On Co-training Online Biometric Classifiers

Himanshu S. Bhatt, Samarth Bharadwaj, Richa Singh, Mayank Vatsa
IIIT Delhi, India
{himanshub, samarthb, rsingh, mayank}@iiitd.ac.in

Afzel Noore, Arun Ross
West Virginia University, USA
{afzel.noore, arun.ross}@mail.wvu.edu

Abstract

In an operational biometric verification system, changes in biometric data over a period of time can affect the classification accuracy. Online learning has been used for updating the classifier decision boundary. However, this requires labeled data that is only available during new enrolments. This paper presents a biometric classifier update algorithm in which the classifier decision boundary is updated using both labeled enrolment instances and unlabeled probe instances. The proposed co-training online classifier update algorithm is presented as a semi-supervised learning task and is applied to a face verification application. Experiments indicate that the proposed algorithm improves the performance both in terms of classification accuracy and computational time.

1. INTRODUCTION

A biometric verification system typically uses a classifier to determine if the unlabeled probe data matches with the labeled gallery data. The performance of such a classifier is affected by the intra-class and inter-class dynamics as biometric data is acquired over a period of time [21]. New information that can affect the biometric data distribution (e.g. match scores) is available from two fronts: (1) new subjects enrolling into the biometric system (labeled data) and (2) previously enrolled subjects interacting with the system and providing new probes (unlabeled data). New enrolments can lead to variations in genuine and impostor score distributions while probe images may introduce wide intra-class variations (due to temporal changes). To maintain the performance and to accommodate the variations caused due to new enrolments and probes, biometric systems generally require re-training. Since re-training with existing and new information in *batch mode* requires a huge amount of time, it is not pragmatic for large scale applications. However, if the classifiers are not re-trained, then the verification performance can be compromised.

Online learning [18] and *co-training* [7] are used to update the classifiers in real time and make them scalable. These paradigms can also be used for updating biometric classifiers. Intuitively,

- labeled information from newly enrolled individuals can be used to update the classifier in incremental-decremental learning mode, also known as online learning. Since corresponding labels ("genuine" or "impostor") are available during enrolment, classifier update using online learning can be viewed as a supervised learning approach.
- unlabeled information obtained at probe level can be used to update the classifier using co-training. In the co-training framework, two classifiers evolve by co-training each other using unlabeled probe information. If the first classifier confidently predicts the class (genuine or impostor) for an instance, while the second classifier is unsure of its classification decision, then this data instance is added to re-train the second classifier with the pseudo label assigned by the first classifier.

If we incorporate both the paradigms, then updating a biometric classifier can be posed as a semi-supervised learning [9] task that seamlessly exploits unlabeled data in addition to the labeled data.

In the literature, incremental (online) learning approaches for principle component analysis [18] and linear discriminant analysis [22] have shown the effectiveness of this paradigm. Kim *et al.* [14] have shown that online learning algorithms can be used for biometric score fusion in order to resolve the computational problems with increasing number of users. Singh *et al.* [21] have proposed an online learning approach for updating a face classifier. Their results show that the performance of online SVM classifiers is comparable to the batch mode counterpart. Further, online SVM classifiers have a significant advantage of reduced re-training time using only the new sample points to update the decision boundary.

In co-training, as proposed by Blum and Mitchell [7], two classifiers that are trained on separate views (features), co-train each other based on their confidence in predicting the labels. Nonetheless, success of a co-training framework is susceptible to various assumptions. Blum and Mitchell [7] showed that two classifiers should have sufficient individual accuracy and should be conditionally independent of each other. Later, Abney [2] showed that weak dependence between the two classifiers can also guarantee successful co-training. Wang and Zhou [23] also reported the sufficient and necessary conditions for success of a co-training framework.

Though co-training has been used in several computer vision applications, in the biometrics literature, use of unlabeled data for updating the system has been mainly restricted to biometric template updates. Jiang and Ser [13] proposed a method to improve fingerprint templates by merging and averaging minutiae from multiple samples of a fingerprint. Ryu *et al.* [20] also proposed a method to update the fingerprint templates by appending new minutiae from the query fingerprint with the gallery fingerprint template. Balcan *et al.* [5] developed a method to address the problem of person identification in low quality web-camera images. They formulated the task of person identification in web-camera images as a graph-based semi-supervised learning problem. Roli *et al.* [19] designed a biometric system that uses co-training to address the temporal variations in a face and fingerprint based multimodal system. Liu *et al.* [15] proposed to retrain the Eigenspace in a face recognition system using the unlabeled data stream. Recently, Poh *et al.* [17] performed a study on the goal of semi-supervised learning where they focused on some of the challenges and research directions for designing adaptive biometric systems.

This research focuses on seamlessly improving the performance of a biometric classifier by updating the classifier's knowledge using additional labeled data obtained during new enrolments as well as unlabeled data obtained during probe verification. *The paper presents a framework for co-training biometric classifiers in an online manner*. Specifically, the concepts of co-training and online learning are applied to a support vector machine (SVM) based biometric classifier update scenario. While online learning updates the SVM classifier using labeled enrolment data, co-training updates the SVM decision boundaries with a large number of unlabeled probe examples. The performance of the proposed co-training framework is evaluated in the context of multi-classifier SVM based face verification where it shows improvements in both verification accuracy and computational time.

2. Proposed Co-training Online Framework

Mathematically, for a two classifier biometric verification system, the process is as follows. Two types of data instances are available: a set of labeled data instances, $\{(\mathbf{u}_1, z_1), (\mathbf{u}_2, z_2), ..., (\mathbf{u}_n, z_n)\}$, is available when new users are enrolled into the system and a set of unlabeled data instances, $\{\mathbf{u'}_1, \mathbf{u'}_2, ..., \mathbf{u'}_n\}$, is available during probe verification. Every instance $\mathbf{u}_i$ or $\mathbf{u'}_i$ has two views, $\mathbf{u}_i = \{x_{i,1}, x_{i,2}\}$; here $x_{i,1}$ and $x_{i,2}$ represent the match scores obtained from the two classifiers and the label $z_i \in \{+1, -1\}$ represents the genuine or impostor class. For labeled data instances available during enrolments, classifier c_j predicts the label for every instance: $c_j(x_{i,j}) \rightarrow y_{i,j}$, where $y_{i,j}$ is the predicted label for the i^{th} instance on the j^{th} view, $i = 1, 2, ..., m$, m is the total number of scores generated when a newly enrolled user is compared against the existing gallery and its own multiple samples, and j is the number of views[1] (number of classifiers), $j = 1, 2$. In online learning, classifiers c_1 and c_2 are updated for every incorrect prediction (i.e., when $y_i \neq z_i$) while no action is taken when the instances are correctly classified. For unlabeled instances, classifiers c_1 and c_2 predict labels on the two separate views, $c_1(x_{i,1}) \rightarrow y_{i,1}$ and $c_2(x_{i,2}) \rightarrow y_{i,2}$. Here, $x_{i,1}$ and $x_{i,2}$ are the two views of the i^{th} instance u'_i, and $y_{i,1}$ and $y_{i,2}$ are the corresponding predicted labels. Classifiers are co-trained for a given instance if one classifier confidently predicts the label of the instance while the other classifier is unsure of its prediction.

2.1. Online SVM Classifiers

Let $\{\mathbf{u}_i, z_i\}$ be the set of data instances (scores) where $i = 1, ..., N$, N is the total number of instances, and z_i is the label such that $z_i \in \{+1, -1\}$. Since $\mathbf{u}_i$ represents the two individual views (classifiers), SVMs are trained individually for both the views using $x_{i,j}$ where $j = 1, 2$.

The basic principle behind SVM is to find the hyperplane that separates two classes with the widest margin, i.e., to maximize $w\phi(x_{i,j}) + b = 0$ or equivalently minimize:

$$min_{w,b,\epsilon} \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} \epsilon_i \qquad (1)$$

subject to constraints:

$$z_i(w\phi(x_{i,j}) + b) \geq 1 - \epsilon, \epsilon \geq 0, i \in 1, ..., N \qquad (2)$$

where ϵ are the slack variables, b is the offset of the decision hyperplane, w is the normal weight vector, $\phi(x_{i,j})$

[1] The terms "views" and "classifiers" are used interchangeably because each classifier is trained on a single view and, therefore, there are as many classifiers as there are number of views.

is the mapping function used to map the data space to the feature space, and C is the tradeoff parameter between the permissible error in the samples and the margin. Note that, in this context, input to the two class SVM is match scores with labels $\{+1, -1\}$ representing the genuine and impostor classes. In large scale biometrics applications, retraining the SVM classifiers is computationally expensive. Existing approaches allow the training of SVM in online manner using only the support vectors and new data points. Methods to add or remove one sample at a time to update SVM (in online manner) are proposed in [8], [21] where an exact solution for $N \pm 1$ can be obtained using the N old samples and the one sample to be added or removed.

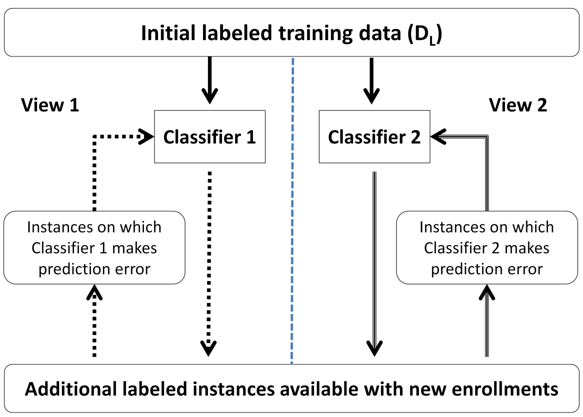

Figure 1. Illustrating the online learning process where each classifier learns from the incorrectly classified instances.

Figure 1 shows the proposed online learning approach when two SVMs are used as biometric classifiers. SVM classifiers for each view/score are first trained on the initial enrolment training data D_L. A unique identification number is assigned to every user being enrolled in the biometric system. Note that, during enrolment, we can store multiple samples from each individual to accommodate intra-class variations and for performing online learning on the SVM classifier. Biometric features of the new user are extracted and compared against the gallery of other individuals to compute the impostor match scores. For genuine match score computation, we use multiple samples captured during enrolment. SVM classifiers are then used to classify each of these match scores as genuine or impostor. In the enrolment stage, labels (ground truth) corresponding to the match scores are compared with the prediction of the classifier. The match scores for which the classifier makes incorrect predictions are used to update the decision boundary of the SVM classifier using online learning [21]. This online learning process is performed for both the classifiers and the two classifiers are updated independently. The online learning algorithm to update the classifiers is described in Algorithm 1.

Algorithm 1 Online Classifier Update

Input: Initial labeled enrolment training data D_L, a set of additional labeled instances $\{u_i, z_i\}$ due to enrolments, $i = 1, 2, ...N$, where N is the number of additional instances. Each instance $u_i = (x_{i,1}, x_{i,2})$ represents two views (or scores).
Iterate: j= 1 **to** number of views (number of classifiers)
Process: Train classifier c_j on j^{th} views of D_L
for $k = 1$ **to** N **do**
  Predict labels: $c_j(x_{i,j}) \rightarrow y_i$
  if $y_i \neq z_i$ **then**
    Update c_j with labeled instance $\{x_{i,j}, z_i\}$
  end if
end for
End iterate
Output: Updated classifier c_1 and c_2.

2.2. Co-training SVM Classifiers

In biometrics, obtaining a large number of labeled examples is a difficult and expensive task. On the other hand, obtaining large scale unlabeled examples is relatively easy. In a semi-supervised co-training framework, a small initial labeled training set is available for training the classifiers and then a large number of unlabeled instances (scores generated during probe verification) are available sequentially once the system is in use. In the proposed framework, co-training is used to leverage the availability of multiple classifiers and unlabeled instances to update the decision boundaries of both the classifiers and account for the wide intra-class variations introduced by the probe set. It assumes the availability of two classifiers trained on separate views where the classifier for each view has sufficient (better than random) classification performance. Further, it is important that the classifiers have low correlation in their match scores. This is because, with low correlation, the two classifiers potentially yield different results. For example, one classifier may correctly classifies the unlabeled instance with high confidence, while the other classifier may make a mistake or may not be confident of the prediction. However, even with limited dependence, the proposed co-training framework can improve the performance of individual classifiers as discussed in [2].

The two classifiers are first trained on an initial small labeled data set. During probe verification, instances (scores) are generated by comparing probe images against the gallery. Unlike online learning, the instances obtained during probe verification are unlabeled. For every query given to the biometric system, both the classifiers are used to classify the instance. Here, each instance has two views, $\mathbf{u}' = \{x_1, x_2\}$ and constitutes the unlabeled set D_U. If one classifier confidently predicts the *genuine* label for the instance while the other classifier predicts the *impostor* label with low confidence, then this instance is added as a labeled re-training sample for the second classifier and vice-versa. In this manner, the co-training framework transforms unlabeled scores into labeled training data to update the classifiers.

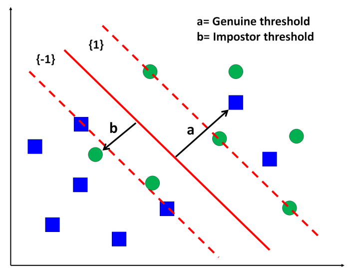

Figure 2. Illustrates the process of computing the confidence of prediction for the SVM classifier.

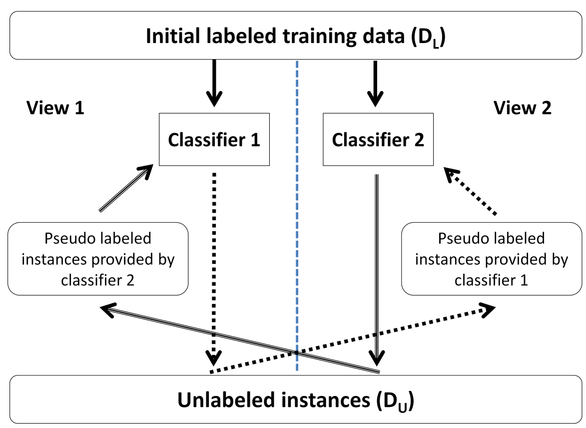

Figure 3. Illustrates the co-training process where each online classifier provides informative labeled instances to the other classifier.

In the co-training approach, as shown in Figure 2, the confidence of prediction by each SVM classifier is measured in terms of distance of the instance from the decision hyperplane. A genuine threshold is computed as the distance of the farthest impostor point that is erroneously classified as a genuine point. An impostor threshold is computed as the distance of the farthest genuine point that is erroneously classified as an impostor. For an instance to be confident enough to lie in the genuine class, its distance from the decision hyperplane should be greater than the genuine threshold. Similarly, for an instance to be confident enough to lie in the impostor class, its distance from the decision hyperplane should be greater than the impostor threshold. Varying the thresholds will change the number of instances on which the co-training is performed. High threshold values imply conservative co-training while smaller values of the threshold will lead to aggressive co-training. The proposed co-training framework is illustrated in Figure 3 and described in Algorithm 2.

Algorithm 2 Co-training

Input: Set of labeled training data D_L, set of unlabeled instances D_U, where each instance $\mathbf{u}' = (x_{i,1}, x_{i,2})$ represents two view/scores.
Process: Train classifier c_j on separate views of D_L. Compute confidence threshold T_j, where $j =$ no of views
for $k = 1$ **to** sizeof(D_U) **do**
  Predict labels: $c_j(x_i) \rightarrow y_{i,j}$; α_j represents confidence of prediction
  if $\alpha_1 > T_1$ & $\alpha_2 < T_2$ **then**
    Update c_2 with labeled instance $\{x_{i,2}, y_{i,1})\}$ & recompute T_2
  end if
  if $\alpha_1 < T_1$ & $\alpha_2 > T_2$ **then**
    Update c_1 with labeled instance $\{x_{i,1}, y_{i,2})\}$ & recompute T_1
  end if
end for
Output: Updated classifier c_1 and c_2.

2.3. Co-training Online SVM Classifiers

The online learning and co-training approaches are extended to propose a framework that simultaneously uses online learning and co-training to update the classifier using labeled and unlabeled data as and when they arrive. The classifiers are initially trained on a small labeled training data set. For every new user being enrolled in the system, online learning is used to update the classifiers using the labeled data generated during enrolment. During probe verification, whenever a user queries the system, co-training is used to update the classifiers using the unlabeled data.

3. Case Study: Multi-classifier Face Verification

To evaluate the effectiveness of the proposed co-training framework, experiments are performed using a multi-classifier face verification application. The case study on multi-classifier face verification comprises of two classifiers trained on separate views (scores) of a face image. Point-based Speeded Up Robust Features (SURF) [6] and texture-based Uniform Circular Local Binary Pattern (UCLBP) [3] are used as facial feature extractors along with χ^2 distance for matching. UCLBP and SURF are used for facial feature extraction because they are fast, discriminating, rotation invariant, and robust to changes in gray level intensities due to illumination variations. Further, selecting point and texture based extractors ensure that the two

Table 1. Constituent face databases used in this research.

Database	Number of subjects	Number of images
AR [16]	119	714
WVU mutimodal [10]	270	3482
MBGC v.2 [1]	446	5468
Caspeal [12]	711	5658
CMU Multi-PIE [4]	287	4828
Total	**1833**	**20150**

views have lower dependence[2]. Two SVM classifiers, one for SURF (*classifier*1) and another for UCLBP (*classifier*2), are trained to classify the scores as $genuine$ or $impostor$. SVM classifiers are then updated using the proposed framework for the labeled and unlabeled instances as and when they arrive. The final classification is obtained by combining the responses from the two updated classifiers using SVM fusion [11].

To analyze the performance on a large database, images from multiple face databases are combined to create a heterogeneous face database of 1833 subjects. The heterogeneous face database comprise of face images with slight pose, expression, and illumination variations. Table 1 provides details about the constituent face databases used in this research. Every subject having six or more samples of face images is selected from these databases. In all the experiments, two images per subject are used in the gallery and the remaining are used as probe. Though each constituent database has large number of images per subject, images exhibiting large pose (> 30 degree), extreme illumination conditions, and occlusion are ignored. Further, face images are geometrically normalized, and the size of each detected face is 196×224 pixels.

3.1. Experimental Protocol

The experimental protocol is designed such that the classifiers are first trained on labeled training data and then variations due to new enrolments and probes are simultaneously learned using online learning and co-training. To update biometric classifiers, a joint adapt-and-test strategy [17] is used which allows for seamlessly adapting and testing. The performance of the proposed framework is compared with batch/offline learning, online learning, and co-training. The following experiments are performed to analyze the performance of the proposed framework.

- For batch learning, the classifiers are trained on all 1833 subjects in batch mode.
- For online learning, the classifiers are initially trained on randomly chosen 600 subjects and then online learning is performed using the remaining 1233 subjects, one subject at a time.
- To evaluate the effectiveness of co-training, two experiments are performed.
  - In the first experiment, the two classifiers are trained on (initial) 600 subjects; however, the gallery comprises of 1833 subjects. The co-training is performed using the probes of all 1833 subjects and this experiment is termed as *co-training*-1.
  - In the second experiment, the classifiers are trained using all 1833 subjects in batch mode and co-training is performed using the probe images. This experiment is referred as *co-training*-2.

[2]In our experiments, SURF and UCLBP had genuine Pearson's correlation of 0.58 and impostor Pearson's correlation of 0.46.

The results are reported based on five-fold non-overlapping random cross validation and verification accuracies are computed at 0.01% false accept rate (FAR).

3.2. Results and Analysis

Figure 4 shows the Receiver Operating Characteristic (ROC) curves for the multi-classifier face verification system. Table 2 summarizes the verification accuracies and computational time for the experiments. The key results and analysis are listed below:

- ROC curves in Figure 4 show modest improvement in the performance of classifiers with the proposed classifier update framework. The framework improves the performance by at least 0.54% compared to batch learning, online learning, and co-training. As mentioned previously, the proposed framework provides a mechanism to seamlessly update the individual classifiers using labeled as well as unlabeled instances. Further, a better classification performance is obtained by combining the decisions from the two classifiers (SVM-fusion) as shown in Figure 4(c).
- The proposed framework provides another benefit in terms of reducing the classifier training time. Table 2 shows that the framework reduces the training time to almost half the time required for batch learning while modestly improving the accuracy.
- It is observed that classification performance of online learning is comparable to that of batch learning. However, online learning provides a great benefit by reducing the training time to one-third. Once the initial training is performed, the classifier is re-trained in a supervised manner using only the instances in which it makes an error and the previous support vectors.
- Co-training provides an improvement in verification accuracy over both batch learning and online learning because the classifiers trained on different scores update each other by providing pseudo labels for the

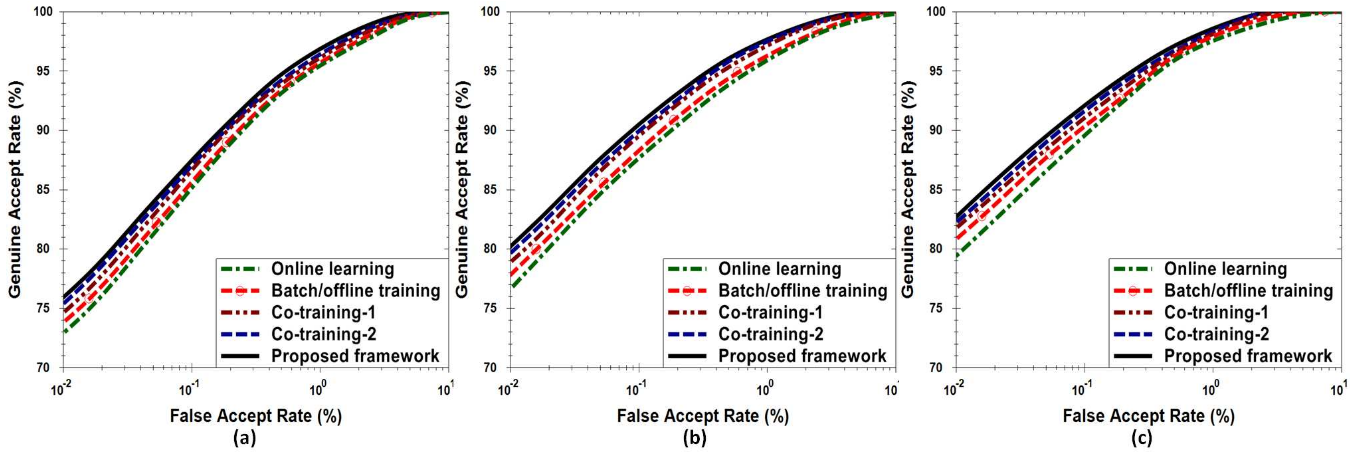

Figure 4. ROC curves showing the comparison between batch/offline training, online leaning, co-training-1, co-training-2 and the proposed framework for (a) Classifier1 (SURF), (b) Classifier2 (UCLBP), and (c) SVM-fusion of the two classifiers.

instances where the other classifier makes an error. This information is not available when classifiers are trained in batch mode or using online learning where each classifier is trained/updated independent of each other. Another reason for such an improvement in individual classifiers is the low correlation between them. Note that correlation between face classifiers for the genuine class is 0.58 and correlation for the impostor class is 0.46. If the correlation between individual classifiers is high, the improvement due to co-training may be limited.

- In co-training-1, since the initial training is done only on 600 subjects, the computational time is lower compared to co-training-2, where classifiers are trained in batch/offline mode on the complete 1833 subjects as reported in Table 2. However, better verification accuracies are achieved by using co-training-2 because classifiers have already seen instances (scores) pertaining to all the users and now they have to incorporate variations due to probe only.

- It can be argued that co-training can be counterproductive because of incorrect pseudo-labeled instances (*negative co-training*). However, it can be avoided by intelligently selecting the confidence thresholds for classifier update. By varying the confidence threshold for a classifier, the number of sample points on which co-training is performed can be controlled. For the proposed framework, *classifier*1 was updated on $34,086$ instances and *classifier*2 was updated on $42,102$ instances using co-training during probe verification. It was observed that 98.86% of the total updates were correct (i.e., pseudo labels obtained using co-training were correct). For online learning, during enrolment, *classifier*1 was updated using 22,145 instances and *classifier*2 was updated using 31,846 instances.

4. Conclusion and Future Work

This research addresses the problem of updating biometric classifiers in order to adapt to variations caused by new enrolments and new probe data. Updating a classifier is posed as a semi-supervised learning task where both labeled as well as unlabeled instances are used. Online learning is used to update the classifier whenever a labeled instance is available and the classifier makes a prediction error on that instance. In addition, co-training is used to update the classifier to continuously improve its performance using the unlabeled instances obtained during probe verification. Further, the proposed co-training online classifier framework provides improvement in the performance both in terms of accuracy and computational time.

The concepts of online learning and co-training are relatively new in biometrics. As future work, the proposed framework can be extended to different stages of a biometric system that require regular updates. We also plan to incorporate the quality of the given gallery-probe pair in computing the confidence of prediction rather than making a decision based only on the distance from the hyperplane.

5. Acknowledgement

The authors thank the reviewers for their constructive feedback. This work was supported by a grant from the US Army Research Laboratory, Contract No. W911NF-10-2-0021.

References

[1] http://www.nist.gov/itl/iad/ig/mbgc-presentations.cfm. 5

Table 2. Verification accuracy and training time of the classifiers trained using different modes.

Mode	Verification Accuracy at 0.01% FAR (%)			Training Time (Minutes)	
	Classifier1	Classifier2	SVM Fusion	Classifier1	Classifier2
Batch/Offline training	73.82	77.88	80.78	98.62	110.84
Online learning	72.96	76.58	79.42	24.32	32.42
Co-training-1	74.64	78.92	81.86	28.25	36.18
Co-training-2	75.48	79.62	82.24	114.75	128.56
Proposed framework	**76.02**	**80.28**	**82.78**	**45.61**	**54.05**

[2] S. Abney. Bootstrapping. In *Proceedings of Annual Meeting on Association for Computational Linguistics*, pages 360–367, 2002. 2, 3

[3] T. Ahonen, A.Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006. 4

[4] S. Baker, T. Kanade, J. Cohn, I. Matthews, and R. Gross. Multi-pie. In *Proceedings of International Conference on Automatic Face and Gesture Recognition*, pages 1–8, 2008. 5

[5] M. Balcan, A. Blum, P. P. Choi, J. Lafferty, B. Pantano, M. R. Rwebangira, and X. Zhu. Person identification in webcam images: An application of semi-supervised learning. In *Proceedings of International Conference on Machine Learning Workshop on Learning from Partially Classified Training Data*, pages 1–9, 2005. 2

[6] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. In *Proceedings of European Conference on Computer Vision*, pages 404–417, 2006. 4

[7] A. Blum and T. Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of Conference on Computational Learning Theory*, pages 92–100, 1998. 1, 2

[8] G. Cauwenberghs and T. Poggio. Incremental and Decremental Support Vector Machine Learning. In *Proceedings of Advances in Neural Information Processing Systems*, pages 409–415, 2000. 3

[9] O. Chapelle, B. Schölkopf, and A. Zien. *Semi-Supervised Learning*. The MIT Press, 1st edition, 2010. 1

[10] S. Crihalmeanu, 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. 5

[11] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun. Discriminative multimodal biometric authentication based on quality measures. *Pattern Recognition*, 38(5):777–779, 2005. 5

[12] W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang, and D. Zhao. The CAS-PEAL large-scale chinese face database and evaluation protocols. Technical report, Joint Research and Development Laboratory, CAS, 2004. 5

[13] X. Jiang and W. Ser. Online fingerprint template improvement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1121–1126, 2002. 2

[14] Y. Kim, K.-A. Toh, and A. Teoh. An online learning algorithm for biometric scores fusion. In *Proceedings of International Conference on Biometrics: Theory Applications and Systems*, pages 1–6, 2010. 1

[15] X. Liu, T. Chen, and S. M. Thornton. Eigenspace updating for non-stationary process and its application to face recognition. *Pattern Recognition*, 36(9):1945–1959, 2003. 2

[16] A. Martinez and R. Benevento. The AR face database. *CVC Technical Report #24*, 1998. 5

[17] N. Poh, R. Wong, J. Kittler, and F. Roli. Challenges and research directions for adaptive biometric recognition systems. In *Proceedings of International Conference on Advances in Biometrics*, pages 753–764, 2009. 2, 5

[18] B. Raducanu and J. Vitri. Online nonparametric discriminant analysis for incremental subspace learning and recognition. *Pattern Analysis and Applications*, 11:259–268, 2008. 1

[19] F. Roli, L. Didaci, and G. Marcialis. Adaptive biometric systems that can improve with use. In *Proceedings of Advances in Biometrics: Sensors, Systems and Algorithms, Springer*, pages 447–471, 2008. 2

[20] C. Ryu, H. Kim, and A. Jain. Template adaptation based fingerprint verification. In *Proceedings of International Conference on Pattern Recognition*, pages 582–585, 2006. 2

[21] R. Singh, M. Vatsa, A. Ross, and A. Noore. Biometric classifier update using online learning: A case study in near infrared face verification. *Image and Vision Computing*, 28:1098–1105, 2010. 1, 3

[22] M. Uray, D. Skocaj, P. M. Roth, and H. Bischof. Incremental lda learning by combining reconstructive and discriminative approaches. In *Proceedings of British Machine Vision Conference*, pages 272–281, 2007. 1

[23] W. Wang and Z. H. Zhou. A new analysis of co-training. In *Proceedings of International Conference on Machine Learning*, pages 1135–1142, 2010. 2