On Comparing Verification Performances of Multimodal Biometrics Fusion Techniques

Romaissaa Mazouni
University of Sidi Bel-Abbes
computer sciences department
Sidi Bel-Abbes; 22000, Algeria

Abdellatif Rahmoun
University of Sidi Bel-Abbes; EEDIS Laboratory
University of Sidi Bel-Abbes, computer sciences department
Sidi Bel-Abbes; 22000, Algeria

ABSTRACT
Fusion of matching scores of multiple biometric traits is becoming more and more popular and is a very promising approach to enhance the system's accuracy. This paper presents a comparative study of several advanced artificial intelligence techniques (e.g. Particle Swarm Optimization, Genetic Algorithm, Adaptive Neuro Fuzzy Systems, etc...) as to fuse matching scores in a multimodal biometric system. The fusion was performed under three data conditions: clean, varied and degraded. Some normalization techniques are also performed prior fusion so to enhance verification performance. Moreover; it is shown that regardless the type of biometric modality , when fusing scores genetic algorithms and Particle Swarm Optimization techniques outperform other well-known techniques in a multimodal biometric system verification/identification.

General Terms
Pattern Recognition, Multimodal biometrics, Artificial intelligece .

Keywords
Adaptive Neuro Fuzzy Systems (ANFIS) , Genetic Algorithm (GA) , Brute Force Search (BFS) , Support Vector Machine (SVM), Unconstrained Cohort Normalization (UCN).

1. INTRODUCTION
The use of biometric technologies such as mathematical analysis of a unique trait such as fingerprint, iris and retina, has been adopted worldwide and on large scale. Most biometric systems are unimodal. However, the best unimodal biometric systems (usually iris and fingerprint scanners) present many drawbacks due to noisy data, intra-classes variations, non-universality, spoof attacks, and unacceptable
error rates , etc...[1]. Most recent works in the literature emphasize on improve multimodal biometric systems performance by fusing modalities at the score level mainly because it offers the best tradeoff between information content and the ease in fusion [2]. The main idea behind using multiple information sources in multimodal biometrics is that biometric modalities are complementary with each others, which makes the system more robust and accurate [3]. Five decision level fusion schemes are performed in this study .These are Particle Swarm Optimization (PSO), Adaptive Neuro Fuzzy Systems (ANFIS), Genetic Algorithm (GA), Brute Force Search (BFS), Support Vector Machine (SVM).The evaluation process is achieved with and without subjecting biometric scores to a normalization process. Multimodal biometrics are the best solution so far to make biometric systems more accurate. Integration of evidence at the matching score level is the most common approach because it offers the best tradeoff between information content and the ease in fusion [2]. The main idea behind using multiple information sources in multimodal biometrics is that each modality is complementary to the others; this makes the system more robust and accurate [3]. For sake of comparison, five decision level fusion schemes are performed in this paper .These are Particle Swarm Optimization (PSO), Adaptive Neuro Fuzzy Systems (ANFIS), Genetic Algorithm (GA), Brute Force Search (BFS), and Support Vector Machine (SVM). The evaluation process is achieved with and without subjecting biometric scores to a normalization process using Unconstrained Cohort Normalization (UCN). The performance measure used throughout this work is the Equal Error Rate (EER). The rest of the paper is organized as follows. Section 2 shows roughly the theoretical aspects of the different techniques. Section 3 describes the experimental investigations. An overall conclusion is presented in Section 4.

2. FUSION TECHNIQUES
2.1 Brute Force Search (BFS)

This technique is used in the case of two matchers only. The approach is based on the following equation [4].

$$u = x_1 w + x_2(1-w) \quad (1)$$

where u is the fused score, xi is the i-th normalized matcher score, w is a weighting factor in the range [0,1]. w is calculated heuristically using an exhaustive search in order to minimize the Equal Error Rate (EER) [3].

2.2 Support Vector Machine (SVM)
Support Vector Machine (SVM) is an effective classification technique which can be used in a two-class problem (Imposter, Genuine). A Linear SVM aims to find the optimal separating hyper plane that should classify not only the development data , but also unknown test data [5].This hyper plane has the following equation.

$$W_T x + b = 0 \quad (2)$$

Where W is a vector in and b is a constant. There are many possible linear classifiers that are able to separate the data, but the one that maximizes the margin will generalize better than other possible separating hyper planes.

2.3 Adaptive Neuro Fuzzy Inference System (ANFIS)

The combination of techniques of fuzzy logic and neural networks suggests the idea of transforming the burden of designing fuzzy logic systems to the training and learning of connectionist structure, and the learning capabilities to the fuzzy logic systems. In turns, fuzzy logic systems provide to the neural networks a structural framework with high level fuzzy IF and THEN rule thinking and reasoning [6].ANFIS is one of the most popular hybrid neuro-fuzzy inference expert systems [7]. ANFIS has a similar structure to a multilayer feed forward neural network, but the links in an ANFIS only indicate the flow direction of signals between nodes and no weights are associated with the links. ANFIS architecture consists of five layers of nodes. Out of the five layers, the first and the fourth layers consist of adaptive nodes while the second, third and fifth layers consist of fixed nodes. The adaptive nodes are associated with their respective parameters, and get duly updated with each subsequent iteration.

2.4Genetic Algorithm (GA)

Genetic Algorithms have proven their capability of search for optimum solutions in multi-dimensional space without worrying about local minima relying on an elite preservation strategy. They have been intensively investigated in the last decades in optimization problems, and several variants have been proposed in the literature [8]. the fundamental idea is to maintain a population of individuals that evokes overtime through a process similar to birth and natural selection, where each population is a potential solution of the problem. The competition based on fitness leads to the formation of new fitter individual [4]. A simple form of the genetic algorithms is summarized as follows:

* Generate random population of n chromosomes(suitable solutions for the problem)
w0i , i = (w1,w2) i = 1..N
where N : size of population
* Evaluate the fitness f(x) of each chromosome x in the population

$$EER(u = w_1 x_1 + w_2 x_2) \quad (3)$$

* Create a new population by repeating following steps until the new population is complete
+Select two parent chromosomes from a population according to their fitness
+With a crossover probability cross over the parents to form a new offspring. If no crossover was performed, offspring is an exact copy of parents.
+With a mutation probability mutate new offspring at each locus (position in chromosome).
+Place new offspring in a new population wij
* Test: If the end condition is satisfied, stop and return the best solution in current population
* Loop: Go to step 2.
J is the maximum number of iterations.

2.5 Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is a population based technique, inspired by social behavior of bird flocking or fish schooling.PSO shares many similarities with Genetic Algorithms (GA).However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, each single solution is a "bird" (particle). All of particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the flying of the particles. The particles y through the problem space by following the current optimum particles[9]. PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations using the following equations:

$$v[\,] = v[\,] + c_1 * rand() * (pbest[\,] - present[\,]) + c_2 * rand() * (gbest[\,] - present[\,]) \quad (4)$$

$$present[\,] = present[\,] + v[\,] \quad (5)$$

v[ ]: is the particle velocity. present[ ]: is the current particle. pbest[] and gbest[ ] are defined as stated before. rand () is a random number in the range (0,1). c_1, c_2 are learning factors, usually $c_1 = c_2 = 2$

2.6 Unconstrained Cohort Normalization (UCN)

In this paper, the fusion techniques, (SVM, ANFIS, BFS, PSO, GA) are applied on the biometric scores with and without subjecting them to an effective score normalization method. Unconstrained Cohort Normalization (UCN) [3]. It has been proved in [10] that UCN helps improve the robustness of multimodal biometrics. With UCN client individuals scores are adjusted without any prior knowledge of the degradation level of data. It facilitates the suppression of the individual biometric scores for imposters in relation to those for the clients due to these two characteristics, the EER of individual scores are reduced and so does the final score.UCN is a useful mean for appropriately adjusting the individual biometric scores for a client [10]. The effect of degradation can be significantly reduced because these are reflected in L(x) where:

$$S_{ucn} = \log(P(x/\lambda_T) - 1/K \sum_{k=1}^{K} L(x) \quad (6)$$

$$L(x) = Log(p(x)) \quad (7)$$

Having:

$$p(x) = [\prod_{k=1}^{K} P(x/\lambda_T)]^{1/k} \quad (8)$$

Where: K is the Cohort size, $P(x/\lambda_T)$ is the probability for the observed test data given the target model and $P(x/\lambda_k)$ is the probability for the observed test data given the cohort models.

3.EXPERIMENTAL INVESTIGATION AND RESSULTS

This section discusses the results of fusing face and voice scores, using some matching scores databases under different data conditions. Five fusion schemes are used in this paper as said earlier (BFS, ANFIS,GA, Linear SVM, PSO). Each of these methods is used along with the MinMax normalization method and with and without using UCN.

3.1 Clean Data

The datasets considered for the face and the voice modalities in this investigation are extracted from the XM2VTS and TIMIT databases respectively. Using these biometric datasets , a total of 140 clients, 19460 development imposters ([140*(140-1)]) and 19460 test imposters, development data comprises 19600 scores where 140 scores are genuine scores[3].The results for the testing experiments are presented in Equal Error Rates (EERs). Table 1 resumes

the EER values shown in the Det-curves of the first and the second figure of each fusion technique where the first column shows the method, the second column shows the error rate for each scheme without subjecting scores using the UCN, and the third column contains the error rates of all techniques where the scores were normalized using the UCN. As observed in table 1 (Figure 1 and Figure2), multimodal systems outperform the unimodal ones. These results do not show huge difference between different techniques. Classical techniques (BFS) and Artificial intelligence techniques (GA, ANFIS, PSO,SVM). This closeness of accuracy rates is due to the use of clean data which is an ideal situation where it is possible that all techniques give acceptable results . It is also clear that the population based methods (PSO ,GA) gave the best results which is due to their ability to quickly scan a large search space ,the ability to converge toward the best solution for the PSO and the ability to converge away from bad circumstances for the GA. The use of UCN has resulted in reduction of the verification EERs for the individual modalities and for the fused biometrics using all five methods it is very obvious when comparing the Figures (1 and 2) ,in the second DET-curve the voice , GA, ANFIS, PSO, BFS don't not appear because the EER (FAR=FRR) is equal to 0 . It is also obvious that the SVM gave the worst results among all techniques. This is due to its linear separation nature, the high false acceptances, and the false rejections are slightly high because of the overlap region between the imposter and genuine distributions.

Clean-Clean	-UCN EER%	+UCN EER%
Clean face	3.57	1.42
Clean voice	2.55	0
BFS	0.0976	0
GA	0.0376	0
ANFIS	0.7143	0
SVM	2.78	0.0612
PSO	0.036	0

Table 1. EER of methods using clean data

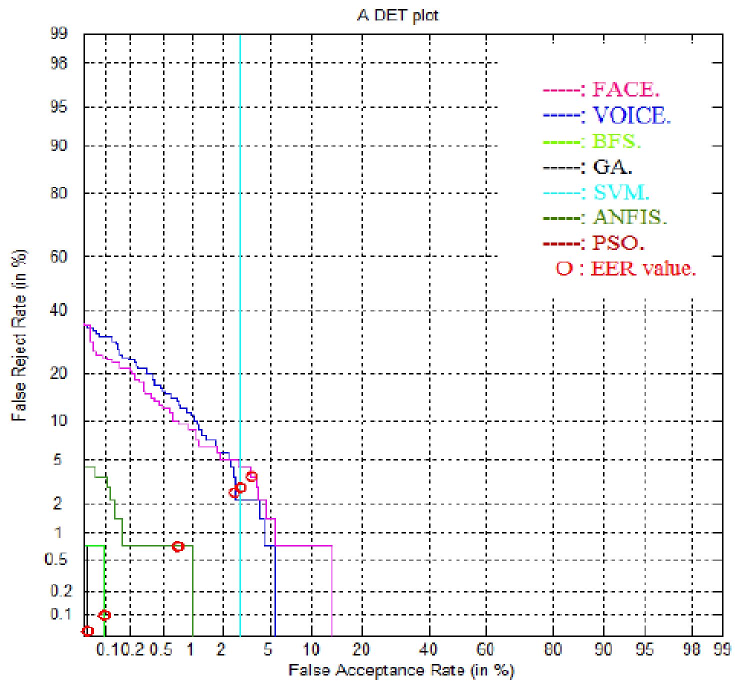

Figure1: Det-Curve for clean data without UCN

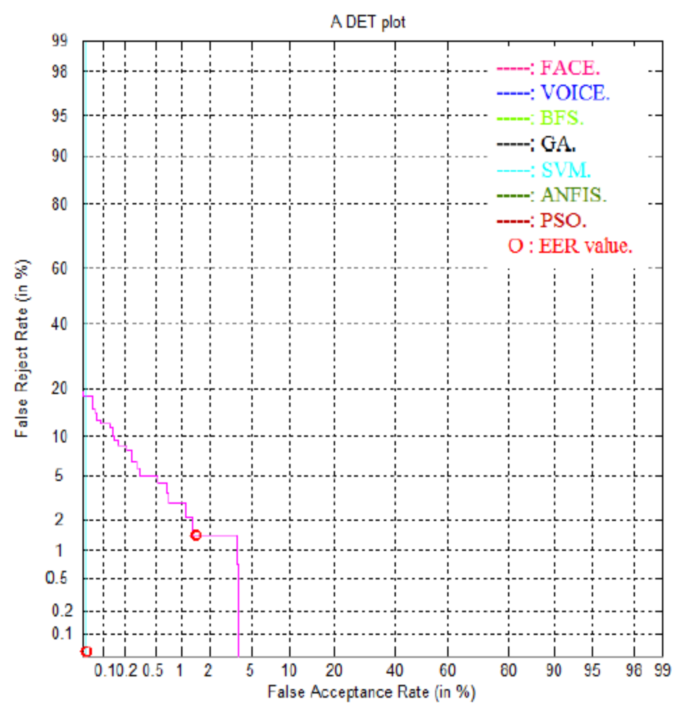

Figure2: Det-Curve for clean data with UCN

3.2 Varied Data

The datasets considered for the face and the voice modalities in this investigation are extracted from the XM2VTS (clean face images) and from the NIST (degraded speech) databases respectively[3]. Using these biometric datasets ,a total of 140 clients, 19460 development imposters ([140*(140-1)]) and 19460 test imposters, development data comprises 19600 score where 140 scores are genuine scores. The results for the verification experiments are presented in Equal Error Rates (EERs) in table.2 . It is obvious that at this stage having exactly the same face error rates the overall accuracy (EERs) of the system decreased This is due to the relatively degraded speech scores. As in table 2, the accuracy (EERs) of individual modalities are less than accuracy of the two modalities fused, using one of the five fusion techniques. Furthermore, multimodal systems outperform unimodal system even under varied data. The results in table 2 show that the use of artificial intelligence(IA) techniques (ANFIS, GA, PSO,SVM) improves the overall accuracy. It is also obvious that the linear-SVM has the worst EER, which is due to the use of relatively degraded voice data where the genuine and imposter scores distribution overlap. It is hard to effectively separate these using a linear hyper-plane. This is also due to the non-use of weights for different modalities in the SVM to lower the effect of the degraded traits. The same table also shows that the use of UCN reduces the verification EERs for the individual modalities and for the fused biometrics using all five methods as well it appears very clearly speacially in the DET-curves (3 and 4) . It is also obvious that the GA and the ANFIS gave the best results .The GA work well under noisy data and do not require prior knowledge. The ANFIS requires few parameters to set, so the user's experience does not much affect the system's accuracy as for other techniques.

Clean-Deg	-UCN EER%	+UCN EER%
Clean face	3.57	1.42
Degraded voice	31.4286	10.7143
BFS	3.11	0.6989
GA	2.8571	0.6423
ANFIS	2.85	0.56
SVM	4.72	2.8670
PSO	3.0319	0.7143

Table 2. EER of methods using varied data

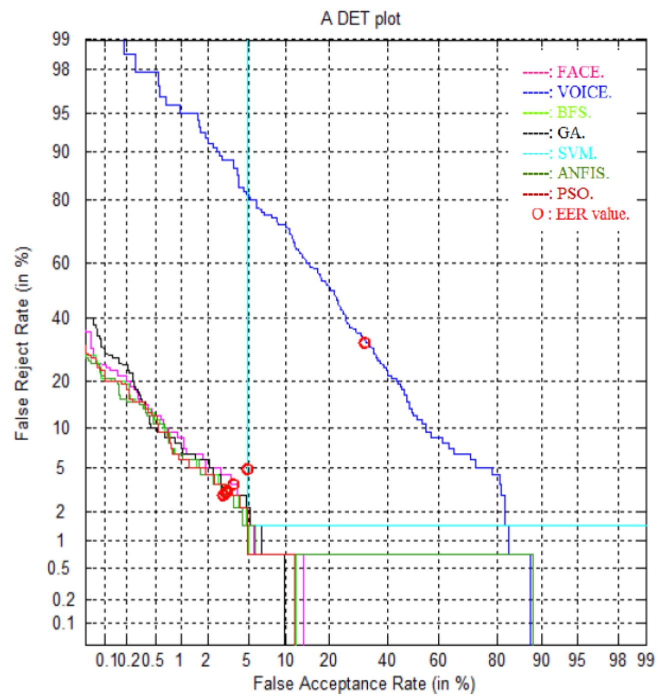

Figure3 : Det-Curve for varied data without UCN

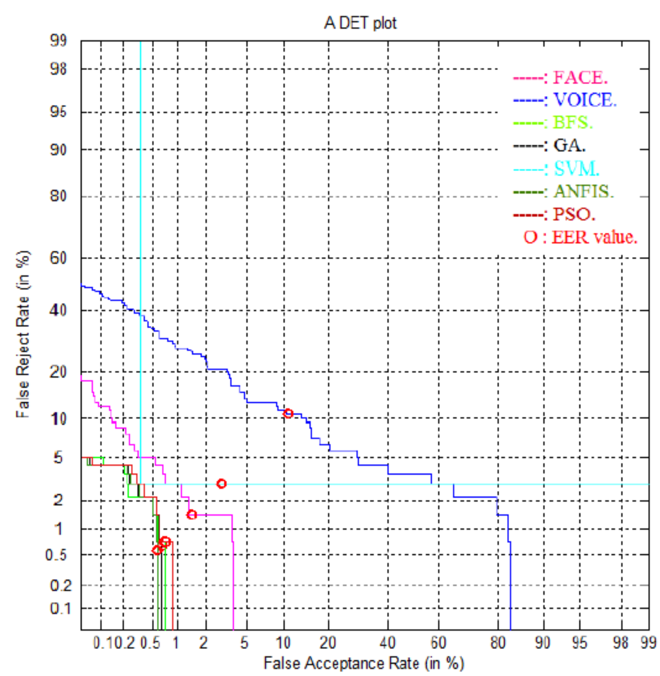

Figure4 : Det-Curve for varied data with UCN

3.3 Degraded Data

In this section, we experiments the effectiveness of UCN and some different fusion techniques (BFS, PSO, GA,ANFIS, SVM) in enhancing the reliability of multi- modal fusion when the biometric datasets are both degraded . The datasets considered for the face and the voice modalities in this investigation are extracted from the BANCA(degraded face images) and from NIST (degraded speech)databases respectively [3]. Using these biometric datasets , a total of 52 chimerical identities consisting of 26 clients and 26 imposters. The face recognition images are obtained based on images captured in four sessions. Development score dataset is formed for the experiments. This consists of 104 (i.e. 4*26) and 2600 (i.e. 4*26*(26-1)) score tokens from the same users and impostors (including cross-users) respectively. The corresponding score tokens used in the testing phase are also 104 (i.e. 4*26) and 2600 (i.e.4*26*(26-1)) respectively. The results for the verification experiments in this section are presented in Equal Error Rates (EERs). Table 3 resumes the EER values shown in the Det-curves of the fifth and the sixth figure of each fusion technique where the first column shows the method. The second column indicates the error rate for each scheme without subjecting scores to the UCN. The third column indicates the error rates of all techniques where the scores were normalized using the UCN. It is a obvious that the linear-SVM have the worst (disastrous) EER (an unacceptable accuracy) which is due to the use of degraded voice and face data where the genuine and imposter scores distributions of the two modalities intensively overlap. It is very hard (even impossible) to effectively separate them using a linear hyper-plane. It is also due to the non-use of weights for different modality in the SVM to lower the effect of the more degraded traits. The same table also shows that the use of UCN has resulted in a reduction of the verification EERs for the individual modalities and for the fused biometrics using all five methods as well. It is obvious that the population based techniques (PSO, GA) gave very good results; we can also notice that the BFS gave also acceptable results.

Deg-Deg	-UCN EER%	+UCN EER%
Degraded face	46.19	43.26
Degraded voice	26.92	23.92
BFS	27	22.1154
GA	27.03	22
ANFIS	29.34	23.07
SVM	78.2308	59.2308
PSO	27.2692	26.22

Table 3. EER of methods using degraded data

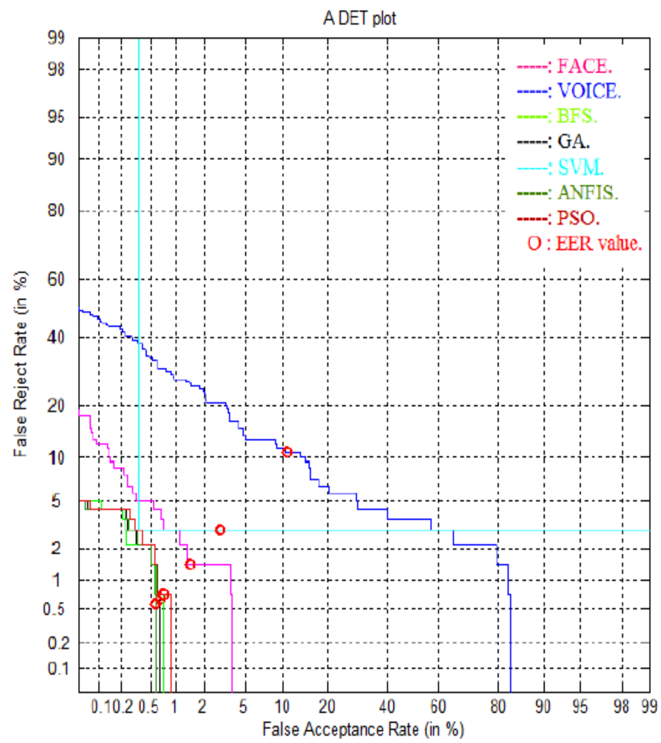

Figure5: Det-Curve for degraded data without UCN

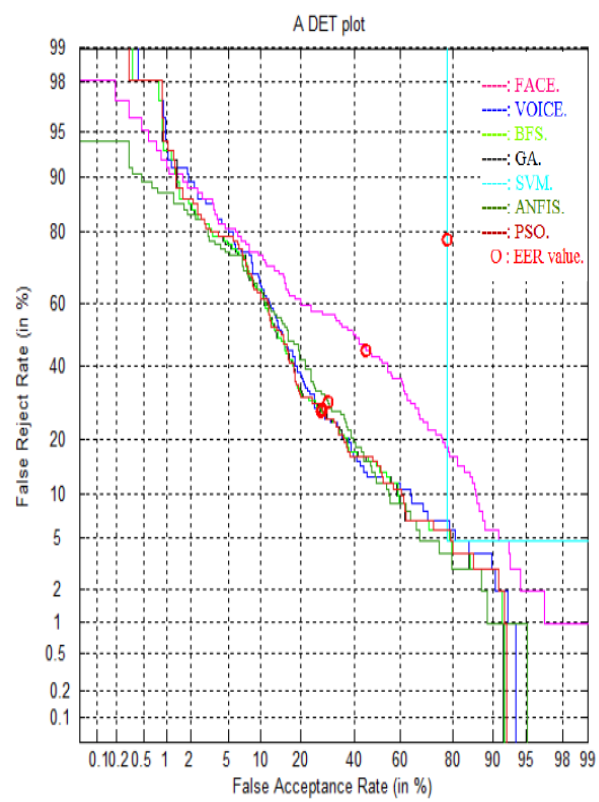

Figure6: Det-Curve for degraded data with UCN

4. CONCLUSION

This paper has presented experiments that performed a comparison between five fusion techniques for score level fusion in multimodal biometrics. These techniques are Brute Force Search (BFS), Genetic Algorithm (GA), Particle Swarm Optimization (PSO) Adaptive Neuro Fuzzy Inference System (ANFIS), Support Vector Machine (SVM). The fusion of scores was preceded by a preprocessing phase where the scores were firstly normalized using the Min-Max normalization, then subjected to the Unconstrained Cohort Normalization that led to an impressive improvement of performance of the fusion with all the techniques and under all three data conditions assuring the effectiveness of UCN under any condition. We can also clearly notice that the Genetic Algorithm (GA) and the Particle Swarm Optimization (PSO) outperform the other techniques especially under the degraded condition, due to their ability to scan a large searching space. In addition, the Adaptive Neuro Fuzzy System (ANFIS) has also good performance, which is also close to the GA and PSO results because of the use of neural networks to choose the best parameters for getting a suitable Fuzzy Inference System . It is also noticed that the SVM has the worst result especially under the degraded conditions because of the great overlap between the genuine distribution and the imposter distribution. This huge overlap prevents any linear hyper plane from separating genuine and imposter scores, without having a large False Acceptance Rate (FAR) and a large False Rejection Rate (FRR) which logically leads to a large Equal Error Rate (EER).

Future work will focus on enhancing fusion multimodal biometrics performance due to their overwhelming importance in security applications. Efforts will emphasize on taking advantage of some hybrid intelligent systems that combine neuro fuzzy techniques, along with UCN method and Gas or PSO for optimal fusion schemes.

5. REFERENCES

[1] Arun Ross, Anil K.jain, 2004. Multimodal Biometrics: An overview ,in Proc of 12th European Signal Processing Conference, Sep 2004 , pp.1221-1224.

[2] Fawaz Alsaade, ,2008 . Score-Level Biometric Fusion,Phd thesis at Hertfordshire University.

[3] F. Alsaade ,A. Rahmoun ,2009. A Method to enhance multimodal biometrics using neural networks and genetic algorithms ,In Signal and Image Processing(SIP 2009).

[4] F. Alsaade, A. Rahmoun , M. Zahrani, 2010 . On Improving Multimodal Biometrics Verification Using Genetic Algorithms,In E-MEDISYS 10 Conference Program .

[5] Sabra Dinerstein , Jonathan Dinerstein , Dan Ventura, 2010 . Robust Multi-Modal Biometric Fusion via Multiple SVMs, In ICB '09 Proceedings of the Third International Conference on Advances in Biometrics ,pp.743-752.

[6] Nina Taheri Makhsoos,Reza Ebrahimpour,Alireza Hajiany, 2009 . Face Recognition Based on Neuro-Fuzzy System. In IJCSNSInternational Journal of Computer Science and network Security,Vol.9 No.4 , pp.319-326.

[7] Pejman Tahmasebi, Ardeshir Hezarkhani, 2010. Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation: Case Study,Sarcheshmeh Porphyry Copper Deposit, In Australian Journal of Basic and Applied Sciences, vol 4,pp.408-420.

[8] Castillo, V. C., El Debs, M. K. E., Nicoletti,M.C, 2007. Using a modified genetic algorithm to minimize the production costs for slabs of precast prestressed concrete joists. Engineering Applications of Artificial Intelligence, vol 20 , Issue 4, pp.519-530

[9] Kaylan. Veeramachani, Lisa.Ann Osadciw, Pramod K Varshney, 2003. Adaptive Multimodal Biometric Fusion Algorithm Using Particle Swarm. In Syracuse,NY , pp.13244-1240.

[10] F. Alsaade,A. M. Ariyaeeinia,A. S. Malegaonkar,M. Pawlewski ,S. G. Pillay, 2008. Enhancement of multimodal biometric segregation using unconstrained cohort normalization, In Pattern Recognition , Vol 41 Issue 3 , pp.814-820.