

A Framework for Quality-based Biometric Classifier Selection

Himanshu S. Bhatt, Samarth Bharadwaj, Mayank Vatsa, Richa Singh
IIIT Delhi, India

{himanshub, samarthb, mayank, rsingh}@iiitd.ac.in

Arun Ross, Afzel Noore
West Virginia University, USA

{arun.ross, afzel.noore}@mail.wvu.edu

Abstract

Multibiometric systems fuse the evidence (e.g., match scores) pertaining to multiple biometric modalities or classifiers. Most score-level fusion schemes discussed in the literature require the processing (i.e., feature extraction and matching) of every modality prior to invoking the fusion scheme. This paper presents a framework for dynamic classifier selection and fusion based on the quality of the gallery and probe images associated with each modality with multiple classifiers. The quality assessment algorithm for each biometric modality computes a quality vector for the gallery and probe images that is used for classifier selection. These vectors are used to train Support Vector Machines (SVMs) for decision making. In the proposed framework, the biometric modalities are arranged sequentially such that the stronger biometric modality has higher priority for being processed. Since fusion is required only when all unimodal classifiers are rejected by the SVM classifiers, the average computational time of the proposed framework is significantly reduced. Experimental results on different multimodal databases involving face and fingerprint show that the proposed quality-based classifier selection framework yields good performance even when the quality of the biometric sample is sub-optimal.

1. Introduction

Multibiometrics-based verification systems use two or more classifiers pertaining to the same biometric modality or different biometric modalities. As discussed by Woods *et al.* [19], there are two general approaches to fusion: (1) classifier fusion and (2) dynamic classifier selection. In classifier fusion, all constituent classifiers are used and their decisions are combined using fusion rules [10], [14]. On the other hand, in dynamic selection, the most appropriate classifier or a subset of specific classifiers is selected [8], [16]

for decision making. In the biometrics literature, classifier fusion has been extensively studied [14], whereas dynamic classifier selection has been relatively less explored. Marcialis *et al.* [11] designed a serial fusion scheme for combining face and fingerprint classifiers and achieved significant reduction in verification time and the required degree of user cooperation. Alonso-Fernandez *et al.* [3] proposed a method where quality information was used to switch between different system modules depending on the data source. Veeramachaneni *et al.* [17] proposed a Bayesian framework to fuse decisions pertaining to multiple biometric sensors. Particle Swarm Optimization (PSO) was used to determine the “optimal” sensor operating points in order to achieve the desired security level by switching between different fusion rules. Vatsa *et al.* [15] proposed a case-based context switching framework for incorporating biometric image quality. Further, they proposed a sequential match score fusion and quality-based dynamic selection algorithm to optimize both verification accuracy and computational cost [16]. Recently, a sequential score fusion strategy was designed using sequential probability ratio test [2]. Though existing approaches improve the performance, in general, it is necessary to capture all biometric modalities prior to processing them.

This research focuses on developing a dynamic selection approach for a multi-classifier biometric system that can yield high verification performance even when operating on moderate-to-poor quality probe images. The case study considered in this work has two biometric modalities (face and fingerprint) and two classifiers per modality. It is generally accepted that the quality of a biometric sample is an important factor that can affect matching performance. Therefore, the proposed approach utilizes image quality to dynamically select one or more classifiers for verifying if a given gallery-probe pair belongs to the genuine class or the impostor class. Experiments on a multimodal database involving face and fingerprint, with variations in probe qual-

ity, suggest that the proposed approach provides significant improvements in recognition accuracy compared to individual classifiers and the classical sum-rule fusion scheme.

2. Quantitative Assessment Algorithm

In the proposed approach, different quality assessment techniques are used to generate a composite quality vector for a given biometric sample. The quality vector used in this study comprises of four quality attributes (scores): *no-reference quality*, *edge spread*, *spectral energy*, and *modality specific image quality*. Details of each quality attribute are provided below:

- **No-reference quality:** Wang *et al.* [18] used blockiness and activity estimation in both horizontal and vertical directions in an image to compute a no-reference quality score. Blockiness is estimated by the average intensity difference between block boundaries in the image. Activity is used to measure the effect of compression and blur on the image. These individual estimates are combined to give a composite no-reference quality score.
- **Edge spread:** Marziliano *et al.* [7] used edge spread to estimate motion and off-focus blurriness in images based on edges and adjacent regions. Their technique computes the effect of blur in an image based on the difference in image intensity with respect to the local maxima and minima of pixel intensity in every row of the image.
- **Spectral energy:** It describes abrupt changes in illumination and specular reflection [13]. The image is tessellated into several non-overlapping blocks and the spectral energy is computed for each block. The value is computed as the magnitude of Fourier transform components in both horizontal and vertical directions.
- **Modality specific image quality:** Along with the above mentioned general image quality attributes, the quality assessment algorithm also computes “usability” quality measures specific to each biometric modality.

Face quality: For face images, pose is a major covariate that determines the usability of the face image. Even a good quality face image may not be useful during recognition due to pose variations. Pose is estimated based on the geometric relationship between face, eyes, and mouth. Depending upon the yaw, pitch and roll values of the estimated pose, a composite score is computed for denoting face quality.

Fingerprint quality: For fingerprint images, Chen *et al.* [5] measured the quality of ridge samples by computing the Fourier energy spectral density concentra-

Table 1. Range of quality attributes over the images used in this research.

Face images	
Quality attribute	Range
Spectral Energy	[1.09, 1.34]
No reference quality	[12.43, 13.50]
Edge spread	[8.51, 16.88]
Pose	[302.31, 466.12]
Fingerprint images	
Quality attribute	Range
Spectral Energy	[0.96, 1.15]
No reference quality	[8.10, 11.50]
Edge spread	[3.94, 6.68]
Global entropy	[0.91, 1.16]

tion in particular frequency bands. Such a measure is global in nature and encodes the overall quality of fingerprint ridges. This quality measure, referred to as *global entropy*, is used in this work.

For a given image, a quality vector comprising of the four aforementioned quality scores is generated. Table 1 shows the range of values obtained by the quality attributes over the face-fingerprint images used in this research (details are available in Section 4.2). The *spectral energy* is considered good if its value is close to 1. For *no reference quality*, higher the value better is the quality of image. For a frontal face image, the value of *pose* attribute is 400. Therefore, a face is right aligned if *pose* is less than 400, otherwise, the face is aligned to the left. For *edge spread*, lower the value better is the quality of image. For *global entropy*, higher the value better is the quality of the fingerprint image. For a given gallery-probe pair, the quality vector of both gallery and probe images are concatenated to form a quality vector of eight quality scores represented as $Q = [Q_g, Q_p]$, where Q_g and Q_p are the quality vectors of gallery and probe images, respectively.

3. Quality Driven Classifier Selection Framework

The proposed framework utilizes the quality vector for classifier selection. As shown in Figure 1, in a face-fingerprint bimodal setting, the individual modalities are processed sequentially. It starts from the strongest modality such that the system has higher chances of correctly classifying the gallery-probe pair using the first biometric modality and obviating the need for processing the second modality. Since classifier selection can also be posed as a classification problem, Support Vector Machine (SVM) is used for classification. One SVM is trained for each biometric modality to select the best classifier for that modality using quality vectors. In this paper, the classifier selection frame-

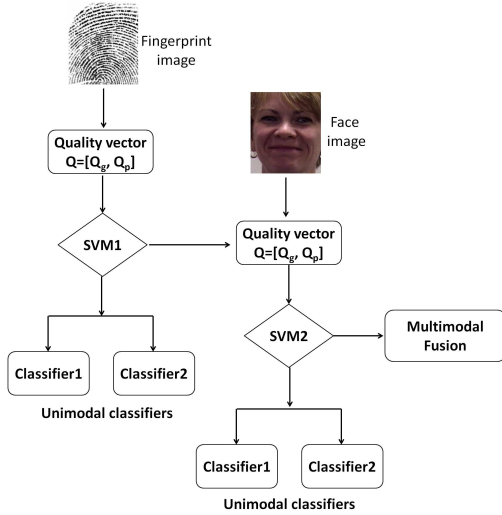


Figure 1. Illustrating the proposed quality based classifier selection framework for face-fingerprint biometrics.

work is presented for a two-classifier two-modality setting involving face and fingerprint. However, the framework can be easily extended to accommodate more choices as it provides the flexibility to add new biometric modalities and to add/remove classifiers for each modality. The framework is divided into two stages: (1) training the SVMs and (2) dynamic classifier selection for probe verification.

3.1. SVM Training

The SVM corresponding to each biometric modality is trained independently using a labeled training database.

Training SVM for Fingerprints: *SVM1* is trained for three classes using the labeled training data $\{x_{1i}, y_{1i}\}$. Here, input $x_{1i} = [Q_g, Q_p]$ is the quality vector of the i^{th} gallery-probe fingerprint image pair in the training set and the output $y_{1i} \in \{-1, 0, +1\}$. The labels are assigned based on the match score distribution of genuine and impostor scores and the likelihood ratio of the two fingerprint classifiers. As shown in Figure 2, for each modality, *distance* scores are computed using the training data and the two fingerprint verification algorithms. If the impostor score computed using *classifier1* is greater than the maximum genuine score (confidently classified as impostor) or if the genuine score computed using *classifier1* is less than the minimum impostor score (confidently classified as genuine), the $\{-1\}$ label is assigned to indicate that *classifier1* can correctly classify the gallery-probe pair. Label $\{0\}$ is assigned when the impostor score computed using *classifier2* is greater than the maximum genuine score (confidently classified as impostor) or when the genuine score computed using *classifier2* is less than the minimum impostor score

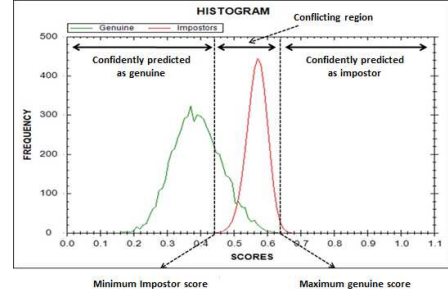


Figure 2. Illustrating the process of assigning labels: the genuine-impostor match score distribution are used to assign labels of input gallery-probe quality vector $Q = [Q_g, Q_p]$ during SVM training.

(confidently classified as genuine). If the score lies within the conflicting region for both the verification algorithms, the $\{+1\}$ label is assigned which signifies that for the given gallery-probe pair, the individual fingerprint classifiers is not able to classify the gallery-probe pair and that another modality, i.e. face, is required. If both the verification algorithms correctly classify the gallery-probe pair based on the score distribution, then the likelihood ratio is used to make a decision (genuine or impostor). The quality vector of the gallery-probe pair is assigned the label corresponding to the verification algorithm that classifies it with higher confidence (based on the accuracy computed using training samples). Under Gaussian assumption, the likelihood ratio is computed from the estimated densities $f_{gen}(x)$ and $f_{imp}(x)$ as $LR(x) = f_{gen}(x) / f_{imp}(x)$.

Training SVM for Face: Similar to *SVM1*, *SVM2* is also a three-class SVM trained using the labeled training data $\{x_{2i}, y_{2i}\}$, where, $x_{2i} = [Q_g, Q_p]$ is the quality vector of the i^{th} gallery-probe face image pair in the training set. The labels are assigned in a similar manner as *SVM1*. The only variation here is with the $\{+1\}$ label. If the score lies within the conflicting region for both the face verification algorithms, the $\{+1\}$ label is assigned which signifies that for the given gallery-probe pair, the individual classifiers are not able to classify the gallery-probe pair and that match score fusion is required.

3.2. Classifier Selection for Verification

During verification, the trained SVMs are used to select the most appropriate classifier for each modality based only on quality. The biometric modalities are used one at a time and the second modality is selected only when the individual classifiers pertaining to the first modality are not able to classify the given gallery-probe pair.

The quality vectors of gallery-probe pair for the first modality is computed and provided as input to the trained *SVM1*. Based on the quality vector, *SVM1* makes the prediction. If *SVM1* predicts that one of the classifiers of the

first modality can be used to correctly classify the given gallery-probe pair, then the framework selects the classifier predicted by *SVM1*. Otherwise, the quality vector for the gallery-probe pair corresponding to the second modality is computed and provided as input to *SVM2*. If *SVM2* predicts that one of the classifiers of the second modality can correctly classify the gallery-probe pair, then the framework selects the classifier predicted by *SVM2*. Otherwise, if both SVMs predict that the individual classifiers of both the modalities are unable to classify the gallery-probe pair, the sum rule-based score level fusion of the classifiers across both modalities is used to generate the final score. It should be noted that since the SVMs are based only on the quality of the gallery-probe pair, the framework does not require computing the scores for all the modalities and classifiers.

4. Experimental Results

To evaluate the effectiveness of the proposed framework, experiments are performed on two different multimodal databases using two face classifiers and two fingerprint classifiers. Details about the feature extractors and matchers used for each modality, database, experimental protocol, and key observations are presented in this section.

4.1. Unimodal Algorithms

Fingerprint: The two fingerprint classifiers used in this study are the NIST Biometric Image Software (NBIS)¹ and a commercial² fingerprint matching software. NBIS consists of a minutiae detector called MINDTCT and a fingerprint matching algorithm known as BOZORTH3. The second classifier, a commercial fingerprint matching software, is also based on extracting and matching minutiae points.

Face: The two face classifiers used in this research are Uniform Circular Local Binary Pattern (UCLBP) [1] and Speeded Up Robust Features (SURF) [4]. UCLBP is a widely used texture-based operator whereas SURF is a point-based descriptor which is invariant to scale and rotation. χ^2 distance measure is used to compare two UCLBP feature histograms and two SURF descriptors.

4.2. Database

The evaluation is performed on two different databases. The first is the WVU multimodal database [6] from which 270 subjects that have at least 6 fingerprint and face images each are selected. For each modality, two images per subject are placed in the gallery and the remaining images are used as probes.

To evaluate the *scalability* of the proposed approach, a large multimodal (chimeric) database is used. The WVU

¹<http://www.nist.gov/itl/iad/ig/nbis.cfm>

²The license agreement does not allow us to name the software in any comparative study.

Table 2. Parameters of noise and blur kernels used to create the synthetic degraded database.

Type	Parameter
Gaussian noise	$\sigma = 0.05$
Poisson noise	$\lambda = 1$
Salt & pepper noise	$d = 0.05$
Speckle noise	$v = 0.05$
Gaussian blur	$\sigma = 1$
Motion blur	angle 5° & length 1-10 pixels
Unsharp blur	$\alpha = 0.1$ to 1

multimodal database consists of fingerprint images from four fingers per subject. Assuming that the four fingers are independent, a database of 1068 virtual subjects with six or more samples per subject is prepared. For associating face with fingerprint images, a face database of 1068 subjects is created containing 446 subjects from the MBGC Version2 database³, 270 subjects from the WVU database [6], 233 from the CMU MultiPIE database [9], and 119 subjects from the AR face database [12].

4.3. Experimental Protocol

In all the experiments, 40% of the subjects in the database are used for training and the remaining 60% are used for performance evaluation. During training, the SVMs are trained as explained in Section 3.1. The 40%-60% partitioning was done five times (repeated random subsampling validation) and verification accuracies are computed at 0.01% false accept rate (FAR). Two experiments are performed as explained below:

Experiment 1: In this experiment, with two biometric modalities (face and fingerprints) and four classifiers, the proposed quality-based classifier selection framework selects the most appropriate unimodal classifier to process the gallery-probe pair based on the quality. In this experiment both gallery and probe images are of good quality (unaltered/original images).

Experiment 2: In this experiment, the quality of probe images is synthetically degraded. A synthetic poor quality database is prepared where probe images are corrupted by adding different types of noise and blur as shown in Figure 3. Table 2 shows the parameters of noise and blur kernels used to create the synthetic database. Experiments are performed for each type of degradation introduced in both fingerprints and face images. It should be noted that for experiment 2, training is done on good quality gallery-probe pairs and performance is evaluated on non-overlapping subjects from the synthetically corrupted database.

³<http://www.nist.gov/itl/iad/ig/mbgc-presentations.cfm>

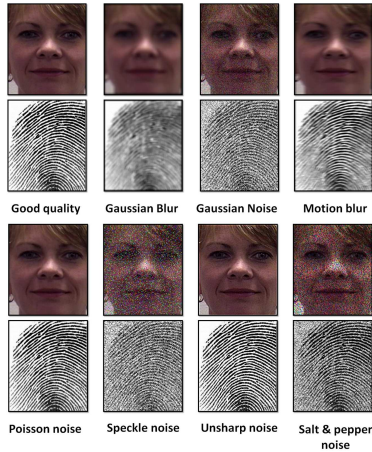


Figure 3. Sample images from the database that are degraded using different types of noise and blur.

Input image	Quality attributes	Selected classifier
 Gallery Probe (a)	1.03, 8.44, 8.15, 0.91	Fingerprint classifier1
 Gallery Probe (b)	1.00, 8.39, 5.58, 0.99	Fingerprint classifier2
 Gallery Probe (c)	1.26, 12.95, 14.51, 342.22	Face classifier2

Figure 4. Sample decisions of the proposed algorithm when (a) Fingerprint classifier 1 is selected, (b) Fingerprint classifier 2 is selected, and (c) Face classifier 2 is selected.

4.4. Results and Analysis

Figure 4 illustrates sample decisions of the proposed algorithm. Figures 5 and 6 show the Receiver Operating Characteristic (ROC) curves for experiment 1. Table 3 summarizes the verification accuracy for different types of degradations introduced in the probe set. The key results are listed below:

- ROC curves in Figures 5 and 6 show that for experiment 1, the proposed quality-based classifier selection framework outperforms the unimodal classifiers and sum-rule fusion by at least 1.05% and 1.57% on the WVU multimodal database and the large scale chimeric database, respectively.

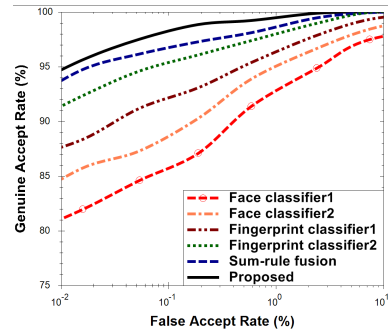


Figure 5. ROC curves of the individual classifiers, sum-rule fusion and the proposed quality based classifier selection framework on the WVU multimodal database with good gallery-probe quality.

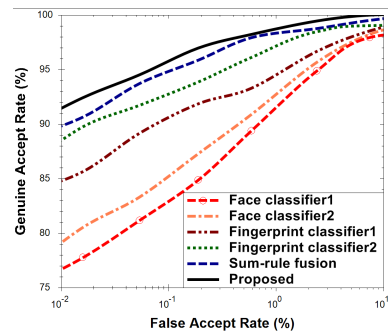


Figure 6. ROC curves of the individual classifiers, sum-rule fusion and the proposed quality based classifier selection framework on the large scale chimeric database with good gallery-probe quality.

- It is observed that when the quality of probe images is degraded, the performances of individual classifiers are affected. However, the quality-based classifier selection framework still performs better than individual classifiers and sum rule fusion. This improvement is attributed to the fact that the proposed framework can dynamically determine when to use the most appropriate single classifier and when to perform fusion based on the quality of gallery-probe image pairs. Table 3 reports the performance of all the algorithms when probe images are of sub-optimal quality.
- In experiment 1 with the WVU database, 27.95% gallery-probe pairs were processed by fingerprint classifier1 - NBIS, 25.33% pairs with fingerprint classifier2 - commercial matcher, 18.99% with face classifier1 - UCLBP, and 15.51% with face classifier2 - SURF. The remaining 12.19% pairs were processed using weighted sum rule fusion. Similarly for the

chimeric database, 31.45% gallery-probe pairs were processed by fingerprint classifier1 - NBIS, 32.12% pairs with fingerprint classifier2 - commercial matcher, 15.32% with face classifier1 - UCLBP, and 13.56% with face classifier2 - SURF. The remaining 7.55% pairs were processed using weighted sum rule fusion.

- Unlike parallel fusion, the proposed framework does not require computing the image quality scores for all modalities up-front. Image quality scores for each modality are computed only when the framework needs to make a decision for that particular modality. Moreover, if one of the biometric modalities cannot be captured, the framework can easily skip that unimodal classifier.
- The proposed classifier selection is about two times faster than the match score fusion algorithm. The time to process a probe using dynamic classifier selection (including quality assessment and feature extraction) is much lesser than the time to process a probe using the sum rule fusion (including processing both the modalities and four classifiers).
- The major advantage of the proposed quality based classifier selection framework is that it can be easily extended to include other biometric modalities, unimodal classifiers and fusion rules.

5. Conclusion

This paper presents a dynamic classifier selection framework for multibiometric systems. The sequential design of the classifier selection framework allows it to process each biometric modality in sequence using the quality of the gallery-probe pair. Since the stronger modality is typically selected for verification, it also offers reduction in computational time. This work establishes the utility of dynamic classifier selection in the context of biometrics.

6. Acknowledgement

This work was supported by a grant from the U.S. Army Research Laboratory Contract No. W911NF-10-2-0021.

References

- [1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Trans. on PAMI*, 28(12):2037–2041, 2006. 4
- [2] L. Allano, B. Dorizzi, and S. Garcia-Salicetti. Tuning cost and performance in multi-biometric systems: A novel and consistent view of fusion strategies based on the sequential probability ratio test (sprt). *Pattern Recognition Letters*, 31:884–890, 2010. 1
- [3] F. Alonso-Fernandez, J. Fierrez, D. Ramos, and J. Gonzalez-Rodriguez. Quality-based conditional processing in multi-biometrics: Application to sensor interoperability. *IEEE Trans. on SMC-A*, 40(6):1168–1179, 2010. 1
- [4] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded up robust features. In *European Conference on Computer Vision*, pages 404–417, 2006. 4
- [5] Y. Chen, S. Dass, and A. Jain. Fingerprint quality indices for predicting authentication performance. In *Proceedings of Audio and Video-Based Biometric Person Authentication*, pages 160–170, 2005. 2
- [6] S. Crialmeanu, A. Ross, S. Schuckers, and L. Hornak. A protocol for multibiometric data acquisition, a protocol for multibiometric data acquisition, storage and dissemination. Technical report, 2007. 4
- [7] P. M. Frederic, F. Dufaux, S. Winkler, T. Ebrahimi, and G. Sa. A no-reference perceptual blur metric. In *Proceedings of International Conference on Image Processing*, pages 57–60, 2002. 2
- [8] G. Giacinto and F. Roli. Methods for dynamic classifier selection. In *Proceedings of International Conference on Image Analysis and Processing*, pages 659–664, 1999. 1
- [9] R. Gross, I. Matthews, J. F. Cohn, T. Kanade, and S. Baker. Multi-PIE. In *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 2008. 4
- [10] J. Kittler, M. Hatef, R. Duin, and J. Matas. On combining classifiers. *IEEE Trans. on PAMI*, 20(3):226–239, 1998. 1
- [11] G. Marcialis, F. Roli, and L. Didaci. Personal identity verification by serial fusion of fingerprint and face matchers. *Pattern Recognition*, 42(11):2807–2817, 2009. 1
- [12] A. Martinez and R. Benevento. The AR face database. *CVC Technical Report #24*, 1998. 4
- [13] N. Nill and B. Bouzas. Objective image quality measure derived from digital image power spectra. *Optical Engineering*, 31(4):813–825, 1992. 2
- [14] A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of multi-biometrics*. 1st edition, 2006. 1
- [15] M. Vatsa, R. Singh, and A. Noore. Context switching algorithm for selective multibiometric fusion. In *Proceedings of International Conference on Pattern Recognition and Machine Intelligence*, pages 452–457, 2009. 1
- [16] M. Vatsa, R. Singh, A. Noore, and A. Ross. On the dynamic selection of biometric fusion algorithms. *IEEE Trans. on IFS*, 5(3):470–479, 2010. 1
- [17] K. Veeramachaneni, L. Osadciw, and P. Varshney. An adaptive multimodal biometric management algorithm. *IEEE Trans. on SMC-C*, 35(3):344–356, 2005. 1
- [18] J. Wang, Y. Shang, G. Su, and X. Lin. Age simulation for face recognition. In *Proceedings of International Conference on Pattern Recognition*, pages 913–916, 2006. 2
- [19] K. Woods, J. Kegelmeyer, W.P., and K. Bowyer. Combination of multiple classifiers using local accuracy estimates. *IEEE Trans. on PAMI*, 19(4):405–410, 1997. 1

Table 3. The performance of the proposed quality-based classifier selection algorithm in comparison with unimodal classifiers and the sum-rule fusion algorithm on different databases. Verification accuracies are reported at 0.01 false accept rate (FAR).

Type of Noise Added	Algorithm	WVU Multimodal Database	Large scale Chimeric Database
		Accuracy	Accuracy
Good Quality Probe	Face classifier1	82.08%	77.48%
	Face classifier2	84.96%	79.04%
	Fingerprint classifier1	87.84%	84.83%
	Fingerprint classifier2	91.34%	88.68%
	Sum-rule Fusion	93.86%	89.84%
	Proposed	94.91%	91.41%
Gaussian Noise	Face classifier1	77.58%	72.48%
	Face classifier2	80.86%	76.86%
	Fingerprint classifier1	81.14%	79.74%
	Fingerprint classifier2	86.74%	82.74%
	Sum-rule Fusion	88.36%	84.74%
	Proposed	90.41%	87.81%
Gaussian Blur	Face classifier1	81.36%	75.16%
	Face classifier2	83.14%	79.64%
	Fingerprint classifier1	87.42%	81.28%
	Fingerprint classifier2	90.74%	84.36%
	Sum-rule Fusion	91.58%	87.43%
	Proposed	92.62%	89.52%
Speckle Noise	Face classifier1	76.48%	73.82%
	Face classifier2	79.63%	76.41%
	Fingerprint classifier1	81.42%	80.38%
	Fingerprint classifier2	84.35%	82.43%
	Sum-rule Fusion	86.22%	84.28%
	Proposed	87.58%	85.76%
Salt and Pepper Noise	Face classifier1	77.56%	74.16%
	Face classifier2	79.38%	77.42%
	Fingerprint classifier1	84.54%	81.18%
	Fingerprint classifier2	87.36%	84.38%
	Sum-rule Fusion	89.12%	86.24%
	Proposed	90.16%	87.32%
Motion blur	Face classifier1	78.23%	75.48%
	Face classifier2	80.16%	78.86%
	Fingerprint classifier1	86.32%	83.64%
	Fingerprint classifier2	88.48%	85.47%
	Sum-rule Fusion	89.74%	86.24%
	Proposed	90.81%	87.14%
Poisson Noise	Face classifier1	76.35%	74.28%
	Face classifier2	79.28%	77.63%
	Fingerprint classifier1	85.32%	82.51%
	Fingerprint classifier2	87.44%	84.29%
	Sum-rule Fusion	88.94%	85.74%
	Proposed	89.65%	86.35%
Unsharp Noise	Face classifier1	76.64%	73.48%
	Face classifier2	79.85%	77.86%
	Fingerprint classifier1	83.85%	83.64%
	Fingerprint classifier2	86.48%	85.38%
	Sum-rule Fusion	88.41%	86.49%
	Proposed	89.72%	87.14%