

Kinship Verification using Color Features and Extreme Learning Machine

Xiaoting Wu^{1,2}, Xiaoyi Feng², Elhocine Boutellaa¹, Abdenour Hadid^{1,2*}

¹Center for Machine Vision and Signal Analysis

University of Oulu

Oulu, Finland

firstname.lastname@oulu.fi

²School of Electronics and Information

Northwestern Polytechnical University

Xi'an, China

fengxiao@nwpu.edu.cn

Abstract—Kinship verification from faces is a challenging task that is attracting an increasing attention in the recent years. The proposed methods so far are not robust enough to predict the kin between persons via facial appearance only. The initial studies using deep convolutional neural networks (CNN) have not shown their full potential as well, mainly due to limited training data. To mitigate this problem, we propose a new approach to kinship verification based on color features and extreme learning machines (ELM). While ELM aims to deal with small size training sets, color features are proven to provide significant enhancement over gray-scale counterparts. We evaluate our proposed method on three benchmark and publicly available kinship databases, namely KinFaceW-I, KinFaceW-II and TSKinFace. The obtained results compares favorably against some state-of-the-art methods including those based on deep learning.

Keywords—kinship verification; color features; extreme learning machine

I. INTRODUCTION

Automatic kinship verification from faces is a relatively young research topic within face analysis. It refers to the problem of verifying whether two persons share a kin relationship or not based only on the facial visual appearances. This is useful in many applications such as finding missing children, family photo organization, social media analysis, etc.

A key question in kinship verification from faces is that what are the facial parts that exhibit the kin relation the most. In other words, what are the most shared facial features between family members? This question has been studied from psychological perspectives (e.g. [1] [2]), suggesting that the eyes may bear more kin information than other facial parts.

These psychological studies on human kinship perception from facial appearances inspired several computer vision researchers to develop automatic kinship verification methods or even to compare the performance of humans versus machines on this task, claiming that some state-of-the-art automatic methods can outperform human kinship perception [3].

Among pioneer attempts on automatic kinship verification is the work of Fang *et al.* [4] in 2010. Since then,

many methods have been proposed in the literature with some relative success. Even the most recent studies using deep convolutional neural networks (CNN) have not shown their full potential due to the limited training data. This makes the topic of kinship verification from facial appearances an exciting and open research problem attracting more and more attention.

In this present work, we propose to tackle the kinship verification challenge by extracting color texture features and using Extreme Learning Machines (ELM) for classification. Our approach is motivated by two observations: (i) color texture features are proven to provide significant enhancement over gray-scale counterparts and (ii) ELM seems to deal better than deep neural architectures when facing small size training sets.

For fair comparison against previous works, we extensively evaluate our proposed method on three benchmark and publicly available kinship databases, namely KinFaceW-I [5], KinFaceW-II [5], and TSKinFace [6]. The obtained results are quite encouraging and compare favorably against some state-of-the-art methods including those based on deep neural architectures.

II. RELATED WORK

As mentioned above, among the first attempts for automatic kinship verification from facial images is the one described in the work of Fang *et al.* [4]. Both local and global features were considered and they used forward selection methodology to find the most discriminative inherited facial features. A classification accuracy of 70.67% was reported on a small test set of image pairs using K-Nearest-Neighbors. Since then, several attempts have been made to tackle the problem of kinship verification from facial images with some relative successes.

Among these contributions, the use of metric learning has resulted in very good performance as can be seen in the popular work of Lu *et al.* [5], which was later extended by Yan *et al.* [7]. Lu *et al.* [5] proposed neighborhood repulsed metric learning (NRML) for kinship verification which aims to learn a distance metric that pulls image pairs without kin relations as far as possible and minimize the distance between image pairs with kin relations. Inspired by this approach, Yan *et al.* [7] proposed discriminative multi-

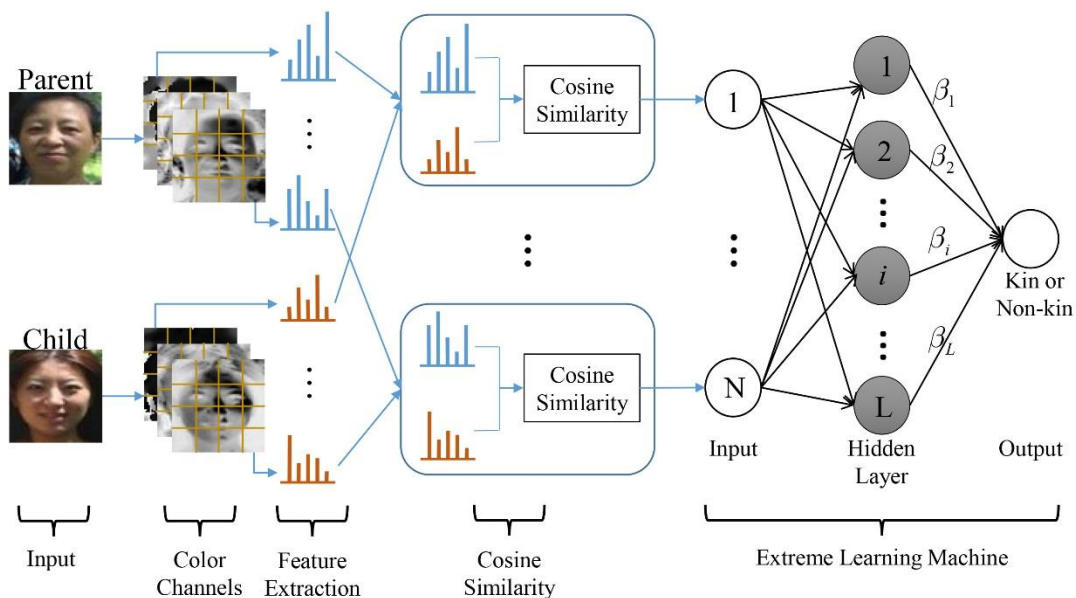


Figure 1. Overview of the proposed approach for kinship verification.

metric learning (DMML) method by involving multiple different features.

More recently, inspired by the impressive success of deep learning in computer vision, many researchers applied deep learning to kinship verification. For instance, Zhang *et al.* [8] proposed a deep convolutional neural network (CNN) architecture for kinship verification from face images. The input of their architecture is the pair of RGB images. Li *et al.* [9] trained a Siamese CNN structure to extract features and employed a similarity metric to measure the kin resemblance between two face images. Boutellaa *et al.* [10] exploited face videos and combined both deep features and handcrafted spatio-temporal features. It is worth nothing that deep learning based approaches require large amounts of training data and hence might not be applicable in every situation.

We propose a new kinship verification framework extracting texture features from color face images and using extreme learning machines to classify the similarities between the pairs of face features as kin related or not. Our approach is motivated by the proven usefulness of color texture features for kinship verification in one hand [11], and by the ability of extreme learning machines (in contrast to deep learning approaches) to deal with small training data sets.

III. PROPOSED METHOD

Fig. 1 depicts an overview of our proposed approach. The input is a pair of two color face images e.g. a parent and a child. We convert these images into different color spaces and encode the facial texture in each channel. We compute the cosine similarities between the texture features in each color channel. 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 color texture features

Most of the proposed methods in the literature for automatic kinship verification have mainly focused on analyzing only gray-scale face images, hence discarding color information. More recently, color texture features have shown better performance than gray-scale counterparts for kinship verification (e.g. [11]). When considering the color information, the problem usually consists in learning a discriminating color space where the classification (e.g. kinship verification in our case) becomes more affordable compared to the gray-scale space.

As suggested in [11], we consider the HSV color space. HSV has uncorrelated information in every channel. This is in contrast to RGB color space in which its R, G and B channels may contain redundant color information. Therefore, we first convert the RGB images into the HSV color space. Then, we extract the binarized statistical image features (BSIF) [12] from each channel separately. Each image is divided into several blocks to encode the facial structure. The features from each block are concatenated into an enhanced histogram.

B. Computing cosine distances

The color texture features extracted from each pair of two images should be combined into a single value. Our experiments exploring different combination strategies (e.g. concatenation, differences, etc.) and different distance measures (Euclidean, cosine, Manhattan, Hamming, etc.) showed that cosine similarity metric [13] provides the best performance. The cosine similarity between two feature vectors X and Y is given by:

$$\text{sim}(X, Y) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \cdot \|\vec{y}\|} \quad (1)$$

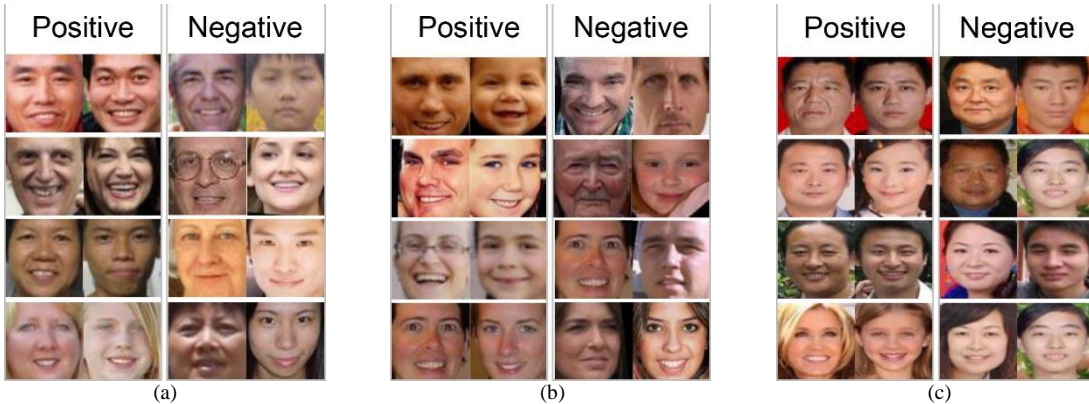


Figure 2. Sample images from (a) KinFaceW-I, (b) KinFaceW-II and (c) TSKinFace kinship databases. From top to bottom of the positive samples, rows correspond to father-son (FS), father-daughter (FD), mother-son (MS) and mother-daughter (MD) relations.

C. Classification using extreme learning machines

An Extreme Learning Machine (ELM) [14] is a single hidden layer network which has been shown to perform better and faster than SVM in some classification problems [15]. The output of an ELM network with L hidden neurons can be represented 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. $X_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T$ is the input vector with the ground truth $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$. 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,

$$H(W_1, \dots, W_L, b_1, \dots, b_L, X_1, \dots, X_L) = \begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \dots & g(W_L \cdot X_1 + b_L) \\ \vdots & \dots & \vdots \\ g(W_1 \cdot X_N + b_1) & \dots & g(W_L \cdot X_N + b_L) \end{bmatrix}_{N \times L} \quad (4)$$

and T is the target output:

$$T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}_{N \times m} \quad (5)$$

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, $\hat{\beta}$ can be calculated as:

$$\hat{\beta} = H^{-1} \cdot T \quad (6)$$

IV. EXPERIMENTAL ANALYSIS

A. Experimental data

For experimental evaluation, we considered two commonly used kinship databases: KinFaceW-I, KinFaceW-II [5] as well as the TSKinFace [6] database. These three databases are composed by four types of kin relations namely father-son (FS), father-daughter (FD), mother-son (MS) and mother-daughter (MD) relations. The face images in these three databases are normalized into 64×64 pixels and are aligned according to the location of the eyes. In KinFaceW-I, there are 156, 134, 116 and 127 pairs of images for FS, FD, MS and MD relations, respectively. In KinFaceW-II, there are 250 pairs of images for each kin relation. TSKinFace database is composed by tri-subjects of face images. As TSKinFace was originally formed of triple relations (two parents and one child), we convert each triplet into two pairs: father-child and mother-child. The database contains 513 Father-Mother-Daughter (FM-D) triplets and 502 Father-Mother-Son (FM-S) triplets. Fig. 2 depicts some samples of positive and negative kin face pairs from each database.

B. Experimental setup

The number of the positive and negative pairs used in the experiments is the same for each relation on the three databases. 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 KinFaceW-I and KinFaceW-II. In the case of TSKinFace database, we randomly generate the negative pairs and folds. For the color texture features, we extracted color-BSIF as

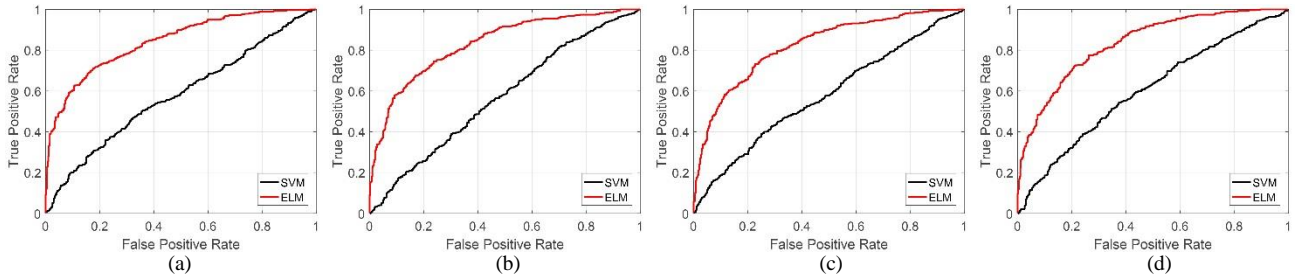


Figure 3. ROC curves of ELM and SVM on TSKinFace database for (a) FS, (b) FD, (c) MS and (d) MD relations

TABLE I. KINSHIP VERIFICATION ACCURACIES (IN %) OF THE ELM AND SVM ON KINFACEW-I, KINFACEW-II AND TSKINFACE DATABASES

Database	Method	FS	FD	MS	MD
KinFaceW-I	SVM	63.2	59.6	57.0	62.4
	ELM	70.0	64.2	73.0	77.2
KinFaceW-II	SVM	59.8	61.2	60.6	56.6
	ELM	78.6	73.6	81.0	79.6
TSKinFace	SVM	57.5	55.0	53.9	52.1
	ELM	78.5	76.0	76.5	76.2

TABLE II. KINSHIP VERIFICATION ACCURACIES (IN %) OF THE PROPOSED APPROACH AND THE STATE-OF-THE-ART METHODS ON KINFACEW-I AND KINFACEW-II DATABASES.

Method	KinFaceW-I				KinFaceW-II			
	FS	FD	MS	MD	FS	FD	MS	MD
PDFL[16]	73.5	67.5	66.1	73.1	77.3	74.7	77.8	78.0
DMML[7]	74.5	69.5	69.5	75.5	78.5	76.5	78.5	79.5
NRML[5]	72.5	66.5	66.2	72.0	76.9	74.3	77.4	77.6
MultiviewSSL[3]	82.8	75.4	72.6	81.3	81.8	74.0	75.3	72.5
SSML[17]	81.7	75.3	71.4	77.9	82.4	78.6	79.8	77.9
SPML-P[18]	81.1	75.7	73.2	75.7	82.4	77.6	76.6	76.2
Proposed	70.0	64.2	73.0	77.2	78.6	73.6	81.0	79.6

TABLE III. KINSHIP VERIFICATION ACCURACIES (IN %) OF THE PROPOSED APPROACH AND THE STATE-OF-THE-ART METHODS ON TSKINFACE DATABASE.

Method	FS	FD	MS	MD
NRML[5]	74.8	70.0	72.2	71.3
Sparse Group Lasso[19]	69.1	66.8	68.7	67.9
ITML[20]	75.6	70.5	72.1	70.7
LMNN[21]	72.7	69.8	71.5	70.1
Proposed	78.5	76.0	76.5	76.2

this has shown to perform better than color-LBP and color-LPQ [11]. The dimensionality of each face block feature is reduced using PCA before computing the cosine similarities. For ELM, the number of neurons in the hidden layer is an important parameter. The number is determined empirically and set to 40.

C. Results and analysis

In the first set of experiments, we compared the performance of our proposed approach using extreme learning for classification against SVM. SVM is one of the most used classification approaches for kinship verification in the previous works [1] [10]. We run the experiments on the four relations of the three databases. The results of these experiments are reported in Table I. The ROC curves comparing ELM and SVM are provided in Fig. 3 for the

TSKinFace database. As can be noticed from the table and the figure, the performance of ELM is much better than that of SVM in all cases. The results also show that the margin between the SVM and ELM accuracies is significant in all cases (10-20%).

Our proposed method is compared against some recent state-of-the-art methods in Table II on KinFaceW-I and KinFaceW-II databases and in Table III on TSKinFace database. Note that some of these methods, such as MultiviewSSL, use combination of different features to describe a face image. Some other methods are based on deep learning. On the KinFaceW-I database, our method gives the best performance on MS subset. For KinFaceW-II, our approach gives the best results for two subsets: MS and MD. On the larger TSKinFace database, our approach yields in the best results for all the four kinship subsets. These results are promising and demonstrate that our proposed approach is competitive compared to recent methods for kinship verification.

V. CONCLUSION

We proposed a new approach for kinship verification from face images using color texture features and extreme learning machines. The proposed approach is shown to perform well with limited amount of training data sets. The obtained results are comparable with those of recent state of the art methods.

This work is by no mean complete. In future, we plan to (i) gain insights into why the method is not working well in some kin relations, (ii) evaluate the proposed method on other classification tasks with limited amount of training data, (iii) extend the proposed framework to handle face video sequences instead of still images, and (iv) collect a large kinship database and make it available for the research community.

ACKNOWLEDGMENT

This work was partially funded by China Scholarship Council and Academy of Finland.

REFERENCES

- [1] M. F. Dal Martello and L. T. Maloney, "Where are kin recognition signals in the human face?," *Journal of Vision*, vol. 6, no. 12, pp. 2–2, 2006.
- [2] M. F. Dal Martello and L. T. Maloney, "Lateralization of kin recognition signals in the human face," *Journal of vision*, vol. 10, no. 8, pp. 9–9, 2010.
- [3] X. Zhou, H. Yan, and Y. Shang, "Kinship verification from facial images by scalable similarity fusion," *Neurocomputing*, vol. 197, pp. 136–142, 2016.
- [4] R. Fang, K. D. Tang, N. Snavely, and T. Chen, "Towards computational models of kinship verification," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 1577–1580, IEEE, 2010.
- [5] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou, "Neighborhood repulsed metric learning for kinship verification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 2, pp. 331–345, 2014.
- [6] X. Qin, X. Tan, and S. Chen, "Tri-subject kinship verification: Understanding the core of a family," *IEEE Transactions on Multimedia*, vol. 17, no. 10, pp. 1855–1867, 2015.
- [7] H. Yan, J. Lu, W. Deng, and X. Zhou, "Discriminative multimetric learning for kinship verification," *IEEE Transactions on Information forensics and security*, vol. 9, no. 7, pp. 1169–1178, 2014.
- [8] K. Zhang, Y. Huang, C. Song, H. Wu, and L. Wang, "Kinship verification with deep convolutional neural networks," in *British Machine Vision Conference*, pp. 148.1–148.12, 2015.
- [9] L. Li, X. Feng, X. Wu, Z. Xia, and A. Hadid, *Kinship Verification from Faces via Similarity Metric Based Convolutional Neural Network*. Springer International Publishing, 2016.
- [10] E. Boutellaa, M. B. López, S. Ait-Aoudia, X. Feng, and A. Hadid, "Kinship verification from videos using spatiotemporal texture features and deep learning," in *International Conference on Biometrics*, pp. 1–7, 2016.
- [11] X. Wu, E. Boutellaa, M. B. López, X. Feng, and A. Hadid, "On the usefulness of color for kinship verification from face images," in *Information Forensics and Security (WIFS), 2016 IEEE International Workshop on*, pp. 1–6, IEEE, 2016.
- [12] J. Kannala and E. Rahtu, "BSIF: Binarized statistical image features," in *Pattern Recognition (ICPR), 2012 21st IEEE International Conference on*, pp. 1363–1366, IEEE, 2012.
- [13] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Asian Conference on Computer Vision*, pp. 709–720, Springer, 2010.
- [14] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Networks*, vol. 61, pp. 32–48, 2015.
- [15] Q. He, X. Jin, C. Du, F. Zhuang, and Z. Shi, "Clustering in extreme learning machine feature space," *Neurocomputing*, vol. 128, pp. 88–95, 2014.
- [16] H. Yan, J. Lu, and X. Zhou, "Prototype-based discriminative feature learning for kinship verification," *IEEE Transactions on cybernetics*, vol. 45, no. 11, pp. 2535–2545, 2015.
- [17] Y. Fang, Y. Y. S. Chen, H. Wang, and C. Shu, "Sparse similarity metric learning for kinship verification," in *Visual Communications and Image Processing (VCIP), 2016*, pp. 1–4, IEEE, 2016.
- [18] H. Liu and C. Zhu, "Status-aware projection metric learning for kinship verification," in *Multimedia and Expo (ICME), 2017 IEEE International Conference on*, pp. 319–324, IEEE, 2017.
- [19] R. Fang, A. C. Gallagher, T. Chen, and A. Loui, "Kinship classification by modeling facial feature heredity," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 2983–2987, IEEE, 2013.
- [20] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic metric learning," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 209–216.
- [21] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *Journal of Machine Learning Research*, vol. 10, no. Feb, pp. 207–244, 2009.