MASKDUL: DATA UNCERTAINTY LEARNING IN MASKED FACE RECOGNITION

Libo Zhang, Weiming Xiong, Ku Zhao, Kehan Chen, Mingyang Zhong*
College of Artificial Intelligence, Southwest University, Chongqing, China

Background

- Masked Face Recognition (MFR) is a challenging task due to the aggravated uncertainty caused by mask occlusion.
- Deterministic point embedding models are limited in representing uncertainty.
- Data Uncertainty Learning (DUL) methods effectively represent uncertainty and reduce the adverse effects of the noises.

Findings

- The masked face tends to be regarded as noise due to the mask occlusion, which weakens its optimization.
- The large representation difference between face and masked face results in a dispersed intra-class distribution.

Contributions

- This is the first time to apply data uncertainty learning mechanism to MFR.
- The proposed MaskDUL includes a dual-stream network with face and masked face branches, and a joint optimization strategy based on data uncertain learning.
- MaskDUL realizes ideal sample distribution, in which high-quality faces and masked faces of same class are close to center, while real noises are far away from center.

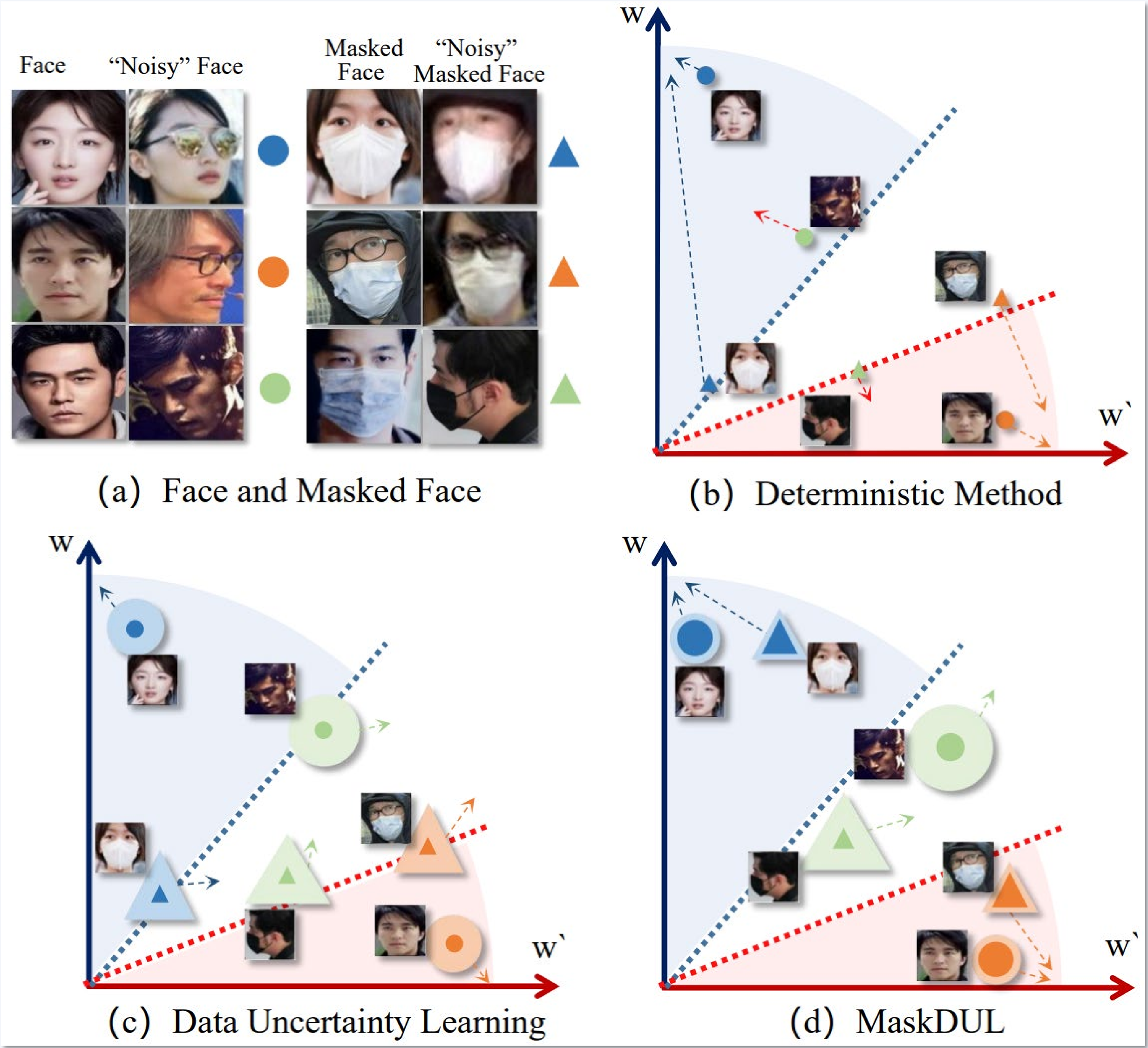

Figure 1 Visualization of the research thread. MaskDUL solves the problems found, and realizes the ideal sample distribution.

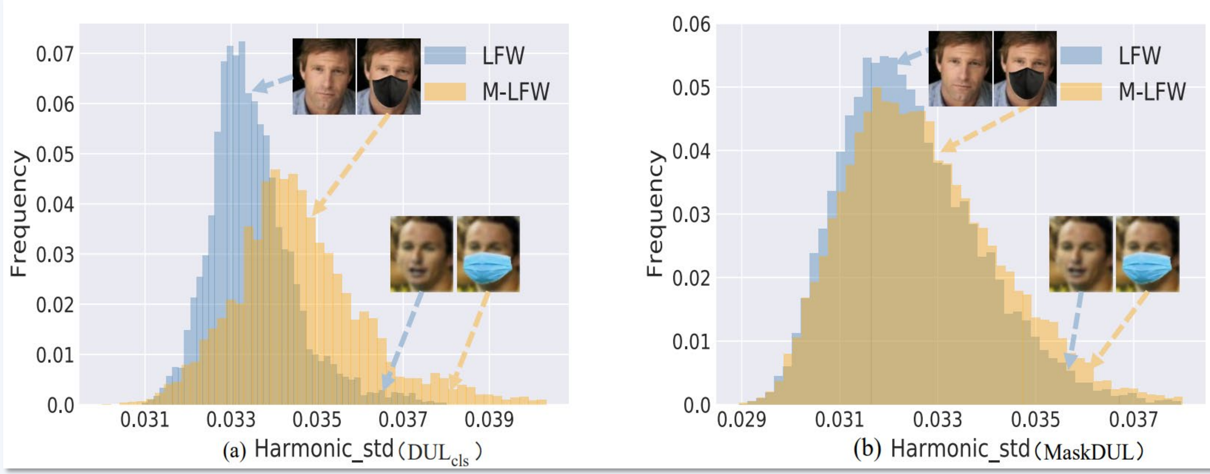

Figure 2 DUL method tends to treat masked face as noise, and learn wrong uncertainty. MaskDUL learns more accurate uncertainty.

Methods

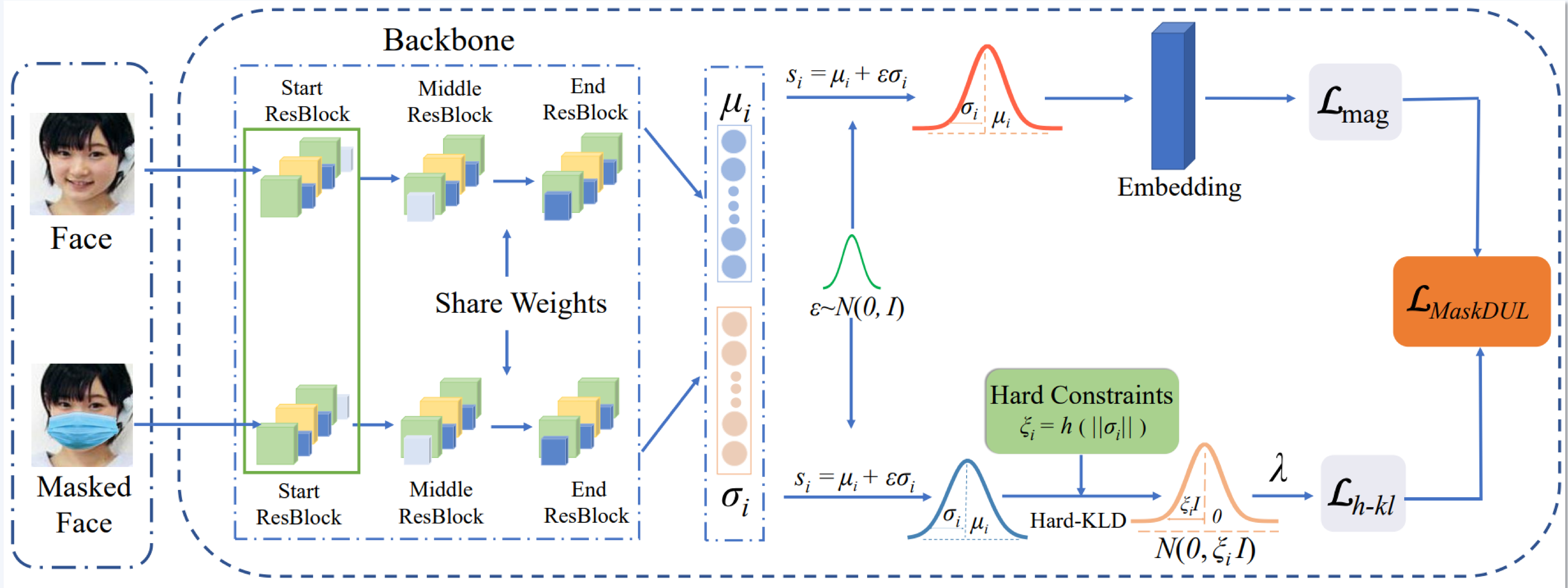

Figure 3 MaskDUL includes a dual-stream network with face and masked face branches, and the optimization of data uncertainty learning based on a hard variance constraint H-KLD.

The output of two branches includes identity feature μ (mean) and uncertainty σ (variance). They are used to calculate loss function $\mathcal{L}_{mag}$ of intra-class distribution optimization. Hard-KLD imposes a hard constraint on variance uncertainty $\|\sigma\|$ to calculate $\mathcal{L}_{h-kl}$. MaskDUL dynamically adjust the learning preference, and gain the optimal effect through $\mathcal{L}_{MaskDUL}$:

$$\mathcal{L}_{MaskDUL} = \mathcal{L}_{Mag} + \lambda \mathcal{L}_{h-kl}$$

where λ is a trade-off parameter for H-KLD.

Hard Kullback-Leibler Divergence

The formula $\mathcal{L}_{h-kl}$ of hard variance constraint:

$$\mathcal{L}_{h-kl} = KL\left[N\left(z_i \mid \mu_i, \sigma_i^2\right) \| N\left(\epsilon \mid 0, (\xi I)^2\right)\right]$$
$$= -\frac{1}{2}\left(1 + \log\frac{\sigma^2}{\xi^2} - \frac{\mu^2}{\xi^2} - \frac{\sigma^2}{\xi^2}\right).$$

Magnitude-based Loss Function

The formula $\mathcal{L}_{mag}$ of angular margin constraint:

$$\mathcal{L}_{Mag} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\cos\left(\theta_{y_i}+m\left(a_i\right)\right)}}{e^{s\cos\left(\theta_{y_i}+m\left(a_i\right)\right)}+\sum_{j\neq y_i}e^{s\cos\theta_i}}+\lambda_g' g(a_i)$$

Table 1 Masked face verification and identification accuracy of various methods. MaskDUL achieves the SOTA results on all benchmarks.

Category	Methods	Verification					Identification			
		Limited		Unconstrained	Large-age		MFR2		RMFRD	
		MFR2	M-LFW	RMFRD	M-CALFW	M-AgeDB_30	Rank-1	Rank-5	Rank-1	Rank-5
FR	ArcFace	89.37	86.63	64.81	76.36	75.11	86.37	88.65	53.12	62.28
	MagFace	93.64	91.02	70.97	83.41	81.50	90.62	92.36	63.86	72.19
DUL	PFE	91.13	88.69	68.03	79.78	77.93	87.37	89.73	57.93	66.12
	ProbFace	91.37	89.02	68.56	80.01	77.22	87.01	89.02	58.65	66.93
	DUL-cls	92.32	69.43	89.64	81.35	79.66	88.11	90.78	60.68	68.65
	DUL-GM	92.97	89.61	68.98	81.26	78.99	88.93	90.85	61.54	70.23
OFR	PDSN	95.76	94.91	74.24	87.84	85.19	93.64	95.03	67.43	76.77
PFR	DFM	96.61	94.52	76.89	86.33	86.43	93.33	95.51	67.99	76.21
MFR	DCR	97.27	94.98	77.87	89.86	88.02	95.78	96.97	68.62	77.02
	MFCosFace	97.35	95.17	78.09	90.02	89.22	95.94	97.98	70.31	79.54
	UPA	98.22	96.12	79.55	89.17	88.37	96.72	98.73	71.03	80.11
	MaskDUL	**99.01**	**96.32**	**80.23**	**91.03**	**90.02**	**97.15**	**99.05**	**72.14**	**80.73**

Experiments

- MaskDUL outperforms baselines on all benchmarks, and the performance improvement is 0.66%~1.11% higher than other advanced methods.
- DUL_{cls} tends to assign large variance to masked faces, and *DUL+Mag* makes the overall variance of masked face smaller and more concentrated.
- MaskDUL limits the variance of masked face, and retains some true noises with large variance, thus distinguishing the high-recognizable masked face from the true noises in the intra-class distribution.

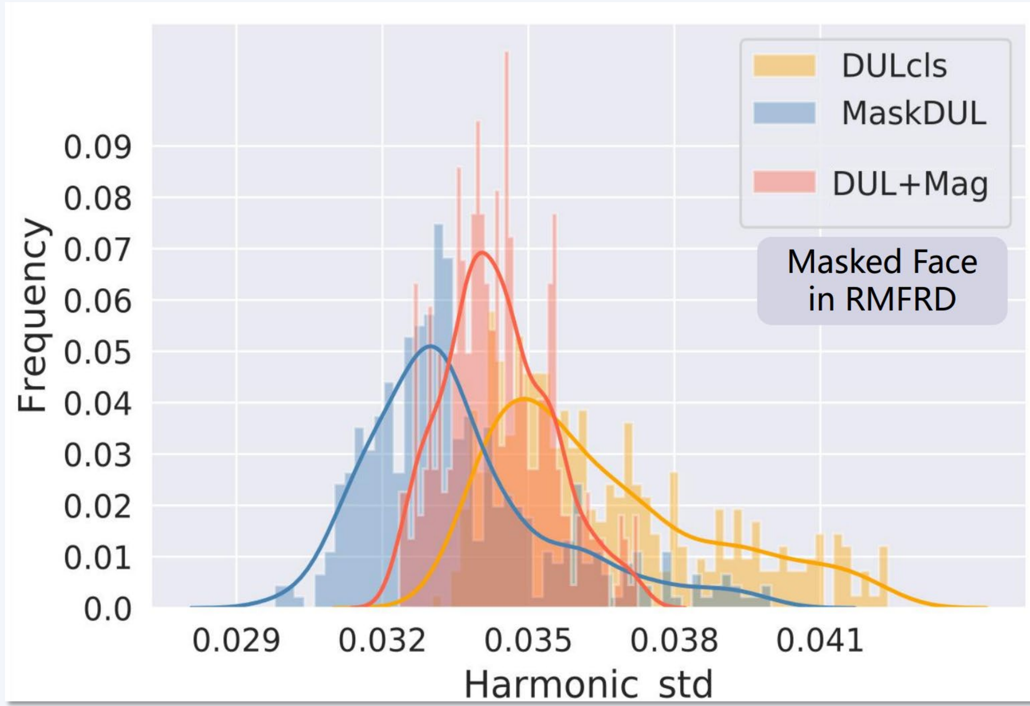

Figure 4 Visualization of modules effectiveness in ablation study.

Conclusions

- MaskDUL is a novel two-stream convolutional network with a Hard Kullback-Leibler Divergence and magnitude-based loss function.
- It explores sample uncertainty and intra-class distribution learning in MFR, which enables the model to learn more accurate uncertainty representations and construct a compact intra-class distribution.
- Finally, comprehensive experiments on MFR tasks prove the advancement of MaskDUL.