

# A Family Relationship Identification Network based on Metric Learning

Decai Li<sup>1</sup>, Chengqun Yi<sup>1</sup>

<sup>1</sup>School of Automation and Information Engineering, Sichuan University of Science & Engineering, Zigong, China.

---

## Abstract

Most of the existing methods for family relationship identification can only identify single parent kinship (father-son relationship, father-daughter relationship, mother-child relationship, mother-daughter relationship), and these methods have poor identification effects on the identification samples with large age gap and different genders, such as father and daughter or mother and child. In order to solve these problems, a new method is proposed to recognize the relationship between children and their parents simultaneously. This paper proposes a model of parental relationship recognition based on metric learning to improve the accuracy of family relationship recognition between children and parents. Based on the biological genetic relationship between children and parents, a multi-linear parallel network was designed to integrate the genetic characteristics of children and parents. Using the advantages of discriminant metric learning and canonical correlation analysis in data processing, feature information beneficial to family relationship recognition is extracted from facial features containing multiple human identities, which can be used to identify whether there is blood relationship between children and parents to improve the accuracy of recognition. Experimental results show that the proposed method is more effective in the recognition of the family relationship between children and parents.

## Keywords

Family Relationship Identification; Metric Learning; Deep Learning.

---

## 1. Introduction

Human faces are the most important component that people use to recognize each other. Humans can recognize each other quickly and easily through faces, and because facial features are not affected by light conditions and postures, faces are still human dynamic recognition methods. In addition to the current soft biometric attributes (such as gender, age, and race), you can also explore whether a face belongs to a certain group of people in appearance and has some similarity with a specific group of people [1], a person's closest relationship is his family. Because of genetic similarity, there are facial similarities between family members. It is worth noting that the ability of the human brain to process facial similarity is achieved by combining the existing neural computing structure [2]. Humans use face similarity to identify their relatives and family members, and adjust their behavior accordingly [3-4]. Family relationship recognition refers to the task of training a machine to recognize whether a pair of faces are relatives or non-relatives based on the features extracted from facial images, and to determine the exact type or similarity of this relationship. Automatic family relationship verification and recognition is an interesting field of investigation, which can be applied to discover human social relationships and activities, find missing children/parents, annotate images, create family trees, automatically manage photos, enrich social media applications [5], and perform images Retrieval and labeling, etc.

In the field of computer vision recognition, convolutional neural networks have achieved great success in face recognition, age recognition, expression recognition and other recognition tasks, but poor performance in family relationship recognition. This is because most of the existing family relationship recognition models use all the features that can be extracted by the convolution layer to calculate the similarity, and the recognition accuracy is affected by the non-related features of kinship. Although some family relationship recognition models [6-9] minimize the influence of unrelated characteristics of kinship as much as possible in the recognition process, most of these models ignore the biological genetic relationship between children and parents, so they still perform poorly in the face of challenging samples. Although the existing models for the identification of the relationship between children and parents take into account the biological genetic association between children and parents, they do not solve the problem of age and gender gap between children and parents in the identification process, so the identification effect of some samples with large differences is not good.

In order to improve the recognition accuracy of the family relationship between children and parents, this paper proposes a family relationship recognition network based on metric learning that can effectively solve the gaps in the family relationship between children and parents. Network based on metric learning), as shown in Figure 1. FRINM, which is used to process the facial features of children and parents, is a multi-linear network with a twin architecture [10-11], which uses the biological genetic relationship between children and parents to improve family relationships about Recognition accuracy. Its design method is to use the difference in distance between different features to extract features that are only related to family relationships through the discriminative measurement learning[12] port, and reduce the impact of age, gender and other non-related features of family relationships on the recognition process; The canonical correlation analysis is used to calculate the covariance matrix of the similarity between the children and the parents, and the similarity between the children and the parents is integrated to realize the recognition of the family relationship between the parents. At the same time, in order to solve the limitation of the small amount of facial data of complete family members in actual training, further improve the family For the recognition accuracy of the relationship, FRINM first trains the convolutional layer on the public large-scale face database to ensure that the convolution part has robust feature extraction capabilities; then, the model is trained on the single-parent relationship-related features of the relationship. Extraction ability and recognition ability; Finally, use the data set of parental relationship to fine-tune FRINM to achieve the improvement of parental relationship recognition ability.

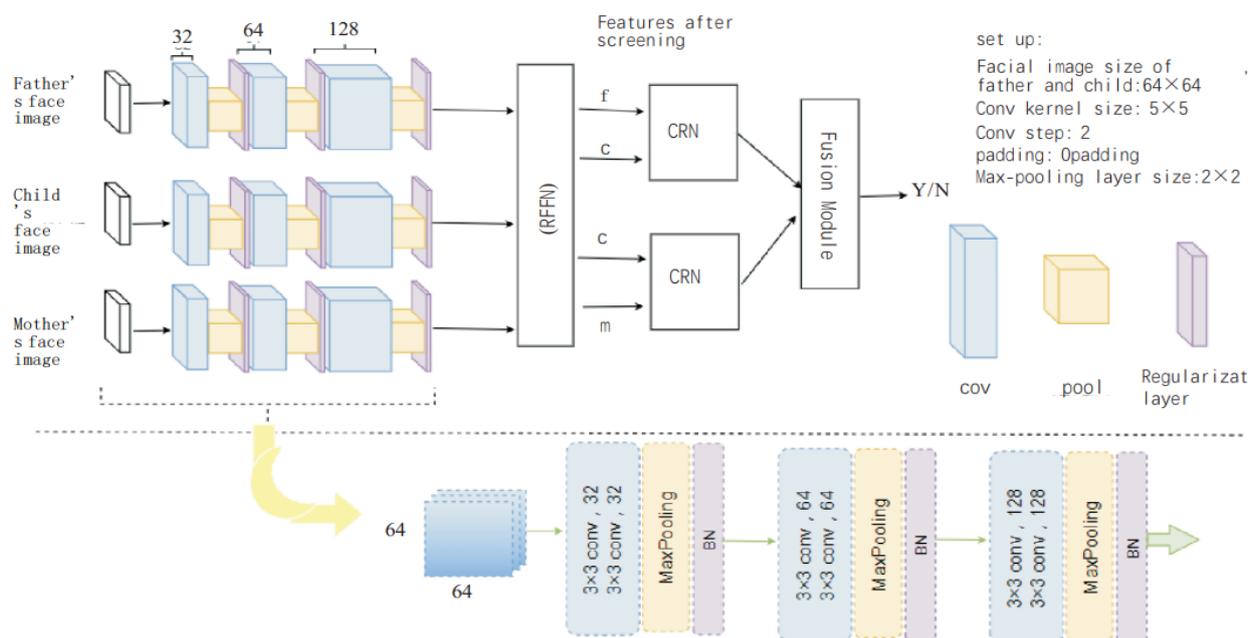


Fig. 1 FRINM's network structure

The main contributions of this paper in the field of family relationship recognition are as follows:

- 1) The multi-linear recognition network of twin architecture can more effectively use the genetic relationship between children and parents to improve the accuracy of family relationship recognition;
- 2) Using metric learning to extract features that are only related to family relationships from the convolutional feature map for identification, which can greatly reduce the influence of the gap between the age and gender of factor women and parents on the recognition results, and improve the accuracy of family relationship recognition.

## 2. Design FRINM

### 2.1 Feature selection module RFFN

Assuming that  $S = \{(x_i, y_i | l_i) | i = 1, 2, \dots, N\}$  pairs have a certain relationship with N, the dataset is divided into positive and negative relationship samples, where  $\varphi = \{(x_i, y_i) | l_i = 1\}$  denotes that the positive sample pair has a similar relationship with  $(x_i, y_i)$ , and  $\varphi = \{(x_i, y_i) | l_i = -1\}$  denotes that the negative sample pair has no similar relationship with  $(x_i, y_i)$ . Correlation feature extraction network (RFFN) aims to filter out the features unrelated to the correlation of the sample, and use the correlation features to distinguish the sample pairs. In order to ensure that the selected features are used as the kinship features, the extracted features are used as the input of the recognition model, so that the recognition accuracy is greater than the preset threshold when the kinship is used as the recognition target. Compared with other relationships, for example, when gender or age is the recognition target, the recognition accuracy is less than the preset threshold.

The goal of RFFN is to use the Mahalanobis distance between different feature vectors to select the features related to the recognition task. Assuming that two data  $x_i \in R^n$  and  $y_i \in R^n$  are given, the Mahalanobis distance [13] calculation formula is:

$$d_A(x_i, y_i) = \sqrt{((x_i - y_i)^T \times A \times (x_i - y_i))} \quad (1)$$

The semi-positive symmetric covariance matrix  $A \in R^{n \times n}$  is the weight parameter of the network. The matrix singular value decomposition is used to simplify the solving process of matrix A, then matrix A is equivalent to  $A = \omega^T \omega$ , and the new Mahalanobis distance formula is:

$$d_A(x_i, y_i) = \sqrt{((x_i - y_i)^T \omega^T \omega (x_i - y_i))} = \sqrt{((\omega \times x_i - \omega \times y_i)^T (\omega \times x_i - \omega \times y_i))} \quad (2)$$

Therefore, solving the Mahalanobis distance of  $x_i \in R^n$  and  $y_i \in R^n$  can be understood as solving the Euclidean distance  $d_E(\omega \times x_i, \omega \times y_i)$  after they are converted by the two-layer nonlinearly activated full connection layer with weight  $\omega \in R^{n \times m}$ , where  $\omega \in R^{n \times m}$  is the weight of the middle full connection layer. In order to extract the relevant characteristics in the sample, the sample is processed under the measurement threshold. The condition is: when the given sample  $(x_i, y_i)$  belongs to the set  $\varphi$ ,  $d_E(\omega \times x_i, \omega \times y_i) \leq \tau$  otherwise the optimization objective function of the  $d_E(\omega \times x_i, \omega \times y_i) > \tau$  full connection layer is:

$$\underset{\omega}{\operatorname{argmax}} J(\omega) = \sum_{x_i, y_i} g\left(1 - l_i(\tau - d_E(\omega \times x_i, \omega \times y_i))\right) + \frac{\lambda}{2} \sum_{m=1}^M ((\omega^m)_F^2 + (b^m)_2^2) \quad (3)$$

$g(x) = \frac{1}{\beta} \log(1 + e^{\beta x})$  is a logical loss function,  $\|\omega^m\|_F$  is the Frobenius norm of the matrix,  $\lambda$  is the regularization parameter,  $\beta$  is the learning rate.

### 2.2 Similarity analysis module CRN

For the data set  $S = \{x_i, y_i | i = 1, 2, \dots, N\}$ , in order to identify the similarity between each group of objects  $(x_i, y_i)$ , a correlation analysis model CRN is designed, as shown in Figure2. The premise of identifying the relationship between the targets is to extract the features related to the recognition task, so by using the previously described RFFN is used to extract the relevant features  $(x'_i, y'_i)$  of the object sample  $(x_i, y_i)$  and analyze their correlation.

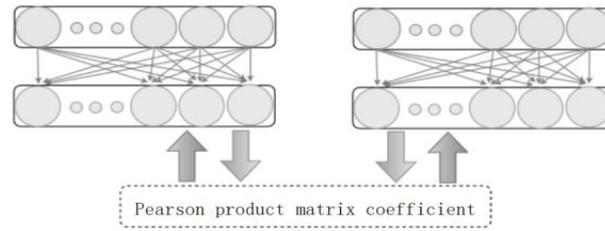


Fig. 2 Similarity Analysis Module (CRN)

In the process of solving the similarity of eigenvectors, considering their internal correlation, the principal component analysis method is used to find the uncorrelated linear combination of each two sets of vectors 291 and discuss their correlation. First, find the first in the two sets of variables Pair linear combinations and make them have the greatest correlation; then find the second pair of linear combinations in the two sets of variables, so that they have the greatest correlation, and they are independent of the first set of linear combinations. Repeat this step and find other linear combinations. Combination, where

$$\begin{cases} u_i = a_{i1}x'_{i1} + a_{i2}x'_{i2} + \dots + a_{im}x'_{im} \\ v_i = b_{i1}y'_{i1} + b_{i2}y'_{i2} + \dots + b_{im}y'_{im} \end{cases} \quad (4)$$

The similarity of the identified object can be obtained from the following equation:

$$\rho(x_i, y'_i) = \rho(u_i, v_i) = \frac{(u_i, v_i)}{\sqrt{D(u_i)D(v_i)}} = \frac{\sum_{i=1}^n u_i \times v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \quad (5)$$

The optimization objective function of CRN is

$$argmax_{a,b} J(a, b) = \sum (1 + \log(1 + \rho(x'_i y'_i) + \varepsilon)). \quad (6)$$

### 2.3 FRINM identifies the network

Given a data set containing multiple data objects for each group of samples  $D=X_i, Y_i, Z_i \mid i=1,2,\dots,N$  (for example, X for father, Y for child, Z for mother), by FRINM identifying each group of objects  $(X_i, Y_i, Z_i)$ . Firstly, the facial images of father, mother and child were input, and the features were extracted through the convolution part and their feature vectors  $(x_i, y_i, z_i)$  were output. Secondly, Use Feature screening module RFFN to select the characteristics related to the recognition of the kinship of parents, mothers and children and mark them as  $(x'_i, y'_i, z'_i)$ ; Again, use the CRN network to calculate the similarity between the father and the son and the mother and son, labelled as  $C_{(x'_i, y'_i)}$  and  $C_{(z'_i, y'_i)}$ ; finally, the correlation coefficient  $C_{(x'_i, y'_i)}$  and  $C_{(z'_i, y'_i)}$  are combined to obtain the final result  $S_{tri}$ , namely

$$S_{tri} = P_{(x_i, y_i)} \times C_{(x'_i, y'_i)} + P_{(z_i, y_i)} \times C_{(z'_i, y'_i)} \quad (7)$$

Among them

$$P_{(x_i, y_i)} = \frac{\exp(C_{(x'_i, y'_i)})}{\exp(C_{(x'_i, y'_i)}) + \exp(C_{(z'_i, y'_i)})}$$

Represents the probability that  $Y_i$  is more like  $X_i$ , and

$$P_{(z_i, y_i)} = \frac{\exp(C_{(z'_i, y'_i)})}{\exp(C_{(x'_i, y'_i)}) + \exp(C_{(z'_i, y'_i)})}$$

Represents the probability that  $Y_i$  is more like  $Z_i$ .

### 3. Experiments on age

This section adopts a staged training method to train the model. First, use recognized face data to train the feature extraction ability of the convolutional network; second, use the kinship data to train

the network to screen and screen the kinship correlation features. Similarity recognition; finally, the recognition results of different branch networks are combined and the body model is fine-tuned. In the experiment, the first step is to train and use 202,599 images with complex backgrounds and different pixel resolutions. The gender, expression, posture, and age are different The Celeba database of face images; the second step is to train and use the KinfaceW-I and KinfaceW-I family member face data sets containing 4 kinds of kinship (father-child, father-daughter, mother-child, mother-daughter); finally, use the complete family relationship The TSKinFace fine-tunes the weight of the volume model.

Use migration learning to train the network 6 times. After standardizing the original training set, the data set is divided into 5 equal parts using the cross-validation method, of which 3 parts are used for model training, 1 part is used for model verification, and 1 part is used for model testing. The first step of training FRINM is to extract the features of the input image, use three convolutional layers with the same weight to extract the image features, and map the extracted features belonging to the three input images to the same feature space; the second step is to filter and Features related to kinship recognition. In order to reduce the impact of parents and children on the accuracy of kinship recognition due to differences in age and gender, the RFFN model is introduced to select useful features; the third step is to identify feature similarity, and CRN calculates differences. The correlation coefficient between the feature vectors is used to obtain the similarity between the features; finally, the similarity between the father and the child, and the mother and the child is combined as the kinship recognition result.

In order to control the influence of a single condition on the experiment, the experiment aimed at the influence of age gap on the accuracy of kinship recognition: the data set sample pair is divided into 4 parts with the age gap between parents and children as the standard. The average age difference is 20~25-35~40. And count the recognition accuracy of the relationship between father and son, mother and daughter of the same sex, and use the average recognition accuracy value as the accuracy value of the corresponding age gap. As shown in the figure 3, the greater the age gap, the more accurate the recognition of the relationship Low. [8,9,13-17]

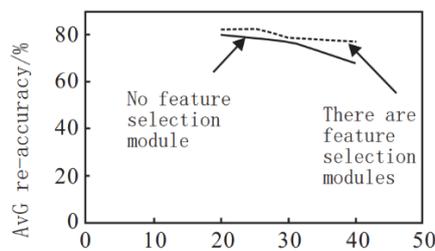


Fig. 3 The effect of age difference on family relationship

#### 4. Conclusions

The FRINM design process proposed in this paper is as follows:

- 1) Drawing on other kinship recognition methods. Introduce a discriminative metric learning method, by measuring the distance of facial features, selecting facial features that contain basic features related to kinship recognition and filtering irrelevant
- 2) This article draws on the calculation methods of some documents. The advantage of calculating the correlation coefficient of the feature vector is that the internal relationship between the various parts of the feature is considered.
- 3) The use of staged training Method. This article uses the principle of segmented training to learn the weights of each part of FRINM. The advantage of this training method is to ensure that the learned parameters have good robustness.

4) Use the face data of all family members to fine-tune the weights of the model Parameters to improve the accuracy and robustness of recognition.

The successful implementation of the FRINM model is conducive to the next step in the study of the family relationship generation model. Using this model as a discriminator against the generation network can effectively determine whether the new facial image generated has a family relationship with the input object. Family relationship generation The study of the model has further promoted the study of improving the accuracy of family relationship recognition.

## References

- [1] Ghahramani, M., W. Yau and E.K. Teoh, A novel approach for recognising people of the same family. *International Journal of Biometrics*, vol. 5(2013) No.3, p. 229-247.
- [2] Platek, S.M. and S.M. Kemp, Is family special to the brain? An event-related fMRI study of familiar, familial, and self-face recognition. *Neuropsychologia*, vol. 47(2009) No.3, p. 849-858.
- [3] Alvergne, A., et al., Cross-cultural perceptions of facial resemblance between kin. *Journal of Vision*, vol. 9(2009). No.3, p. 23-23.
- [4] DeBruine, L.M., et al., Social Perception of Facial Resemblance in Humans. *Archives of Sexual Behavior*, vol. 37(2008) No.1, p. 64-77.
- [5] Y. Liu. Fortune teller: predicting your career path. in *Proceedings of the AAAI conference on artificial intelligence*. 2016, 30(1).
- [6] J.W. Lu, X.Z. Zhou and Y.P. Tan, et al. Neighborhood repulsed metric learning for kinship verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Piscataway: IEEE, vol.27(2012) No.5, p.331-345.
- [7] S. Xia, M. Shao, Y. Fu, et.al. Kinship verification through transfer learning, *International Joint Conference on Artificial Intelligence*. Barcelona: ACM, vol.9(2011) No.8,p.2539-2544.
- [8] H.B. Yan, J.W. Lu, W.H. Deng, et al. Discriminative multimetric learning for kinship verification, *IEEE Transactions on Information Forensics and Security*, Vol. 9(2014) No.7, p. 1169-1178.
- [9] H.B. Yan, J.W. Lu and X.Z. Zhou, et al. Prototype-based discriminative feature learning for kinship verification, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 11(2015) No.11, p. 2535-2545.
- [10] Vijay Kumar B G, Carneiro G, Reid I. Learning local image descriptors with deep Siamese and triplet convolutional networks by minimizing global loss functions, 2016 *IEEE Conference on Computer Vision and Pattern Recognition*. CVPR, vol.2(2016) No.7,p.5385-5394.
- [11] Melekhov I, Kannala J, Rahtu E. Siamese network features for image matching, *The 23rd International Conference on Pattern Recognition*. Piscataway: IEEE, vol.2(2016) No.2,p.378-383.
- [12] Hoffer E, Ailon N. Deep metric learning using triplet network, *Similarity-Based Pattern Recognition*. Cham: Springer International Publishing, vol.1(2015) No.1,p.84-92.
- [13] S.M. Xiang, F.P. Nie, C.S. Zhang, et al. Learning a Mahalanobis distance metric for data clustering and classification, *Pattern Recognition*, vol. 41(2008) No.12, p. 3600-3612.
- [14] J.L. Hu, J.W. Lu, J.S. Yuan, et al. Large margin multi-metric learning for face and kinship verification in the wild, *Asian Conference on Computer Vision*. Singapore: Springer Link, vol.6(2014) No.2,p.252-267.
- [15] Babagholami-Mohamadabadi B, Roostaiyan S M, Zarghami A, et al. Multi-modal distance metric learning: A Bayesian non-parametric approach, *European Conference on Computer Vision*. Zürich: Springer, vol.1(2014) No.1,p.63-77.
- [16] S.M. Xiang, F.P. Nie, C.S. Zhang, et al. Learning a Mahalanobis distance metric for data clustering and classification, *Pattern Recognition*, vol. 41(2008) No.12, p. 3600-3612.
- [17] Andrew G, Arora R, Bilmes J A, et al. Deep canonical correlation analysis, *International Conference on Machine Learning*. Georgia: JMLR.org, vol.7(2013) No.5,p.1247-1255.
- [18] Li. Y, J.B. Zeng, J. Zhang, et al. KinNet: Fine-to-coarse deep metric learning for kinship verification[C]. *Proceedings of the 2017 Workshop on Recognizing Families In the Wild*. New York: ACM, vol.1(2017) No.1,p.13-20.