

Kinship Verification based Deep and Tensor Features through Extreme Learning Machine

Oualid Laiadi^{1,4}, Abdelmalik Ouamane², Abdelhamid Benakcha³, Abdelmalik Taleb-Ahmed⁴ and Abdenour Hadid⁵

¹ Laboratory of LESIA, University of Biskra, Algeria

² University of Biskra, Algeria

³ Laboratory of LGEB, University of Biskra, Algeria

⁴ IEMN DOAE UMR CNRS 8520 Laboratory, Polytechnic University of Hauts-de-France, France

⁵ Center for Machine Vision and Signal Analysis, University of Oulu, Finland

Abstract—Checking the kinship of facial images is a difficult research topic in computer vision that has attracted attention in recent years. The methods suggested so far are not strong enough to predict kinship relationships only by facial appearance. To mitigate this problem, we propose a new approach called Deep-Tensor+ELM to kinship verification based on deep (VGG-Face descriptor) and tensor (BSIF-Tensor & LPQ-Tensor using MSIDA method) features through Extreme Learning Machine (ELM). While ELM aims to deal with small size training features dimension, deep and tensor features are proven to provide significant enhancement over shallow features or vector-based counterparts. We evaluate our proposed method on the largest kinship benchmark namely FIW database using four Grandparent-Grandchild relations (GF-GD, GF-GS, GM-GD and GM-GS). The results obtained are positively compared with some modern methods, including those that rely on deep learning.

I. INTRODUCTION

Kinship verification models consists in checking if two persons are belonging to the same family or not is termed kinship (or family) verification. Automatic kinship verification aiming to discover computational models to decide whether two persons are from the same family or not and purely based on patterns such as voices, faces and gaits. The automatic kinship verification systems define their inputs by two faces (Face A and Face B) and the predictable output is a decision whether Person A is with relation with a family member (father, mother, brother, sister etc.) of Person B or not. Many applications can be beneficial e.g. for forensics, finding missing children, social media comprehension and image annotation. Though a DNA test is the most precise way for kinship verification, it regrettably cannot be used in many situations such as in video surveillance.

The few existing works on kinship verification essentially share similar face features as in face recognition. This involves for instance the use of shallow features LBP (Local Binary Patterns), LPQ (Local Phase Quantization) and HOG (Histograms Of Gradients) features for inputs to SVMs (Support Vector Machines) for verification of kin relation [7], [2]. Such methods work better under some limited face image variations (in terms of image resolution, illumination, blur etc.) but always to suffer under unconstrained environment or to generalize to unseen data. However, the

very recently developments in machine learning suggest that highest performance can be obtained from learned features e.g. based on deep learning methods [13] instead of shallow features e.g. LBP, LPQ and HOG.

In this work, we develop a novel framework to kinship verification from face images using deep and tensor features based extreme learning machine. Especially, we used ELM method to fuse metrics (scores) of VGG-Face for deep features and two tensors (BSIF-tensor and LPQ-tensor) features extracted by Multilinear Side-Information based Discriminant Analysis (MSIDA) method, we called our framework Deep-Tensor+ELM. We report our preliminary experimental investigations on the FIW benchmark using four relations, Grandfather-Granddaughter (GF-GD), Grandfather-Grandson (GF-GS), Grandmother-Granddaughter (GM-GD) and Grandmother-Grandson (GM-GS) face subsets showing very promising performance compared to state-of-the-art methods.

II. PROPOSED FRAMEWORK

Figure 1 depicts an overview of our proposed approach. The input is a pair of two face images e.g. a grandparent and a grandchild. We extract features from these images into deep and tensor features of each pair. We compute the cosine similarities between the deep and tensor features to encode the final metric (fused score). These similarities are fed to an extreme learning machine (ELM) classifier. The ELM classifier is trained to predict whether the two persons are kin related or not.

A. Extracting Deep and Tensor Features

Many methods suggested in the literature on automatic verification of kinship have focused mainly on analyzing deep features only, thus ignoring tensor information. Recently, tensor features have shown better performance than their deep counterparts to verify kinship relevance (e.g. [1]). When considering the tensor information, the problem usually consists in learning a discriminating metric where the classification (e.g. kinship verification in our case) becomes more affordable compared to the deep features. As suggested in [5], we consider the score fusion methodology for kinship verification. Deep and tensor metrics show a complementarity which extracted by ELM method. This is in contrast

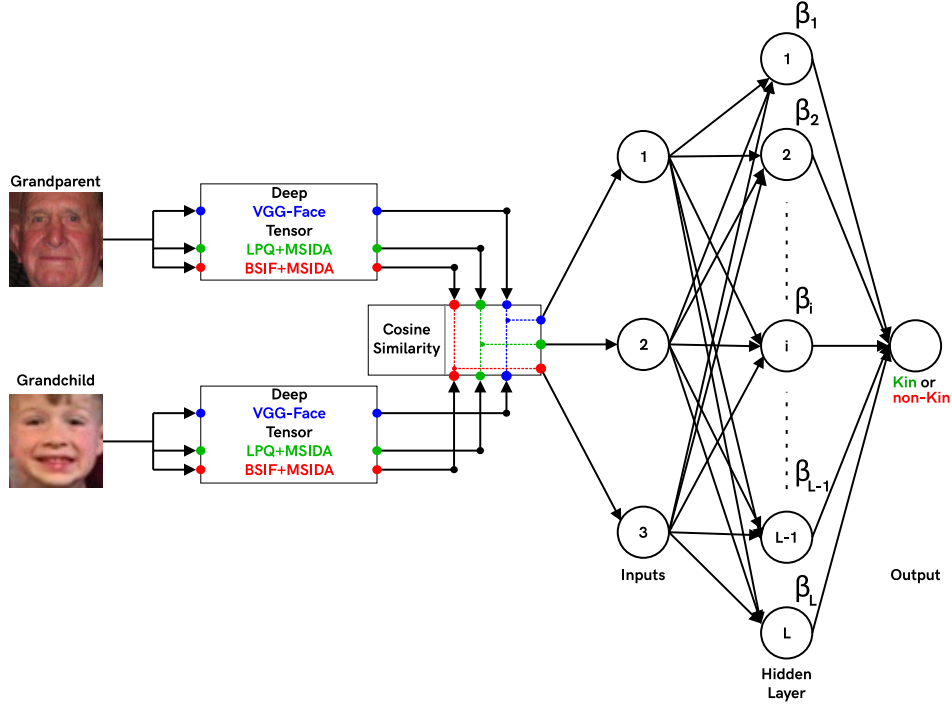


Fig. 1: Overview of the proposed deep-tensor+ELM framework for kinship verification.

to use deep or tensor information separately. Therefore, we extract the deep features using VGG-Face [9] method and tensor features (LPQ-tensor & BSIF-tensor) using MSIDA [1] method.

B. Cosine similarity metrics

The deep and tensor features extracted from each pair of two images should be integrated into a single value. Our experiments examine different metric measures (cosine, Euclidean, Hamming, Manhattan, etc.) demonstrated that cosine similarity metric [8] provides the best performance. The cosine similarity between two feature vectors \mathbf{X} and \mathbf{Y} is given by:

$$\cos(\mathbf{X}, \mathbf{Y}) = \frac{\mathbf{X}^T \mathbf{Y}}{\|\mathbf{X}\| \|\mathbf{Y}\|} \quad (1)$$

Where $\|\cdot\|$ is the Euclidean norm. A high value of the produced score means a high probability that \mathbf{X} and \mathbf{Y} are same family.

C. Classification using Extreme Learning Machine

An Extreme Learning Machine (ELM) [4] contains a single hidden layer in their network which has been shown to perform superior and faster than SVM classifier in some classification problems [3]. The output of an ELM network with L hidden neurons can be defined as:

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = o_j, j = 1, \dots, N \quad (2)$$

where β_i is the weight between the hidden layer and output layer, and $g(x)$ is the activation function.

$X_i = [x_{i_1}, x_{i_2}, \dots, x_{i_n}]^T$ is the input vector with the ground truth $t_i = [t_{i_1}, t_{i_2}, \dots, t_{i_m}]^T$. W_i and b_i are the weight and bias of the hidden layer. One key feature of ELM is to randomly set both W_i and b_i to speed up the training process. The distances between the ground truth and actual output $\sum_{j=1}^N \|o_j - t_j\|$ should be minimized. The output weights are optimized by minimizing the approximation in 3:

$$H \cdot \beta = T \quad (3)$$

where H is the randomly generated hidden layer output matrix,

$$H(W_1, \dots, W_L, b_1, \dots, b_L, X_1, \dots, X_L) = \begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \dots & g(W_L \cdot X_1 + b_L) \\ \vdots & \dots & \vdots \\ g(W_1 \cdot X_N + b_1) & \dots & g(W_L \cdot X_N + b_L) \end{bmatrix}$$

and T is the target output:

$$y = \begin{bmatrix} T_1^T \\ T_2^T \\ \vdots \\ T_N^T \end{bmatrix} \quad (4)$$

The optimization procedure in ELM can be reduced to computing the Moore-Penrose inverse of H , determined at the beginning of the training, rather than optimizing β using gradient descent algorithm by tuning the parameters in an iterative algorithm as in deep architectures. Thus, β can be calculated as:

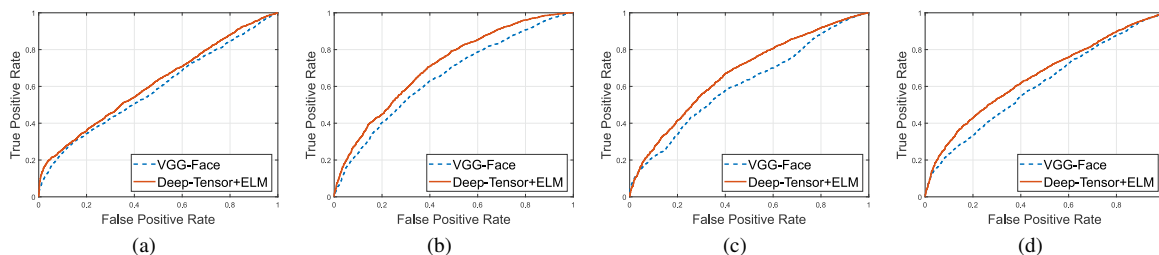


Fig. 2: ROC curves of different methods (VGG-Face and Deep-Tensor+ELM) on FIW database obtained on (a) GF-GD set, (b) GF-GS set, (c) GM-GD set and (d) GM-GS set.

TABLE I: Averaged verification accuracy scores (%) on FIW database for GF-GD, GF-GS, GM-GD and GM-GS relations.

Method	GF-GD	GF-GS	GM-GD	GM-GS	Mean
ResNet+CF [11]	66.37	66.45	64.81	64.39	65.51
SphereFace [6]	66.07	66.36	64.58	65.40	65.60
ResNet+SDMLoss [12]	65.89	65.12	66.41	64.90	65.58
Deep-Tensor+ELM (Our)	68.36	68.18	70.14	67.77	68.61

$$\hat{\beta} = H^{-1} \cdot T \quad (5)$$

III. EXPERIMENTAL ANALYSIS

For experimental evaluation, we considered the largest FIW [10] kinship database using: four relations, Grandfather-Granddaughter (GF-GD), Grandfather-Grandson (GF-GS), Grandmother-Granddaughter (GM-GD) and Grandmother-Grandson (GM-GS) face subsets. In GF-GD subset, there are 7,078 pairs of images for positive and negative relations. In GF-GS subset, there are 4,830 pairs of images for positive and negative relations. In GM-GD subset, there are 6,512 pairs of images for positive and negative relations. In GM-GS subset, there are 4,614 pairs of images for positive and negative relations.

A. Experimental Setup

The number of the positive and negative pairs used in the experiments is the same for each relation on the four subsets. We use five-fold cross validation strategy for the evaluation. We report the mean accuracy over the five folds. The negative pairs and folds are predefined for the all four relations. For the deep and tensor features, we extracted VGG-Face, LPQ-tensor and BSIF-tensor as this has shown to perform better than shallow or vector-based methods [1], [9]. The tensor features are performed by MSIDA[1] method. For LPQ-tensor, the window size is $l = \{1, 2, 3\}$ for each face sample. For BSIF-tensor, we use eight filters with different sizes $w = \{3, 5, 7\}$ for each face sample. For ELM, the number of neurons in the hidden layer is an important parameter. The number is determined empirically and set to 10.

B. Results and Analysis

We run the experiments on the four relations of the FIW database. The results of these experiments are reported in Table I. The ROC curves comparing Deep-Tensor+ELM and VGG-Face are provided in Figure 2 for the four relations from FIW database. As can be noticed from the figure, the performance of Deep-Tensor+ELM is much better than that of VGG-Face in all cases.

Our proposed method is compared against some recent state-of-the-art methods in Table I. Note that some of these methods, such as ResNet+SDMLoss, use combination of different features to describe a face image. Some other methods are based on deep learning. On the four relations of FIW database, our approach yields in the best results for all the four kinship subsets. These results are promising and demonstrate that our proposed approach performs better than the recent methods for kinship verification.

IV. CONCLUSION

In this paper, we presented an effective approach based on deep and tensor features to the problem of kinship verification. To achieve a low dimensional and discriminative subspace, we proposed the Deep-Tensor+ELM method. The experimental evaluation shows the superiority of our method. The best results of our approach are obtained by fusing scores of three multi-view deep and tensor features learned with the proposed Deep-Tensor+ELM method. These results outperform the state of the art on four Grandparent-Grandchild relations from FIW database. Furthermore, these results point out to the need of using deep/tensor features for kinship verification. As future work, we plan to investigate the complementarity of more features description for face representation with the proposed method (Deep-Tensor+ELM).

REFERENCES

- [1] M. Bessaoudi, A. Ouamane, M. Belahcene, A. Chouchane, E. Boutellaa, and S. Bourenane. Multilinear side-information based discriminant analysis for face and kinship verification in the wild. *Neurocomputing*, 329:267 – 278, 2019.
- [2] R. Fang, K. D. Tang, N. Snavely, and T. Chen. Towards computational models of kinship verification. In *2010 IEEE International Conference on Image Processing*, pages 1577–1580, Sept 2010.
- [3] Q. He, X. Jin, C. Du, F. Zhuang, and Z. Shi. Clustering in extreme learning machine feature space. *Neurocomputing*, 128:88 – 95, 2014.
- [4] G. Huang, G.-B. Huang, S. Song, and K. You. Trends in extreme learning machines: A review. *Neural Networks*, 61:32 – 48, 2015.
- [5] O. Laiadi, A. Ouamane, E. Boutellaa, A. Benakcha, A. Taleb-Ahmed, and A. Hadid. Kinship verification from face images in discriminative subspaces of color components. *Multimedia Tools and Applications*, Dec 2018.

- [6] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song. Sphereface: Deep hypersphere embedding for face recognition. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul 2017.
- [7] J. Lu, J. Hu, V. E. Liong, X. Zhou, A. Bottino, I. U. Islam, T. F. Vieira, X. Qin, X. Tan, S. Chen, S. Mahpod, Y. Keller, L. Zheng, K. Idrissi, C. Garcia, S. Duffner, A. Baskurt, M. Castrillón-Santana, and J. Lorenzo-Navarro. The fg 2015 kinship verification in the wild evaluation. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, volume 1, pages 1–7, May 2015.
- [8] H. V. Nguyen and L. Bai. *Cosine Similarity Metric Learning for Face Verification*, pages 709–720. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [9] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, 2015.
- [10] J. P. Robinson, M. Shao, Y. Wu, H. Liu, T. Gillis, and Y. Fu. Visual kinship recognition of families in the wild. In *IEEE Transactions on pattern analysis and machine intelligence*, 2018.
- [11] J. P. Robinson, M. Shao, H. Zhao, Y. Wu, T. Gillis, and Y. Fu. Recognizing families in the wild (rfiw): Data challenge workshop in conjunction with acm mm 2017. In *Proceedings of the 2017 Workshop on Recognizing Families In the Wild, RFIW '17*, pages 5–12, New York, NY, USA, 2017. ACM.
- [12] S. Wang, Z. Ding, and Y. Fu. Cross-generation kinship verification with sparse discriminative metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2018.
- [13] H. Yan, J. Lu, W. Deng, and X. Zhou. Discriminative multimetric learning for kinship verification. *IEEE Transactions on Information Forensics and Security*, 9(7):1169–1178, July 2014.