Local-Adaptive Face Recognition via Graph-based Meta-Clustering and Regularized Adaptation

Wenbin Zhu[1*] Chien-Yi Wang[2*] Kuan-Lun Tseng[2] Shang-Hong Lai[2] Baoyuan Wang[3]
[1]Microsoft Cloud and AI [2]Microsoft AI R&D Center, Taiwan [3]Xiaobing.AI
{wenzh, chiwa, kutseng, shlai}@microsoft.com zjuwby@gmail.com

1. Adapt to Different Races

1.1. Evaluation on All Pairs in $\mathbb{S}_T$

We also evaluate different end-to-end LaFR models in race protocols using "all pairs" in the testing set $\mathbb{S}_T$. The ROC curves and the corresponding False Non-Match Rate (FNMR) @ False Match Rate (FMR) = { 1e-5, 1e-4, 1e-3 } are shown in Figure 2 and Table 1, respectively. The result in the last row of each adaptation protocol which adopts ground truth (GT) labels from $\mathbb{S}_A$ serves as the upper bound of end-to-end unsupervised LaFR methods. Our proposed "Meta-GCN + RCT" method for LaFR is superior to other methods in all protocols.

Race	Label	Loss	1:1 Verification FNMR FMR=1e-5	FMR=1e-4	FMR=1e-3
African	$\mathbb{S}_B$-GT(pre-trained)		72.35	58.89	42.05
	Distance-based [1]	RCT	99.04	97.43	90.02
	GCN [2]	RCT	58.25	42.03	24.69
	Meta-GCN	RCT	**52.27**	**36.28**	**20.28**
	$\mathbb{S}_T$-GT	RCT	<u>31.92</u>	<u>18.48</u>	<u>8.34</u>
Asian	$\mathbb{S}_B$-GT(pre-trained)		67.26	52.80	34.38
	Distance-based [1]	RCT	99.79	99.44	98.07
	GCN [2]	RCT	67.92	50.34	30.05
	Meta-GCN	RCT	**59.78**	**43.47**	**25.07**
	$\mathbb{S}_T$-GT	RCT	<u>33.81</u>	<u>19.40</u>	<u>8.77</u>
Indian	$\mathbb{S}_B$-GT(pre-trained)		43.42	29.58	15.86
	Distance-based [1]	RCT	46.24	32.25	19.19
	GCN [2]	RCT	41.99	28.55	16.09
	Meta-GCN	RCT	**36.39**	**23.03**	**12.00**
	$\mathbb{S}_T$-GT	RCT	<u>26.29</u>	<u>15.51</u>	<u>7.20</u>

Table 1. Comparison between different LaFR methods on racial adaptation protocols using all pairs in the testing set. The performance is measured by FNMR@FMR={1e-5, 1e-4, 1e-3}, the lower the better.

Figure 1. Mean Face Images captured by 4 different infrared cameras (the first two capture 940nm wavelength, and the other two 850nm). We can observe obvious appearance shifts between these infrared sensors because of the differences in both ISP (Image Signal Processor) and wavelength.

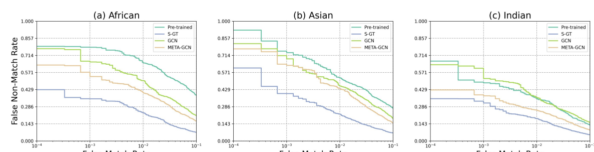

Figure 2. The ROC curves of deployed models evaluated on all pairs of (a) African (b) Asian (c) Indian testing sets.

Sensor	Methods	F_P	F_B
IR-A	Distance-based [1]	0.9845	0.9926
	GCN [2]	0.9888	0.9942
	Meta-GCN	**0.9903**	**0.9966**
IR-B	Distance-based [1]	0.9831	0.9963
	GCN [2]	0.9868	0.9966
	Meta-GCN	**0.9893**	**0.9978**
IR-C	Distance-based [1]	0.9865	0.9931
	GCN [2]	0.9905	0.9950
	Meta-GCN	**0.9953**	**0.9977**
IR-D	Distance-based [1]	0.8461	0.9258
	GCN [2]	0.9082	0.9687
	Meta-GCN	**0.9236**	**0.9703**

Table 2. Comparison of face embedding clustering performance on 4 sensor adaptation protocols. Two clustering metrics: Pairwise F-score (F_P) and Bcubed F-score (F_B) are reported.

2. Adapt to Different Sensors

2.1. Face Embedding Clustering

Figure 1 shows the mean face of each different sensors we used in our experiments. As we can see, there is obvious appearance shift among them. The clustering per-

[*] indicates equal contribution.

formance of sensor adaptation protocols is shown in Table 2. From the results, we can observe that the clustering F-score is high compared with racial protocols, which means the clustering task is relatively easy on datasets with fewer identities. However, it still demonstrates that our proposed "Meta-GCN" clustering method can generalize better in the unseen local environment and achieve better F-score than previous distance-based [1] and GCN [2] methods.

References

[1] 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 *ICCV*, 2019.

[2] Lei Yang, Dapeng Chen, Xiaohang Zhan, Rui Zhao, Chen Change Loy, and Dahua Lin. Learning to cluster faces via confidence and connectivity estimation. In *CVPR*, 2020.