Dive into the Resolution Augmentations and Metrics in Low Resolution Face Recognition: A Plain yet Effective New Baseline

*Xu Ling, Yichen Lu, Wenqi Xu, Weihong Deng, Yingjie Zhang, Xingchen Cui, Hongzhi Shi, Dongchao Wen**

lingxu421@gmail.com; wendongchao@inspur.com

inspur 浪潮

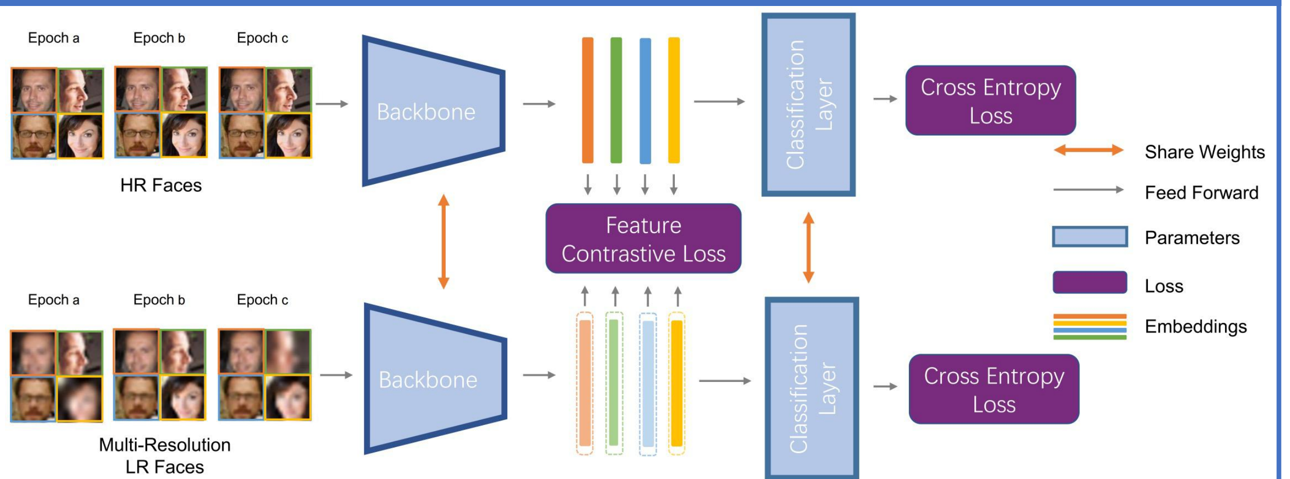
Introduction

Due to the huge domain gap between Low Resolution(LR) and High Resolution(HR) faces and very limited LR faces data in the wild, the performance of the Low Resolution Face Recognition (LRFR) tasks is still poor. However, LRFR tasks are again widely seen in real-world scenarios.

We dive into the two most important parts in LRFR with such a concise architecture to further reveal the huge potential of each:

- **The Multi-Resolution Augmentation**
- **The Feature Contrast Loss**

With the help of both, we argue our method as a new plain yet effective baseline of LRFR. Code is available at: https://github.com/hurricanelx/LRFR.

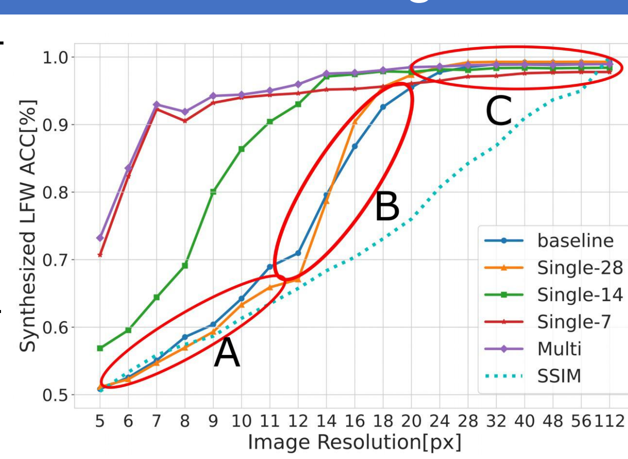


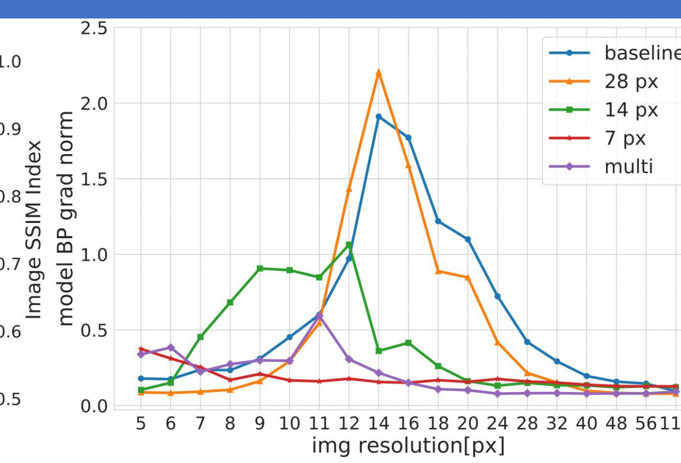
Architecture of our method

The Multi-Resolution Augmentation

We use the synthetic LFW accuracy at different resolutions and the backpropagation gradient norm of different models for different resolution inputs to study the impact of different resolutions on performance. As a result, we address two critical questions about multi-resolution augmentation: which resolutions should be augmented and why they work. The findings are as follows:

- The loss of information increases rapidly as the resolution decreases.
- The introduction of some resolutions will make it less difficult to learn at lower resolutions.
- The difficulty of samples should be considered as A: Extremely hard, B:Hard and C:Semi-hard.





With these findings and our experiment results, we propose to combine three different difficulties in training, using three representative resolutions of 7px, 14px, and 20px for augmentation.

The Feature Contrast Loss

Due to the **huge domain gap** between LR and HR domains, the feature contrast loss need to be carefully designed.We thus propose a novel Feature Constrast Loss L_{LogExp} which **incorporates the advantages of** both L_1 and L_2.

$$L_{LogExp} = \frac{1}{D}log(1+\sum_{i=1}^{D}(e^{|x_i-y_i|}-1)$$

$$\frac{\partial L_{LogExp}}{\partial x_i} = \frac{1}{D}\cdot\frac{(\pm 1)\cdot e^{|x_i-y_i|}}{1+\sum_{i=1}^{D}(e^{|x_i-y_i|}-1)}$$

Also, just like both L_1 and L_2 could be generalized as L_p , L_{LogExp} in general form is:

$$L_{LogExp} = \frac{1}{p\cdot D}log(1+\sum_{i=1}^{D}(e^{(|x_i-y_i|)^p}-1)$$

Besides the advantages of both L_1 and L_2, the L_{LogExp} function could dynamically adjust the gradient magnitude according to the ratio of errors between dimensions.Thus, it is possible to focus on the dimensions with larger errors during the training process to improve this part first.

Reference

ST-M: Knoche M, Hörmann S, Rigoll G. Image resolution susceptibility of face recognition models[J]. arXiv preprint arXiv:2107.03769, 2021.

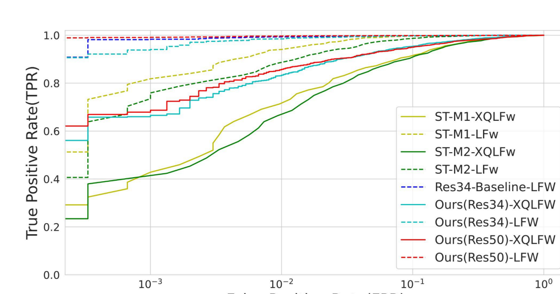
RPCL-DCR: Li P, Tu S, Xu L. Deep rival penalized competitive learning for low-resolution face recognition[J]. Neural Networks, 2022.

Experiment Results

We achieve SOTA results on SCface and XQLFW and balanced result on TinyFace.

Methods	$d1$	$d2$	$d3$	Avg
DCR (2018)	73.3	93.5	98.0	88.27
TCN (2019)	74.6	94.9	98.6	89.37
DATL (2019)	76.2	96.9	98.1	90.40
DAQL (2019)	77.3	96.6	98.1	90.66
RAN (2020)	81.3	97.8	98.8	92.63
DDL(Res50) (2020)	86.8	98.3	98.3	94.47
NPT-Loss(Res50) (2022)	85.7	*99.1*	99.1	94.63
RPCL-DCR (2022)	90.4	98.0	98.0	95.47
Ours(w/o FT)	73.0	96.0	98.8	89.27
Ours(Res34)	*91.8*	99.0	*99.3*	*96.67*
Ours(Res50)	**92.5**	**99.3**	**99.5**	**97.08**

Results on SCface



ROC curve on XQLFW and LFW

The ablation study results are as follows:

Methods	SCface w/o FT				XQLFW	LFW
	$d1$	$d2$	$d3$	Avg		
Baseline	40.3	92.8	98.5	77.2	68.95	**99.42**
Aug	71.5	92.8	96.3	86.8	84.33	98.65
MAug	67.0	95.3	98.0	86.8	90.89	98.98
PAug+Ours(1)	60.3	95.3	**98.8**	84.8	90.35	99.02
Aug+L_1	67.3	94.5	94.8	85.5	85.23	98.57
Aug+Ours(1)	70.0	94.8	96.5	87.1	87.78	98.60
MAug+Ours(1)	73.0	**96.0**	**98.8**	**89.3**	**92.18**	99.03
Aug+L_2	71.8	90.0	92.8	84.8	84.87	98.50
Aug+Ours(2)	72.5	92.3	94.3	86.3	86.88	98.63
MAug+Ours(2)	**74.0**	95.5	97.8	*89.1*	*91.98*	*99.08*

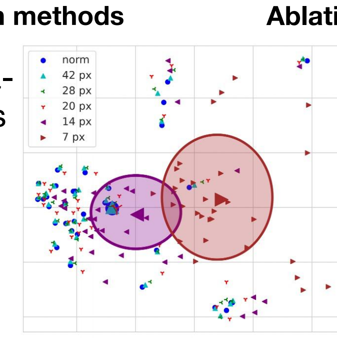
Ablation Study on methods

Ratio	SCface w/o FT				XQLFW	LFW
	$d1$	$d2$	$d3$	Avg		
Baseline	40.3	92.8	98.5	77.2	68.95	**99.42**
MAug only	67.0	95.3	98.0	86.8	90.89	98.98
1:1:0	**73.8**	94.0	96.8	88.2	**92.43**	98.68
0:1:1	71.5	95.8	98.3	88.5	83.52	99.05
1:0:1	65.3	94.8	97.8	86.0	90.53	99.02
1:1:1	72.5	96.5	98.3	89.1	91.45	98.88
1:2:1	72.8	95.0	98.0	88.6	91.68	98.82
2:1:1	72.8	**96.8**	97.3	88.9	92.05	98.75
1:1:2	73.0	96.0	**98.8**	89.3	92.18	99.03

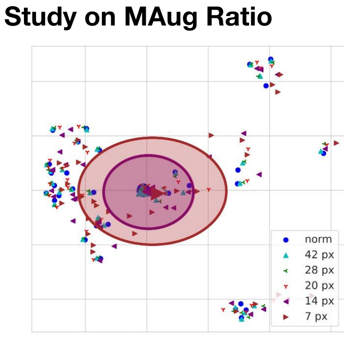
Ablation Study on MAug Ratio

The feature visualization experiment is as follows.

These experiments verify the effectiveness of our method.



MAug + L1



MAug + Ours LogExp