Heatmap-Guided Balanced Deep Convolution Networks for Family Classification in the Wild.

Decky Aspandi, Oriol Martinez and Xavier Binefa
Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain

Abstract— **Automatic kinship recognition using Computer Vision, which aims to infer the blood relationship between individuals by only comparing their facial features, has started to gain attention recently. The introduction of large kinship datasets, such as Family In The Wild (FIW), has allowed large scale dataset modeling using state of the art deep learning models. Among other kinship recognition tasks, family classification task is lacking any significant progress due to its increasing difficulty in relation to the family member size. Furthermore, most current state of-the-art approaches do not perform any data pre-processing (which try to improve models accuracy) and are trained without a regularizer (which results in models susceptible to overfitting). In this paper, we present the Deep Family Classifier (DFC), a deep learning model for family classification in the wild. We build our model by combining two sub-networks: internal Image Feature Enhancer which operates by removing the image noise and provides an additional facial heatmap layer and Family Class Estimator trained with strong regularizers and a compound loss. We observe progressive improvement in accuracy during the validation phase, with a state of the art results of 16.89% is obtained for the track 2 in the RFIW2019 challenge and 17.08% of familly classification task on FIW dataset.**

I. INTRODUCTION

Kinship recognition using Computer Vision techniques, a.k.a visual kinship recognition, has attracted the interest of the research community in recent years [11], [23]. This is due to the wide range of applications such as searching for a missing human, human trafficking, or for recent issues of refuge crisis [23]. The main goal of kinship recognition is to train a model to recognize the kin or Non-Kin relationship between different individuals from facial features[11]. Currently, kinship recognition can be split in two tasks: kinship verification which tries to predict if there is a relationship between two individuals; Family classification, which aims to map certain individuals to specific family classes.

The introduction of The large kinship datasets i.e. Family In The Wild (FIW) [23]- has accelerated the development of kinship recognition solutions and allowed the utilization of powerful deep learning models that reach state of the art results, especially for kinship verification [32]. In spite of the current progress, the problem is still open, which is specially true in the case of Family classification, where difficulty increases as the number of family members (and therefore the diversity of classes) [23]. For such reasons, family classification has been overlooked in the literature compared to kinship verification.

Because of its challenging nature, in this work we focus on family classification by improving the main limitation of the current approaches that deal with it: 1) no data

understanding or pre-processing is done (which has been known to improve the overall model accuracy [9], [15]) and 2) no strong regularization techniques applied (which may hinder the model learning [2], [26]). Specifically, we propose a deep learning-based model that integrates a strong regularizer in model learning and incorporates the appropriate data pre-processing via an internal image enhancer. The internal image enhancer is based on a Deep Image Denoiser build with Auto-Encoder architecture [8] coupled with a Facial Heatmap Estimator built with the Inception-Resnet architecture [27]. As final pipeline, we adopt the state-of-the-art Squeeze and Excitation Network [13] trained using compounding losses of center loss [29] with sample weight normalization. The main contributions of this paper are:

1) We introduce a deep learning based image enhancement stage that normalizes the input and produces an accurate heatmap layer to guide the feature learning.
2) We show progressive improvement of our model accuracy to justify the effectiveness of our learning approach.
3) We show state-of-the-art results on the family classification of [23] and second track of RFIW 2019 challenge.

II. RELATED WORK

Kinship recognition, where the goal is to identify the relationship between individuals, is an active research field in physiology [10], biology [3] and recently computer vision [21], [17]. Early attempts to automatically model these type of relationships were made by extracting hand-crafted features such as Local Binary Pattern (LBP)[17], Scale Invariant Feature Transform (SIFT) [33], Gabor Filter [31], and Histogram Of Gradient (HOG) [35] which were later used to train a Support Vector Machine(SVM) classifier [11] or to be clustered with techniques such as K-Means clustering[17]. Generally, these models were limited to identify the relationship accross a single generation e.g Father Daughter (F-D) or Sibling-Sibling (S-S) [12], [30]. However, recent progress with larger datasets such as KFW [17] or FIW [23], [22] jointly with the end to end deep learning models, allows for the identification accross multiple generation kinship relationships with state of the art accuracy [32], [28].

Simultaneously, other Kinship Recognition tasks, like family classification, have been overlooked in the literature. Thus lacking any similar progress. Family classification, which aims to identity the set of members to family classes, is another important task given the better simulation of the
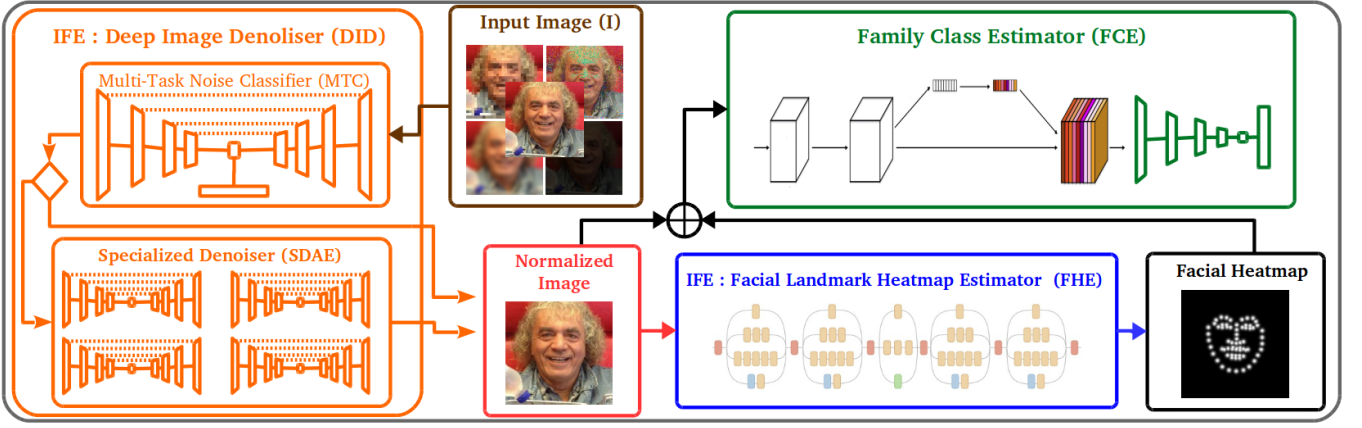
Fig. 1. Overview of our models pipeline

inheritance biological process [11], and becomes more challenging as the number of family increase [23]. The first work on this task [11] used SIFT features from several facial parts and evaluatedon a relatively small Family101 dataset. The introduction of larger datasets, such as FIW[24], has allowed the development of more expressive deep learning based architecture. Convolutional network with VGG[19], Resnet, and CenterFace[29] currently hold the state-of-the-art accuracy[23]. In spite of the success of these models, there is still a lack a significant progress compared to the kinship verification, as typically, straight-forward learning strategy is adopted by performing cross domain transfer learning with plain image as input. Such an approach neglects significant data pre-processing, which has been shown to benefit the models performance [9], [15] and is prone to overfit on account of their large networks size [2], [26]. We tackle these shortcomings in our approach by putting the emphasizes on the image feature enhancement technique and inclusion of proper training strategy using strong regularizers and compounding loss.

III. METHODS

Our Deep Family Classifier (DFC) consists of two main sub-networks, each responsible for their specific task: Image Feature Enhancer (IFE) and Family Class Estimator (FCE). Given an individual with corresponding image input (**I**), the IFE sub-network normalizes the image by minimizing the known image noise, and it also provides additional facial heatmaps as an added layer to guide the feature learning[6]. These spatial features are then passed into the DFC to estimate the family class c of the corresponding family member I.

$$c(I) = DFC(I) = FCE(IFE(I)) \quad (1)$$

Figure 1 visualizes the overall structure of our pipeline.

A. Image Feature Enhancer (IFE)

We argue that the final estimation of the overall model will be improved by using the enhanced image input. We introduce two main sub-modules to perform a two-stage of enhancement : Deep Image Denoiser(DID), which normalizes the input image, and Facial Heatmap Estimator(FHE) which generates an additional facial heatmap layer.

$$IFE(I) = concat(I, FHE(DID(I))) \quad (2)$$

The end result of IFE is the concatenated facial heatmap layer with three RGB layers of normalized input, which are then fed onto the FCE sub-network.

1) Deep Image Denoiser (DID): In this setup, we normalize the input image by minimizing the known noise. To do this, we build our image normalizer by using internal Multi-Task noise Classifier(MTC) and Specialized Internal image Denoising Auto-Encoder ($SDAE$) which operate in tandem.

$$DID(I) = SDAE_{\phi^2}(MTC_{\phi^1}(I), I) \quad (3)$$

We adopt the Hourglass shaped Auto-Encoder Architecture (DAE) with skip connection [18], [8] as the base architecture of both MTC and $SDAE$. We train the DID to clean synthetically perturbed image with four classes of noise: First is image down-sampling, second is gaussian blurring, third is gaussian noise, and fourth is color scaling [36]. The first task of DID is to estimate the noise class c_n as follows:

$$MTC_{\phi^1}(c_n|I) = f_{pr}(c, \frac{e^{W_{\phi^1_k}^{MTC} \odot conv_{\phi^1_k}^{MTC}(I)}}{\sum_{k=1}^{K} e^{W_{\phi^1_k}^{MTC} \odot conv_{\phi^1_k}^{MTC}(I)}}), \quad (4)$$

$$f_{pr}(x, S) = \begin{cases} 1, \ if \ x = argmax\{s_1, s_2..s_K\} \\ 0, \ otherwise \end{cases}, \quad (5)$$

where W^{MTC} is multinomial bottleneck regression layers parameterized by Φ^1, **S** is the set of scores for all K, $conv$ is the set of convolutional layers of MTC, and the f_{pr} is the indicator function. If noise is present (i.e. $c_n > 0$), $SDAE$, which consists of specialized denoisers $SDAE = \{DAE_1, DAE_2...DAE_K\}$ will use specific DAE class to remove the noise c_n. Otherwise, the denoising process is skipped to preserve the characteristics of the image.

$$SDAE_{\phi^2}(c_n, I) = DAE_{\phi^2_{c_n}}(I), \quad (6)$$

where Φ_c^2 is the parameter learned for specific DAE and specific noise class c_n.

2) Facial Heatmap Estimator: FHE augments the input features by introducing an additional facial heatmap layer centered around the landmark estimates of l_t [6]. To do this, we first obtain the facial landmark points l_t using the FLL part of Recurrent Tracker [4], which consist of Inception-Resnet (res) [27] and a regression layer of weight matrix W^{FLL} parameterized by Φ^3 :

$$\mathbf{l}_t = FLL_{\Phi^3}(X) = W_{\Phi^3}^{FLL} \odot res_{\Phi^3}(X) \tag{7}$$

Given the l_t, FHE produces the facial heatmap layers by activating the circular area following 2D Gaussian distribution centered on each landmark points on l_t with radius r.

$$FHE(I), by\ g(l_t, I) = f_{l_t} : I \rightarrow I_h \tag{8}$$

$$f_{l_t}(I) = 0_{x_i \pm a, y_i \pm b} + f_x(x_i, y_i | \mu, \sigma)\ for\ x_i, y_i \in l_t \tag{9}$$

with a and b satisfy :

$$(x - a)^2 + (y - b)^2 = r^2 \tag{10}$$

Where μ denotes the center of each facial landmarks, σ is the specified radius r, and f_x is the 2D normal probability density function.

B. Family Class Estimator (FCE)

1) Model Definition: We build our final pipeline of FCE using the state-of-the-art Squeeze and Excitation Network (SENET) [13] which has been pre-trained on the recently published VGGFace2 [7] facial dataset [1]. This is used to estimate the family class label c. Specifically, we replace the first 2D convolutional layer to accommodate the additional facial heatmap input channel and add an additional 524 bottleneck regression layer for the corresponding 524 family classes. The FCE consists of relatively similar operations of MTC:

$$FCE_{\phi^4}(c|I) = f_{pr}(c, \frac{e^{W_{\phi_k^4}^{FCE} \odot SENET_{\phi_k^4}^{FCE}(I)}}{\sum_{k=1}^{K} e^{W_{\phi_k^4}^{FCE} \odot SENET_{\phi_k^4}^{FCE}(I)}}) \tag{11}$$

2) Compounding Loss (CL): To improve the convergence on training FCE sub-network, we propose to add the compounding loss based on the center loss [29] and the sample weight normalization, along with standard cross entropy loss. The center loss is used to minimize intra-class variations and the weight normalization factor accommodates the unbalanced classes of the training dataset. The final loss of the FCE is as follows :

$$\mathbf{L_{FCE}} = \lambda_{sig}\ell_{sig} + \lambda_{ctr}\ell_{ctr} \tag{12}$$

$$\ell_{sig} = \sum_{i=1}^{N} \frac{n_i}{N} log \frac{e^{W_{\phi_k^4}^{FCE} \odot SENET_{\phi_k^4}^{FCE}(I)}}{\sum_{k=1}^{K} e^{W_{\phi_k^4}^{FCE} \odot SENET_{\phi_k^4}^{FCE}(I)}} \tag{13}$$

$$\ell_{ctr} = \frac{1}{2} \sum_{i=1}^{m} \|x_i - ctr_{y_i}\|, \tag{14}$$
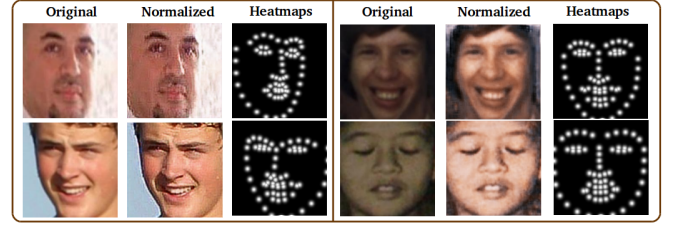
[1] https://github.com/ox-vgg/vgg_face2

Fig. 2. Examples of normalized input with facial heatmap layer of FIW dataset

Where the λ is the regularizing coefficient for each term, ctr_{yi} denotes the y_i-th class center of embedding features, n_i is the total instances of class i, and N is the normalization factor for the total of estimated classes, i.e 524.

C. Model training and setup

1) Training IFE: Considering that we do not have the noise-free image as ideal reference in the original FIW dataset, we resort to the cross-dataset capability of our model by training the IFE independently using 300-W [25] and Menpo [34] dataset. Specifically, we used standard ℓ^2 loss, learning rate of 0.001, and ADAM[14] optimizer.

An example of the normalized image from FIW dataset jointly with its corresponding facial heatmaps layer can be seen on the Figure 2. Notice that IFE model is able to normalize the input according to the image conditionss : the model sharpens the image in if is too blurry (column 1), and it scales up the image intensity in case of low illumination(column 2). Furthermore, the estimated facial heatmaps are too visually accurate.

2) Training FCE: It is well known that very deep models can be too expressive and may lead to overfiting if trained without any regularizer [2]. To avoid this, we incorporate two regularization strategies : (1) we perform data augmentation procedure of $-45°$ to $45°$ degree rotations, random-centered cropping and horizontal flipping; (2) we add a series of dropout layer [26] with high probability value of 0.5. During training, we set the value of each λ to 0.5, r by 1% the value of the facial bounding box norm, learning rate of 0.0001 with weight decay every 1000 iterations and trained using SGD [5] optimizer. Finally, we used the FIW training dataset provided within the challenge for training.

3) Technical Specifications: All model were collectively implemented using both Tensorflow[1] and Pytorch [20]. Trainings was done using five Titan X GPUs and took over three days for each model. Our trained models are available on our github page [2].

IV. EXPERIMENTS

A. Experiment Settings

We evaluated the performance of our model on the Second Track of Recognize Families in the Wild 2019 Challenge (RFIW2019) which deals with the family classification task. Additionally, we also provide comparison on the original

[2] https://github.com/deckyal/DFC

Fig. 3.   Examples of families member on the FIW[23] dataset

Fig. 4.   Examples of correctly classified families

TABLE I

THE ACCURACY OF OUR MODELS ON THE VALIDATION DATASET ON THE SECOND TRACK OF RFIW2019

Networks	Train. Acc	Val. Acc.
$B1(FCE)$	**100**	4.52
$B2(FCE + R)$	95.5	10.89
$B3(FCE + R + IFE)$	97.3	13.73
$DFC(FCE + R + IFE + CL)$	97.8	**14.28**

TABLE II

TEST RESULT ON THE ORIGINAL FIW DATASETS

Networks	Acc.
VGG-Face, fc7 (4,096D)+one-vs-rest SVMs [23]	3.04
VGG-Face, replaced softmax (564D)+fine-tuned [23]	10.42
ResNet-22 + softmax (564D) [23]	14.17
SphereFace(564D)[23]	13.86
ResNet-22 + CF (512D) + softmax (564D)[23]	16.18
DFC (Ours)	**17.08**

familiy classification task on original FIW dataset [23]. The goal of family classification is to classify a family membership of hold-out individual from 524 classes of family. Several samples of family in the FIW dataset can be seen on the Figure 4.

For the second track of RFIW2019 challenge, the training samples are provided for each family class with the remaining unseen data used for both validation and test phase. This ensures the generalization of the model on the real unseen test samples [24]. While on the original FIW family classification task, five cross validation scheme is employed with at least 4 family members used for training and 1 for testing[23]. On all scenario, precision score is used for the evaluation of the models accuracy.

B. Result and Discussion

1) Ablation Studies: We first present the ablation studies on the validation dataset on the second track of RFIW2019 challenge to see our models accuracy under different settings : 1) Baseline-1 (B1) consisting of FCE sub-network trained directly with data training without any pre-procesing; 2) Baseline-2 (B2) Similar to $B1$ but includes the regularization strategy (R); 3)Baseline-3 (B3) Extension of $B2$ but using the enhanced data from IFE instead of original image as input; 4) DFC, our full model which follow all pipeline described in methodologies including Compounding Loss(CL). We can observe the progressive improvement of the overall accuracy of our model in the Table I.

Specifically, we found that a direct approach using only FCE without any regularizer results in over-fitting due to our models size. This fact is indicated by perfect training accuracy but poor validation result. Adding the regularizers remedies this problem producing a better validation accuracy ($B2$) with the sacrifice of slightly lower training accuracy.
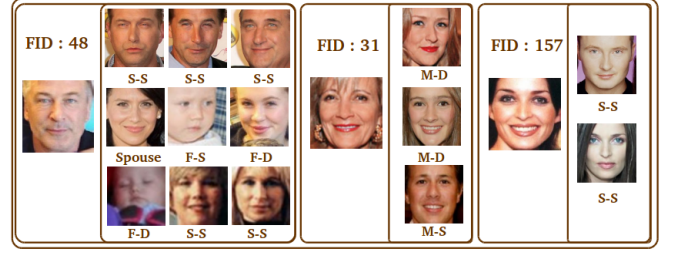
The model accuracy of $B3$ further increased by using the internal IFE proving the benefit of pre-processing strategy. Finally, we found the further accuracy improvements using the compounding losses to our final DFC model.

The examples of correctly predicted family of individual can be seen on the Figure 4. In these examples, we found that our model is not only able to relate the individuals with existing multiple relationship such as on the first column, but also correctly predict the family class where only few kin-relationship available. This occurs in the second column where only Mother-Daughter(M-D)/Son(M-S) relationship exists, and on third column where only single relationship exists, i.e Sibling-Sibling(S-S). Note that, less available relationships also signifies fewer training samples which suggests the benefit of our sample weight normalization.

2) Result on FIW dataset and 2nd Track of RFIW2019 Challenge: Finally, we present the results of our model, both tested on the original five cross validation scheme in the work of [23] (Table II) and also on the test dataset of second track of RFIW 2019 challenge (table III). We can see the state-of-the-art accuracy of our model on the original FIW challenges, compared with current baselines [23], which includes the deep VGG-based [19] model and SphereFace [16].

V. CONCLUSIONS

In this paper, we present our state-of-the-art deep learning based model for family classification in the wild. Different from other approaches, our model puts emphasis on the data pre-processing with image normalization, additional features

TABLE III

TEST RESULT ON THE SECOND TRACK OF RFIW2019

Networks	Acc.
DFC (Ours)	**16.89**

input, and multiple regularizers of compounded training loss to improve the model learning. We utilized Denoising Autoencoder to reduce the variability on the dataset and generate an additional facial heatmap layer on top of the normalized input to guide the feature learning. Furthermore, we trained the state-of-the-art Squeeze Excitation Network for family class estimation using a compound loss of the center loss and sample weight normalization. Finally, the dropout and data augmentations technique used to regularize the training process. On the validation phase, we see the progressive improvement confirming the benefit of each our approach with the test on FIW dataset and the second track of RFIW2019 shows the state-of-the-art accuracy of our model. In the future, we seek to improve the models and to employ on other kinship recognition tasks.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: a system for large-scale machine learning. In *OSDI*, volume 16, pages 265–283, 2016.
[2] E. Ahmed, M. Jones, and T. K. Marks. An improved deep learning architecture for person re-identification. In *IEEE CVPR*, pages 3908–3916, 2015.
[3] A. Alvergne, C. Faurie, and M. Raymond. Father–offspring resemblance predicts paternal investment in humans. *Animal Behaviour*, 78(1):61–69, 2009.
[4] D. Aspandi, O. Pujol, F. Sukno, and X. Binefa. Fully end-to-end composite recurrent convolution network for deformable facial tracking in the wild. In *IEEE FG 2019*, In Press.
[5] L. Bottou. Large-scale machine learning with stochastic gradient descent. In *in COMPSTAT*, 2010.
[6] A. Bulat and G. Tzimiropoulos. How far are we from solving the 2d & 3d face alignment problem? (and a dataset of 230,000 3d facial landmarks). *2017 IEEE ICCV*, pages 1021–1030, 2017.
[7] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *IEEE FG*, 2018.
[8] C. R. A. Chaitanya, A. S. Kaplanyan, C. Schied, M. Salvi, A. Lefohn, D. Nowrouzezahrai, and T. Aila. Interactive reconstruction of monte carlo image sequences using a recurrent denoising autoencoder. *ACM Trans. Graph.*, 36(4):98:1–98:12, July 2017.
[9] D. CireşAn, U. Meier, J. Masci, and J. Schmidhuber. Multi-column deep neural network for traffic sign classification. *Neural networks*, 32:333–338, 2012.
[10] M. F. Dal Martello and L. T. Maloney. Lateralization of kin recognition signals in the human face. *Journal of vision*, 10(8):9–9, 2010.
[11] R. Fang, A. C. Gallagher, T. Chen, and A. Loui. Kinship classification by modeling facial feature heredity. In *20th IEEE ICIP*, pages 2983–2987. IEEE, 2013.
[12] R. Fang, K. D. Tang, N. Snavely, and T. Chen. Towards computational models of kinship verification. In *17th IEEE ICIP*, pages 1577–1580. IEEE, 2010.
[13] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In *IEEE CVPR*, pages 7132–7141, 2018.
[14] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.
[15] N. Liu, M. Zhang, H. Li, Z. Sun, and T. Tan. Deepiris: Learning pair-wise filter bank for heterogeneous iris verification. *Pattern Recognition Letters*, 82:154–161, 2016.
[16] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song. Sphereface: Deep hypersphere embedding for face recognition. In *IEEE CVPR*, volume 1, page 1, 2017.

[17] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou. Neighborhood repulsed metric learning for kinship verification. *IEEE T Pattern Anal*, 36(2):331–345, 2014.
[18] X. Mao, C. Shen, and Y.-B. Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 2802–2810. Curran Associates, Inc., 2016.
[19] O. M. Parkhi, A. Vedaldi, A. Zisserman, et al. Deep face recognition. In *bmvc*, volume 1, page 6, 2015.
[20] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. 2017.
[21] X. Qin, X. Tan, and S. Chen. Tri-subject kinship verification: Understanding the core of a family. *IEEE T Multimedia*, 17(10):1855–1867, 2015.
[22] J. P. Robinson, M. Shao, Y. Wu, and Y. Fu. Families in the wild (fiw): Large-scale kinship image database and benchmarks. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 242–246. ACM, 2016.
[23] J. P. Robinson, M. Shao, Y. Wu, H. Liu, T. Gillis, and Y. Fu. Visual kinship recognition of families in the wild. *IEEE T Pattern Anal*, 2018.
[24] 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*, pages 5–12. ACM, 2017.
[25] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *IEEE ICCVW*, pages 397–403, Dec 2013.
[26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*, 15(1):1929–1958, 2014.
[27] C. Szegedy, S. Ioffe, and V. Vanhoucke. Inception-v4, inception-resnet and the impact of residual connections on learning. *CoRR*, abs/1602.07261, 2016.
[28] S. Wang, J. P. Robinson, and Y. Fu. Kinship verification on families in the wild with marginalized denoising metric learning. In *12th IEEE FG 2017*, pages 216–221. IEEE, 2017.
[29] Y. Wen, K. Zhang, Z. Li, and Y. Qiao. A discriminative feature learning approach for deep face recognition. In *ECCV*, pages 499–515. Springer, 2016.
[30] S. Xia, M. Shao, and Y. Fu. Kinship verification through transfer learning. In *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.
[31] S. Xia, M. Shao, J. Luo, and Y. Fu. Understanding kin relationships in a photo. *IEEE T Multimedia*, 14(4):1046–1056, 2012.
[32] C. Xiong, L. Liu, X. Zhao, S. Yan, and T.-K. Kim. Convolutional fusion network for face verification in the wild. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(3):517–528, 2016.
[33] H. Yan, J. Lu, and X. Zhou. Prototype-based discriminative feature learning for kinship verification. *IEEE T Cybernetics*, 45(11):2535–2545, 2015.
[34] S. Zafeiriou, G. Trigeorgis, G. Chrysos, J. Deng, and J. Shen. The menpo facial landmark localisation challenge: A step towards the solution. In *IEEE CVPRW 2017*, pages 2116–2125, July 2017.
[35] X. Zhou, J. Hu, J. Lu, Y. Shang, and Y. Guan. Kinship verification from facial images under uncontrolled conditions. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 953–956. ACM, 2011.
[36] Y. Zhou, D. Liu, and T. Huang. Survey of face detection on low-quality images. In *13th IEEE FG 2018*, pages 769–773, May 2018.