Kinship Representation Learning with Face Componential Relation - Supplemental Material -

Wen-Tai Su*
Novatek Microelectronics Corp.

Min-Hung Chen*
NVIDIA

Chien-Yi Wang*
NVIDIA

Shang-Hong Lai*
National Tsing Hua University

Trista Chen
Microsoft Corp.

1. Supplementary

This supplementary material provides more details including: 1) the detailed distribution of the FIW dataset (Sec. 1.1), 2) the additional evaluation results of full comparison of all 11 kinship categories (Sec. 1.2), 3) some hard sample cases in practical kinship recognition protocol (Sec. 1.3), 4) the visualization result on kinship similarity (Sec. 1.4), and 5) the limitations and some failure cases (Sec. 1.5).

1.1. Data Distribution of the FIW Dataset

The FIW dataset includes 11 kin relationship types as: a) *Siblings*: Brother-Brother (BB), Sister-Sister (SS), and Sister-Brother (SIBS); b) *Parent-Child*: Father-Daughter (FD), Mother-Daughter (MD), Father-Son (FS), and Mother-Son (MS); c) *Grandparent-Grandchild*: GFGD, GFGS, GMGD, and GMGS, with the same naming convention as above. In Tables 1 and 5 of the main paper, we mainly focus on the first 7 kinship relationships since the Grandparent-grandchild categories contain much smaller data by an order of magnitude, as shown in Fig. 1.

1.2. Full Comparison of All Kinship Categories

To supplement the results of Tables 1 and 5 of the main paper, we provide a full comparison for the FIW dataset of all 11 kinship categories in Table 1 and Table 2, including the addition of *Grandparent-Grandchild* categories: GFGD, GFGS, GMGD, and GMGS. Since the Grandparent-Grandchild categories have only one-tenth of the data of the other categories, there is not enough data for model training and inference. This is the potential reason why our FaCoRNet has sub-optimal performance in the Grandparent-Grandchild categories.

*Work was done during Microsoft

1.3. Hard Sample Cases in Practical Protocol

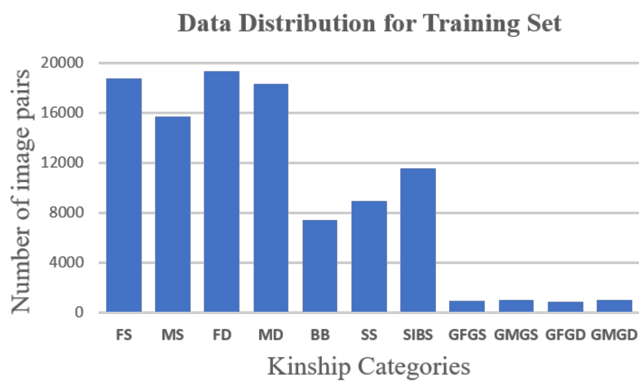
The result of *FaCoRNet* in the MD case has a significant improvement of 2.4 percent (0.818 → 0.842) from the standard to the quality-filtered protocol (see Table 2 (b)), showing that the MD cases include a large amount of low-quality face images in the standard protocol. On the other hand, MD has slightly lower recognition accuracy than FS, and we conjecture that it is due to the challenging MD cases caused by makeup and coverings as shown in Fig. 2. Moreover, the accuracy of the SIBS case decreases after selecting high-quality face images. The main reason is that SIBS has less data than other kinship categories as illustrated in Fig. 1.

1.4. Visualization on Kinship Similarity

We also show the kinship correlation map between two families as shown in Fig. 3. Each family includes 4 categories and each category includes 2 samples (i.e., 2 $\mathbf{F_1}$). We calculate the cosine similarity across each sample, including the same and different families. The kinship categories from same family members (i.e., the same subscript: Father ($\mathbf{F_1}$), Mother ($\mathbf{M_1}$), Son ($\mathbf{S_1}$), and Daughter ($\mathbf{D_1}$)) have higher cosine similarity score. In contrast, the kinship categories from different family members have smaller cosine similarity scores (i.e., the Father-Son relationship from $\mathbf{F_1}$ and $\mathbf{S_2}$).

1.5. Limitations

We enumerate some of the failure cases in Fig. 4. Among these hard samples, some face pairs are captured in different scenarios and ages, resulting in large variants in illumination, expression, and pose; except for the low-quality face image case as mentioned above, some of them are obscured and covered (i.e., bread and wearing glasses), making the kinship recognition challenging.



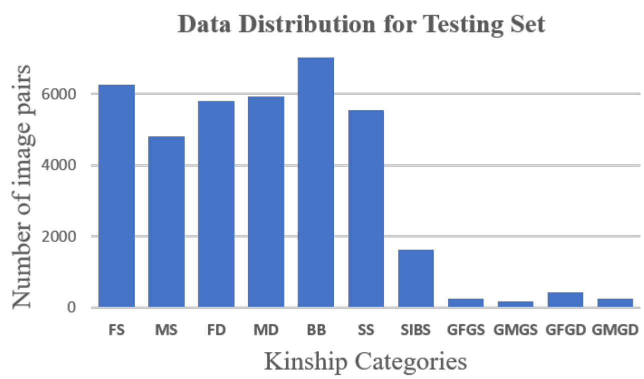


Figure 1. Illustration of the data distribution on the FIW dataset. The x-axis is the 11 kinship categories and the y-axis is the number of image pairs, respectively. The left and right figures represent the data distribution for the training and testing set, respectively.

Method	BB	SS	SIBS	FD	MD	FS	MS	GFGD	GMGD	GFGS	GMGS	AVG.
(a) Pre-trained model: ArcFace [1]												
Stefhoer† [2]	0.660	0.650	0.760	0.770	0.770	0.800	0.780	0.700	0.640	0.730	0.600	0.740
DeepBlueAI† [4]	0.770	0.770	0.750	0.740	0.750	0.810	0.740	0.720	0.670	0.730	0.680	0.760
Ustc-nelslip† [7]	0.750	0.740	0.720	0.760	0.750	0.820	0.750	0.790	0.760	0.690	0.670	0.760
Vuvko† [6]	0.800	0.800	0.770	0.751	0.780	0.810	0.740	0.780	0.760	0.690	0.690	0.780
Contrastive [8]	0.803	0.829	0.794	0.753	0.803	0.823	0.751	0.754	0.740	0.702	0.592	0.793
FaCoRNet (Ours)	0.820	0.833	0.810	0.773	0.804	0.826	0.788	0.774	0.706	0.702	0.587	0.806
(b) Pre-trained model: AdaFace [3]												
Contrastive [8] (naive)	0.630	0.776	0.731	0.663	0.687	0.736	0.687	0.722	0.665	0.669	0.525	0.728
Contrastive [8]	0.821	0.831	0.798	0.766	0.806	0.828	0.767	0.756	0.725	0.669	0.626	0.802
FaCoRNet (Ours)	0.832	0.836	0.824	0.795	0.818	0.848	0.802	0.799	0.684	0.690	0.575	0.820

Table 1. The state-of-the-art performance comparison of *Kinship Verification* on FIW dataset in all 11 kinship categories by two pre-trained backbones: (a) ArcFace [1] and (b) AdaFace [3]. †The results are from [5].

Method	BB	SS	SIBS	FD	MD	FS	MS	GFGD	GMGD	GFGS	GMGS	AVG.
(a) Standard Protocol												
Contrastive [8]	0.803	0.829	0.794	0.753	0.803	0.823	0.751	0.754	0.740	0.702	0.592	0.793
FaCoRNet (Ours)	0.832	0.836	0.824	0.795	0.818	0.848	0.802	0.799	0.684	0.690	0.575	0.820
(b) Quality-Filtered Protocol (Quality Score > 0.5)												
Contrastive [8]	0.800	0.817	0.772	0.739	0.784	0.836	0.786	0.791	0.737	0.703	0.691	0.792
FaCoRNet (Ours)	0.836	0.838	0.784	0.784	0.842	0.862	0.815	0.810	0.686	0.715	0.651	0.826

Table 2. Performance comparison of kinship on FIW dataset in all 11 kinship categories for two quality-filtered protocols: (a) standard protocol: use all image pairs without filtering; (b) quality-filtered protocol: select the image pairs with the pair quality scores larger than 0.5, which is more practical in real-world scenarios. AdaFace is used as the pre-trained model here.

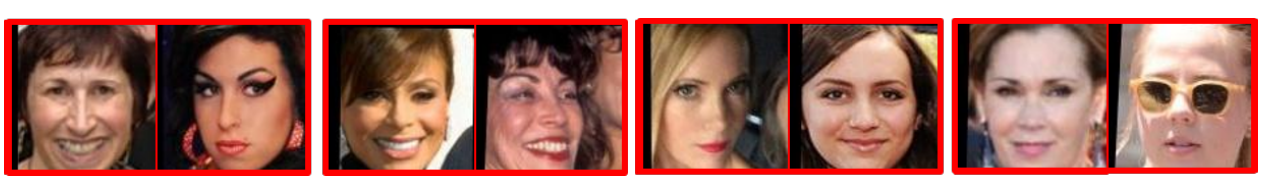


Figure 2. Illustration of the hard samples in the Mother-Daughter (MD) case. This figure shows that the face has makeup, glasses, etc., which makes it challenging to identify the kin relation.

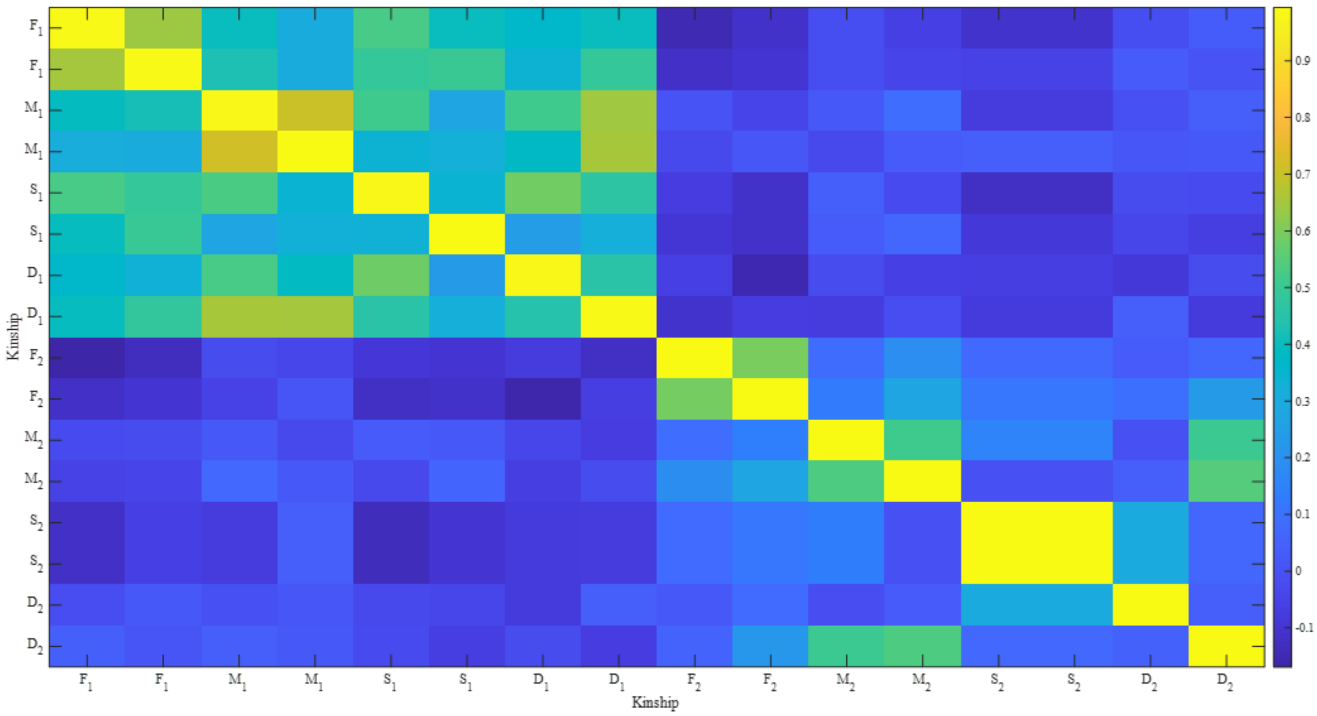


Figure 3. Visual analysis of the correlation map with two family groups (subscript 1 and 2) and 4 kinship categories.

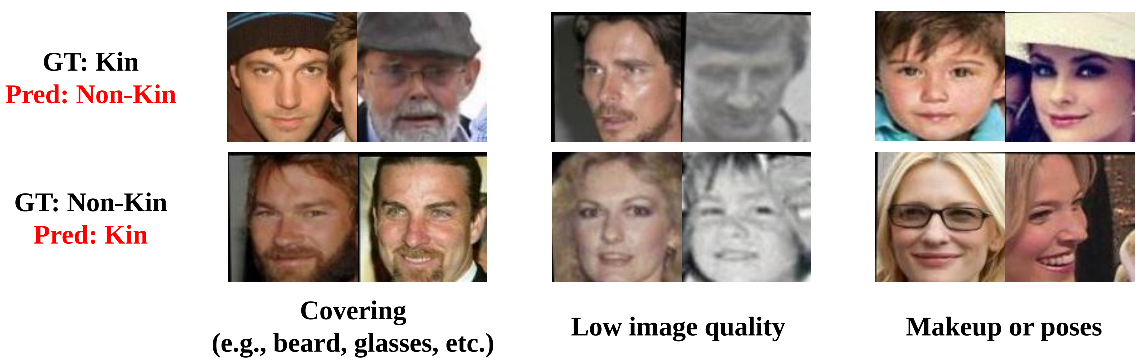


Figure 4. Illustration of failure cases on FIW dataset include: The left side shows the cases where the face is covered by a beard, glasses, etc; The middle shows low-quality images of a face; The right-hand side shows the face with makeup and different poses.

Besides, in the case of low-quality face images, as long as one image in the pair is of poor quality, the recognition result will be seriously affected. Finally, extreme poses also cause difficulty for kinship recognition due to less face component information.

References

[1] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4690–4699, 2019. 2

[2] Stefan Hörmann, Martin Knoche, and Gerhard Rigoll. A multi-task comparator framework for kinship verification. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 863–867. IEEE, 2020. 2

[3] Minchul Kim, Anil K Jain, and Xiaoming Liu. Adaface: Quality adaptive margin for face recognition. In *Proceedings of*

the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 18750–18759, 2022. 2

[4] Zhipeng Luo, Zhiguang Zhang, Zhenyu Xu, and Lixuan Che. Challenge report recognizing families in the wild data challenge. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 868–871. IEEE, 2020. 2

[5] Joseph P Robinson, Ming Shao, and Yun Fu. Survey on the analysis and modeling of visual kinship: A decade in the making. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8):4432–4453, 2022. 2

[6] Andrei Shadrikov. Achieving better kinship recognition through better baseline. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 872–876. IEEE, 2020. 2

[7] Jun Yu, Mengyan Li, Xinlong Hao, and Guochen Xie. Deep fusion siamese network for automatic kinship verification. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 892–899. IEEE, 2020. 2

[8] Ximiao Zhang, XU Min, Xiuzhuang Zhou, and Guodong Guo. Supervised contrastive learning for facial kinship recognition. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, 2021. 2