IET Biometrics

ORIGINAL RESEARCH PAPER

Optimized multi-biometric enhancement analysis

Saliha Artabaz[1] (ORCID) | **Layth Sliman[2,3]** | **Karima Benatchba[1]** | **Mouloud Koudil[1]**

[1]Laboratoire de Méthodes de Conception de Systèmes LMCS, Ecole nationale Supérieure d'Informatique ESIOued-Smar, Algiers, Algeria

[2]Ecole d'ingénieur généraliste en informatique et technologies du numérique Efrei, Villejuif, France

[3]IBISC Laboratory, University of Évry Val d'Essonne, Évry, France

Correspondence

Saliha Artabaz, Laboratoire de Méthodes de Conception de Systèmes LMCS, Ecole nationale Supérieure d'Informatique ESI, Oued-Smar, Algiers, Algeria.
Email: s_artabaz@esi.dz

Abstract

A multi-biometric system uses different modalities to identify individuals more accurately. The authors analyse fusion efficiency of a significant number of multi-biometric fusion schemes. To do so, the study applies different functions that are generated using genetic programming (GP) on the 2000 multi-biometric instances produced by the fusion of different biometric matching scores. The functions are represented using a tree of arithmetic operations and are used for fusion at score level. First, genetic programming is implemented on the XM2VTS score database. The GP optimizes the half total error rate of fused matching scores. Then, a comparative study is performed based on our experiments on matching scores of different biometric baseline systems provided by the bio-secure database. This database provides 24 streams that we use to generate 2000 multi-biometric combinations. These multi-biometric instances combine matching scores of different instances, sensors and traits. To assess the quality of the fused scores and the quality of performing biometric baseline systems, we use weighted functions based on user-specific and group-specific normalization. Then, we propose a hybrid cat swarm optimization (CSO) based on the average-velocity inertia-weighted CSO and the normal mutation strategy-based CSO to compute the weights of the selected functions for the fused biometric systems. Finally, we present the statistical significance tests to confirm that the proposed functions outperform the existing functions based on arithmetic rules, normalization fusion and evolutionary algorithms.

1 | INTRODUCTION

Multi-biometrics addresses several unimodal biometric systems drawbacks. Mostly, they are applied to reduce matching errors [1] since the research experiments show that combining different traits reduces both false positive and false negative errors [2–5]. With recent user-friendly applications like patient identification, healthcare monitoring [6, 7], mobile and smartphone identification for sensitive information and services protection [8, 9], biometrics is dealing with new challenges like variable configurations and adaptive modalities combinations [10, 11] in addition to complex and interactive environment constraints that induce high-level data confidentiality risks [12]. Providing similar contextual conditions in the framework testing, which allows different combinations and fusion methods testing, is highly desirable. Thus, the biometric system can adapt a fusion method according to the acquired modalities and successful features extraction.

Multi-biometrics fusion is applied at different levels: signal level (namely image level), feature level, score level, decision level and rank level. The most used among these in the literature is the score level [13, 14]. This level has many advantages:

- Scores are much easier to fuse than features [3].
- The score level provides more information than the decision level [3].
- They are easier to assess from existing uni-biometric systems.
- The fusion at this level provides better accuracy than fusion at other levels [15, 16].

Fusion at score level considers different factors or parameters that must be optimized to ensure the success of the applied fusion strategy. On one hand, the authors in [17] propose an innovative scheme to fuse scores by combining the weighted mean and quasi-arithmetic mean without any learning process, while [18] uses consecutively a suite of derivatives on training

data to estimate the vector score integration adapted to the multi-class problem that minimizes the probability of error. On the other hand, many evolutionary methods are applied to optimize these metrics with a learning process like differential evolution [19, 20], particle swarm optimization (SO) [21] and quasi-convex optimization [22]. SO is suitable for different optimization applications. The use of SO is motivated by its ability to emerge intelligence using collaboration between individuals that has limited or no intelligence [23]. This is a good opportunity to solve problems of high dimensionality that are complex to resolve by brain perception. Cat swarm optimization (CSO) is one of the most widely used and is inspired by the natural cat behaviour that models exploration and exploitation using an innovative technique with respect to the trade-off between the two phases [24]. The authors in [24] describe the different applications of CSO in computer vision and biometric recognition and confirm through their survey and performance evaluation that CSO outperforms PSO, Genetic Algorithm (GA), DE and other methods applied in different fields. Their results prove that CSO is a competitive algorithm that leads to the optimal fitness value for different problems. To the best of our knowledge, this method is not yet applied in score fusion factors computation like weighting. Therefore, we propose a CSO algorithm proposed to compute the fusion weights that are combined with the user-specific quality measure on pre-computed score fusion functions from genetic programming (GP).

The fusion at score level considers scores resulting from different traits, instances, sensors and algorithms. Therefore, to identify the best system adapted to the situation, we must consider all these parameters. Indeed, the experiments conducted in [16] achieved different gain ratios for different combination methods. The experiments showed that a successful fusion depends on many factors including the availability of data for training and matchers, the accuracy of the matchers, the correlation of the scores and the chosen fusion. Consequently, assessing the efficiency of a fusing strategy requires testing the strategy on different score combinations before concluding whether the fusion method can outperform individual accuracy of the baseline systems. However, most of the research found in the literature compares fusion methods relying on their own baseline systems [3, 15, 16]. However, this cannot affirm whether the enhancement occurs as a result of the fusion or the used baseline systems' performances. Therefore, for a predefined fusion strategy, we can search the best biometric combinations that work better than the fused baseline systems.

Here, we propose to combine GP with the proposed CSO algorithm to generate different weighted fusion functions represented in a tree of arithmetic rules. Then, we compare our proposal to the fusion strategies of main proposals in the literature such as the use of norms [5], DE, PSO and BSA [19–21, 25]. Our first version and application of GP is proposed in [26] whose extended experiments of the best trees are conducted on the Biosecure score database [27]. Then, we perform multiple simulations using GP to get closer to the optimized fitness value, within the defined number of generations, for the maximum number of individuals. These simulations allow us to select multiple trees that optimize the score fusion. These trees are used to benchmark and analyse fusion

on an unlimited number of biometric output combinations that are constructed thanks to the scores provided by the Biosecure score database [27]. This database includes non-correlated baseline systems of different traits and a quality assessment of the processed signals. The quality assessment must consider the acquired signal and the quality of the biometric system that processes it. The first evolve for each sample, where the second can be evaluated to provide the scalar that indicates its performance and reliability. Therefore, we aim to combine baseline systems weighting based on an evolutionary method with quality assessment computed from the baseline systems quality.

The available quality measure gives a scalar indicator based on different sources of quality [28] such as the trait character, imaging and environment. We use the measures, provided by the Biosecure database, to weight the fusion functions. The proposed approach outperforms the baseline biometric subsystems. Then, a comparative study is conducted between the best-computed fusion functions that are obtained from the GP. Then, the statistical analysis of the tested multi-biometric scores fusion is described allowing us to show the ones which reduce the baseline biometric systems' errors for instance the equal error rate (EER) with or without the hybrid cat swarm optimization (HCSO) weighting.

In Section 2, we present the score level fusion applied in multi-biometrics. Section 3 describes GP used for generating the proposed functions, the HCSO algorithm applied to compute baseline score weighting and the tested databases. Then, we display and analyse the different proposed functions used to fuse multiple combinations of baseline biometric systems in Section 4. This leads us to propose a useful platform that gathers an unlimited number of multi-biometric fused scores to find the best combinations. Finally, we conclude and propose some perspectives of the work performed.

2 | MULTI-BIOMETRICS AND SCORE LEVEL FUSION

Multi-biometrics are gaining popularity since they address unimodal biometric system weaknesses such as non-universality, noisy signals, low accuracy, and spoofing attacks. It is well-known that multi-biometrics is mostly applied to reduce uni-biometric systems errors according to the different studies published by researchers in this field [13]. The uni-biometric systems are becoming less and less reliable to protect confidential data at a high level. Authors in [12] demonstrates the effectiveness of fusion applied on Support Vector Machine (SVM) classifiers outputs, using the Dempster-Shafer evidence theory, that can provide a reliable biometric in complex interactive environments. In addition, fusion provides effective systems which combined with template protection can preserve and enhance system performances against spoofing attacks [29].

Fusion of multiple traits faces many challenges to ensure performances' precision and avoid biased results. These challenges include the multiple data sources incompatibility, matchers scores normalization and the noise, which influence systems performances, bringing false positive or false negative authentication.

TABLE 1 Review of fusion methods

Modalities	Methods	References
Multi-spectral palmprint	Different t-conorm compared to sum, SVM	[8]
Iris, face	GA features selectionSVM score fusion	[9]
Speech, lipreading	Average	[30]
NIST BSSR1 multi-biometric score database (fingerprint and face), face, iris dataset	Dezert-Smarandache theory DSmT	[15]
Palm/phalanges print	Sum, product, min, max, hamacher t-norm, frank t-norm, yager's t-norm	[5]
Face, palmprint, signature, speech	Sum rule compared to the unimodal system	[31]
Fingerprint, iris, left ear and right ear	DE, exponent control, and Kernel mapping	[19]
Fingerprint and voice	DE and proportional conflict redistribution	[20]
Face and voice	PSO, Dempster Shafer and Belief functions	[21]
Two face databases	Quasi-convex optimization	[22]
Iris, finger vein and fingerprint	Evolutionary backtracking search optimization algorithm	[25]

Abbreviation: GA, Genetic Algorithm.

TABLE 2 Some publicly available databases

Database	Modalities	Tools
BSSR1 Score database	Face, fingerprint	None
FRGC Image database	Still face, multiple face images, 3D images	None
XM2VTS Score database	Face, speech	Baseline systems with evaluation protocol
Biosecure Score database	Face, fingerprint, Iris	Baseline systems, multiple streams of multiple scenarios with evaluation protocol

Recognition in multi-biometric fusion schemes combines traits of different sources (multiple samples of modalities or traits acquired with different sensors and processed with different algorithms) at different levels (sensor, feature, score, rank, or decision):

- Sensor level: combines multiple images or signals obtained from the same sensor or different sensors. This level can be considered in signal reconstruction (multiple images of the same evidence) or fusion.
- Feature level: combines multiple feature vectors. The combination consists of fusion or selection of features from multiple sources.
- Score level: combines matching scores of different matchers or classifiers.
- Decision level: combines decisions made by different biometric systems.
- Rank level: combines ranking results of different biometric systems.

2.1 | Score level fusion

The score level fusion is the most cited and used fusion [8, 9, 15, 30, 31], which is due to its low computational complexity. In addition, comparative studies have shown that fusion at the score level outperforms other levels [15]. The proposition of [19] discusses the use of differential evolution, which has a similar structure to GA, in parameters tuning to reduce the overlapping of genuine and impostor distribution. Also, authors in [20] use DE to find the confidence factors of the belief assignments that reduce the weighted error rate. The belief assignments representing the transformed scores are then fused using proportional conflict redistribution (PCR). The same factors are tuned using PSO in [21], and the fusion is performed using the Dempster Shafer and PCR. PCR is also used to resolve conflicting beliefs from different classifiers whose performance is optimized with the evolutionary backtracking search optimization algorithm that searches the best confidence factors that boost or remove classifiers scores [25]. Also, there is another method which optimizes the fusion process using the quasi-convex optimization [22] that requires domain-specific knowledge on feature similarity score distributions. Table 1 describes some of the works proposed in the literature.

The comparative study in [30] shows that the score level slightly surpasses both feature and decision level. However, due to the limited number of tests, the work does not provide a significance test to confirm the score level superiority. A recent work proposes a unique blend of belief assignment and decision-making methods in Dezert-Smarandache theory framework [15]. They tested fusion of different datasets where

only one contains two traits (NIST BSSR1), and the two other databases (FRGC, IRIS dataset) are used for multi-algorithm and multi-sample fusion (Table 2).

Using multiple databases in multi-biometric fusion analysis can be suitable to compare their performances. However, fusion cannot guarantee performance enhancement. The used modalities that belong to different persons are not dependent. Despite this, multiple works perform their experimental study on combined databases to demonstrate the effectiveness of their fusion proposal. The use of standard multimodal databases remains the best option according to the availability of the chosen modalities.

2.2 | Normalization and weighting

Normalization and weighting system's outputs are necessary to preserve and improve the provided system accuracy. Normalization is usually used to transform scores of different sources into a common domain. Recent researches [5, 8] use normalization methods as fusion operations. They combine couples of scores in flow. Their results seem to be interesting when compared to SVM and Sum rule, but need further investigations [8]. On one hand, normalization is needed to combine heterogeneous scores outputs of uni-biometric systems. The parameters of normalization methods are computed in the training step and are tested in the validation step. On the other hand, the main problem that encounters the biometric evaluation is the disparity between scores of samples belonging to the same class: genuine or impostor. This causes misidentification of the real class and results in false acceptance and false rejection [32]. The client-specific normalization is a good solution that can suffer from the lack of samples. Therefore, authors in [32] discuss the usefulness of group-specific normalization methods to overcome this problem.

The quality measurement [33] provides a predictor of the system performances. Sample quality is defined as 'scalar quantity that is related monotonically to the performance of biometric matchers' [33].

Many recent investigations [2, 34–37] are interested in the sample quality. They use the sample quality scalars to get a well-adjusted fusion function. Another application considers the system quality to indicate to the fusion approach the relevance of each uni-biometric system. This quality evaluation allows the fusion process to provide an adapted combination according to the identified biometric signals for each authentication [38]. Here, the main issue is to preserve accuracy even if the signal quality changes from an authentication to another. This can be interpreted as considering the quality scalar as a quantification of the evidence degradation due to noisy signals. Authors in [37] use an image reconstruction module that optimizes the false acceptance rate based on the quality measurement of the processed signal. Otherwise, it is important to notice that the perception of a supervised biometric system is out-of-date in recent user-friendly applications that require more convenience [39]. Therefore, quality evaluation is crucial to establish a reliable multi-biometric system that is adapted to

TABLE 3 HCSO parameters

HCSO Parameter	Value
Number of iterations	itMax = 50
Population size	popSz = 100
Decision variables	dimns = 8; (+function number)
Mixture ratio (seeking cats)	MR = 0.8
Dimensions to be changed	CDC = 0.4
Dimensions range	SRD = 1
Number of copies	SMP = 10
Weight	w = [0.9, 0.4] (itMax values)

the variation of the processed data [34]. To summarise, the studied works related to this field employ signal or biometric system quality to ensure the consistency of the biometric decision. Other researchers consider the quality evaluation of the biometric system reliability [35].

Improving system accuracy using a fusion method on a limited number of multi-biometric schemes does not attest its efficiency. Indeed, we can reach improved accuracy using a fusion strategy thanks to the baseline system combination. Here, the authors aim to experiment different functions to fuse different scores outputs of uni-biometric systems and verify this assumption. Therefore, the authors propose a useful platform [26] to test the predefined fusion functions and classify the different biometric system combinations according to their accuracy enhancement. Thereby, the platform gives the best combinations and fusion methods according to the required security level (high, medium, and low). In addition, it determines if the used fusion functions give equivalent accuracy for different multi-biometric combinations. That ensures that the adopted fusion gives steady performances in different situations. It has been proven that typical operations are the best ones in combining limited number of uni-biometric systems [14]. Therefore, the authors aim to propose different fusion functions based on these operations, and study and compare between them on different fused uni-biometric systems. Further, the authors analyse the impact of normalization and quality integration on performances of the studied multi-biometric combinations.

3 | MATERIALS AND METHODS

Choosing a fusion method in score level is very crucial to enhance the provided multi-biometric system performances. The work of the authors provides performance analysis of different fusion functions (see Table 8) that combine elementary operations like sum, product, min and max along many multi-biometric scores combinations generated from the score database. To accomplish that, we start by explaining the GP applied previously [26] and enhanced to be tested and compared on a new and different database. Then, we define the proposed HCSO to reduce errors of a subset of multi-biometric that we

| Data sets | Lausanne protocols | | |
	LP1	LP2	Fusion protocols
LP Train client accesses	3	4	NIL
LP Eval client accesses	600 (3 × 200)	400 (2 × 200)	Fusion dev
LP Eval impostor accesses	40,000 (25 × 8 × 200)		Fusion dev
LP Test client accesses	400 (2 × 200)		Fusion eva
LP Test impostor accesses	112,000 (70 × 8 × 200)		Fusion eva

TABLE 4　The Lausanne and fusion Protocols of the XM2VTS score database [14]

Modality	Referenced biometric systems
Fingerprint	NIST Fingerprint system
Still face	Omniperception Affinity SDK face detector; LDA-based face verifier
Iris	A variant of Libor Masek's iris system

TABLE 5　Modalities and their biometric systems in the Biosecure score database [27]

analyse through the analysis performed here. After that, we describe the databases that we use in our experimental study. Finally, we discuss the proposed experiments.

3.1 | GENETIC PROGRAMMING

We use GP to get a subset of optimal functions among the constructed ones that are represented using a tree structure. This section explains the used GP simulations that conduct us to preliminary results discussed in [26]. The aim of the authors is to re-explore the proposed optimization and test a diversity of functions that have good performances. The mutation and crossover operations allow exploring search space of different trees by modifying operations on the selected nodes. As we perform tests on a score database, we can vary the number of leaves of the initial population to match the number of input scores of the tested score database.

Figure 1 recapitulates the GP parameters. Figure 2 summarizes the applied GP to get the best trees or functions that we use for fusion and test on the score database. All the population trees are tested using the evaluation protocol offered by the XM2VTS database. As we perform testing on different databases, we give details of the train and test data later.

As shown in Figure 2, in each simulation, we start with the initial population, generated randomly or from the previous simulation by varying the death nodes (between 2 and 8) and the number of nodes that corresponds to a predefined interval for the number of leaves (at least 8 leaves that support the input scores). Afterwards, we evaluate the half total error rate (HTER) of each tree that combines the database scores. Then, according to the fitness of the population or the number of generations, the process continues with tree transformations until it reaches the best population. In our case, we succeed to optimize all the population to get a mixture of functions with an HTER in a reduced interval. Here we present details of the GP used to explore the search space:

Algorithm 1: Genetic Programming

```
Result: Give an optimized solution
Generate the initial population pop;
foreach tree t of the population pop do computeFitness(t);
repeat
   /* fitness is computed using this formula
      max(avg(hter),max(hter)-min(hter)            */
   /* bestList is an ordered list that includes trees whose
      fitness is less than 0.05; selectedList includes
      selected trees with Roulette Wheel              */
   SortList(pop,bestList);
   RouletteWheelSelection(pop, selectedList);
   /* mutation                                        */
   nbtrees ⟵ rand(nbmax);
   for i ⟵ 1 to nbTrees do
      prob ⟵ rand(1);
      tree ⟵ RandomTreesSelection(bestList);
      if prob < mutationprob then
         mutate a number of nodes of tree with operations from
           operation list;
      else
         permute operations between couples of nodes of tree;
      end
   end
   /* crossover                                       */
   nbCrossover ⟵ rand(nbmax);
   for i ⟵ 1 to nbCrossover do
      treeA ⟵ RandomTreesSelection(selectedList);
      treeB ⟵ RandomTreesSelection(selectedList);
      repeat
         apply crossover between couples of nodes of treeA et treeB;
      until ratio of crossed nodes reached;
   end
   foreach tree t of the population pop do computeFitness(t);
until best fitness or max generations reached;
```

3.2 | Hybrid CSO (AVICSO and NMCSO)

Now, we combine average-inertia weighted CSO (AICSO) [40] with normal mutation strategy-based CSO (NMCSO) [41]. The first introduces a balance between the global and the local search where the second reinforces the global search to avoid premature convergence to local optima. In our case, we only apply the average on the velocity using this formula:

TABLE 6 Data details of the Biosecure score database [25]

Data sets		Number of match scores per subject	
		Development	Evaluation
Session 1	Genuine	1	1
	Impostor	103*4	206*4
Session 2	Genuine	2	2
	Impostor	103*4	126*4

TABLE 7 Comparison of GP optimization with state-of-the-art results

	DE [19]	DE [20]	PSO [21]	Our GP
XM2VTS AUC	0.99981	-	-	0.99991
XM2VTS HTER	-	1.119	-	0,1
XM2VTSEER	-	-	1.887	0,11

Our GP outperforms DE and PSO applied on the same database.

TABLE 8 Non exhaustive list of functions generated on the XM2VTS score database

	Function
Fct1	avg (S$_1$,avg(min(min(S$_2$,S$_3$)+S$_4$*S$_5$,S$_6$),min(S$_7$,S$_8$)));
Fct2	S$_1$+avg(min(avg(S$_2$+S$_3$, min(S$_4$,S$_5$)),S$_6$),min(S$_7$,S$_8$));
Fct3	S$_1$+avg(max(avg(S$_2$+S$_3$, avg(S$_4$,S$_5$)),S$_6$),min(S$_7$,S$_8$));
Fct4	S$_1$+avg(min((min(S$_2$, S$_3$))+S$_4$+S$_5$,S$_6$]),min(S$_7$,S$_8$));
Fct5	avg(max(max(S$_1$, S$_2$),(S$_3$+((S$_4$*S$_5$)-S$_6$))),avg(S$_7$,S$_8$));
Fct6	max(avg(S$_1$, S$_2$),S$_3$)+avg(S$_4$,S$_5$)+min(S$_6$,S$_7$)+S$_8$
Fct7	avg(S$_1$, S$_2$)*S$_3$+avg(S$_4$,S$_5$)+min(S$_6$,S$_7$)+S$_8$
Fct8	S$_1$+avg(max(avg(S$_2$+S$_3$, S$_4$+S$_5$),S$_6$),min(S$_7$,S$_8$));
Fct9	avg(avg(min(S$_1$, min(max(S$_2$,S$_3$),S$_4$)),S$_5$), min(avg(S$_6$,S$_7$), S$_8$));
Fct10	avg(avg(S$_1$, avg(S$_2$,S$_3$)),min(S$_4$,S$_5$)+S$_6$+S$_7$+S$_8$);

$$x_{i+1} = x_i + \frac{(v_{i+1} + v_i)}{2} \qquad (1)$$

Now, the problem is defined as the search of scores weights and fusion functions that optimize the EER of the studied multi-biometric system. The modification of the used function by each cat of the CSO population occurs when we detect that the current function does not reduce the EER.

CSO algorithm is based on two phases that simulate the cat's behaviour which are seeking mode and tracing mode. To find the optimal weights, we based our optimization on the development set and test these weights on the evaluation set. The proposed HCSO takes the following steps to get the optimal solutions:

$$x_{i+1} = x_i + (SRD * x_i * N(0,1)) \qquad (2)$$

Step 1 Initialize all parameters of seeking and tracing mode and evaluate the fitness of N cats represented by their decision variables that are the function fusion chosen randomly combined to our scores weights initialized randomly in the defined velocity range. The weights are modified in steps 3 and 4, while the function number is controlled to diversify solutions if we do not find solutions that reduce the function error.

Step 2 Classify the cats into seeking and tracing mode according to the mixture ratio (MR) that defines the ratio of cats in seeking mode.

Step 3 Go through the seeking mode and update the copies (added according to seeking memory pool SMP-1) of parent cats with normal mutation, instead of traditional random mutation, and SRD (seeking range of the selected dimension) parameter that defines the mutative ratio. The normal updates the cats positions that are randomly distributed according to the range of the resting cats. The used formula is as follows:

The CDC (counts of dimension to change) dimensions define the ratio of random dimensions that are modified.
Then evaluate the copies and select the next position for each cat.

Step 4 Go through tracing mode by updating dimensions of the cats in movement according to w that defines an inertia value to the velocity formula defined by:

$$V_{k,d} = w_{iter} * V_{k,d} + r * c * \left(X_{best,d} - X_{k,d}\right) \qquad (3)$$

The decreasing value of w enables one to maintain a global search at the beginning and gradually move to local search. This updated velocity for each dimension d is then added to the position of cat$_k$.

Step 5 Check bounds of the solution and explore other functions for the best solution if there is no fitness improvement after a defined number of iterations. In that case, we propose to enhance the global search by exploring other bad solutions. Then, combine cats of seeking and tracing mode and repeat step 2 to step 5 until reaching the fixed number of iterations.

The HCSO algorithm is proposed to test the generated multi-biometric instances and improve those that need improvement. Also, it allows us to bring out the best functions and confirm or not our function ranking according to non-optimized weighting scores. The chosen parameters are described in Table 3:
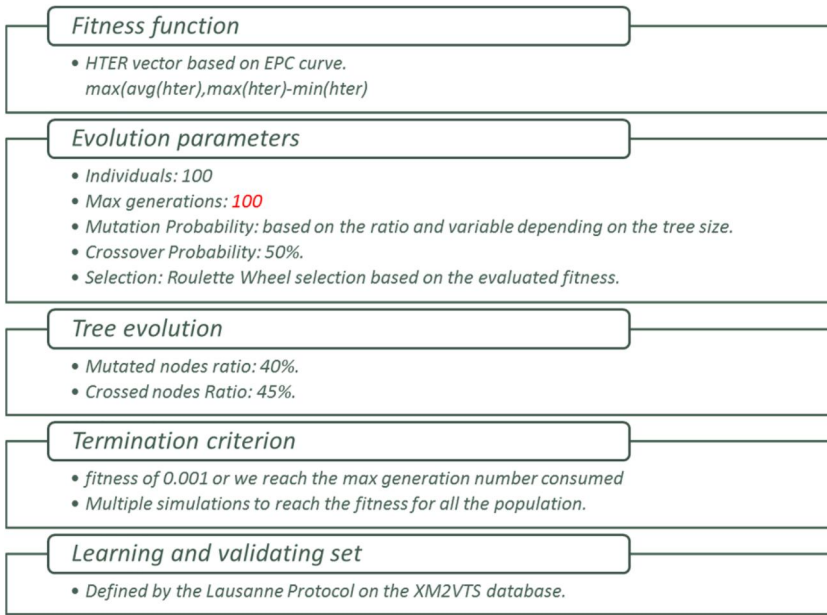
Fitness function
- HTER vector based on EPC curve.
  max(avg(hter),max(hter)-min(hter)

Evolution parameters
- Individuals: 100
- Max generations: *100*
- Mutation Probability: based on the ratio and variable depending on the tree size.
- Crossover Probability: 50%.
- Selection: Roulette Wheel selection based on the evaluated fitness.

Tree evolution
- Mutated nodes ratio: 40%.
- Crossed nodes Ratio: 45%.

Termination criterion
- fitness of 0.001 or we reach the max generation number consumed
- Multiple simulations to reach the fitness for all the population.

Learning and validating set
- Defined by the Lausanne Protocol on the XM2VTS database.

FIGURE 2 Genetic programming on the XM2VTS score database

3.3 | Databases

First, we need to present the databases used in our experiments. We choose these databases as they are constructed with respect to the biometric database protocols. Furthermore, they are more suitable for our comparative study as they give a great number of scores computed from different biometric systems.

3.3.1 | XM2VTS score database

We use the XM2VTS score database [14], where the scores are computed based on the publicly available XM2VTS face

and speech database. These scores give access to the provided baseline systems with their documentation for further experiments. Publishing score results of these baseline systems is important for researchers to compare them in more detail with their obtained scores. However, fusion, in this context, is less considered as it is more useful to use scores for combining on common settings instead of a simple comparison of outputs. This database provides well-defined protocols (train and test data details are described in Table 4) and a set of evaluation tools such as the DET curve, expected performance curve (EPC), and the HTER significance test, which we use in our fitness evaluation as an optimization metric.
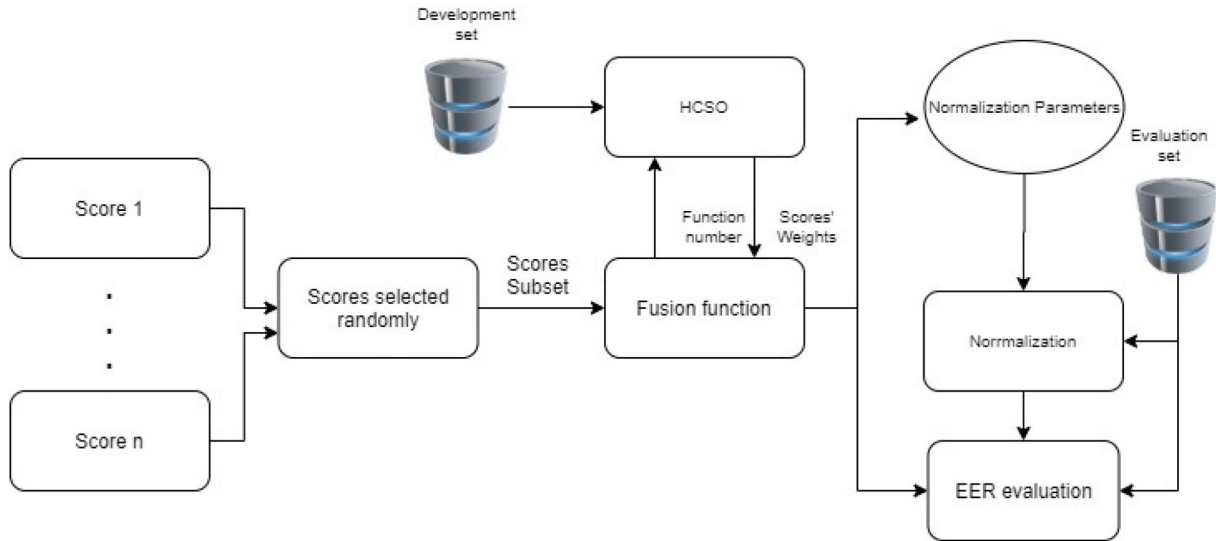
FIGURE 3 Multi-biometric score fusion evaluation

3.3.2 | Biosecure score database

The Biosecure score database is a benchmark of quality-dependent, client-specific, and cost-sensitive fusion algorithms. It contains fingerprint, face and Iris modalities acquired with different sensors and scenarios. Baseline systems are evaluated using biometric verification systems cited in Table 5 [27]. The benchmark contains 24 streams of matching scores from two sessions divided into development set and evaluation set as represented in Table 6.

The database simulates a real test environment for fusion schemes such as scores of cross-devices, quality evaluation of acquired traits and failures due to sensors or unmeasurable quality. In addition, it provides tools for analysing system performances based on the processed data.

3.4 | Multi-biometric optimization analysis

Here, we aim to analyse the impact of fusion on the performances of biometric systems. Thus, we need to generate different combinations of existing biometric systems as shown in Figure 3. Then, we apply multiple fusion strategies on the selected scores. This selection, carried out on n provided scores, gives a subset of scores that are used in the fusion process. We use the development set to compute normalisation parameters. Then, we test each multi-biometric combination using the evaluation set. The random selection provides 2000 multi-biometric instances and covers different modalities and traits' instances that tests sensor cross-matching. We use the same fusion functions tested in [26] in the extended comparative study, and we compare them with typical fusion operations, normalization functions and evolutionary methods applied in score fusion.

First, we use GP results obtained previously from the XM2VTS [26]. Figure 4(a) shows the graph oscillations that

illustrate the progress of GP from the first to the last simulation. It shows the percentage of minimal HTER obtained for the population at each generation. Figure 4(b) illustrates the fitness average according to the number of crossover and mutation. The GP provides a set of the best trees that represent a fusion function of elementary operations. Figure 5 shows the enhancement of the average HTER with GP simulation along 100 generations using the crossover that evolves the population to the optimal value and the roulette selection that enlarges the exploration. The mutation allows avoiding premature convergence. As a result, the HTER average is reduced to a range of [0, 5%] in the last experiments [26] that decreases under 1% using two successive simulations. After that, we start our experimental study according to the following steps:

1. First, a set of trees is selected from our previous simulation. The selection was conducted randomly along all generations to get a set of generations according to the HTER. Then we select different trees of each one.
2. We use these trees to fuse matching scores provided by different biometric systems of the Biosecure database.
3. We use the development set to compute the min-max standard deviation applied in normalization.
4. First, we apply a user-specific weighting using the normalized database quality measures giving the sample fidelity to the claimed ID, multiplied by the standard deviation of min-max for each score (computed from the development set). We compute the weight as follows:

$$\left(\frac{1}{2} * \left(\left(x_t - \overline{x}\right)^2 + \left(x_q - \overline{x}\right)^2\right)^{1/2}\right)^{-1} * \text{min\_scores,max\_scores} \tag{4}$$

where x_t is the template quality (claimed id) and x_q is the query quality (true id). This is a measure of dissimilarity between the two measurements.
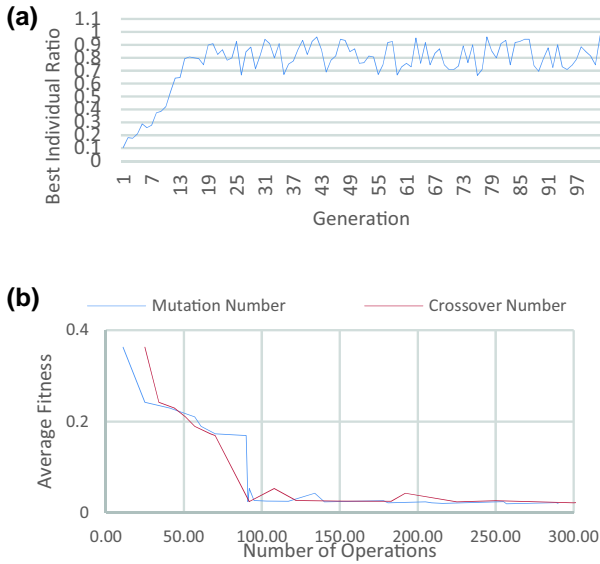
(a)

(b)

FIGURE 4   Impact of mutation and crossover number on fitness average [26]
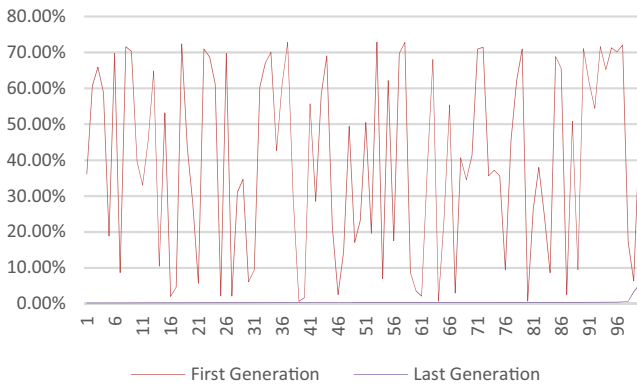

FIGURE 5   The Fitness of initial and last generations on the XM2VTS score database after only two simulations

5. Second, we use Frobenius norm for the HCSO algorithm to rescale the distribution of client and impostor using a group-centred strategy [32] that aggregates the score matching of different samples belonging to the same group. In our context, we consider only two groups that are genuine and impostor. The normalization is the norm 2 between the two vectors and is computed from the development set and applied on the evaluation set.

4 | EXPERIMENTS AND DISCUSSION

In the experiments conducted, we get the optimized population illustrated in Figure 5 (the last population). As shown in this figure, the error range [0.67, 72.89]% of the initial population is reduced after the first simulation to [0.3, 0.87]%, then [0.1, 0.34]%. We get the best trees from the generated populations and use them in the next experimental study. We

compare our Area Under Curve and HTER achieved in the XM2VTS score database with two evolutionary methods that use the same evaluation metrics (see Table 7).

4.1 | Experimental study

In the experimentation, we aim to test different configurations to analyse the impact of using the same functions and varying the fused uni-biometric scores. The study is essential as it demonstrates that fusion depends on the fusion function and the scores that are fused [1]. Also, we compare the proposed functions and the HCSO function to other typical operations, the fusion normalization proposed in the literature [5], and the applied evolutionary algorithms for fusion weighting [21, 25]. We use the algorithm presented here:

```
Result: Comparison of multibiometric systems
Generate n Multibiometric systems of 8 baseline systems pop;
Tree selection using random sampling method;
for i ← 1 to n do
    Test baseline systems;
    Test usual fusion operators;
    /* functions selected from GA simulation on XM2VTS score
       database                                              */
    for j ← 1 to nbfunctions do
        Test fusion of baseline scores;
        Test fusion of weighted function on normalized scores;
    end
    Functions comparison with usual operators;
    Significance test of obtained results
end
```

We perform the experiments on the Biosecure score database [27], a collection of matching scores computed from the Biosecure biometric databases that proposes a quality evaluation of the captured signals based on the features richness, environmental effects and signal properties. The quality is a crucial information about the studied fusion schemes to discard undesired influence of high scores provided by poor quality traits. We select a subset of fx10functions that provides low HTER error in the XM2VTS database that is representative of groups of functions with similar HTER.

The selected functions are shown in Table 8. Based on these functions, we search the optimized weights using the proposed HCSO algorithm that explores them to find the best one. The experiments are done on MATLAB 20a. The HCSO optimization is performed on two PCs: Intel i5 processor, 8 GB RAM, and Intel i7 processor, 16 GB RAM, and takes three days for 1000 multi-biometric scores using parallel execution with 6 MATLAB workers.

Figure 6 illustrates the statistics (MIN, MAX, AVG and STD) of the cited functions that are tested using the evaluation set of the score database. These functions are classified according to the variance between the two sessions (see Table 6) to measure the function performance variation. In this ranking,
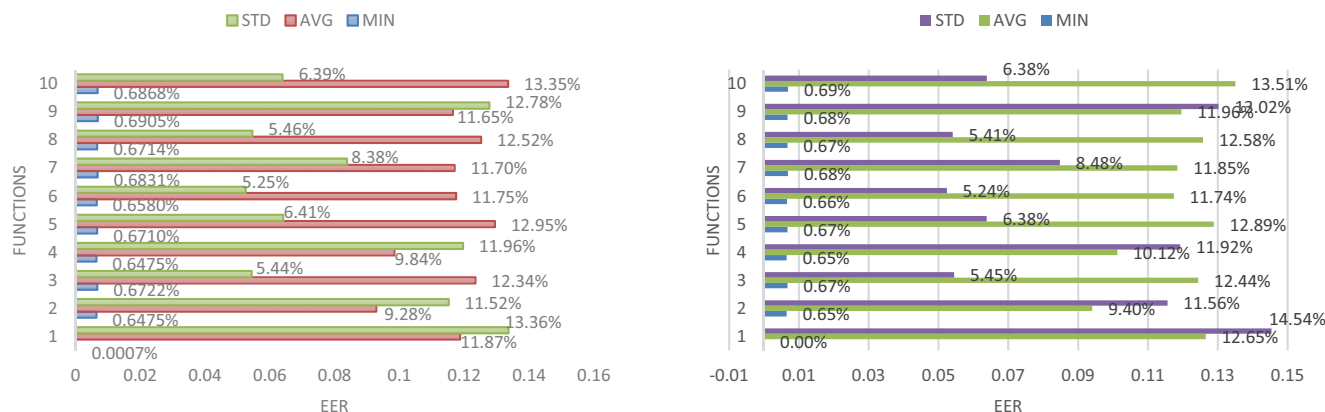
FIGURE 6 MIN, AVG and STD equal error rate of the tested fusion functions on 1000 and 2000 multi-biometric combinations

TABLE 9 Functions ranking results according to statistics and HCSO results

Function	Ranking with 1000 instances				Average	Mean of 2000 instances	HCSO rank
	Rank1Min	Rank2Max	Rank3Average	Rank4Std			
Fct1	1	10	6	9	6.5	7.25	1
Fct2	4	7	1	7	4.75	4.5	10
Fct3	2	2	4	1	2.25	4.5	5
Fct4	5	8	2	8	5.75	5.5	9
Fct5	7	1	3	3	3.5	5.75	2
Fct6	6	6	7	5	6	2.25	3
Fct7	8	4	9	6	6.75	6.25	6
Fct8	3	3	5	2	3.25	4.5	4
Fct9	10	9	10	10	9.75	7.25	7
Fct10	9	5	8	4	6.5	7.25	8
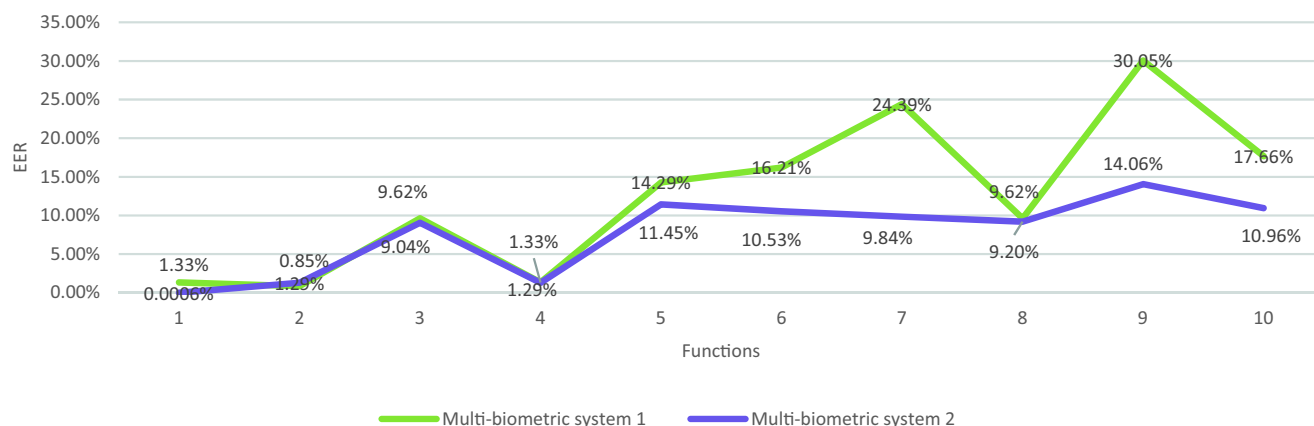
FIGURE 7 Comparison between two of the best combinations of biometric scores from two simulations with 1000 and 2000 multi-biometric instances

we compute the sum of ranking considering that Session 2 is more important as recommended in the database documentation. We use the sum of these ranks to show the best functions. As a result, Function 3, 8, 5 and 2 whose averages of ranks is less than 5 are the best ones as illustrated in Table 9. The HCSO algorithm gives a different ranking by selecting the best function according to the EER and the fused matching scores. In fact, function 5 and 8 remains the most recurrent with functions 1 and 6.

While the new ranking that considers the 2000 combinations is similar for 4 functions with the previous simulation of 1000 combinations, the HCSO algorithm ranks 6 functions
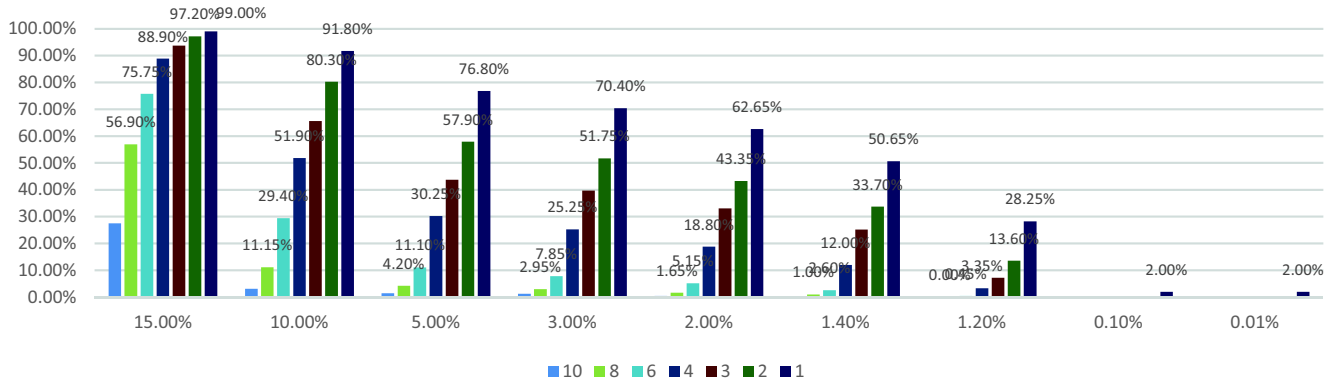
FIGURE 8 Cumulative ratio of multi-biometric fusion schemes according to the number of functions and fixed maximum EER

TABLE 10 EER statistics of compared fusion functions with t-norms and evolutionary methods

Method	Multi-biometric instances	EER Std average	EER average	Min EER
Hamacher t-norm [5]	419	13.52	21	0.13
Frank t-norm [5]	13	12.3	25.12	0.19
PSO [21]	-	-	-	1.887
BSA [25]	-	-	-	1.57
Our functions without HCSO weighting	1000	[5.25, 13.36]	[9.28, 13.35]	0.0006
Our functions with HCSO weighting	1000	8.27%	4.31%	0.00004

Abbreviation: BAS, Backtracking search optimization algorithm
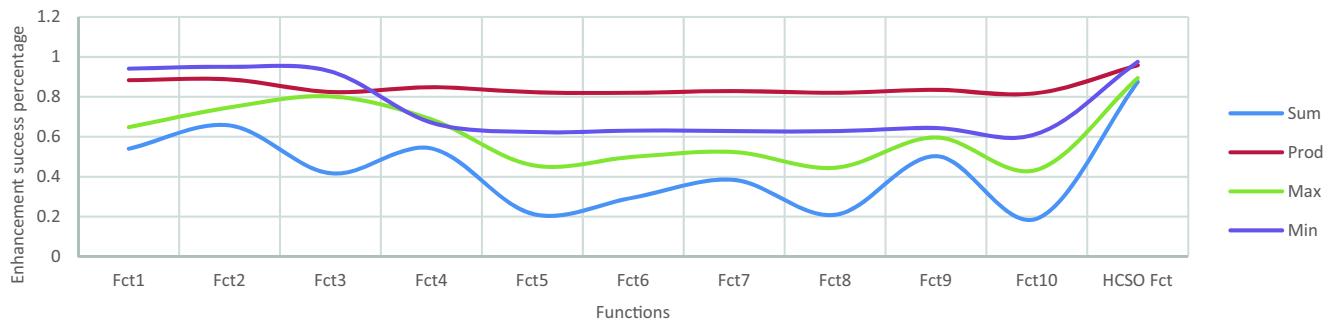
FIGURE 9 Percentage of enhanced multi-biometric scores combinations using our fusion functions comparing with usual fusion operations

similarly. We consider that the two rankings are similar if the difference is less than 1.

Function 3 is the best function with EER less than 33%. In Figure 7, we can observe that functions 1 and 2 give better result for the best multi-biometric fusion from two simulations and provide the lowest EER – 0.0006%. Our results remain coherent because function 2 is one of the best-ranked functions. However, it is the less selected by the HCSO algorithm for 2.40% of the multi-biometric instances. In Figure 8, we present statistics of the tested fusion functions on multiple multi-biometric fusion schemes. It illustrates the percentage of enhanced multi-biometric performances under the corresponding EER on the x-axis for each sub-set of functions. As we decrease the number of functions that must provide a

reduced EER, the number of multi-biometric instances increases. For example, six different functions enhance the EER of 7.85% of the tested ones under 3% while 70.40% of them have only one function with the same EER. These results seem interesting as they demonstrate the influence of the fused baseline outputs scores on the resulted EER. Nevertheless, we provide 40 multi-biometric fusion schemes that are optimised to an EER less than 0.001%. This is not verified if we increase the number of functions which explain the decreasing ratio until 0%. In addition, the simulation demonstrates a significant enhancement compared to operations: sum, product, min, max. The proposed functions provide an EER less than typical operations for 89.5% of the 2000 multi-biometric combinations considering only the quality weighting.
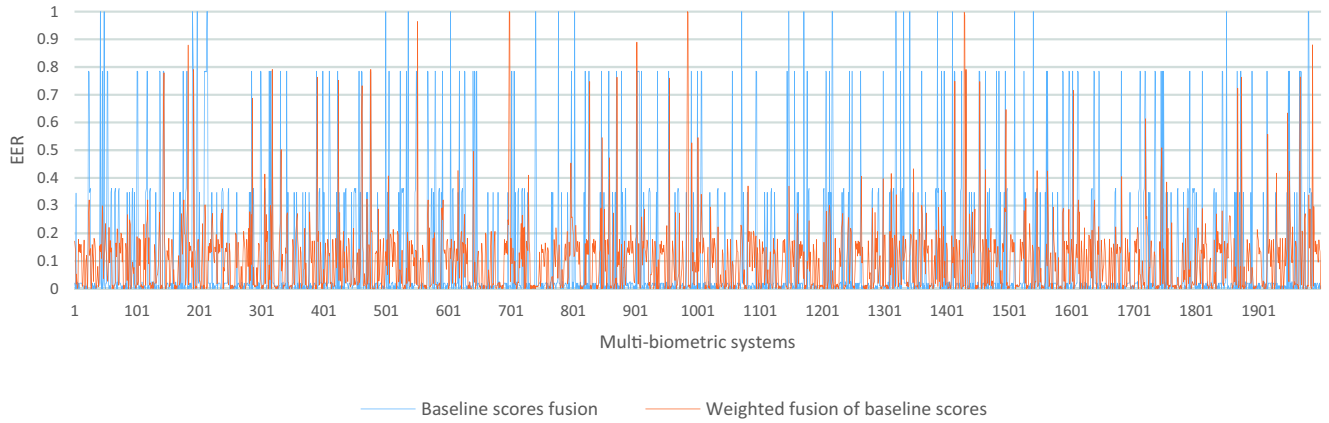
FIGURE 10 EER of normalized and weighted fused normalized scores computed for 2000 multi-biometric combinations of the evaluation set
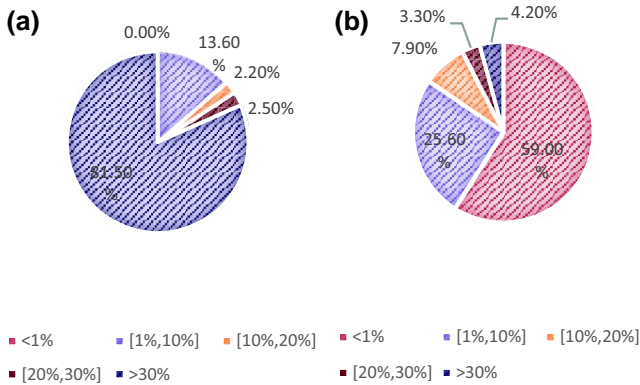

FIGURE 11 (a) Multi-biometric combinations percentages according to their EER after one iteration and (b) EER of the multi-biometric combinations on the evaluation set after 50 iterations of the HCSO algorithm

FIGURE 12 Comparison of EER using user-specific quality weighting with the HCSO weighting

We compare our results to the recently used Hamacher t-norm and Frank t-norm normalization methods that we implement on the tested data of the Biosecure database. Also, Table 10 gives EER comparison with the state-of-the-art results. The HCSO weighting optimization gives the lowest EER average. As shown in Table 10, only 50% of the tested multi-biometric scores are fused successfully in the best case. Our functions outperform the two best norms that give good results in [5], PSO [21], and Backtracking search optimization algorithm [25]. The proposed HCSO algorithm enhances the selected function using the score weighting and reaches more than 87.5% EER reduction compared to typical operations (see Figure 9). After 50 iterations, more than half of the multi-biometric combinations have less than 1% EER as demonstrated in Figure 11.

The performed experimental study confirms the assumption that having good results after fusion does not assert the robustness of the chosen fusion method. Here, we get a significant STD for some functions like f1, as seen in Figure 6. This means that the resulted functions depend on the selected scores of the biometric systems. The robustness of the chosen baseline systems used in fusion plays a crucial role in
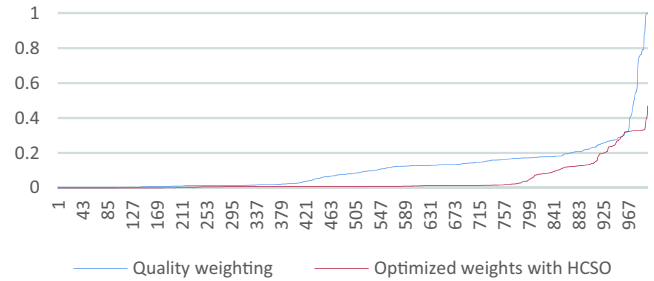
performance enhancement. Therefore, comparing different fusion functions based on the same data and baseline systems is required to affirm whether fusion has improved baseline system performances or not. Besides, testing different scenarios is always important because the ideal case does not exist. Generally, system failure is caused by system weaknesses to face some special scenarios, which are not considered, like sensor failure or insufficient image quality that affect decision-making. Hence, we must deal with these drawbacks to prevent system failure in a real environment. The chosen scores database is built upon some of these scenarios. In the experiments conducted, we test quality weights that we propose based on score quality available in the Biosecure database. We compare it to a normalized score without employing weights. The results are presented along 2000 multi-biometric combinations in Figure 10. As shown, normalization provides the lowest values. However, if we omit zero-values, mostly caused by missing values in the database, we can see that quality weighting provides a steady range and reduces significant EER fluctuation. However, max peaks are caused by missing quality scalars for some scores in the testing database.

Figure 12 compares between the EER of 1000 combinations using user-specific quality weighting and the added HCSO weights. Combining the two quality measures for baseline scores and those related to the signal quality is more performing as demonstrated by these results.

In conclusion, searching for the best combination, instead of proving the effectiveness of a unique process, remains

TABLE 11 One-way ANOVA test summary

Source of variations	Sum of squares	Freedom degrees	Squares average	F	P-value	Critical F value
Between Groups	169.8343658	13	13.06418198	562.1476607	0	1.720855479
Within groups	324.8687359	13979	0.023239769			
Total	494.7031017	13992				

**TABLE 12** Significance of the *t*-test comparing between the proposed functions and the typical fusion operations

	Sum	Product	Max	Min
Fct 1	0.041	0.0002	4.74E-17	3.17E-170
Fct 2	1.87E-17	1.28E-09	6.09E-50	4.39E-188
Fct 3	0.056	0.0006	6.84E-63	2.89E-175
Fct 4	2.52E-12	2.33E-08	1.59E-41	1.02E-184
Fct 5	0.037	0.005	1.53E-29	1.50E-169
Fct 6	5.01E-07	6.73E-05	5.54E-71	8.09E-180
Fct 7	0.0005	8.95E-05	2.63E-38	7.83E-178
Fct 8	0.322	0.001	1.03E-60	9.50E-175
Fct 9	0.011	0.0001	1.66E-20	1.50E-173
Fct 10	1.50E-06	0.0165	9.41E-31	1.88E-168

important in future studies. Here, we optimize performances to reach the best multi-biometric systems on two simulations illustrated in Figure 7. This multi-biometric system gives nearby errors for the ten functions on session 1 of the evaluation set. The EER is in the range of [0.0002 6.88]%. However, the best functions with small variance (less than 1.5%) between the two sessions are 1, 2 and 4 in this case. With the HCSO optimization, function 1 is the most selected for 24.30% of the tested multi-biometric combinations.

4.2 | Significance test

To prove that the proposed functions outperform the typical fusion operations, we use the t-test to approve the significance of the variation between the two sets. First, we use a paired test to analyse the average difference in one direction. We get a t-value1.63 less than the critical t-value 1.64. So, the hypothesis of significant improvement is assumed ($P < 0.1$). The proposed fusion is significantly better than typical fusion operations with 90% confidence interval. Second, we use a paired test to analyse average difference assuming that the two variances are different. The test is almost successful ($P = 0.05 < 0.1$ and t value = 1.58 < 1.64).

Next, we perform a significance t-test between pairs of our proposed functions and the four typical operations. Results are significant for all pairs as seen in Table 12 except for function 8 whose average is close to sum's average. Our functions perform better than all typical operations do.
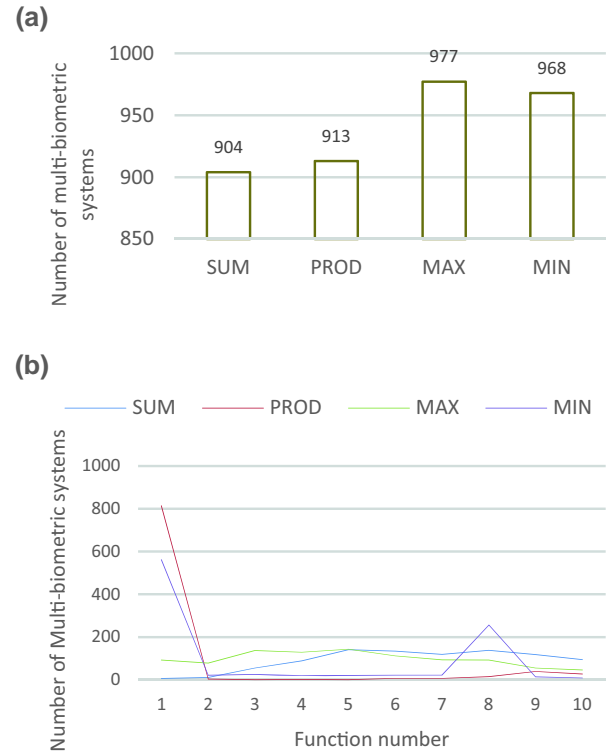
FIGURE 13 (a) Number of multi-biometric scores combinations where at least one function outperforms the typical operations and (b) corresponding number of multi-biometric combinations and number of succeeded functions

The comparison between our functions and the typical fusion operator is illustrated in Figure 13. It represents the number of multi-biometric combinations, from 1000 tested combinations with the proposed functions that outperform the typical operations. Fct2 is the function that yields the best enhancement comparing to the Sum operator, while it reaches proximate percentage to Fct1 comparing to Prod and Min operations. [Fct4-Fct10] gives approximately the same results that outperform Max and Prod ones for more than 60% of multi-biometric combinations.

In summary, the percentage of success exceeds 90.4% compared to the typical operations. Figure 13 describes the total number of multi-biometric combinations for which we find at least one function that outperforms sum, prod, max and min. Also, the number of functions that succeeds in each multi-biometric system is numbered, where each series represents one of the typical functions, in Figure 13(b). The results of max and sum approach the min and prod, respectively.

TABLE 13 Best multi-biometric fusion details

	Face	Fingerprints with same sensors	Fingerprints with crossed sensors	Iris
1	Webcam (low resolution)	Thermal: right/left thumb, right index,	Left index, left middle finger	-
	LDA-based face verifier	Optical: left thumb, right index	NIST fingerprint system	
2	Webcam (low resolution)	Thermal: right thumb, right index	Left index, left middle finger	-
	LDA-based face verifier	Optical: Right/left thumb, right index.		
3	Webcam (low resolution)	Optical: right middle finger	Thermal: left thumb	Left eye
	LDA-based face verifier	Thermal: right/left index, right middle finger		
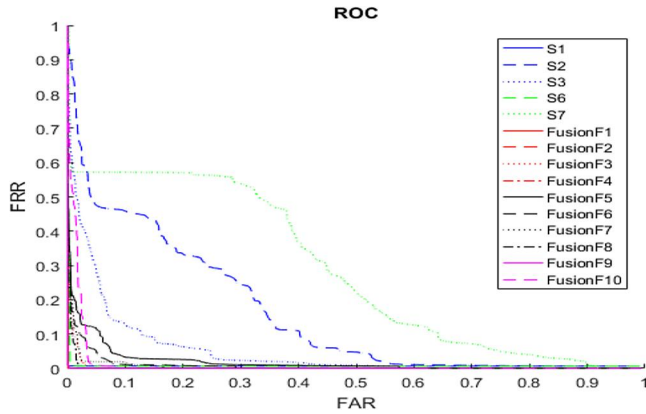
FIGURE 14 ROC curves (EER = 0.8% for Function 2) of all the used functions compared to a subset of fused scores and baseline systems

Finally, we use the ANOVA test to confirm that the results of our experiments are statistically significant. The H0 hypothesis of the ANOVA test is that the tested functions do not optimize the EER, whereas the H1 hypothesis states that the tested functions optimize the EER.

Table 11 shows the result of the ANOVA test; the probability of H0 hypothesis is very low (*p*-value = 6.59e-32). It confirms that the results of our experiments are statistically significant, and the tested functions optimize EER.

4.3 | Multi-biometric fusion

In the experiments, we reach an EER under 0.013% and we get three multi-biometric fusion schemes that optimize the EER to the range of [0.0006%, 1.30%] from baseline scores with EER in the range of [1%, 99.55%]. This means that we reach a range improvement of 98%. Our functions outperform the results obtained in [42] using Random Subspace of AdaBoost (RS-ADA) for fusion which reaches an EER of 1.98%.

Table 13 describes the fused traits in these multi-biometric schemes. The traits are face, multiple instances of fingerprint, multiple captures using different sensors and iris. Furthermore, the fusion is done on scores—the result of matching between fingerprint query and template from different sensors. This confirms the usability and significance of quality evaluation for weighting scores in such cases. Figure 14 gives the evaluation

of the 10 fusion functions on one of the best multi-biometric fusion, on two tested sessions. Function 2 is the best function in this case as demonstrated previously in Table 9 using our ranking method. The HCSO reduces the EER of function 1 from 1.33% to 0.9%.

The example illustrated in Figure 14 shows that fusion allows reducing errors caused by divergent scores. Indeed, function 2, 5 and 8 give already good results and optimize Area Under Curve. For instance, the fused scores represent respectively the Face (CANON), Face (CANON with Flash), Iris and two fingerprints taken with different devices as described in the database. The other scores are filtered for dummies values included in the tested database.

5 | CONCLUSIONS AND FUTURE WORK

Here, the authors propose a new multi-biometric fusion scheme based on optimized and weighted functions. The fusion functions are generated using the GP based on arithmetic operations. The score weighting optimization is performed using the HCSO based on average-inertia weighted CSO and Normal Mutation CSO. The authors conduct the experimentation on the Biosecure score database to compare different combinations of the provided scores and find the best solutions that carry out a fixed range of EER. The proposed hybrid CSO optimizes fusion score weights and reach at least 80% enhancement of the tested multi-biometric combinations. The significance test confirms the improvement of our weighted scores fusion compared to typical fusion operations and weights' optimization using evolutionary algorithms applied on similar databases or approaching modalities. With different baseline systems, having probably divergent performances, the system improvement cannot be ensured by combining their scores. Therefore, associating weights to scores is essential in multi-biometric fusion. Quality weighting applied to user scores and classifier scores brings additional information about the used sensor, user behaviour and classifier performances. The experimental study compares a significant number of multi-biometric combinations instead of a limited number. Therefore, this allows analysing more expressively how much fusion and quality integration in the biometric process is important to get better results and

confirm the effectiveness of each function to improve multiple configurations rather than a configured one. This can be used in the different contexts to verify whether the merit of fusion methods outstrips the performances of baseline systems. In future work, it has been planned to test another variant of CSO to reduce the time cost using parallel CSO and combine it with the current version.

ACKNOWLEDGEMENTS

ORCID

Saliha Artabaz   https://orcid.org/0000-0002-9673-1507

REFERENCES

1. Singh, M., Singh, R., Ross, A.: A comprehensive overview of biometric fusion. Inf. Fusion. 52, 187–205. https://doi.org/10.1016/j.inffus.2018.12.003
2. Dwivedi, R., Dey, S.: Score-level fusion for cancelable multi-biometric verification. J. Pattern Recogn. Lett. (2018). https://doi.org/10.1016/j.patrec.2018.04.022
3. Dass, C., et al.: A principled approach to score level fusion in multimodal biometric systems. In: Lecture Notes in Computer Science 3546, 1049–1058 Springer, Berlin (2005)
4. Piciucco, E., Maiorana, E., Campisi, P.: Biometric Fusion for Palm-Vein-Based Recognition Systems. In: Communications in Computer and Information Science, pp. , 766, 18–28.TIWDC Springer, Cham (2017)
5. Gopal, S.S.: Accurate human recognition by score-level and feature-level fusion using palm–phalanges print. Arabian J. Sci. Eng. 43(2), 543–554. Springer (2018)
6. Gavrilova, M.L., et al.: Multi-modal motion-capture-based biometric systems for emergency response and patient rehabilitation. In: Kisku, D., Gupta, P., Sing, J. (eds.) Design and Implementation of Healthcare Biometric Systems, pp. 160–184 (2019)
7. Wadhera, T., et al.: Pre-Clinical ASD screening using multi-biometrics-based systems. In: Kisku, D., Gupta, P., Sing, J. (eds.) Design and Implementation of Healthcare Biometric Systems, pp. 185–211.IGI Global, Hershey, PA (2019)
8. Afzal, M., Srivastava, S.: Score Level Fusion of Multispectral Palmprint with Triangular Conorm. Intl. J. Emerging Res. Manag. Technol. 6(6), 88–91 (2017)
9. Bouzouina, Y., Hamami, L.: Multimodal biometric: iris and face recognition based on feature selection of iris with GA and scores level fusion with SVM. In: Proceedings of bioengineering for smart technologies (BioSMART). 1–7. IEEE IEEE, Paris, France (2017)
10. Sivakumar, S.: Biometric authentication techniques and its future. In: Kumar, A. (ed.) Biometric Authentication in Online Learning Environments, pp. 122–149 (2019)
11. Brostoff, G.: How AI and biometrics are driving next-generation authentication. Biometr. Technol. Today (6), 7–9 (2019)
12. Wang, Y., et al.: The research of identity authentication based on multiple biometrics fusion in complex interactive environment. J. Netw. Intel. 4(4), 124–139 (2019)
13. Ross, A., Nandakumar, K., Jain, A.K.: Introduction to multi-biometrics. In: Patrick, R., Arun, A. (eds.) Handbook of Biometrics, Jain, Anil, Flynn Springer, Boston (2008)
14. Poh, N., Bengio, S.: Database, protocol and tools for evaluating score-level fusion algorithms in biometric authentication. J. Pattern Recogn. 39(2), 223–233 (2006)
15. Sharma, R., Das, S., Joshi, P.: Score-level fusion using generalized extreme value distribution and DSmT, for multi-biometric systems. J. IET Biometr. (2018)
16. Ulery, B., et al.: Studies of Biometric Fusion. National Institute of Standards and Technology United States Department of Commerce (2006)
17. Herbadji, A., et al.: Weighted quasi-arithmetic mean based score level fusion for multi-biometric systems. IET Biometr. 9(3), 91–99 (2020)
18. Safont, G., Salazar, A., Vergara, L.: Vector score alpha integration for classifier late fusion. Pattern Recogn. Lett. 136, 48–55 (2020)
19. Mukherjee, S., et al.: Differential evolution based score level fusion for multimodal biometric systems, pp. 1–7.IEEE Orlando, FL, USA (2014)
20. Mezai, L., Hachouf, F.: Score level fusion algorithm using differential evolution and proportional conflict redistribution rule. IEEE, Budapest (2016). 001051-001056
21. Mezai, L., Hachouf, F.: Score-level fusion of face and voice using particle swarm optimization and belief functions. IEEE Trans. Human-Machine Syst.. 45(6), 761–772 IEEE, (2015)
22. Gong, Y., Chen, J., Su, G.: Quasi-convex optimization of metrics in biometric score fusion, pp. , 998–1003.Sixth international conference on image and graphics, Anhui (2011)
23. Sun, W., et al.: A Survey of Using Swarm Intelligence Algorithms in IoT. Sensors 20(5), 1420(2020). http://dx.doi.org/10.3390/s20051420
24. Ahmed, A.M., Rashid, T.A., Saeed, S.A.M.: Cat Swarm Optimization Algorithm: A Survey and Performance Evaluation. Computational Intelligence and Neuroscience 2020, 1–20 (2020). https://doi.org/10.1155/2020/4854895
25. Walia, G.S., et al.: Robust multimodal biometric system based on optimal score level fusion model. Expert Systems with Applications 116, 364–376 (2019). http://dx.doi.org/10.1016/j.eswa.2018.08.036
26. Artabaz, S., Layth, S.: Study of score fusion and quality weighting in the bio-secure DS2 Database. Soft Computing in Data Science, 1100, 252–261 Springer, Singapore (2019)
27. Poh, N., Bourlai, T., Kittler, J. A multimodal biometric test bed for quality-dependent, cost-sensitive and client-specific score-level fusion algorithms. Pattern Recognition 43(3), 1094–1105 (2010). http://dx.doi.org/10.1016/j.patcog.2009.09.011
28. Tabassi, E., Grother, P.: Biometric sample quality. Encyclopedia of Biometrics. Springer, Boston MA (2015)
29. Wang, Z., et al.: Template protection based on chaotic map and DNA encoding for multimodal biometrics at feature level fusion. J. Inform. Hiding Multimedia Signal Process., January. 10(1), 1–10 (2019)
30. Lip, C.C., Ramli, D.A.: Comparative study on feature, score and decision level fusion schemes for robust multi-biometric systems. In Frontiers in Computer Education, Sambath, S.and Zhu, E. (eds.), Springer, pp. 941-948 Berlin, Heidelberg (2012)
31. Supreetha Gowda, H.D, Hemantha Kumar, G., Imran, M.: Combination of physiological and behavioral modalities based on score level fusion In: Proceedings of electrical computer and communication technologies (ICECCT), pp. 1–5. IEEE, Coimbatore (2017)
32. Poh, N., KittlerRattani, J.A., Tistarelli, M.: Group-specific score normalization for biometric systems. computer society conference on computer vision and pattern recognition - Workshops, pp. 38–45.IEEE, San Francisco, CA (2010)
33. Grother, P., Tabassi, E.: Performance of biometric quality measures. Proc. IEEE Trans. Pattern Anal. Machine Intel. 29(4), 531–543 (2007)
34. Fierrez-Aguilar, J., et al.: Discriminative multimodal biometric authentication based on quality measures. J. Pattern Recogn. 38(5), 777–779 (2005)
35. Kabir, W., Ahmad, M.O., Swamy, M.N.S.: Normalization and weighting techniques based on genuine-impostor score fusion in multibiometric systems. J. IEEE Trans. Inform. Forensics Security, 13(8), 1989–2000 (2018)
36. Zhang, D., Lu, G., Zhang, L.. Finger-Knuckle-Print Verification with Score Level Adaptive Binary Fusion. Advanced Biometrics. Springer, Cham (2018)

37. Theofanos, M.: Biometrics systematic uncertainty and the user. In: Proceeding of biometrics: theory applications and systems. 1–6. IEEE Crystal City (2007)

38. Brown, D., Bradshaw, K.: A Dynamically Weighted Multi-Modal Biometric Security System. Proceedings of Southern Africa tele-communication networks and applications, 254–258 South Africa (2016)

39. Alonso-Fernandez, F.: Biometric sample quality and its application to multimodal authentication systems. Doctoral dissertation. Dept. Signals Systems and Radiocommunications. Universidad Politecnica de Madrid (2008)

40. Orouskhani, M., Mansouri, M., Teshnehlab, M.: Average-inertia weighted cat swarm optimization. In: Proceedings of the international conference in swarm intelligence. pp. 321–328. Springer, Chongqing, China (June 2011)

41. Pappula, L., Ghosh, D.: Cat swarm optimization with normal mutation for fast convergence of multimodal functions. Appl. Soft Comput. 66, 473–491 (2018)

42. Lumini, A., Nanni, L.: Overview of the combination of biometric matchers. Information Fusion. 33, 71–85 (2017). https://www.sciencedirect.com/science/article/abs/pii/S1566253516300446