Aggregating Partial Least Squares Models for Open-set Face Identification

Samira Silva[1], Filipe Costa[2], William Robson Schwartz[1]

[1]*Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil*
[2]*CPqD – Image and Speech Processing Management, Campinas, Brazil*
{samirasilva, william}@dcc.ufmg.br, fcosta@cpqd.com.br

***Abstract*—Face identification is an important task in computer vision and has a myriad of applications, such as in surveillance, forensics and human-computer interaction. In the past few years, several methods have been proposed to solve face identification task in closed-set scenarios, that is, methods that make assumption of all the probe images necessarily matching a gallery individual. However, in real-world applications, one might want to determine the identity of an unknown face in open-set scenarios. In this work, we propose a novel method to perform open-set face identification by aggregating Partial Least Squares models using the one-against-all protocol in a simple but fast way. The model outputs are combined into a response histogram which is balanced if the probe face belongs to a gallery individual or have a highlighted bin, otherwise. Evaluation is performed in four datasets: FRGCv1, FG-NET, Pubfig and Pubfig83. Results show significant improvement when compared to state-of-the art approaches regardless challenges posed by different datasets.**

I. Introduction

Face Recognition is a natural task performed daily and effortlessly by human being. However, the necessity of recognizing large sets of faces either in challenging scenarios or in a short time has led to the emergence of research and development of computer systems able to automatically recognize face images.

According to [1], there are three different tasks in face recognition relying upon which scenario it will be required: verification, identification and watch list. The *verification* task consists in comparing two face images in order to determine whether or not they belong to the same person. In face *identification* task, we compare a query face against multiple faces in the enrollment database to associate the face identity to one of those gallery individuals. In the *watch list* task, also known as open-set recognition task, the goal is more than finding the most likely identity for a probe face. This task also considers the possibility of the query face does not belong to any individual enrolled in the gallery. In other words, the query face may belong to an unknown individual. Generally, a filtering method is applied basing on a confidence threshold which delimits the minimum similarity score necessary to consider a query as belonging to a known individual. Queries which do not attain that minimum value are then rejected (not identified).

Numerous challenges are associated to face recognition, mainly in the image acquisition in unconstrained scenarios, and several methods to overcome these problems have been proposed [2]–[5]. Besides, the scenario in which the application is immersed may also imposes other challenges. Most studies in the literature are focused on solving problems related to the aforementioned challenges considering a closed-set scenario. In identification task, for instance, they assume that all probe face images belong to a known individual. However, in a real-world application, this assumption may not be correct.

Different from most approaches in the literature, we propose in this thesis approaches for combining a set of classification models trained employing the *one-against-all* protocol, aiming at identifying face images in an open-set scenario. These models are based on Partial Least Squares (PLS) [6], which operates weighting features from images in order to discriminate throughout different classes, dealing satisfactorily with high-dimensional data and overcoming the frequent problem of lack of samples. We present an effective and simple technique for face recognition that handles unknown subjects, namely open-set face identification task or watch list. Specifically, we want to identify a face image from a gallery/known individual and reject if it belongs to a non-gallery/unknown individual. We demonstrate that information produced by a set of simple *one-against-all* PLS models can be combined to reveal not only whether a sample is known but also its identity. Finally, we also proposes a novel manner to compute the threshold that establishes whether a face image belongs to a known subject or not which improves literature results.

This paper is structured as follow. Section II presents a brief review of literature, analyzing methods regarding two tasks, watch list and open-set recognition in different scenarios (object, cameras, handwritten digit, etc). The proposed methodology is detailed in Section III and a vast number of experiments is reported in Section IV. Finally, we conclude in Section V, presenting final remarks and future works.

MSc. Thesis. The authors thank the financial suport of the Coordination for the Brazilian National Research Council - CNPq (Grant #311053/2016-5), the Minas Gerais Research Foundation - FAPEMIG (Grants APQ-00567-14, RED-00042-16, PPM-00540-17) and the Coordination for the Improvement of Higher Education Personnel - CAPES (DeepEyes Project).

II. Literature Review

According to Google Scholar[1], almost a thousand of face recognition algorithms were published in 2016. It demonstrates that it still is a problem not solved, mainly in large-scale and unconstrained scenarios. This section provides a brief review on the literature concerning watch list and open-set recognition, the former focus on face identification in open-set scenarios and the latter tackles open-set recognition works in general, not necessarily in face recognition.

A. Watch List

Watch list consists in first detecting whether a face image belongs to a gallery individual and then determining its identity. One of the scenarios in which several approaches fail is when there are variations on the number of samples in the training stage. The work of [7] formulates the open-set face recognition problem as a multiple verification task using SVMs. Their experiments demonstrate that an increase in the gallery size can impact negatively the approach performance. Besides, classification with SVM may not perform effectively when only a low amount of training data is available. On the other hand, [8] represents a face image as a linear combination of training sample images. This representation is precise only if a sufficient number of samples are available in the training set for the correct testing class.

Various works to perform open-set identification in literature are based on the confidence of a previous stage. [9] proposes the examination of a media collection to perform watch-list, while [10] address the problem of open-set recognition in web-scale datasets using a combination of a subset for images from the dictionary of faces. Both approaches relies on a accurate preprocessing stage to perform well which could be a problem.

Some approaches in literature focus on solving the watch list problem using PLS as the work herein proposed. [11] propose the Partial Least Squares Hashing (PLSH), a scheme combining hashing functions and classification methods in order to predict when probe samples belong to gallery set based on the methods proposed by [12], [13]. They combine the responses of a determined number of PLS or FNC (Fully Connected Network) models trained with the whole gallery in order to increment a vote list which, after being thresholded, can decide whether a face image belongs to the gallery or not. However, they only evaluate the former step of watch list in which known/unknown samples are detected. The identification of samples predicted as known is not performed and evaluated. Furthermore, the fact that each models is trained with half of gallery in each set, positive and negative, builds models less discriminative than using the one-against-all protocol, as in the proposed work.

Finally, more recent works employ the deep learning concepts in the watch list task. [14] state the problem of deep face recognition in open-set scenario as a problem in which face features have smaller maximal intra-class distance than minimal inter-class distance considering a determined metric. They propose a new softmax loss function to a CNN which is employed to feature extraction. Cosine distance combined with nearest neighboor classifier provides a score to accept or reject a sample. Besides the fact that CNNs need a huge amount of data in training, the computation of nearest neighboor is also an expensive task, since it is needed to be computed between probe image and all gallery subjects.

[1]https://scholar.google.com

B. Open-set Recognition

Aside from face identification, open-set recognition, in general, has been studied and employed in many applications. Several approaches employ the Support Vector Machine (SVM) paradigm to solve the open-set recognition problem. The work of [15], with base on the traditional 1-class and binary SVMs, presents the “1-vs-set machine” approach, aims at sculpting a decision space from the marginal distances of a 1-class or binary SVM with a linear kernel. Thus, instead of one plane marking the decision boundary, as in the traditional SVM, two planes are defined to minimize a determined equation in a greedy way and then a refining step is applied. Its extension proposed in [16] defines the term “open space risk” in order to solve the same problem. Besides, [17] proposes an approach to solve the problem of open-set source camera attribution also based on SVM, which can be seen as an extension of the work of [15].

Different from the majority of methods in open-set recognition, [18] handle the possibility of at testing time seeing a sample from an unknown class and adding it as a new class to the model. This problem is introduced and formalized by them as “Open World Recognition”. As the aforementioned works, they also manage the open space risk.

Visual recognition problems have taken advantage of deep networks in many aspects. They can be mostly useful in applications where human being are not effective in recognizing since they consider parts in images which are unmeaningful for humans. The development of an open-set deep network was only first approached in the work of [19], introducing a new layer to handle unknown classes. Neural networks provide impressive accuracy but also present drawbacks related to few samples. Training a neural network with few samples may cause overfitting in the data. Other aspect with respect to neural network training is the also expensive computational cost.

III. Methodology

This section describes the approach proposed in this work with the purpose of performing watch list. It is inspired by the approach developed by [11], which also proposes an embedding of PLS classifiers. However, we present modifications in many aspects to promote improvements in terms of accuracy, while requiring a lower number of models, reducing both the computational time and the memory required for store the models. Also, different from several works in the literature, we present herein the complete procedure for identifying a person in an open-set scenario, which is also a difference compared to the work proposed in [11]. We divide the procedures into three stages which are presented and detailed as follows.

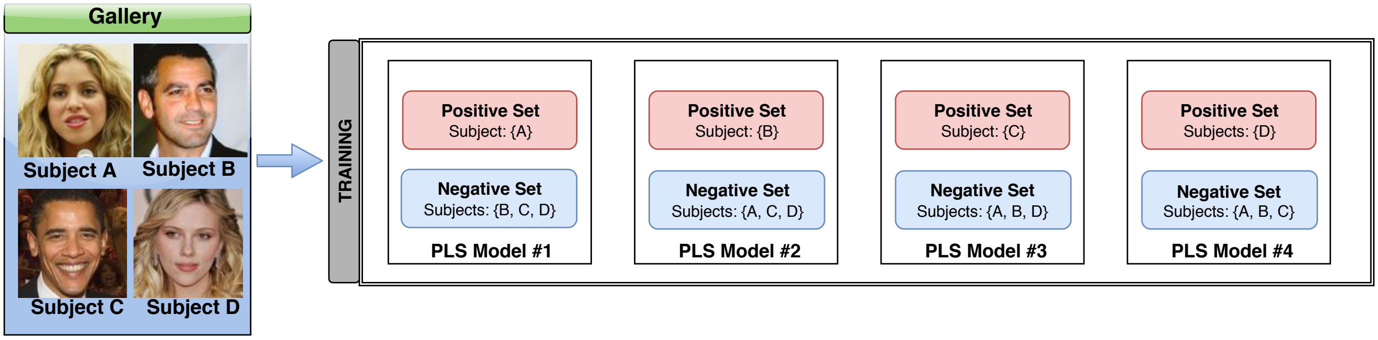


Fig. 1: The training stage of the proposed approach. Face images of gallery subjects are used to learn a set of PLS classifiers. Each model utilizes a different subject as positive set against a negative set constituted by the remaining subjects. Models generation is finished in the moment we have a model per gallery subject.

A. Representation and Partitioning Stage

Initially, subjects of the dataset are randomly split into two balanced sets: known (gallery) subjects and unknown (non-gallery) subjects. Samples from known subjects are also equally partitioned into training and testing data. Features are then extracted for all dataset samples using the VGGFace CNN descriptor [20]. Thus, the input of the training stage consists in features extracted only from known subject samples. On the other hand, the test stage receives features extracted from both known and unknown subject samples.

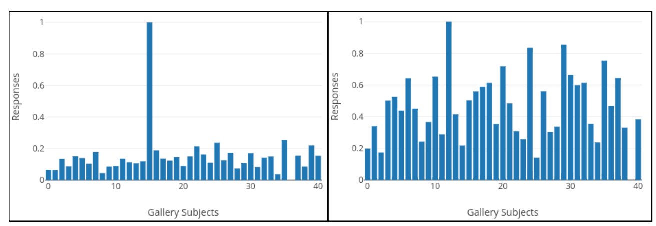
B. Training Stage

Let gallery subjects be represented by S=$\{s_1, s_2, s_3, s_4, ..., s_{n-1}, s_n\}$, where n is the number of subjects in gallery. We train a PLS model per subject s_i with features extracted from gallery samples in training data as follows. A PLS_i model is a classifier trained with samples of subject s_i in the positive set and samples of the remaining subjects $\{s_1, s_2, .., s_{i-1}, s_{i+1}, s_{i+2}, ..., s_{n-1}, s_n\}$ in the negative set. In other words, each PLS_i classifier models a specific gallery subject s_i against the remaining gallery subjects. We perform the training of PLS_i models until we have one model per subject s_i. Figure 1 depicts the training stage of the proposed approach.

C. Test Stage

We start the test stage by creating a vote-list v with length equals to n, the gallery size, replete of zeros. Considering all PLS_i models have already been trained, a query face image is presented to each of them and their response values are added to the position i of the vote-list, v_i, which corresponds to the identity in the positive set of the corresponding model. In other words, each model provides a score that encodes the similarity between the probe and the subject in the positive set used to train the model. It can be high if probe matches the positive set or low if probe does not match the positive set.

After projecting a probe image onto all PLS_i classifiers, if the vote list has a highlighted bin, probe probably belongs to the subject corresponding to it. Otherwise, probe does not match any individual, being then predicted as unknown. Higher responses are provided by models trained with the probe identity in the positive set. Lower responses are obtained by models trained with the non-matching identities in the positive set. If a known probe is presented to the models, only the model using its identity in the positive set would provide high response. Figure 2 exemplifies the vote-list for both cases, known *(a)* and unknown *(b)* probe samples. Finally, we



(a) Gallery probe (b) Non-Gallery probe

Fig. 2: The vote-list for known and unknown probe images from FG-NET dataset. The vote-list for known subjects *(a)* presents a position with value much higher than the remainder. The vote-list for unknown subject *(b)* is balanced since probe image is not similar to any gallery subject in particular.

threshold the vote-list with the purpose of identifying whether it presents a position with a highlighted value, and therefore, probe is predicted as known. Otherwise, probe is rejected as unknown. The method we use to perform this procedure is described in Section IV. Figure 3 summarizes the test stage of the proposed approach.

Taking into account a probe image is predicted as belonging to a gallery subject, we identify this sample using PLS with the one-against-all protocol. For a fair comparison, this identification approach is fixed and only the stage of rejecting or accepting a face image may be variated to all compared approaches.

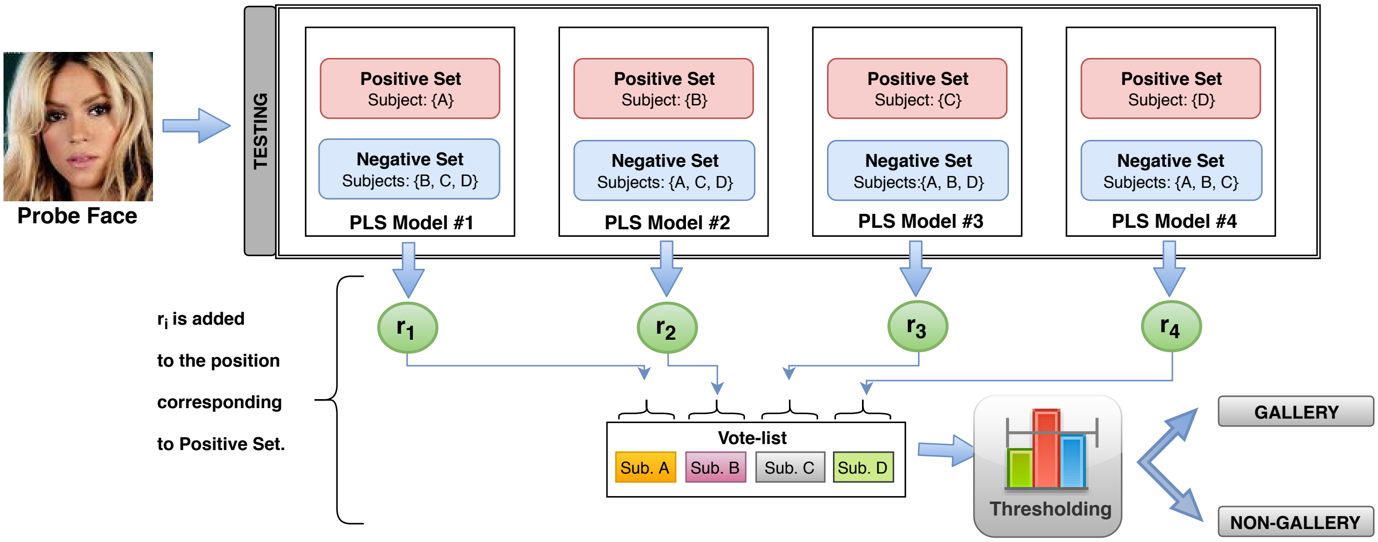


Fig. 3: The test stage of the proposed approach. We first project a test sample onto the set of PLS models learned in training stage. Each model provides a response r_i related to its similarity to the corresponding positive set. Then, we add the response r_i to the vote-list position corresponding to the positive set of model i. We repeat this proceeding for all models. Finally, a threshold applied in the normalized vote-list predicts probe image as known or unknown.

IV. Experiments

This section provides the datasets, evaluations metrics and protocol used, as well as an analysis of the experiments performed with the proposed approach and literature methods.

A. Datasets

To validate the effectiveness of our method, we select datasets with different characteristics mostly associated to the acquisition process. They vary from frontal images taken in total controlled scenarios to images in the wild collected from Internet, presenting variations including illumination, pose and expression. We evaluate our method on four datasets: FRGCv1 [21], Pubfig [22], Pubfig83 [23] and FG-Net Aging [24]. The first is a well-known face recognition dataset and the others are more recent and unconstrained datasets.

B. Evaluation Metrics

When it comes to metrics for open-set face recognition, there is not a worldwide consensus. Thus, in this work, we choose two metrics to evaluate different aspects in recognition: the Receiver Operating Characteristic (ROC) and the Open-set Receiver Operating Characteristic (Open-set ROC). The former captures information on how accurate an approach is in terms only of detecting a gallery probe sample, i.e., determining whether a probe sample belongs to the gallery. The latter evaluates both the detection of a gallery sample and its correct identification.

C. Evaluation Protocol

The proposed approach will be compared with five different approaches: 1-class and Binary 1-vs-set M. [15], the 1-class and Binary SVM [25] baselines, and PLSH [11]. PLSH needs the number of models to build as a parameter. In our experiments, we report the results using two different numbers: 500, which is the maximum value evaluated in their paper, and the number of known subjects, which is the same number of models of the proposed approach. We call them PLSH-500 and PLSH-Subj, respectively.

In our experiments, we call *background set* some extra non-overlapping data to be added to the negative set, aiming at helping to reinforce the differences betweeen negative and positive sets. The methods which employs this set are the proposed approach (Our-Bk), the 1-class and the Binary 1-vs-set M. [15]. To FRGC experiment four, FRGC experiment two and Pubfig, data with similar acquisition characteristics are considered to background set.

The *threshold* determines whether a probe face image belongs to a gallery subject or not. It is employed in the vote-list described in Section III-C. We provide an evaluation of four different thresholds to find out the one that better impacts our algorithm. Thresholds 1, 2 and 3 are proposed by [11]. We propose Threshold 0 as follows:

$$\tau_0 = \mathcal{Z}_{S_1} - AVG(\mathcal{Z}_{S_2}, ..., \mathcal{Z}_{S_{p+1}}), p = \lceil 0.10 \times |\mathcal{Z}| \rceil, \quad (1)$$

where $Z = \{\mathcal{Z}_{S_1}, ..., \mathcal{Z}_{S_n}\}$ is the normalized vote-list for a probe image, sorted in ascending order. The value of p defines the proportion of Z considered. The function $AVG(X)$ computes the average of the values in X. The idea is to capture the relation of the top scorer s_1 with the succeeding subjects. p is set to 10% of the total of known subjects since it presents the best results. If the probe image belongs to a known subject, the difference between them would be high.

The remaining thresholds (1, 2 and 3) are based on the ratio between the top scorer and the succeeding subjects.

Figure 4 displays the ROC curve using the proposed approach on FG-Net dataset for each threshold aforementioned. As depicted, the threshold described in Equation 1 provides better performance than the remainder, confirming our assumption that the difference between the top scorer value and the averaged remainder in vote-list better separates known subjects from unknown when compared to approaches based on the ratio. Therefore, we choose the threshold τ_0 in the remaining experiments.

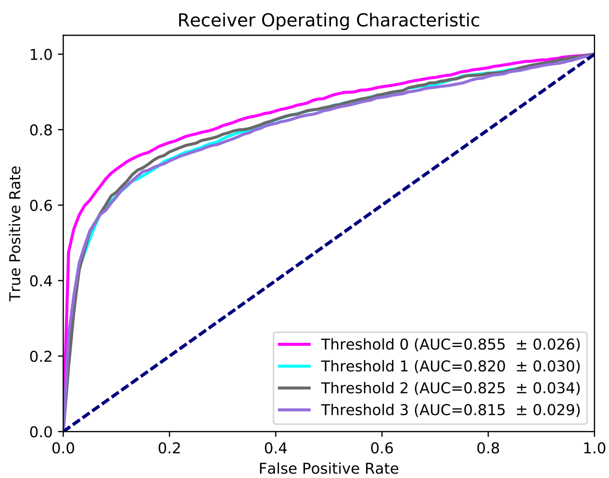


Fig. 4: Average ROC curves for the proposed approach performed on the FG-NET dataset with different thresholds. We repeated this experiment ten times, fixing variable p to 10% of all known subjects.

D. Detection and Identification Evaluation

In this section we present the experimental results reported with ROC and Open-set ROC curves in different datasets. The AUC (Area Under Curve) attained by different approaches is provided to make a comparison regardless the dataset difficulty. Tables I and II summarize the results obtained by the proposed approach and literature methods using AUC of ROC and Open-set ROC curves, respectively.

TABLE I: Summarization of AUC (± standard deviation) of ROC curve for different open-set recognition approaches and datasets.

Dataset	FRGC-SET2	FRGC-SET4	FG-NET	Pubfig83	Pubfig-Eval
1-Class-1-vs-set-M.	0.817 ± 0.044	0.610 ± 0.037	0.630 ± 0.029	0.856 ± 0.015	0.719 ± 0.014
1-Class-SVM	0.525 ± 0.049	0.511 ± 0.028	0.520 ± 0.064	0.550 ± 0.047	0.523 ± 0.026
Binary-1-vs-set-M.	0.970 ± 0.005	0.873 ± 0.022	0.499 ± 0.029	0.833 ± 0.025	0.899 ± 0.018
Binary-SVM	0.994 ± 0.000	**0.913**± 0.023	0.831 ± 0.029	0.973 ± 0.003	0.915 ± 0.005
PLSH-500	0.990 ± 0.003	0.867 ± 0.019	0.820 ± 0.018	0.966 ± 0.003	0.926 ± 0.006
Ours	0.993 ± 0.002	0.882 ± 0.022	**0.855** ± 0.026	**0.975** ± 0.003	0.942 ± 0.003
Ours-Bk	**0.994** ± 0.001	0.905 ± 0.019	0.851 ± 0.019	0.973 ± 0.003	**0.943** ± 0.004

In the comparison with the literature approaches reported in Table I, we can notice that the proposed approaches achieve the best results in all datasets, except by the FRGC-SET4, in which Binary-SVM provides the highest AUC. However,

TABLE II: Summarization of AUC (± standard deviation) of Open-set ROC curve for different open-set recognition approaches and datasets.

Dataset	FRGC-SET2	FRGC-SET4	FGNET	Pubfig83	Pubfig-Eval
1-Class-1-vs-set-M.	0.814 ± 0.044	0.551 ± 0.038	0.540 ± 0.030	0.849 ± 0.014	0.704 ± 0.013
1-Class-SVM	0.524 ± 0.049	0.454 ± 0.031	0.423 ± 0.050	0.543 ± 0.045	0.510 ± 0.025
Binary-1-vs-set-M.	0.966 ± 0.007	0.781 ± 0.031	0.412 ± 0.031	0.824 ± 0.024	0.845 ± 0.017
Binary-SVM	**0.992** ± 0.003	0.777 ± 0.034	0.675 ± 0.031	0.957 ± 0.004	0.858 ± 0.006
PLSH-500	0.988 ± 0.004	0.784 ± 0.028	0.728 ± 0.026	0.957 ± 0.005	0.886 ± 0.008
Ours	**0.992** ± 0.003	0.787 ± 0.037	**0.761** ± 0.032	**0.965** ± 0.004	**0.899** ± 0.005
Ours-Bk	0.991 ± 0.005	**0.788** ± 0.026	0.739 ± 0.027	0.962 ± 0.005	0.895 ± 0.005

Binary-SVM sees unknown subject data in the training stage, different from our methods that use only gallery data in the training phase, which matches a more realistic scenario. Our method with no background overcomes the method with background when the background set does not have common acquisition chracteristics with the main dataset. It occurs in the experiments with FGNET and Pubfig83. In the remaining, the background set helps to improve the AUC of ROC.

In the experiments using the AUC of Open-set ROC curve as metric (Table II), that is, evaluating the complete pipeline of watch list, our approach reaches the best results using all datasets, but the experiment two of FRGC dataset. In this experiment, our approach presents a result similar to the Binary-SVM. As mentioned before, it is justified by the fact that our approaches do not see non-gallery data in training stage. Comparing the two proposed approaches, we have similar values of AUC in most of the datasets, showing that the use of background does not present significant improvement.

E. Evaluation with the Same Number of Models in PLSH Approach

Considering that our work is an extension of the approach proposed by [11], this section presents a direct comparison between our work and PLSH, reducing the number of models in PLSH to the same number of models we build, the number of gallery subjects. Tables III and IV report the results contrasting our method with PLSH-Subj using ROC and Open-set ROC curves, respectively. As we can notice, the proposed approach outperforms PLSH regardless the metric or dataset difficulty. This fact can be explained since modeling a single subject against the remainder is simpler since samples in the positive set present similar features. On the other hand, PLSH trains models considering half of the gallery subject as positive and the remainder as negative, that is, the positive set contains images with different features since they come from different subjects. In this case, building a single model to discriminate the positive set may be a harder task.

TABLE III: Comparison between AUC (± standard deviation) of ROC of the proposed approach and PLSH containing the same number of models we use.

Dataset	FRGC-SET2	FRGC-SET4	FG-NET	Pubfig83	Pubfig-Eval
PLSH-500	0.990 ± 0.003	0.867 ± 0.019	0.820 ± 0.018	0.966 ± 0.003	0.926 ± 0.006
PLSH-Subj	0.970 ± 0.015	0.841 ± 0.026	0.786 ± 0.026	0.917 ± 0.018	0.904 ± 0.013
Ours	0.993 ± 0.002	0.882 ± 0.022	**0.855** ± 0.026	**0.975** ± 0.003	0.942 ± 0.003
Ours-Bk	**0.994** ± 0.001	**0.905** ± 0.019	0.851 ± 0.019	0.973 ± 0.003	**0.943** ± 0.004

TABLE IV: Comparison between AUC (± standard deviation)of Open-set ROC of the proposed approach and PLSH containing the same number of models we use.

Dataset	FRGC-SET2	FRGC-SET4	FGNET	Pubfig83	Pubfig-Eval
PLSH-500	0.988 ± 0.004	0.784 ± 0.028	0.728 ± 0.026	0.957 ± 0.005	0.886 ± 0.008
PLSH-Subj	0.967 ± 0.015	0.753 ± 0.027	0.689 ± 0.033	0.907 ± 0.018	0.863 ± 0.016
Ours	**0.992** ± 0.003	0.787 ± 0.037	**0.761** ± 0.032	**0.965** ± 0.004	**0.899** ± 0.005
Ours-Bk	0.991 ± 0.005	**0.788** ± 0.026	0.739 ± 0.027	0.962 ± 0.005	0.895 ± 0.005

V. Conclusions and Future Works

In this work, we proposed and evaluated an approach to perform watch list using a combination of PLS models trained in an one-against-all protocol. Responses from the probe projection onto the models are aggregated in a vote-list and used to discriminate between known and unknown subjects. PLS presents many advantages when compared to other techniques, such as SVM. It handles feature vectors with high dimension, applications with few and unbalanced samples per class, and presents robust results while keeping a low computational cost.

Experiments were carried out in a diversity of datasets attaining satisfactory results regardless the challenges they present. Improvements in performance were achieved by embedding a background set, data presenting similar aspects to the main dataset, in the negative set of each PLS model. Different from some traditional approaches, we do not make assumption that unknown training data is available. Instead, we provide a robust method using only training samples from known subject that outperforms literature approaches in most of the experiments.

The adoption of a large number of models, as used in the PLSH algorithm, is unnecessary since training only one model per subject in an one-against-all protocol can provide more discriminative models, while keeping a lower number of models. We also contribute with a novel manner to compute the threshold in a vote-list which attains a more accurate performance when compared to other literature techniques. Different from PLSH, we also proposed a complete pipeline to perform open-set face identification.

In future works, we intend to consider an evaluation of the proposed method on datasets with a huge amount of subjects in order to evaluate the impact in the computational cost. Besides, the method used to build each model, PLS, can be replaced by other binary classifiers such as neural-network-based techniques and other classifiers with accurate performance. Lastly, changes in the partitioning of positive and negative sets in a model could be made in order to not use the whole gallery. Future extensions of the proposed work will be submitted to a journal paper.

References

[1] R. Chellappa, P. Sinha, and P. J. Phillips, "Face recognition by computers and humans," *IEEE Computer*, vol. 43, no. 2, pp. 46–55, 2010.

[2] M. Hayat, S. H. Khan, N. Werghi, and R. Goecke, "Joint registration and representation learning for unconstrained face identification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. –.

[3] X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li, "High-fidelity pose and expression normalization for face recognition in the wild," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 787–796.

[4] L. Tran, X. Yin, and X. Liu, "Disentangled representation learning gan for pose-invariant face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 4, no. 5, 2017, p. 7.

[5] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4295–4304.

[6] H. Wold, "Partial least squares," *Encyclopedia of Statistical Sciences*, pp. 581–591, 1985.

[7] H. K. Ekenel, L. Szasz-Toth, and R. Stiefelhagen, "Open-set face recognition-based visitor interface system," in *International Conference on Computer Vision Systems*, 2009, pp. 43–52.

[8] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[9] L. Best-Rowden, H. Han, C. Otto, B. F. Klare, and A. K. Jain, "Unconstrained face recognition: Identifying a person of interest from a media collection," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2144–2157, 2014.

[10] E. G. Ortiz and B. C. Becker, "Face recognition for web-scale datasets," *Computer Vision and Image Understanding*, vol. 118, pp. 153–170, 2014.

[11] R. Vareto, S. Silva, F. Costa, and W. R. Schwartz, "Towards open-set face recognition using hashing functions," in *International Joint Conference on Biometrics*, 2017, pp. 1–8.

[12] C. dos Santos, E. Kijak, G. Gravier, and W. Schwartz, "Learning to hash faces using large feature vectors," in *International Workshop on Content-Based Multimedia Indexing*, 2015, pp. 1–6.

[13] ——, "Partial least squares for face hashing," *Neurocomputing*, vol. 213, pp. 34–47, 2016.

[14] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "Sphereface: Deep hypersphere embedding for face recognition," *arXiv preprint arXiv:1704.08063*, 2017.

[15] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, 2013.

[16] W. J. Scheirer, L. P. Jain, and T. E. Boult, "Probability models for open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2317–2324, 2014.

[17] F. de O. Costa, E. Silva, M. Eckmann, W. J. Scheirer, and A. Rocha, "Open set source camera attribution and device linking," *Pattern Recognition Letters*, vol. 39, pp. 92–101, 2014.

[18] A. Bendale and T. Boult, "Towards open world recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1893–1902.

[19] A. Bendale and T. E. Boult, "Towards open set deep networks," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1563–1572.

[20] O. M. Parkhi, A. Vedaldi, A. Zisserman *et al.*, "Deep face recognition," in *British Machine Vision Conference*, vol. 1, no. 3, 2015, p. 6.

[21] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 947–954.

[22] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *IEEE International Conference on Computer Vision*, 2009, pp. 365–372.

[23] N. Pinto, Z. Stone, T. Zickler, and D. Cox, "Scaling up biologically-inspired computer vision: A case study in unconstrained face recognition on facebook," in *IEEE Conference Computer Vision and Pattern Recognition Workshops*, 2011, pp. 35–42.

[24] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes, "Overview of research on facial ageing using the fg-net ageing database," *IET Biometrics*, vol. 5, no. 2, pp. 37–46, 2016.

[25] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.