KFC: Kinship Verification with Fair Contrastive Loss and Multi-Task Learning

Jia Luo Peng
garynlfd@gapp.nthu.edu.tw

Keng Wei Chang
percyx987654321@gapp.nthu.edu.tw

Shang-Hong Lai
lai@cs.nthu.edu.tw

Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan

Abstract

Kinship verification is an emerging task in computer vision with several potential applications. However, there is a lack of large kinship datasets to train a discriminative and robust model, which is a major limitation for this problem. Moreover, face verification is known to exhibit bias in skin colors and ethics, which has not been fully resolved by previous works. In this paper, we propose a multi-task learning model structure with attention module to improve the accuracy, which surpasses state-of-the-art performance. In addition, our fairness-aware contrastive loss function combined with a debias term and adversarial learning greatly mitigates racial bias, thus significantly improving the fairness. In the experiment, we build a large dataset by combining several existing kinship datasets. Exhaustive experimental evaluation demonstrates the effectiveness and superior performance of the proposed KFC in both standard deviation and accuracy at the same time. Code is available at *https://github.com/garynlfd/KFC*.

1 Introduction

Facial kinship verification[41][27] has been an emerging research area, which can be used for missing children, forensic analysis, multimedia analysis, and personal photos app management (e.g., creating a family tree). Human faces contain not only family information but also racial traits, but racial traits often get ignored while researchers usually focus more on accuracy performance. However, such unfairness has been an infamous problem in AI systems and involves in many social problems in recent years, e.g., healthcare system[33], hiring algorithm[12], recidivism judgement[15], etc.

Many previous works have developed methods for improving fairness in face recognition and face verification. For example, [22][42] proposed fairness-aware loss functions. The former exploits the interrelation between anchor and sample to design a sensitive attribute removing loss function, while the latter uses instance FPR in loss function to constrain bias. [39] implemented post-processing data perturbation without changing their parameters and structures that can hide the information of protected attributes. [11] proposed to include demographic-adaptive layers that make the model generate face representations for every demographic group.
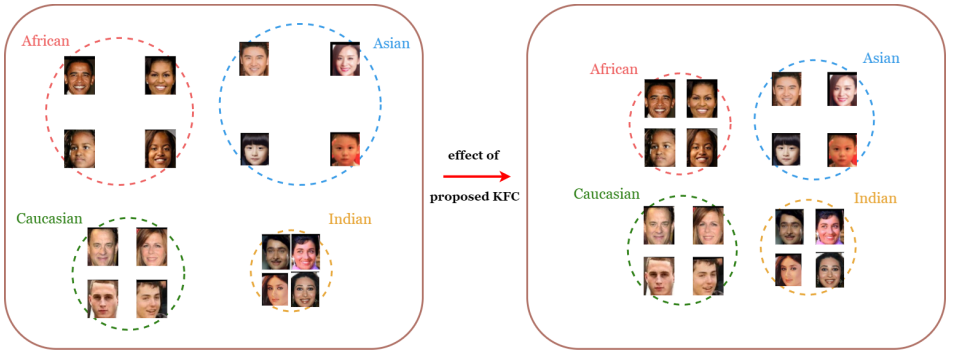
Figure 1: The schematic diagram for improving the fairness in kinship verification. Our method can effectively adjust the intra-class compactness and inter-class discrepancy in the feature space. We mitigate racial bias by balancing four races' intra-class and inter-class angle and making them as consistent as possible.

Kinship verification is a subfield of face verification that involves identifying the degree of relatedness between individuals based on their unique facial features. Siamese neural network[5][4][32][2] is an intuitive network that uses weight-sharing structure to extract features, which is an effective kinship verification network that focuses on similar facial traits. Attention scheme[21][51][43] is another popular method in kinship verification. Since facial features are crucial for determining the relationship between two face images, it is important to make the model focus on the most critical regions.

Based on the above idea, we propose an approach to improve racial fairness while achieving higher accuracy in kinship verification. Furthermore, to deal with fairness issues and small dataset constraint, we combine multiple kinship datasets and label every individual's race, constructing a large kinship dataset with race label, called KinRace dataset. To boost accuracy and fairness at the same time, we propose a fairness-aware loss function in a multi-task learning framework. To improve the kinship verification accuracy, we use attention module that makes model focus on the most representative facial regions for feature representation learning. For fairness, we reverse the gradient of race classification branch to remove the racial information in the feature vector and design a fairness-aware contrastive loss function which can mitigate pairwise bias and significantly decrease the standard deviation in four races.

With the above feature learning schemes, our method successfully finds a good balance for the distribution of the intra-class and inter-class distances for all the races. As shown in Figure 1, the proposed feature learning approach effectively improves racial fairness and equilibrates the compactness for all races. Compared with the previous works, our method firstly has an innovative model structure to tackle the kinship verification task. Secondly, it utilizes two different debias techniques to improve the fairness. Finally, all these methods are integrated to enhance both fairness and accuracy performance on our integrated dataset.

In summary, the main contributions of this paper are three-fold:

- To the best of our knowledge, this is the first work proposed to mitigate bias and achieve state-of-the-art accuracy simultaneously for kinship verification.

- Our fairness-aware contrastive loss function automatically mitigates the pairwise bias

and balances the degree of compactness of every race, which improves racial fairness.

- We build a large kinship dataset with racial labels from several public kinship datasets, which is targeted for the research on the fairness of kinship verification.

2   Related Work
2.1   Kinship Verification

Kinship verification has attracted a number of researchers in the past few years, due to the emergence of datasets[28][29][19][18][9][25], challenges[27], and the potential in real-world applications. Traditional methods used hand-crafted features to extract local texture information. [1] used binary code to extract local binary patterns which are invariant to grey scale transformation. [3] used two texture descriptors (LDP and LPQ) for feature extraction, which produces better performance than the previous approaches. Shallow metric learning method[6] selected critical feature patches and discards trivial ones by incorporating AdaBoost and SVM to create a discriminative model.

Deep learning methods can extract effective features automatically, which largely enhance the accuracy. [44] proposed a feature fusion method that uses discriminative features from backbone network and fuses the features to determine if two face images are with kinship relationship or not. [46] adopted ArcFace[7] as the backbone model pre-trained on MS-Celeb-1M[13] to obtain more representative features. Moreover, they used supervised contrastive loss function to contrast samples against each other, thus enhancing the ability to distinguish the kinship relation.

2.2   Bias Mitigation

There are many ways to alleviate racial bias, e.g., designing a new algorithm, modifying the model structure, proposing a loss function, etc. [24] tried to create fair features by using adversarial learning. Gradient reversal layer reverses the gradient of the classification head for sensitive attributes so that the encoder can generate features that do not include sensitive information. [10] also used adversarial learning to achieve discriminative feature representation, while they disentangle features into four attributes in order to retain the crucial attributes and remove the unfair ones.

In addition, [11] proposed adaptive layer in the model structure to improve robustness of representations for every demographic group, and used automation module to decide whether to utilize adaptive layer in a given layer. [42] modified the softmax loss function to mitigate bias while improving accuracy. They used instance FPR as an extreme case of demographic FPR, which does not need the label of demographic group. [35] focused on identity bias instead of demographic group bias and combined CosFace[36] with bias difference to create a new loss function. By minimizing identity bias without using the sensitive attribute labels, they can improve the fairness between demographic groups.

Inspired by the above works, we combine these two domains and improve fairness and accuracy at the same time. We propose attention mechanism and multi-task model structure, while using adversarial learning and incorporating a fairness term in the loss function.

3   Dataset Construction

We combine six kinship datasets to increase kinship verification dataset size and name the dataset KinRace. There are four small datasets: CornellKin[8], UBKinFace[28][29][19],

KinFaceW-I[13], KinFaceW-II[13], and two large datasets: Family101[9], FIW[25]. Because most datasets only label four kinship types: Father-Son(FS), Father-Daughter(FD), Mother-Son(MS), and Mother-Daughter(MD), we only use these four kinship types in our mixed dataset. For Family101[9], some identities have too many images, so we limit the total number of images for each identity to at most 30. To deal with racial fairness, we manually label every identity's race. Following previous racial datasets[[37], [38]], we have four races in our dataset: African, Asian, Caucasian, and Indian.

Other-race-effect (ORE) is a phenomenon that humans are better at recognizing faces of their race more readily than other races. Due to ORE, our labeling process was done by three different racial annotators (African, Asian, Caucasian). If two or three annotators label the same race to an identity, we take the values as ground truth. If all three annotators have different labeling results, we will not use this identity in our dataset. Table 1 is the data distribution of our KinRace dataset. The total number of our image pairs is the largest among the existing kinship datasets, which means we can have more representative features and more generalizable experimental results. The race distribution is also similar to BUPT-Globalface [38] whose racial distribution is approximately the same as the real distribution of the world's population. For the correctness of experiments, our positive pairs (in the same family) must be the same race, which means we remove those mixed race samples.

	African	Asian	Caucasian	Indian	sum
CornellKin[8]	8	56	72	3	139
UBKinFace[23][29][19]	18	192	173	0	383
KinFaceW-I[13]	19	327	172	4	522
KinFaceW-II[13]	55	96	788	35	974
Family101[9]	5554	6540	82820	25374	120288
FIW(train)[25]	8353	3841	59028	681	71903
FIW(val)[25]	2087	1398	30147	0	33632
FIW(test)[25]	1231	799	9665	97	11792
sum	17325	13249	182865	26194	239633
percent	7.23%	5.53%	76.31%	10.93%	100%

Table 1: dataset race distribution

4 Proposed Method

The core idea of the proposed approach is to mitigate racial bias while improving accuracy in kinship verification task. In this section, we first introduce our model structure in detail and present why it can improve accuracy. Secondly, we analyze our fair contrastive loss function. Finally, we explain why the methods are effective for fairness and explore the reasons why they benefit racial fairness.

4.1 Model Structure

The proposed framework is illustrated in Figure 2. For input images $(I_1, I_2) \in \mathbb{R}^{H \times W \times 3}$, we have middle layer feature maps $(m_1, m_2) \in \mathbb{R}^{H \times W \times C}$ and the outputs of backbone feature embeddings $(e_1, e_2) \in \mathbb{R}^C$. The model has a multi-task structure, the upper row is in charge of kinship verification task while the lower row predicts the race of input images.

In the upper branch, we pass feature maps into attention module to obtain representative features, and then pass them into debias layer to generate a debias term in our loss function. For our attention module in Figure 3, we refer to [51] and modify their structure to reach
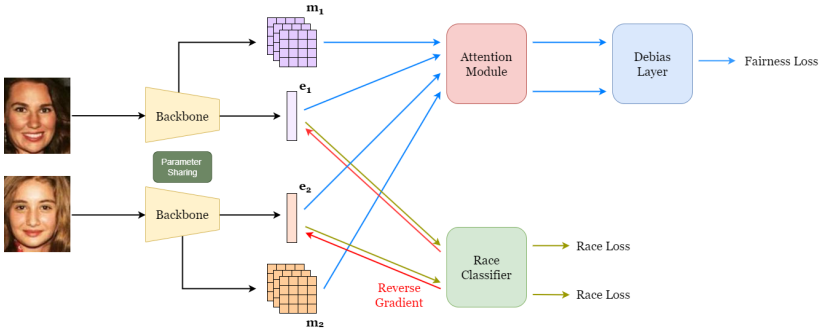
Figure 2: Overview of the proposed KFC model structure. e_1 and e_2 are feature vectors from backbone, and they are used for race prediction with gradient reversal. m_1 and m_2 are feature maps computed from backbone. They are passed to attention module to extract important semantics and then passed to the debias layer to generate debias term in loss function.

better accuracy. We first use average pooling to obtain channel features (s_1, s_2) from feature map (m_1, m_2) and then use 1x1 convolution layer to get feature vectors (s_1', s_2') to create attention map rather than only using one 1x1 convolution layer. This modification can not only improve the computational efficiency for downstream tasks but also allow for the interaction of different channels within a layer. Secondly, we make use of the effective attention impact of CBAM[40]. In the original method[31], they use two layers of 1x1 convolution to obtain feature vectors (c_1, c_2), however, we use CBAM[40] that contains channel and spatial attention, which focuses on the most relevant features of an input and generates more sophisticated features.

Note that we also apply CBAM[40] to our backbone in the last three layers($\mathbb{R}^{7 \times 7 \times 512}$) of ResNet101[14] to enhance the output feature maps and vector. Behind attention module, we pass the concatenation of two feature vectors $(e_1 \cdot c_1, e_2 \cdot c_2)$ to debias layer which is a simple linear layer for generating the debias term in loss function.

In the lower branch, we use feature vectors to predict the race. By using two layers of MLP as race classifier, we calculate cross entropy loss between the output and race ground truth. To alleviate bias, the gradient of the race classifier is reversed to remove racial information in feature vectors and decrease standard deviation. However, if the gradient is not reversed, the model can greatly improve kinship verification accuracy through multi-task learning.

4.2 Loss Function

Supervised contrastive loss[16] learns a similarity metric between samples in different classes, pulling samples from the same class closer and pushing samples from different classes apart. It outperforms other loss functions in high-dimensional feature spaces and focuses on hard samples, leading to higher accuracy. Following prior work[46] in kinship verification task, supervised contrastive loss can be formulated as

$$L = \frac{1}{2n} \sum_{i=1}^{n} [L_c(x_i, y_i) + L_c(y_i, x_i)] \tag{1}$$
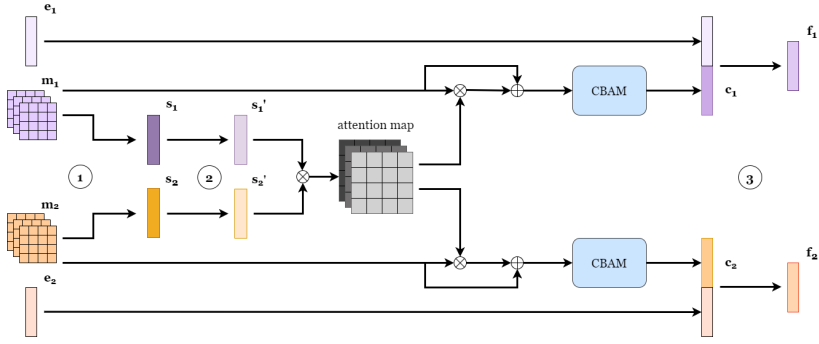
Figure 3: Overview of the proposed attention module. By using the feature maps computed from the two face images to construct an attention map and employing CBAM to extract feature vectors, we concatenate the outputs with the feature embeddings to create the final output features. ①: average pooling. ②: 1x1 Conv with ReLU. ③: 2 layers of 1x1 Conv with ReLU

where

$$L_c(x_i, y_i) = -log \frac{e^{s(x_i,y_i)/\tau}}{\sum_{j=1}^{n}[e^{s(x_i,x_j)/\tau} + e^{s(x_i,y_j)/\tau}]} \tag{2}$$

where $s(x,y)$ is defined as the cosine similarity between x and y, and (x_i, y_i) are positive pairs while $(x_i, x_j), (x_i, y_j)(j \neq i)$ are negative pairs.

[55] proposed a loss function aiming to reduce identity bias in face recognition, rather than demographic bias. They pass two identities' (x_i, x_j) feature maps into a debias layer which is a fully-connected layer to extract final feature maps (f_i, f_j), and then use them to calculate the estimated bias between two identities:

$$cos(M(f_m), M(f_i))^2 - cos(M(f_m), M(f_j))^2 = \varepsilon \tag{3}$$

where $f_m = \frac{1}{2}(f_i + f_j)$ and $M(.)$ is the debias layer, and they claim that if $\varepsilon > 0$ means i has large biases than j, and vice versa. They combine the estimated bias with the cosine value and the margin in CosFace[36], which improves the fairness in face recognition models.

Sparked by the previous works, we propose fairness-aware contrastive loss function:

$$L_{fairness} = -log \frac{e^{(cos(x_i,y_i)-b_i)/\tau}}{\sum_{j\neq i}^{N}[e^{cos(x_i,x_j)/\tau} + e^{cos(x_i,y_j)/\tau}] + e^{(cos(x_i,y_i)-b_i)/\tau}} \tag{4}$$

where b_i is averaging ε in Equation 3 representing the bias between images I_i and other images in the same batch. Following the gradient analysis in [34], if we minus cosine value with bias term b_i, both of the gradients of positive and negative pairs would become larger, which is similar to the margin penalty in CosFace[36]. However, the margin is fixed in CosFace[36], while b_i changes with identity pairs. After subtracting the bias term in loss function, the greater gradients help each pair balances its unfairness situation, which can make the compactness degree of every race as consistent as possible. Furthermore, the temperature τ in the original contrastive loss can further tackle with the hard samples. Combining temperature τ with the debias term, we take into account both accuracy and fairness.

In the race classifier branch, we use cross entropy(CE) loss to train the race classifier:

$$L_{race} = -\sum_{i=1}^{n} t_i log(p_i) \tag{5}$$

where n is the total number of races in the training dataset, t_i is the ground truth label, and p_i is the softmax probability for the ith class. Combining these two loss functions, we have our total loss function given by

$$L_{total} = L_{fairness} + L_{race} \qquad (6)$$

4.3 Gradients of Fair Contrastive Loss Function

Wang and Liu[54] analyzed the gradients with respect to positive samples (x_i, x_i) and different negative samples $(x_i, x_j)(j \neq i)$:

$$\frac{\partial L(x_i)}{\partial cos(x_i, x_i)} = -\frac{1}{\tau}\sum_{k \neq i}P_{i,k}, \quad \frac{\partial L(x_i)}{\partial cos(x_i, x_j)} = \frac{1}{\tau}P_{i,j} \qquad (7)$$

where $P_{i,j}$ is the probability of x_i and x_j being recognized as positive pair:

$$P_{i,j} = \frac{e^{(cos(x_i,x_j))/\tau}}{\sum_{k \neq i}e^{(cos(x_i,x_k))/\tau} + e^{(cos(x_i,x_i))/\tau}} \qquad (8)$$

In our case, we subtract a bias term in positive pair, which makes $P_{i,j}$ as:

$$P_{i,j} = \frac{e^{(cos(x_i,x_j))/\tau}}{\sum_{k \neq i}e^{(cos(x_i,x_k))/\tau} + e^{(cos(x_i,x_i)-b_i)/\tau}} \qquad (9)$$

When bias b_i is positive, $P_{i,j}$ becomes larger, which makes the gradients of positive pair and negative pair in Equation 7 larger, and vice versa.

We have analyzed the average bias value during training, and it is positive. This means that both the gradients of our positive and negative pairs become larger, which makes the training more efficient and helps it converge faster.

4.4 Fairness Mechanism

This work employs two methods for improving fairness: adversarial learning and fair loss function. We use a race classifier in adversarial learning to remove racial information from feature vectors, which decreases standard deviation. However, due to the small dataset size compared to popular benchmark datasets and the difficulty to train a robust network using adversarial training, our results are not as outstanding as the previous works. As a result, we further modify the contrastive loss function to a fairness-aware loss function. As described in subsection 4.2, our loss function can eliminate bias in every two identities which indirectly improves the fairness in every race. Moreover, these two methods can decrease standard deviation independently which is shown in the next section, and the standard deviation becomes smaller while the accuracy is improved when we combine them together.

5 Experiment

5.1 Experimental Setting

Dataset.   In this study, we employ KinRace dataset for training, validation, and testing. The dataset was divided into three sets with no overlapping families, and the ratio of the four races in each set was similar to that in Table 1. We choose four kinship relationship in our dataset: Father-Son(FS), Father-Daughter(FD), Mother-Son(MS), Mother-Daughter(MD).

Face images were resized to 112 x 112 using MTCNN[45] for face detection and alignment during preprocessing.

Implementation Details. We adopt ResNet101[14] with ArcFace[35] as the pre-trained backbone network. The sizes of feature map and feature vector are $\mathbb{R}^{7\times7\times512}$ and $\mathbb{R}^{512}$, respectively. We set the temperature parameter τ to 0.08 based on the analysis in [46]. The models are trained with SGD algorithm, with momentum 0.9 and weight decay 1e-4. We train the networks for 10 epochs with 60000 iterations and set batch size to 25. We train a baseline[46] which won the first place in RFIW 2021[27] to compare with our method. In the experiment below, if we mention **adversarial**, it means we reverse the gradient of race classification like the red line in the indication in Figure 2. If we mention **multi-task**, it means we do not reverse the gradient of race classification, instead we just train the model normally with the green line in the indication in Figure 2.

5.2 Ablation Study

Effect of improving accuracy. We conduct experiments at three techniques that improves accuracy in our methods and compare them with the baseline[46]. The result is shown in Table 2. By using attention mechanism and CBAM to extract informative features and incorporating race classification branch, our model can learn features containing racial information, leading to a nearly 7% improvement in accuracy over the baseline. Additionally, the debias layer helps to rectify bias and make the compactness degree more uniform, resulting in a slight accuracy improvement and reduced standard deviation.

Methods%	African	Asian	Caucasian	Indian	Avg	Std
baseline	82.18	83.71	78.00	80.70	79.08	2.43
KFC(attention)	84.31	85.99	81.59	79.28	81.74	2.96
KFC(attention+multi-task)	85.03	84.39	87.01	78.95	85.82	3.45
KFC(attention+multi-task+debias layer)	85.16	84.94	86.67	81.32	**85.86**	**2.27**

Table 2: Ablation study on accuracy.

Methods%	African	Asian	Caucasian	Indian	Avg	Std
baseline	82.18	83.71	78.00	80.70	79.08	2.43
KFC(adversarial)	78.45	81.51	79.11	76.36	78.88	2.10
KFC(debias layer)	80.35	80.62	78.67	77.35	78.74	1.53
KFC(adversarial+debias layer)(acc best)	81.61	82.68	82.90	82.37	**82.74**	0.56
KFC(adversarial+debias layer)(std best)	81.28	81.29	80.83	80.80	80.88	**0.27**

Table 3: Ablation study on standard deviation. *acc best* refers to the epoch with the highest accuracy on validation dataset, and *std best* refers to the epoch with the lowest standard deviation on validation dataset.

Effect of improving fairness. By reversing the gradient of race classification branch, the feature vector contains less racial information, which reduces the standard deviation slightly but also harms the accuracy. After adopting debias layer in our model, the debias term in loss function evidently mitigates the bias between identities, however, the accuracy goes down as well. Since these two methods can effectively improve fairness, we merge them as our final proposed method. Remarkably, the standard deviation greatly diminishes while boosting the accuracy. Firstly, the feature vector excludes racial information, which benefits from adversarial learning. Secondly, the debias layer becomes more robust because it can generate debias term depending on the most essential facial features while racial traits are removed. Table 3 show that our strategy enhances fairness while maintaining accuracy.

Method	African	Asian	Caucasian	Indian	Avg	Std
Vuvko[30]	71.13	73.32	72.61	76.19	72.96	3.40
Ustc-nelslip[44]	76.05	77.07	75.54	63.98	74.33	6.15
TeamCNU[46]	82.18	83.71	78.00	80.70	79.08	2.43
KFC(multi-task)	85.16	84.94	86.67	81.32	**85.86**	2.27
KFC(adversarial)	81.28	81.29	80.83	80.80	80.88	**0.27**

Table 4: Comparisons with SOTA methods on KinRace dataset

Method	C & YP	C & OP	Avg	Std
LPQ_ML[8]	-	-	73.25	-
StatBIF-SIWEDA-WCCN[9]	75.71	76.92	76.32	-
FAML[6]	78.30	75.00	76.54	-
BC²DA[6]	83.28	**82.69**	83.30	-
KFC(multi-task)	**87.24**	81.00	**84.12**	1.36
KFC(adversarial)	82.71	77.75	80.23	0.06

Table 5: Comparisons with SOTA methods on UB KinFace dataset. This dataset provides child, young and old parent. C & YP means child and young parent, and C & OP means child and old parent.

Method	FD	MD	FS	MS	Std
Vuvko[46]	75.00	78.00	81.00	74.00	-
Ustc-nelslip[46]	76.00	75.00	82.00	75.00	-
TeamCNU[46]	75.00	80.00	82.00	77.0	-
FaCoRNet(ArcFace)[51]	77.30	80.40	82.60	78.80	-
FaCoRNet(AdaFace)[51]	**79.50**	81.80	**84.80**	**80.20**	-
KFC(multi-task)	79.05	**83.61**	84.63	78.25	7.81
KFC(adversarial)	78.81	82.56	81.69	77.43	**5.57**

Table 6: Comparisons with SOTA methods on FIW dataset. Since we only focus on four kinship relations, we compare these four types with SOTA methods.

5.3 Comparisons with SOTA methods

In this section, we compare our method with three other works[30][44][46] that performed well in the RFIW challenge[26][27]. We re-implement these methods using their training settings and evaluate them on our KinRace dataset. Our proposed method outperforms these methods in both accuracy and standard deviation, as shown in Table 4. Our multi-task method achieves almost 13% gain for average accuracy, while our adversarial method reduce the standard deviation to 0.27 which tremendously improves fairness. In Figure 4, we calculate the standard deviation on our KinRace testing set every 10000 iteration on SOTA methods and our proposed KFC. It is obviously that the experimental result well demonstrates fairness of the proposed KFC method, which effectively decreases the standard deviation to a very low level.

To evaluate the generalization of our method, we evaluate our method on other datasets: UB KinFace[28][29][19] and FIW[25]. The results of UB KinFace[28][29][19] is shown in Table 5, our multi-task version outperforms previous works, while the adversarial version diminishes the standard deviation significantly. We follow the standard protocol[28][29][19] with five-fold cross-validation. In Table 6, we compare 4 kinship relations with SOTA methods. Although our accuracy may not be the highest, our results are competitive compared to FaCoRNet(AdaFace)[51], and our backbone is ArcFace, which actually makes our results better than FaCoRNet(ArcFace)[51]. Additionally, our adversarial version reduces the standard deviation compared to the multi-task version, as shown in Table 5 and Table 6, and more details will be provided in supplementary material. In these experiments, we demonstrate that our proposed KFC can perform well on datasets of different sizes, outperforming state-of-the-art methods and effectively decreasing standard deviation.

5.4 Visualization and Analysis on Fairness

To prove our proposed method can genuinely improve fairness, we analyze intra-class and inter-class angle between baseline[46] and our KFC. In Table 7, our KFC balances the intra-class and inter-class angle between four races, which makes the compactness degree comparable, and thus markedly improves fairness. As shown in Figure 5, the feature embeddings
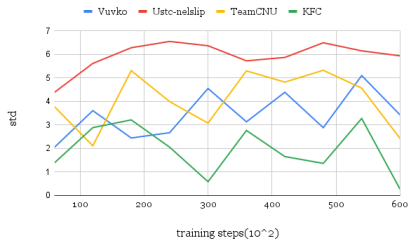
Figure 4: Comparison on standard deviation variations for different methods during training.
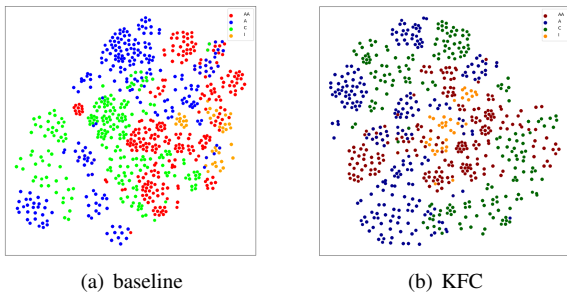

(a) baseline   (b) KFC

Figure 5: t-SNE visualizations. We randomly pick 400 pairs per race from KinRace dataset. AA for African, A for Asian, C for Caucasian, and I for Indian

for each race are distributed more evenly than the baseline[46]. In the baseline graph, Asian and Caucasian groups are mainly on the left while African and Indian groups are mainly on the right. Our adversarial learning approach removes clear boundaries between the four races, providing a more fair solution for kinship verification, supported by our experimental results and feature space analysis.

Methods		African	Asian	Caucasian	Indian	Std
baseline[46]	intra	17.21	18.32	15.36	10.08	3.68
	inter	46.85	52.34	40.85	44.20	4.85
ours	intra	12.21	16.59	15.65	11.36	**2.47**
	inter	41.35	49.93	41.04	42.80	**4.17**

Table 7: Intra-class and inter-class angle comparison. We randomly select 20 families per race from KinRace dataset.

6 Conclusion

This paper focuses on mitigating the issue of racial bias in kinship verification by using fair contrastive loss function and adversarial learning while improving the accuracy by utilizing attention module and multi-task learning. We also provide a kinship dataset KinFace with racial labels to broaden kinship research area to a more diversified scenario. Our experimental results suggest that our combination of model structure, gradient design, and fair loss function can significantly improve the racial fairness for kinship verification while automatically adjust intra- and inter-class angles in feature space, thus leading to standard deviation reduction and accuracy boost simultaneously.

References

[1] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. volume 3021, pages 469–481, 05 2004. ISBN 978-3-540-21984-2. doi: 10.1007/978-3-540-24670-1_36.

[2] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a "siamese" time delay neural network. In *Proceedings of the 6th International Conference on Neural Information Processing Systems*, NIPS'93, page 737–744, San Francisco, CA, USA, 1993. Morgan Kaufmann Publishers Inc.

[3] Abdelhakim Chergui, Salim Ouchtati, Hichem Telli, Fares Bougourzi, and Salah Eddine Bekhouche. Lpq and ldp descriptors with ml representation for kinship verification. 04 2018.

[4] Davide Chicco. *Siamese Neural Networks: An Overview*, pages 73–94. Springer US, New York, NY, 2021. ISBN 978-1-0716-0826-5. doi: 10.1007/978-1-0716-0826-5_3. URL https://doi.org/10.1007/978-1-0716-0826-5_3.

[5] S. Chopra, R. Hadsell, and Y. LeCun. Learning a similarity metric discriminatively, with application to face verification. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 539–546 vol. 1, 2005. doi: 10.1109/CVPR.2005.202.

[6] Lvye Cui and Bo Ma. Adaptive feature selection for kinship verification. In *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pages 751–756, 2017. doi: 10.1109/ICME.2017.8019326.

[7] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. ArcFace: Additive angular margin loss for deep face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):5962–5979, oct 2022. doi: 10.1109/tpami.2021.3087709. URL https://doi.org/10.1109%2Ftpami.2021.3087709.

[8] Ruogu Fang, Kevin D. Tang, Noah Snavely, and Tsuhan Chen. Towards computational models of kinship verification. In *2010 IEEE International Conference on Image Processing*, pages 1577–1580, 2010. doi: 10.1109/ICIP.2010.5652590.

[9] Ruogu Fang, Andrew C. Gallagher, Tsuhan Chen, and Alexander Loui. Kinship classification by modeling facial feature heredity. In *2013 IEEE International Conference on Image Processing*, pages 2983–2987, 2013. doi: 10.1109/ICIP.2013.6738614.

[10] Sixue Gong, Xiaoming Liu, and Anil K. Jain. Jointly de-biasing face recognition and demographic attribute estimation, 2020.

[11] Sixue Gong, Xiaoming Liu, and Anil K. Jain. Mitigating face recognition bias via group adaptive classifier, 2020.

[12] The Guardian. Amazon ditched ai recruiting tool that favored men for technical jobs, 2018. URL https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine. October 11, 2018.

[13] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition, 2016.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

[15] Lauren Kirchner Jeff Larson, Surya Mattu and Julia Angwin. How we analyzed the compas recidivism algorithm, 2016. URL https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. May 23, 2016.

[16] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning, 2021.

[17] Oualid Laiadi, Abdelmalik Ouamane, · Benakcha, Abdelmalik taleb ahmed, and Abdenour Hadid. A weighted exponential discriminant analysis through side-information for face and kinship verification using statistical binarized image features. *International Journal of Machine Learning and Cybernetics*, 12, 01 2021. doi: 10.1007/s13042-020-01163-x.

[18] Jiwen Lu, Junlin Hu, Xiuzhuang Zhou, Yuanyuan Shang, Yap-Peng Tan, and Gang Wang. Neighborhood repulsed metric learning for kinship verification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2594–2601, 2012. doi: 10.1109/CVPR.2012.6247978.

[19] S. Xia M. Shao and Y. Fu. Genealogical face recognition based on ub kinface database. In *Proc. IEEE CVPR Workshop on Biometrics (BIOM)*, 2011.

[20] Moumita Mukherjee and Toshanlal Meenpal. Binary cross coupled discriminant analysis for visual kinship verification. *Signal Processing: Image Communication*, 108:116829, 2022. ISSN 0923-5965. doi: https://doi.org/10.1016/j.image.2022.116829. URL https://www.sciencedirect.com/science/article/pii/S0923596522001229.

[21] Abhilash Nandy and Shanka Mondal. Kinship verification using deep siamese convolutional neural network. pages 1–5, 05 2019. doi: 10.1109/FG.2019.8756528.

[22] Sungho Park, Jewook Lee, Pilhyeon Lee, Sunhee Hwang, Dohyung Kim, and Hyeran Byun. Fair contrastive learning for facial attribute classification, 2022.

[23] Xiaoqian Qin, Dakun Liu, and Dong Wang. A novel factor analysis-based metric learning method for kinship verification. *Multimedia Tools Appl.*, 81(8):11049–11070, mar 2022. ISSN 1380-7501. doi: 10.1007/s11042-022-12032-w. URL https://doi.org/10.1007/s11042-022-12032-w.

[24] Edward Raff and Jared Sylvester. Gradient reversal against discrimination, 2018.

[25] Joseph P. Robinson, Ming Shao, Yue Wu, and Yun Fu. Families in the wild (fiw): Large-scale kinship image database and benchmarks. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 242–246. ACM, 2016.

[26] Joseph P. Robinson, Yu Yin, Zaid Khan, Ming Shao, Siyu Xia, Michael Stopa, Samson Timoner, Matthew A. Turk, Rama Chellappa, and Yun Fu. Recognizing families in the wild (rfiw): The 4th edition. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 857–862, 2020. doi: 10.1109/FG47880.2020.00138.

[27] Joseph P. Robinson, Can Qin, Ming Shao, Matthew A. Turk, Rama Chellappa, and Yun Fu. The 5th recognizing families in the wild data challenge: Predicting kinship from faces, 2021.

[28] J. Luo S. Xia, M. Shao and Y. Fu. Understanding kin relationships in a photo. *IEEE Transactions on Multimedia*, 14(4):1046–1056, 2012.

[29] M. Shao S. Xia and Y. Fu. Kinship verification through transfer learning. In *Proc. International Joint Conferences on Artificial Intelligence (IJCAI)*, pages 2539–2544, 2011.

[30] Andrei Shadrikov. Achieving better kinship recognition through better baseline. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, nov 2020. doi: 10.1109/fg47880.2020.00137. URL https://doi.org/10.1109%2Ffg47880.2020.00137.

[31] Weng-Tai Su, Min-Hung Chen, Chien-Yi Wang, Shang-Hong Lai, and Trista Pei-Chun Chen. Kinship representation learning with face componental relation, 2023.

[32] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014. doi: 10.1109/CVPR.2014.220.

[33] Starre Vartan. Racial bias found in a major health care risk algorithm, 2019. URL https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/. October 24, 2019.

[34] Feng Wang and Huaping Liu. Understanding the behaviour of contrastive loss, 2021.

[35] Fu-En Wang, Chien-Yi Wang, Min Sun, and Shang-Hong Lai. Mixfairface: Towards ultimate fairness via mixfair adapter in face recognition, 2022.

[36] Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu. Cosface: Large margin cosine loss for deep face recognition, 2018.

[37] Mei Wang, Weihong Deng, Jiani Hu, Xunqiang Tao, and Yaohai Huang. Racial faces in the wild: Reducing racial bias by information maximization adaptation network. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[38] Mei Wang, Yaobin Zhang, and Weihong Deng. Meta balanced network for fair face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11): 8433–8448, 2022. doi: 10.1109/TPAMI.2021.3103191.

[39] Zhibo Wang, Xiaowei Dong, Henry Xue, Zhifei Zhang, Weifeng Chiu, Tao Wei, and Kui Ren. Fairness-aware adversarial perturbation towards bias mitigation for deployed deep models, 2022.

[40] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module, 2018.

[41] Xiaoting Wu, Xiaoyi Feng, Xiaochun Cao, Xin Xu, Dewen Hu, Miguel Bordallo López, and Li Liu. Facial kinship verification: A comprehensive review and outlook. *Int. J. Comput. Vision*, 130(6):1494–1525, jun 2022. ISSN 0920-5691. doi: 10.1007/s11263-022-01605-9. URL https://doi.org/10.1007/s11263-022-01605-9.

[42] Xingkun Xu, Yuge Huang, Pengcheng Shen, Shaoxin Li, Jilin Li, Feiyue Huang, Yong Li, and Zhen Cui. Consistent instance false positive improves fairness in face recognition, 2021.

[43] Haibin Yan and Shiwei Wang. Learning part-aware attention networks for kinship verification. *Pattern Recognition Letters*, 128:169–175, 2019. ISSN 0167-8655. doi: https://doi.org/10.1016/j.patrec.2019.08.023. URL https://www.sciencedirect.com/science/article/pii/S0167865519302351.

[44] Jun Yu, Mengyan Li, Xinlong Hao, and Guochen Xie. Deep fusion siamese network for automatic kinship verification, 2020.

[45] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, oct 2016. doi: 10.1109/lsp.2016.2603342. URL https://doi.org/10.1109%2Flsp.2016.2603342.

[46] Ximiao Zhang, Min XU, 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)*, pages 01–05, 2021. doi: 10.1109/FG52635.2021.9666944.