Fact sheet: ECCV 2020 ChaLearn Looking at People 1st Fair Face Recognition Challenge

Please fill out the following sections carefully in a scientific writing style. Then, send the compressed project (in `.zip` format), i.e., the generated PDF, `.tex`, `.bib` and any additional files to `juliojj@gmail.com`, and put in the Subject of the email "ECCVW 2020 FairFaceRec Challenge / Fact Sheets", following the schedule and instructions provided in the Challenge webpage [1] (post-challenge/fact sheets).

I. Team details

- Team leader name: Kun Zhang
- Username on Codalab: zhaixingzi
- Team leader affiliation: None
- Team leader address: No2015, Wangshang Road,Binjiang District, Hangzhou, China
- Team leader phone number: +8618100182694
- Team leader email: m18637959467@163.com
- Name of other team members (and affiliation): Jun Ma, Yilun Hua, Zhubei Ge
- Team website URL (if any): None

II. Contribution details

A. Abstract

Recent advances in face recognition uncover the intrinsic vulnerability of the bias problem. For the sake of reducing bias and making it more robust in face recognition task, we developed a method that combines multiple models's prediction to get a more fair prediction result for different protected groups. These models were constructed with different CNN structures and trained with different domain datasets. And, our proposed approach ranked 5th in the current leaderboard.

B. Introduction and Motivation

Based on our observation, some general face recognition systems make more mistakes for specific skin color and gender when these faces with a large area occlusion or a large pose. To resolve this problem, we propose a practical method to reduce the bias caused by the aforementioned problem. Our method makes full use of the whole face's information with detailed description in section *Detailed method description*. [2] develop an algorithm to mitigate the hidden biases within training data. [3] propose a domain adaptation network to reduce racial bias in face recognition. [4] propose a disentangled representation by training an adversarial autoencoder to extract features that can capture identity discrimination and its complementary knowledge. [5] present a novel de-biasing adversarial network that learns to extract disentangled feature representations for both unbiased face recognition and demographics estimation. Compared with ours, our proposed is more simple to develop and practical.

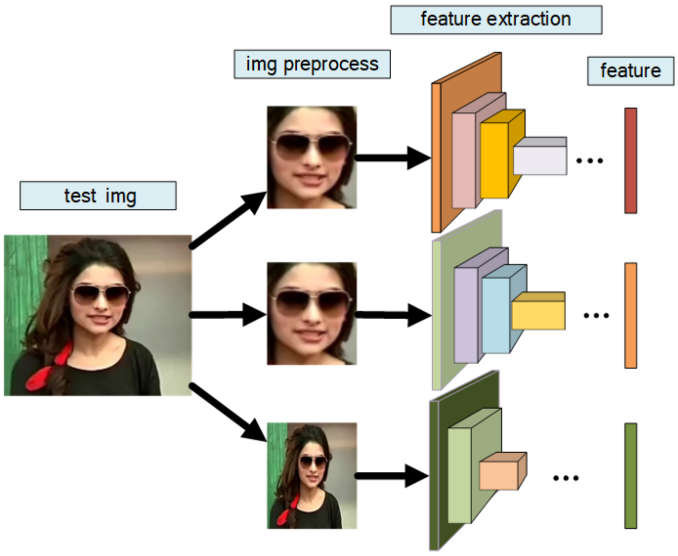

Fig. 1. face feature extration pipline

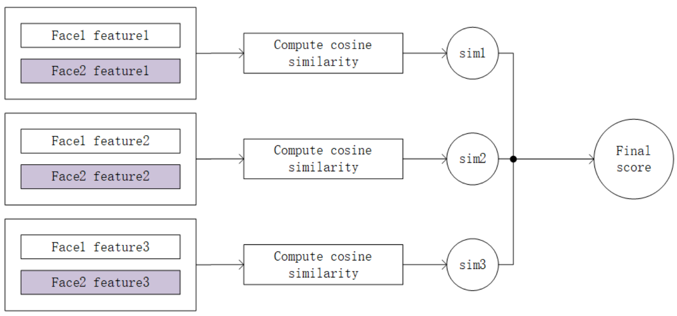

Fig. 2. get prediction score

C. Representative image / workflow diagram of the method

Fig.1 illustrates the pipe line of how to get the features for given ace image. Fig.2 describes that how to get the final prediction score.

D. Detailed method description

Preprocessing. We use two different method to preprocess the training data and test data.
1. All the faces in images and their landmarks are detected by the recently proposed RetinaFace[6]. We use 5 landmarks (two eyes, nose and mouth corners) for similarity transformation. When the detection fails, we simply resize it. All

the faces are cropped to 112×96 gray images. Then, each pixel in gray images is normalized by subtracting 127.5 then dividing by 128.
2. All the images are resized to 112×96 gray images. Then, each pixel in image is normalized, just as method1.

Training data. We use the public datasets, including MS1M-ArcFace[7] and IJB-C[8]. MS1M-ArcFace is preprocessed by method1, namely MS1M-ALIGNED. IJB-C is preprocessed by method1 and method2 respectively, namely IJBC-ALIGNED and IJBC-ORI.

Detailed settings in CNNs. We implement the CNN model using the mxnet[9] library. The model1 and model2 are the same architecture with resnet100. And, the model3 is resnet64. The model1 is trained with MS1M-ALIGNED. For model1, the learning rate starts from 0.1, and divided by 10 at the 100000, 137500, 170000 iterations, stop at 185000 iterations. The model2 is trained with IJBC-ALIGNED. For model2, the learning rate starts from 0.1, and divided by 10 at the 10000, 20000 and 30000 iterations, stop at 40000 iterations. The model3 is trained with IJBC-ORI. For model3, the learning rate starts from 0.1, and divided by 10 at the 10000, 20000 and 30000 iterations, stop at 40000 iterations.

All the three models are trained using arcface loss[10] with batch size of 1024 on 8 GPUs (V100) and used sgd optimizer. For model1, the s is 64 and m is 0.5 in the arcface loss, and weight decay is set to 0.0005. For model2 and model3, the s is 64 and m is 0.7 the arcface loss, and weight decay is set to 0.00005.

Detailed settings in testing. The face feature is taken from the output of the first fully-connected layer. we extract the features for each image and compute the cosine similarity for every pair in test template. For a pair, there are three similarities, because we have three models. After getting the three similarities, we sum up these similarities to get the final similarity score.

E. Challenge results and final remarks

TABLE I
LEADERBOARD: RESULTS OBTAINED BY THE PROPOSED APPROACH.

Phase	Rank	Bias positive pairs	Bias negative pairs	Accuracy
Development	13	0.010022	0.005322	0.984689
Test	5	0.000237	0.000116	0.999698

III. ADDITIONAL METHOD DETAILS

Please reply if your challenge entry considered (or not) the following strategies and provide a brief explanation.

- **Did you use pre-trained models?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use external data?** (X) Yes, ( ) No
  If yes, please detail:
  We used the public dataset MS1M-arcface and IJB-C.
- **Did you use other regularization strategies/terms?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use handcrafted features?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use any face detection, alignment or segmentation strategy?** (X) Yes, ( ) No
  If yes, please detail:
  We use the retina face to do face detecion and alignment. If there are more than one face are detected, we chose the most center one.
- **Did you use ensemble models?** (X) Yes, ( ) No
  If yes, please detail:
  We use three model to get the final pair similarity. These three models are traned with different training data.
- **Did you use different models for different protected groups?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you explicitly classify the legitimate attributes?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you explicitly classify other attributes (e.g. image quality)?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use any pre-processing bias mitigation technique (e.g. rebalancing training data)?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use any in-processing bias mitigation technique (e.g. bias aware loss function)?** ( ) Yes, (X) No
  If yes, please detail:
- **Did you use any post-processing bias mitigation technique?** ( ) Yes, (X) No
  If yes, please detail:

IV. CODE REPOSITORY

Code repository: `https://github.com/ZHAIXINGZHAIYUE/FairFaceCode.git`
Docker repository: zhaixingzhaiyue/mxnetcu90-py2

REFERENCES

[1] ChaLearnLAP. ECCV 2020 ChaLearn Fair Face Recognition Challenge. [Online]. Available: http://chalearnlap.cvc.uab.es/challenge/38/description/

[2] A. Amini, A. P. Soleimany, W. Schwarting, S. N. Bhatia, and D. Rus, "Uncovering and mitigating algorithmic bias through learned latent structure," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 289–295. [Online]. Available: https://doi.org/10.1145/3306618.3314243

[3] M. Wang, W. Deng, J. Hu, X. Tao, and Y. 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.

[4] Y. Liu, F. Wei, J. Shao, L. Sheng, J. Yan, and X. Wang, "Exploring disentangled feature representation beyond face identification," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[5] S. Gong, X. Liu, and A. Jain, "Jointly de-biasing face recognition and demographic attribute estimation," in *In Proceeding of European Conference on Computer Vision*, Virtual, August 2020.

[6] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-shot multi-level face localisation in the wild," in *CVPR*, 2020.

[7] insightface. Web face recognition training datasets. [Online]. Available: https://github.com/deepinsight/insightface/wiki/Dataset-Zoo

[8] C. Whitelam, E. Taborsky, A. Blanton, B. Maze, and P. Grother, "Iarpa janus benchmark-b face dataset," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017.

[9] T. Chen, M. Li, Y. Li, M. Lin, N. Wang, M. Wang, T. Xiao, B. Xu, C. Zhang, and Z. Zhang, "Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems," 2015, cite arxiv:1512.01274Comment: In Neural Information Processing Systems, Workshop on Machine Learning Systems, 2016. [Online]. Available: http://arxiv.org/abs/1512.01274

[10] J. Deng, J. Guo, X. Niannan, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *CVPR*, 2019.