51(3), 1–34 (2018)
48. Conti, V, Vitello, G, Sorbello, F, Vitabile, S.: An advanced technique for user identification using partial fingerprint. In: Proceedings of 7th

International IEEE Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 236–242. IEEE, Los Alamitos (2013)
49. Gupta, K, Walia, GS, Sharma, K.: Quality based adaptive score fusion approach for multimodal biometric system. Appl. Intell. 50, (4), 1086–1099 (2020)
50. Czyżewski, A, Hoffmann, P, Szczuko, P, et al.: Analysis of results of large-scale multimodal biometric identity verification experiment. IET Biom. 8(1), 92–100 (2019)

How to cite this article: Conti V, Rundo L, Militello C, Salerno VM, Vitabile S, Siniscalchi SM. A multimodal retina-iris biometric system using the Levenshtein distance for spatial feature comparison. *IET Biome.* 2021;10:44–64. https://doi.org/10.1049/bme2.12001