

Investigating Deep CNNs Models Applied in Kinship Verification through Facial Images

Abdelhakim Chergui

Electronics Research Laboratory of Skikda (LRES)
Skikda University, Algeria
e-mail: hakim-kom@hotmail.com

Salim Ouchtati

Electronics Research Laboratory of Skikda (LRES)
Skikda University, Algeria
e-mail: ouchtatisalim@yahoo.fr

Sebastien Mavromatis

LIS Laboratory (UMR CNRS 7020)
Aix Marseille University, France
e-mail: sebastien.mavromatis@univ-amu.fr

Salah Eddine Bekhouche

Center for Machine Vision and Signal Analysis
University of Oulu, Finland
e-mail: bekhouchesalah@gmail.com

Jean Sequeira

LIS Laboratory (UMR CNRS 7020)
Aix Marseille University, France
e-mail: jean.sequeira@univ-amu.fr

Abstract—The kinship verification through facial images is an active research topic due to its potential applications. In this paper, we propose an approach which takes two images as input then give kinship result (kinship / No-kinship) as an output. our approach based on the deep learning model (ResNet) for the feature extraction step, alongside with our proposed pair feature representation function and RankFeatures (Ttest) for feature selection to reduce the number of features finally we use the SVM classifier for the decision of kinship verification. The approach contains three steps which are : (1) face preprocessing, (2) deep features extraction and pair features representation (3) Classification . Experiments are conducted on five public databases . The experimental results show that our approach is comparable with existed approaches.

Keywords—Kinship verification; Deep learning; CNNs; ResNet; Rank Features; Ttest; SVM

I. INTRODUCTION

The kinship verification among people mainly concentrated on studying the similarity between human faces. Indeed, there are several types of kinship relationships: father-daughter (F-D), mother-son (M-S), father-son (F-S) and mother-daughter (M-D). The recognition of kinship is a very challenging topic since the face conveys different facial features such as identity, age, gender, and expression. Nowadays, the recognition of these familial relationships has much application such as organizing a photo album, annotating images as well as identifying lost or wanted people. Furthermore, the determination of the kinship is no longer limited exclusively to genetic analysis, but it is now extending to the field of biometrics.

There are many studies have been conducted on kinship verification from facial images which can be categorized based on the type of feature extraction and the similarity algorithms. Fang *et al.* [1] proposed a kinship verification system based on PSM (Pictorial structure model) feature extraction and selection methods and they used KNN for the classification phase, they obtained a promising result on the Cornell KinFace database, and Xia *et al.* [2] used a (TL) transfer learning method with a application on UB KinFace new database.

Another interesting work was proposed by Lu *et al.* [3] where they proposed a neighborhood repulsed metric learning (NRML) method for kinship verification. In addition, they proposed a multi view NRML (MNRML) method to seek a common metric distance in order to better use of the multiple descriptor features, they applied their method on The two datasets which are KinFaceW-I and W-II, Yan *et al.* [4] proposed a discriminative multi metric learning method for kinship verification, they applied their method on two databases: Cornell KinFace and UB Kin database.

Yan *et al.* [5] proposed a new prototype-based discriminative feature learning (PDFL) method for kinship verification, this method aims to learn discriminative mid-level features , they applied their method on both Cornell KinFace and UB Kin databases. Wang *et al.* [6] used the KinFaceW-I and KinFaceW-II datasets they proposed a deep kinship verification (DKV) model by integrating excellent deep learning architecture into metric learning. They employed a deep learning model, Zhou *et al.* [7] proposed an of ensemble similarity learning (ESL), they applied their method on the KinFaceW-I and KinFaceW-II datasets.

Chergui *et al.* [8] proposed approach based on the LBP and BSIF descriptors and the PML features they applied their method on both Cornell KinFace, UB KinFace , KinFace W-I and W-II databases and in other work Chergui *et al.* [9] proposed a Discriminant Analysis for Facial Verification using Color Images they applied their method on three databases Cornell, UB KinFace and family 101. and proposed LTP descriptor with ML face representation and fisher Score selection for kinship verification [10], and they proposed another approach based on the deep features of VGG-FACE descriptor, They applied their approach on five databases (Cornell, UB KinFace, Family 101, Kinface W-I and W-II)[11] .

The remainder of this paper is organized as follows. Section II provides steps of the proposed approach. Section III reviews the used databases and illustrated the experiments and obtained results . Finally, we conclude our work in section IV.

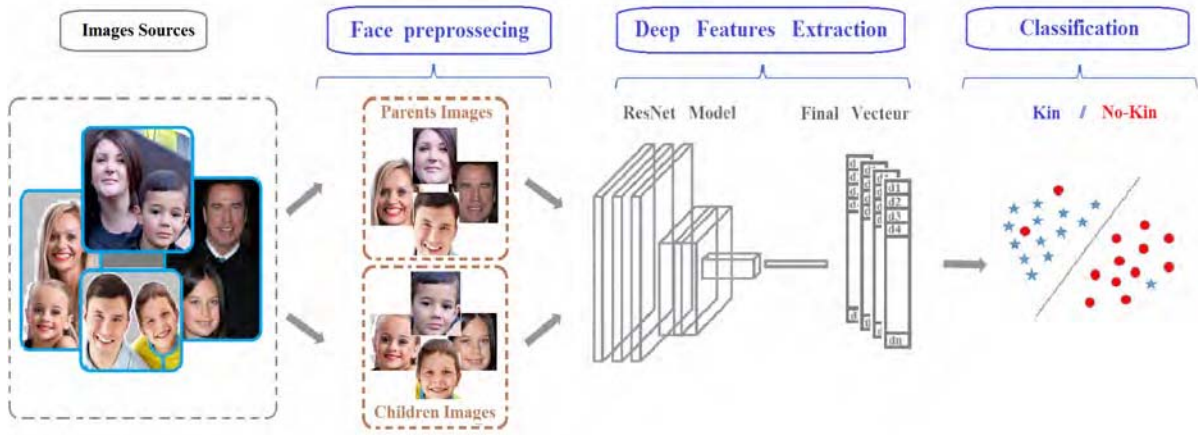


Figure 1. General structure of the proposed Kinship Verification approach.

II. KINSHIP VERIFICATION APPROACH

Our proposed kinship verification approach consists of three stages which are : (1) face preprocessing, (2) deep features extraction and features representation, (3) features selection and Classification. Figure 1 illustrates the general structure of the proposed framework.

A. Face Preprocessing Step

In the face preprocessing stage, we located the facial landmarks using Ensemble of Regression Trees (ERT) algorithm [12]. From these landmarks, we use the coordinates of the center of the two eyes to align the face pose using the same params as in [13], [14]. Figure 2 presents the steps of this phase.

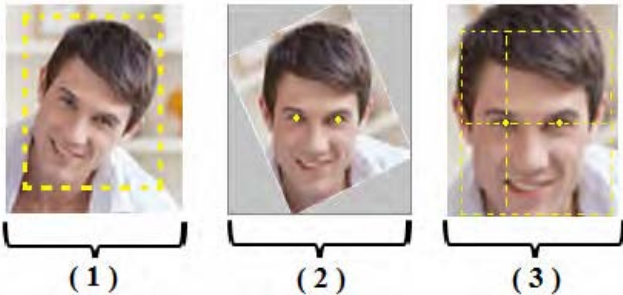


Figure 2. Face preprocessing based on the eyes centers.

B. Deep Features Extraction

1) *ResNet Descriptor*: In recent years, neural networks and deep learning have provided the best solutions to many problems in image recognition, speech recognition, and natural language processing. Deep learning is based on some ideas of how learning happens or could happen in the brain. In fact, the original of deep learning is the neural networks that backs to the 1940s, however, the huge popularity of deep learning started around 2006 ([15], [16]). One of the most popular classes of deep learning is the convolutional neural network (CNN), feed-forward artificial neural networks were the first

and simplest type of neural network that has successfully applied in computer vision tasks. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing [17].

Figure 3 illustrates the general architecture of VGG-Face descriptor.

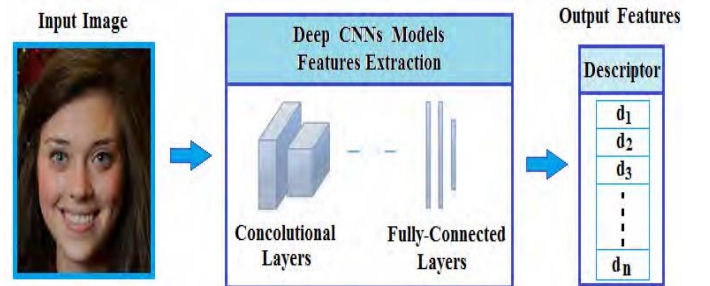


Figure 3. CNN Model descriptor.

The CNN structure consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers. There are many popular CNN architectures that have achieved a high performance on many machine learning tasks. Among these architectures we use ResNet on our work as face descriptor [18]. The deep features are extracted from the layer f1000 (fully connected layer) of this architecture, and the number of these features is 1000. Table I presents the architectures for ResNet, from different models of ResNet : ResNet-18, Resnet-50 and ResNet-101 (a pretrained convolutional neural network). and the building blocks are shown in brackets, with the numbers of blocks stacked. Down sampling is performed by conv3, conv4, and conv5 with a stride of 2. the principle is the same for the ResNet models, the difference in the number of convolution layers using 3x3 filters. The layers 1000 was chosen due to the fact that deeper layers contain higher-level features compared with earlier layers.

TABLE I: ResNet CNN general architecture.

Layer name	Output size	Res-Net 18-layers	Res-Net 50-layers	Res-Net 101-layers	
Conv 1	112 x 112 x 64	7 x 7, 64 stride 2			
Conv 2-x	56 x 56 x 64	3 x 3 max pool, stride 2			
		3 x 3, 64	x 2	1 x 1, 64 3 x 3, 64 1 x 1, 256	x 3
Conv 3-x	28 x 28 x 128	3 x 3, 128	x 2	1 x 1, 128 3 x 3, 128 1 x 1, 512	x 4
		3 x 3, 128			
Conv 4-x	14 x 14 x 256	3 x 3, 256	x 2	1 x 1, 256 3 x 3, 256 1 x 1, 1024	x 6
		3 x 3, 256			
Conv 5-x	7 x 7 x 512	3 x 3, 512	x 2	1 x 1, 512 3 x 3, 512 1 x 1, 2048	x 3
		3 x 3, 512			
Average pool	1 x 1 x 512	7 x 7 average pool			
Fully Connected	1000	512 x 1000 Fully Connectedl			
Softmax	1000	FC 1000			

2) *Pair features representation and normalization*: After extracting the features, we normalize the features of each pair (child / parent) using the Min-Max scaling method. The formula of Min-Max scaling is given below:

$$F_{norm} = \frac{F - F_{min}}{F_{max} - F_{min}} \quad (1)$$

Then this two feature vectors (child / parent) are presented as one feature vector using this proposed formula:

$$F = \frac{F_{child} + F_{parent}}{|F_{child} - F_{parent}|} \quad (2)$$

where F , F_{child} , F_{parent} are the new feature vector, the feature vector of the child and the feature vector of the parent respectively.

C. Classification

1) *Features selection*: For the feature selection phase, we used a feature selection scheme given by TTest. Its weights [19] are based on the absolute value two-sample TTest with pooled variance estimate. The weight is given by:

$$W_{ttest}(i) = \left| \frac{(m_k - m_n)}{\sqrt{\frac{\sigma_k^2}{N_k} + \frac{\sigma_n^2}{N_n}}} \right| \quad (3)$$

where $W_{ttest}(i)$ is the score of feature F_i , \bar{m} is the feature mean, N_i is the number of samples in the kinship class i ($1 \rightarrow \text{kin} / 2 \rightarrow \text{non-kin}$), m_i and σ_i^2 are the mean and the variance of the kinship class in the intended feature.

2) *decision*: Due to the fact that the kinship verification is just two classes problem, so we used binary Support Vector Machine as classifier. The SVM model aims to represent the training features as points in space, then seeks to separate the points of the different classes by a clear gap which must be wide as much as possible, with an RBF kernel. In addition, we used grid search strategy alongside with cross-validation to have the best classifier parameters.

III. EXPERIMENTS AND DISCUSSION

In this section, we present the details of the used five publicly databases (Cornell KinFace, UB Kindatabase, Family 101, KinFace-I, and KinFace-II), then we discuss the results obtained.

A. Databases

In this section, we describe briefly the used databases. Figure 4 shows some images from these databases.

1) *Cornell KinFace (2010)*: The Cornell KinFace database is the first database that have been used for the kinship verification, it was created by Fang *et al.* [1], from the University of Cornell. They collected 150 pairs (300 images). The distribution of the kinship pairs are as follows: 40% (F-S), 22% (F-D), 13% (M-S) and 26% (M-D).

2) *UB KinFace (2012)*: The UB KinFace database was created by Shao *et al.* [20], and it was collected from 400 persons and that produced 600 static images. it's the first database that designed for kinship verification with different ages of children and their parents. This database composed of this kinship relations:(F-S), (F-D), (M-S),(M-D).

3) *Family 101 (2013)*: proposed by Fang *et al.* [21], which containing 14,816 images, with 101 different family trees, 607 individuals, This database includes renowned public families. It has four kinship relations as follow: 213 (F-S), 147 (F-D), 184 (M-S), and 148 (M-D).

4) *KinFaceW-I and KinFaceW-II (2014)*: Lu *et al.* [3] provided the researchers with a databases called KinFaceW-I and KinFaceW-II. The images of KinFaceW-I database were collected from the web and it was captured under uncontrolled environments in terms of gesture, demographic attributes, lighting, background, expressions, and partial occlusion, This daatabase contains 4 kin relations, 156 (F-S), 134 (F-D), 116 (M-S), and 127 (M-D) of kinship pairs.

The images of KinFaceW-II database include some celebrities face images as well as their childrens or parents, this images were collected from the Internet, the KinFaceW-II dataset contains 4 kin relations (F-S), (F-D), (M-S) and (M-D), 250 pairs of kinship images for each kin relation.

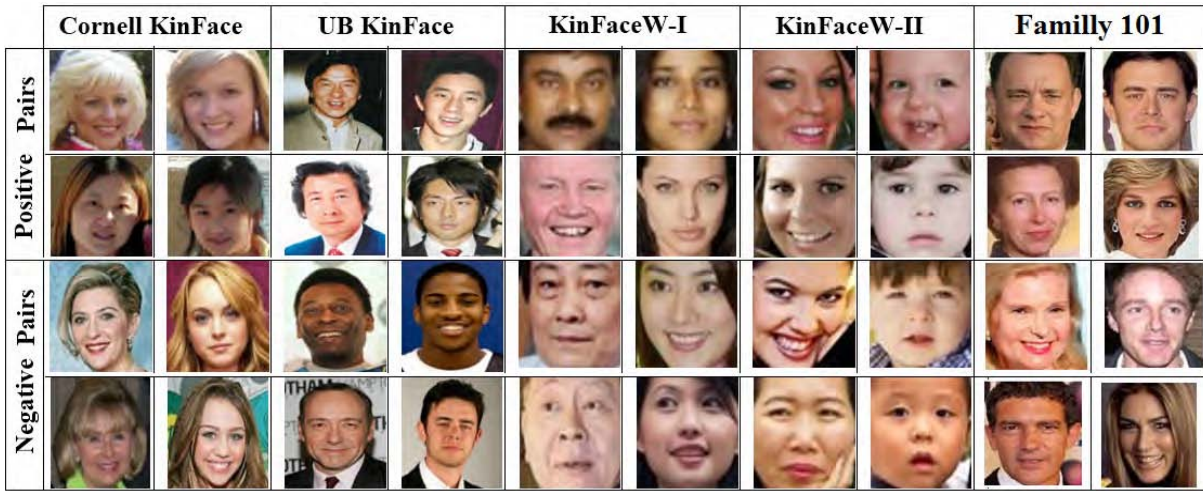


Figure 4. Example of positive and negative pairs from used databases.

B. Experimental Results and Discussion

in this section we conducted a lot of experiments by using a deep learning models (ResNet-18, Resnet-50 and ResNet-101) . We conducted experiments on the five databases that are publicly available (Cornell, UB KinFace, Family101, KinFaceW-I and KinFace-II) to prove the performance of our approach, In what follows we detail the experimental results.

1) *Effect of features extraction models* : In our study, face features are provided by the pre-trained deep convolutional neural networks ResNet, we tested our approach first by using only one model of ResNet, Then we fusion the Models between them by two Models and finally all Models.

From Table II, we can see that the use only one model (ResNet -18, ResNet-50 or ResNet-101) gives results less than the use ResNet-(18+50), ResNet-(18+101) or ResNet-(50+101). our approach gives better results when using ResNet-(18+50 +101).

TABLE II: Performance of our approach with ResNet-18, ResNet-50 and ResNet-101 descriptors and the mixed between them

	Cornell Kin	UB KinFace	Family 101	KinFace W-I	KinFace W-II
ResNet-18	72.45	68.74	67.14	65.41	64.46
ResNet-50	74.74	69.24	68.57	66.23	65.74
Resnet-101	76.86	71.83	69.93	68.32	67.25
ResNet-(18+50)	78.71	72.37	70.54	69.63	70.02
ResNet-(18+101)	80.32	76.68	75.42	73.24	72.45
ResNet-(50+101)	83.35	79.36	78.62	76.86	75.47
ResNet-(18+50+101)	87.16	83.68	82.07	79.76	76.89

According to our experiments, we observe that the ResNet (18+50+101) improves the results which makes the system more efficient.

2) *Number of features used and CPU time* : In this experiment, we sorted the features according to their weight

using Rank Features (Ttest), then we tested different features dimensions (10, 50, 100,1000) for ResNet-18, 50 and 101, and (10, 503000) for ResNet-(18+50+101) to see the effects of it. Table III shows this experiment of five databases. .

TABLE III: Accuracy of our approach as a function of different features dimensions.

Features	ResNet-18	ResNet-50	Resnet-101	Resnet-(18+50+101)
10	65.36%	66.51%	67.45%	72.41%
50	67.47%	67.28%	68.25%	75.36%
100	68.63%	68.47%	70.12%	76.02%
850	69.36%	70.04%	71.23%	86.58%
1000	72.45%	74.74%	76.86%	87.16%
2000	-	-	-	87.28%
3000	-	-	-	87.34%

Figure 5 shows both the accuracy of the kinship verification (Cornell KinFace database results) and the CPU time that is needed for training the model as a function of the selected features number which obtained after using Ttest.

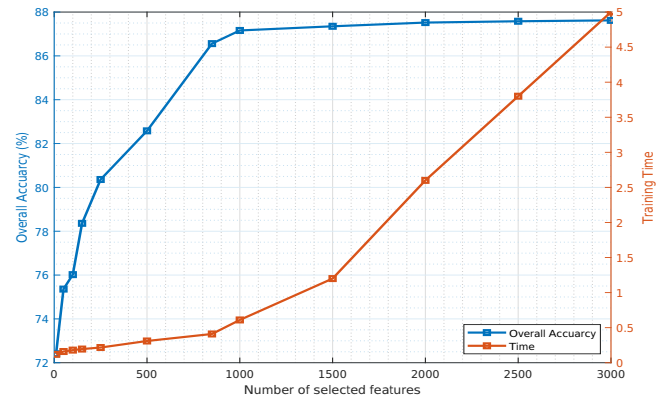
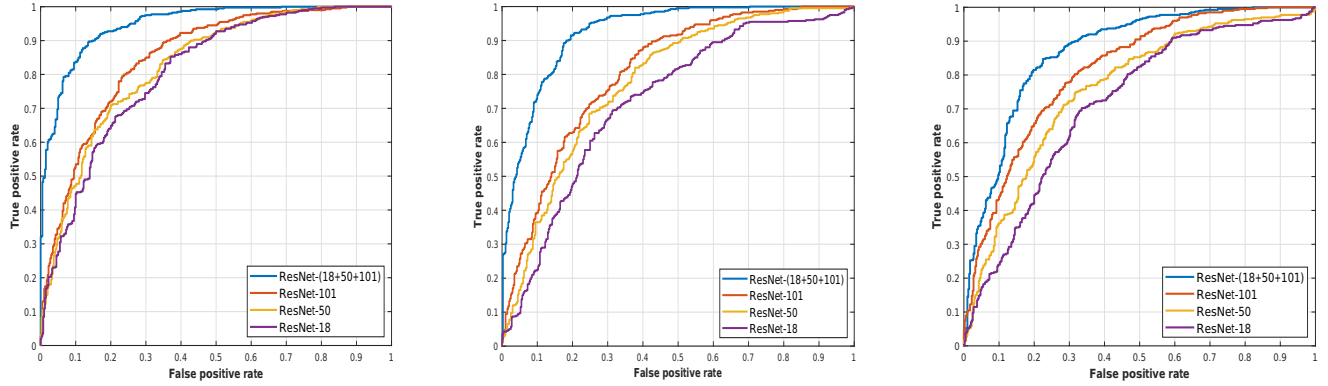


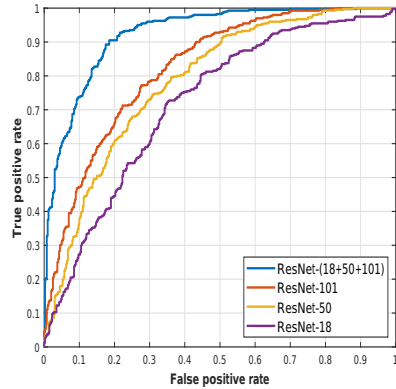
Figure 5. Accuracy of the kinship verification and the CPU time as a function of different features number (Cornell database).



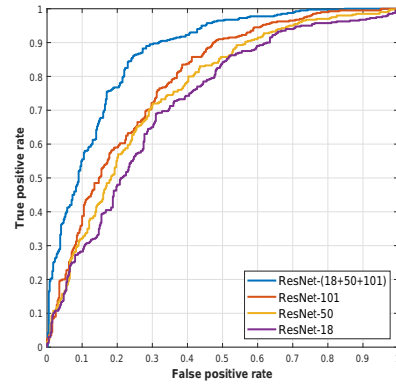
(a) ROC of Cornell Kinface databas

(b) ROC of UB KinFace database

(c) ROC of Family 101 Database



(d) Roc of KinFace W-I Database



(e) ROC KinFace W-II Database

Figure 6. ROC curves of different ResNet Models for five databases.

Through Figure 5 and the table III, the time consuming of the proposed approach increasing linearly with the increase in the number of features. On the other hand, the accuracy curve changes as an exponential increase until we reach 1000 features then it varies as a logarithm function. Based on that, the choice of 1000 features using ResNet-(18+50+10) can be considered as a good trade-off between accuracy and computational complexity.

3) *Relationship accuracy*: We obtained the final results using the ResNet-(18+50+101) for features extraction with the proposed pair feature representation function and Rank-Features for feature selection approach. Table IV shows the performance of the proposed approach with different relationships (F-S), (F-D), (M-S) and (M-D).

TABLE IV: The accuracy (%) on the four relationship of the used database

ResNet-(18+50+101)	F_S	F_D	M_S	M_D	Mean
Cornell KinFace	88.12	86.47	85.74	88.31	87.16
UB KinFace	85.02	83.53	82.38	83.79	83.68
Family 101	82.98	81.76	80.30	83.24	82.07
KinFace W-I	81.11	79.25	78.03	80.65	79.76
KinFace W-II	77.69	76.53	76.21	77.13	76.89

We can note: the easiest relationship to classify is the relationship that the same gender, whether male or female, like (F-S) is easier than (F-D) and (M-D) is easier than (M-S), and hardest relationship to classify is the relationship that is not the same gender.

4) *The receiver operating characteristic* : For the good confirmation of the best results and the performance of our approach, the receiver operating characteristic (ROC) curves of the different ResNet models are shown in Fig.6. This figure illustrates the ROC curves of the classification results on the use five databases.

We can observe the robustness of the CNNs Models using ResNet-18, ResNet-50 and ResNet-101 for deep features extraction and the best result using ResNet-(18+50+101) .

5) *A comparison with other approaches*: Table V illustrates the performance of our proposed kinship verification approach as well as of that of some competing approaches. From this Table, we can see that we got 87.16 %, 83.68%, 82.07%, 79.76% and 76.89 % kinship verification accuracy on Cornell KinFace, UB Kin, Family 101, KinFace-I and KinFace-II respectively. This confirms the robustness of our approach on all the databases which has been used, due to the superiority of accuracy in comparison with the rest of the approaches and CPU time spent in processing.

TABLE V: A comparison of the proposed approach with other kinship verification approaches

Year	Approaches	Databases				
		Cornell KinFace	UB KinFace	Family 101	KinFace W-I	KinFace W-II
2010	PSM[1]	70.67 %	-	-	-	-
2011	TSL[20]	-	69.67 %	-	-	-
2014	PDFL [5]	71.90 %	67.30 %	-	-	-
	MNRML [3]	-	-	-	69.90 %	76.5 %
2015	RSBM [22]	-	-	69.60 %	-	-
	DKV [6]	-	-	-	66.90 %	69.50 %
2016	ESL [7]	-	-	-	74.10 %	74.30 %
	ESL[7]	-	-	-	74.10	74.30
2017	NRCML [23]	-	-	-	65.80 %	65.80 %
2018	LPQ-PML [24]	82.86 %	73.25 %	-	77.20 %	75.98 %
	MHSL[25]	68.40 %	56.20 %	-	75.20 %	75.40 %
2019	KML [26]	81.40 %	75.50 %	-	75.57 %	85.70 %
	Proposed	87.16 %	83.68 %	82.07 %	79.76 %	76.89 %

IV. CONCLUSION

In this paper, we described a novel approach for kinship verification based on the deep learning models (ResNet) for deep features extraction . The experimental results showed that our approach provides better performance than previous approaches. As a future work, we propose to use other deep features models such as AlexNet, ImageNet . Moreover, we envision to fine-tune or train some CNN architectures from scratch. Also, we envision to use another soft biometrics traits such as age and gender to improve the kinship verification accuracy.

REFERENCES

- [1] 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*. IEEE, 2010, pp. 1577–1580.
- [2] S. Xia, M. Shao, and Y. Fu, "Kinship verification through transfer learning," in *IJCAI*, 2011, pp. 2539–2544.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] M. Wang, Z. Li, X. Shu, J. Tang *et al.*, "Deep kinship verification," in *Multimedia Signal Processing (MMSP), 2015 IEEE 17th International Workshop on*. IEEE, 2015, pp. 1–6.
- [7] X. Zhou, Y. Shang, H. Yan, and G. Guo, "Ensemble similarity learning for kinship verification from facial images in the wild," *Information Fusion*, vol. 32, pp. 40–48, 2016.
- [8] A. Chergui, S. Ouchtati, J. Sequeira, S. Bekhouche, and F. Bougourzi, "Kinship verification using bsif and lbp," in *The fourth edition of the international Conference on Signal, Image, Vision and their Applications , SIVA18*, 2018.
- [9] A. Chergui, S. Ouchtati, J. Sequeira, S. Bekhouche, B. F., and A. Benlamoudi, "Discriminant analysis for facial verification using color images," in *Proceeding of the first International Conference on Electrical Engineering ICEEB18*, Dec 2018.
- [10] A. Chergui, S. Ouchtati, J. Sequeira, S. Bekhouche, and H. Telli., "Robust kinship verification using local descriptors," in *Proceeding of Third International Conference On Advanced Technologies And Electrical Engineering ICTAEE18*, Dec 2018.
- [11] A. Chergui, S. Ouchtati, J. Sequeira, S. Bekhouche, F. Bougourzi, and H. Telli., "Deep features for kinship verification from facial images," in *Proceeding of Third International Conference on Advanced Systems and Emergent Technologies IC-ASET 19*, Mars 2019.
- [12] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1867–1874.
- [13] S. E. Bekhouche, A. Ouafi, F. Dornaika, A. Taleb-Ahmed, and A. Hadid, "Pyramid multi-level features for facial demographic estimation," *Expert Systems with Applications*, vol. 80, pp. 297–310, 2017.
- [14] S. E. Bekhouche, "Facial soft biometrics: Extracting demographic traits," Ph.D. dissertation, Faculté des sciences et technologies, 2017.
- [15] G. E. Hinton, "To recognize shapes, first learn to generate images," *Progress in brain research*, vol. 165, pp. 535–547, 2007.
- [16] Y. Bengio, Y. LeCun *et al.*, "Scaling learning algorithms towards ai, larg," *Scale Kernel Mach*, vol. 34, no. 5, 2007.
- [17] Y. LeCun *et al.*, "Lenet-5, convolutional neural networks," URL: <http://yann.lecun.com/exdb/lenet>, 2015.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [19] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification second edition john wiley & sons," *New York*, vol. 58, p. 16, 2001.
- [20] M. Shao, S. Xia, and Y. Fu, "Genealogical face recognition based on ub kinface database," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, 2011, pp. 60–65.
- [21] 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*. IEEE, 2013, pp. 2983–2987.
- [22] 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.
- [23] H. Yan, "Kinship verification using neighborhood repulsed correlation metric learning," *Image and Vision Computing*, vol. 60, pp. 91–97, 2017.
- [24] A. Chergui, S. Ouchtati, H. Telli, F. Bougourzi, and S. Bekhouche, "Lpq and ldp descriptors with ml representation for kinship verification," in *The second edition of the International Workshop on Signal Processing Applied to Rotating Machinery Diagnostics (SIGPROMD'2018)*, April 2018, pp. 1–10.
- [25] X. Qin, D. Liu, and D. Wang, "Heterogeneous similarity learning for more practical kinship verification," *Neural Processing Letters*, vol. 47, no. 3, pp. 1253–1269, 2018.
- [26] X. Zhou, K. Jin, M. Xu, and G. Guo, "Learning deep compact similarity metric for kinship verification from face images," *Information Fusion*, vol. 48, pp. 84–94, 2019.