



Look More Into Occlusion: Realistic Face Fortalization and Recognition With BoostGAN (IEEE TNNLS 2020)

■ 汇报人：赵思桓、刘睿龙、曾强文

2024.5.15



目录

01 研究背景与相关工作

02 实现原理

03 实验



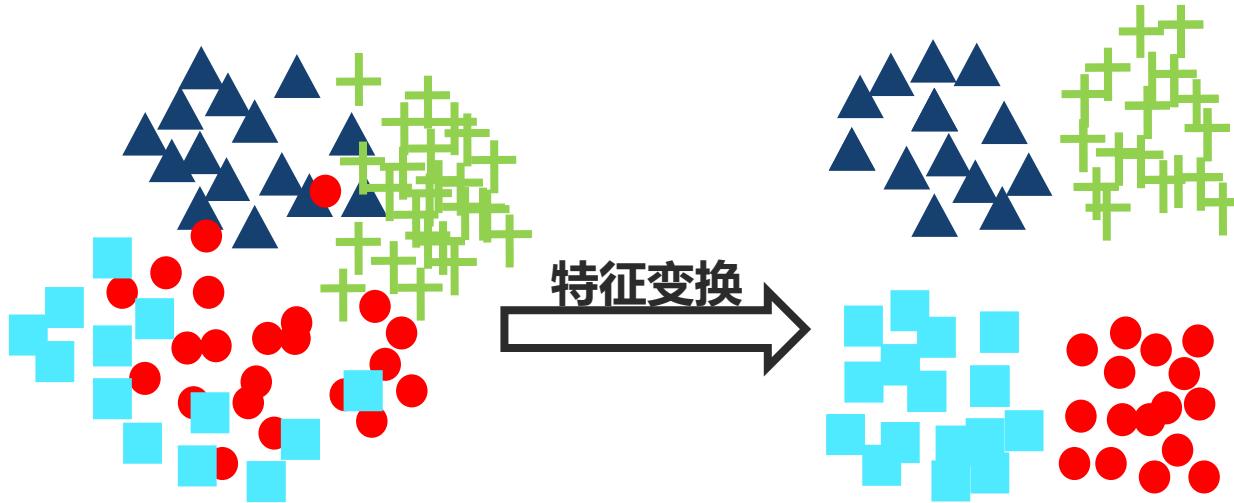
Part
01

研究背景与相关工作

人脸识别



传统方法

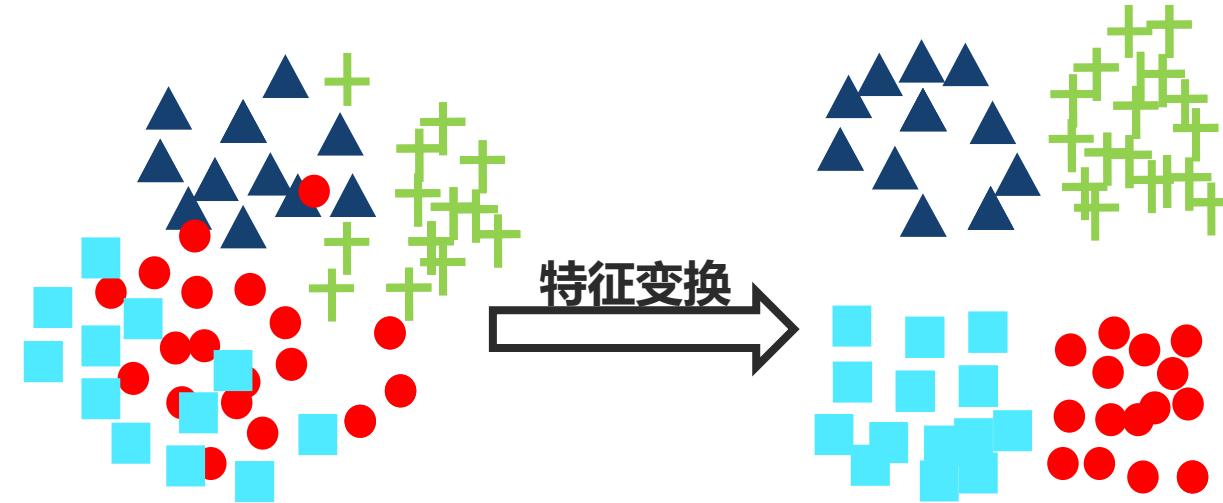


高维人工特征空间

低维判别特征空间

利用Gabor、LBP或SIFT变换，从图像中生成高维的人工特征空间，之后利用PCA、LDA等特征变换方法进行特征变换

深度学习



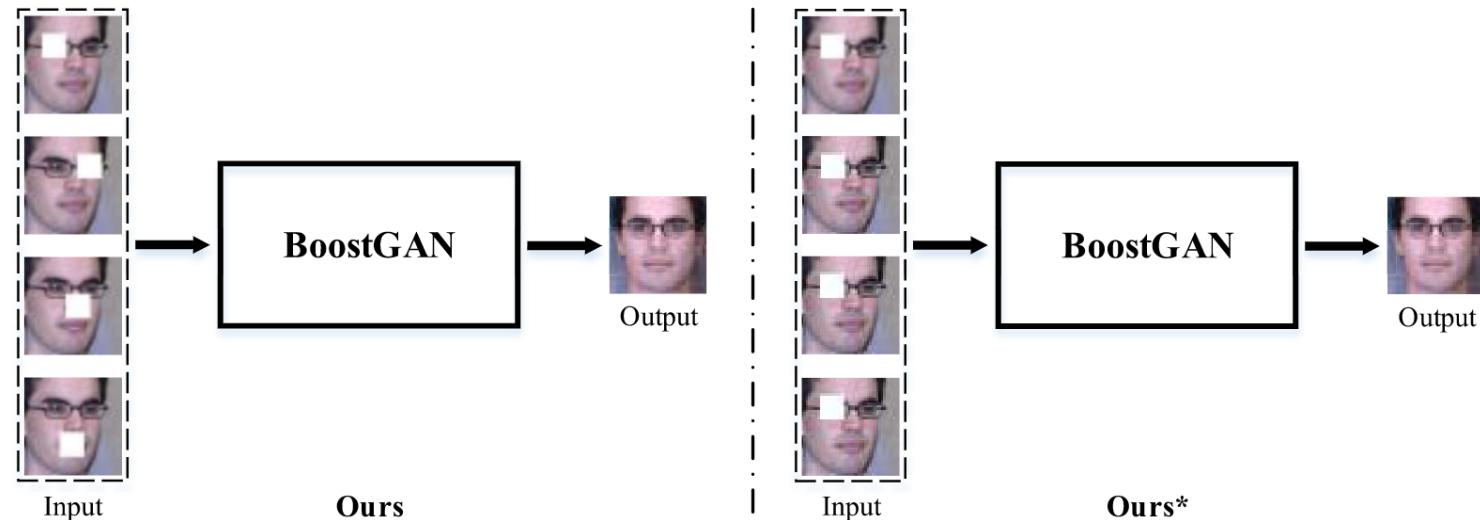
输入图像

低维判别特征空间

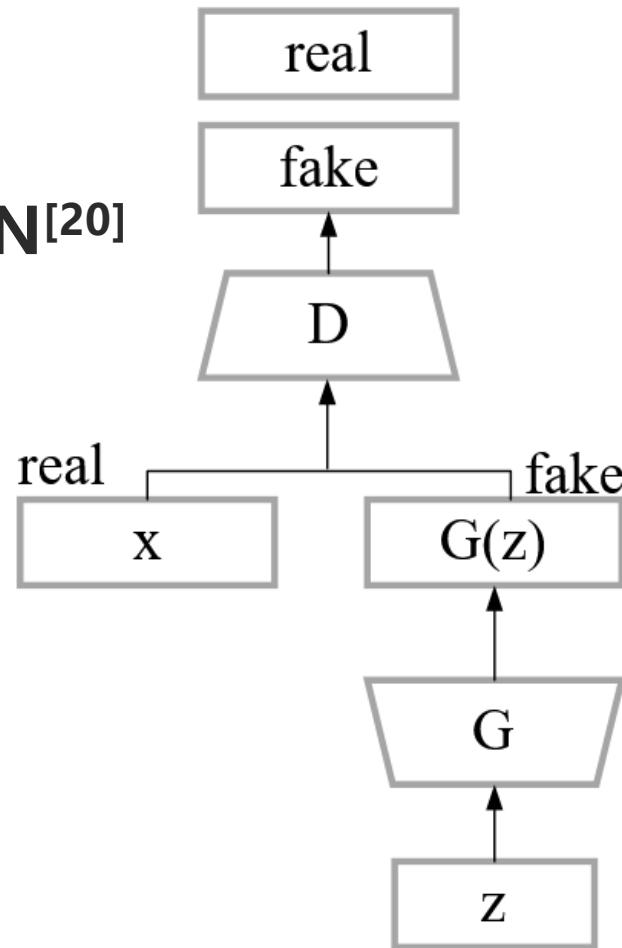
直接从图像中学习特征并生成特征映射。通常使用卷积神经网络实现人脸识别，此外GAN常用于生成逼真的脸帮助识别

研究动机

- 现存的问题：**
1. 遮挡和人脸的姿态会极大影响人脸识别的准确性；
 2. 现存的方法基本上只针对一个点：正面化or人脸补全；
 3. 现存的人脸补全的方法只关注面部结构和逼真的像素细节，而不是身份信息的保留。



BoostGAN目的：实现一个端到端的人脸正面化以及图像补全网络，为高精度人脸识别做好准备。

GAN^[20]

目标函数:

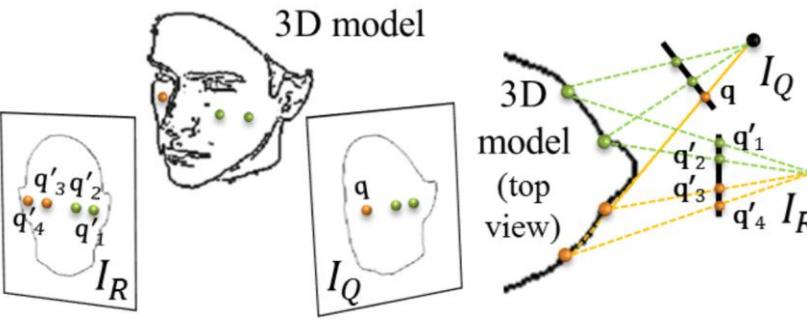
$$\min_G \max_D V(D, G) = E_{\mathbf{x} \sim p_d(\mathbf{x})} [\log D(\mathbf{x})] + E_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]. \quad (1)$$

$$\max_D V_D(D, G) = E_{\mathbf{x} \sim p_d(\mathbf{x})} [\log D(\mathbf{x})] + E_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))], \quad (2)$$

$$\max_G V_G(D, G) = E_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(D(G(\mathbf{z})))] \quad (3)$$

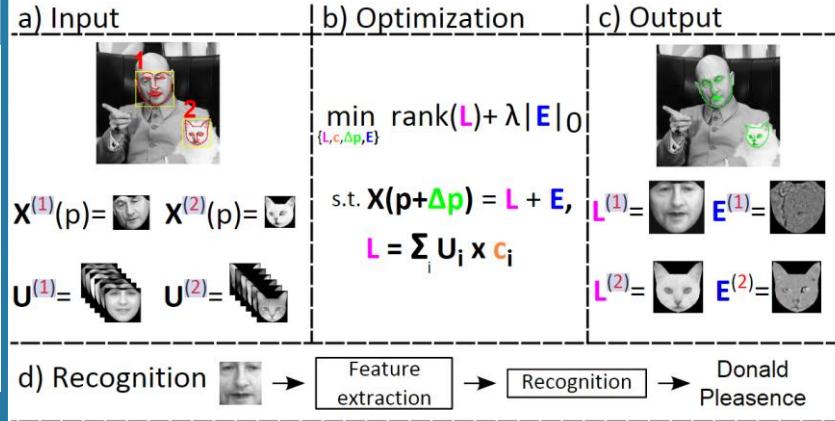
人脸正面化

2-D/3-D局部纹理扭曲



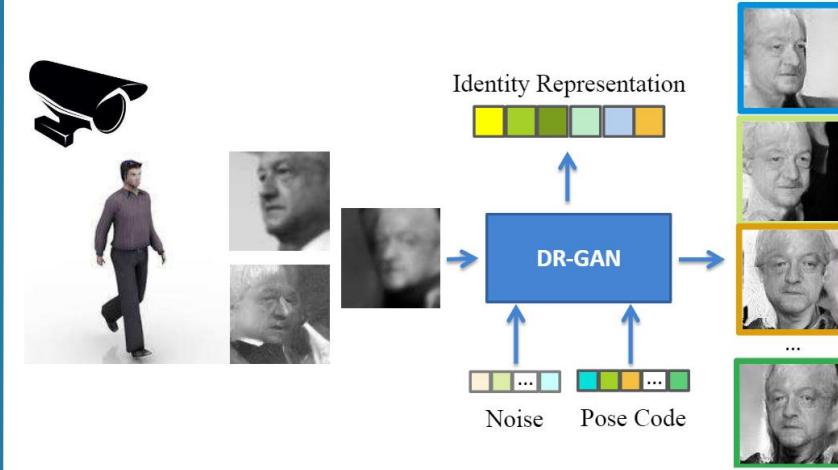
利用三维人脸的平均参考面来计算受试者的正面视图的人脸图像^[21]

统计方法



采用统计方法来计算正面图像^[22]

深度学习



使用GAN来直接生成正面图像^[23]

[21] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in Proc. CVPR, 2015, pp. 4295–4304.

[22] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic, "Robust statistical face frontalization," in Proc. ICCV, 2015, pp. 3871–3879.

[23] L. Tran, X. Yin, and X. Liu, "Disentangled representation learning GAN for pose-invariant face recognition," in Proc. CVPR, 2017, pp. 1415–1424.



人脸正面化：深度学习

DR-GAN^[23]模型用于姿态不变特征表示和识别，旨在从编码器中提取姿态不变特征表示。

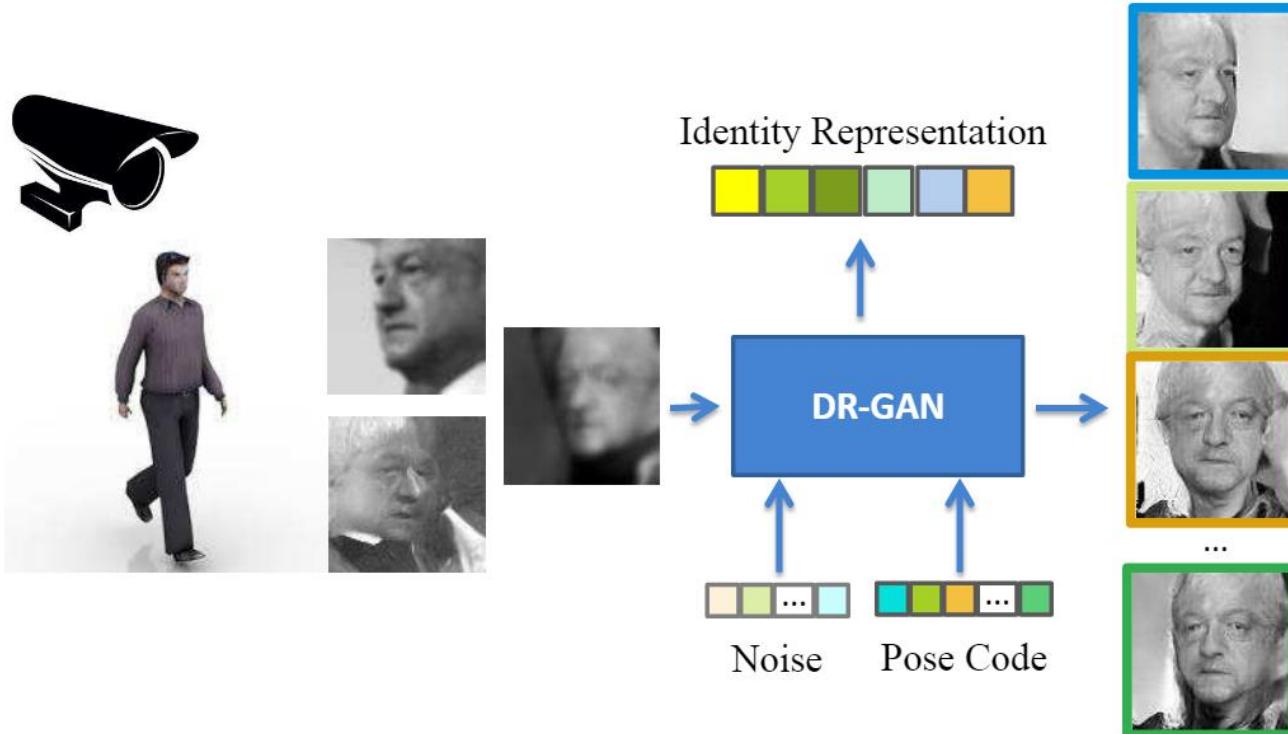
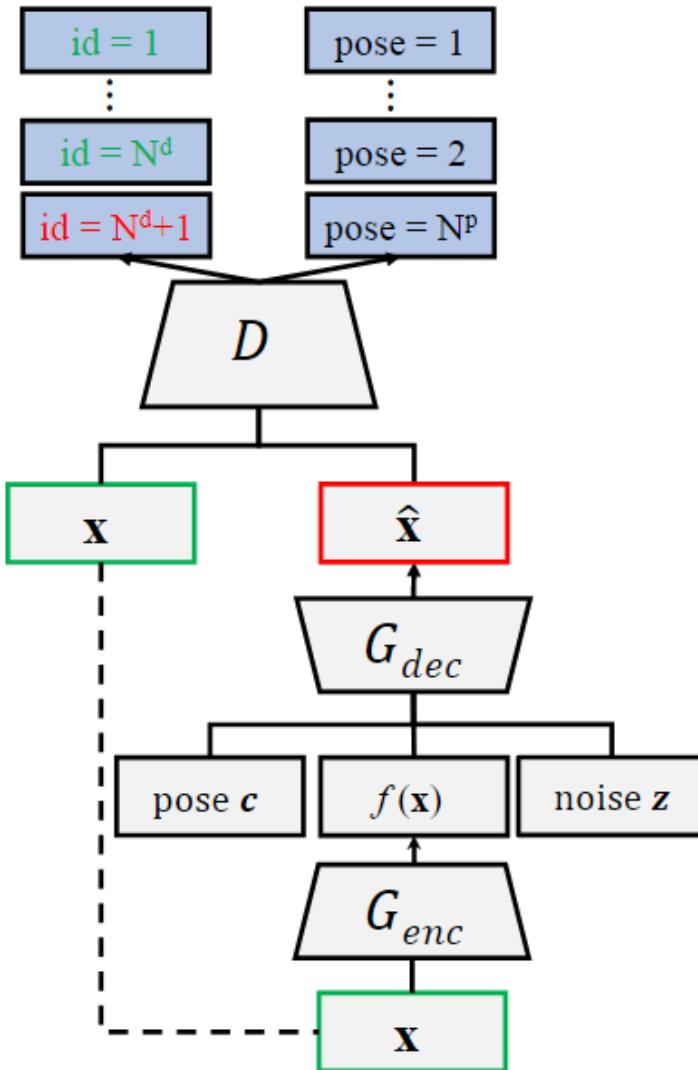


这些方法只考虑姿态，没有考虑遮挡。

TP-GAN^[24]提出通过分别使用全局和局部网络分别处理轮廓人脸，利用网络融合合成最终的正面视图人脸。

- [23] L. Tran, X. Yin, and X. Liu, "Disentangled representation learning GAN for pose-invariant face recognition," in Proc. CVPR, 2017, pp. 1415–1424.
- [24] R. Huang, S. Zhang, T. Li, and R. He, "Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis," in Proc. ICCV, 2017, pp. 2439–2448.

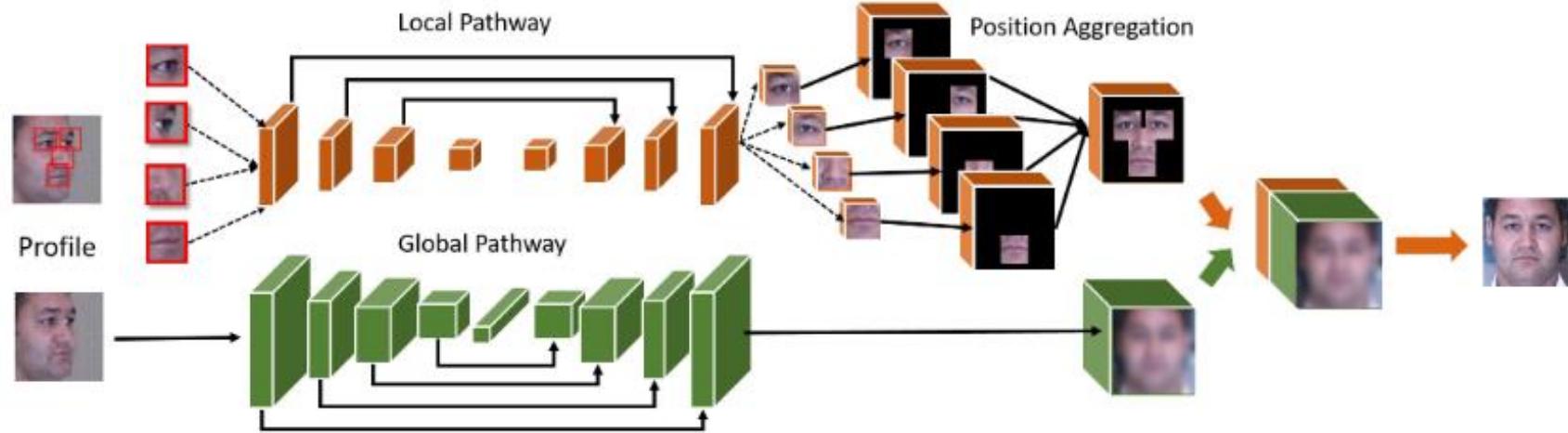
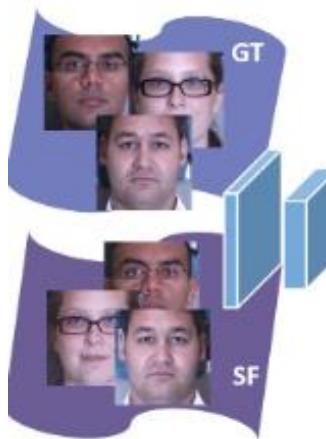
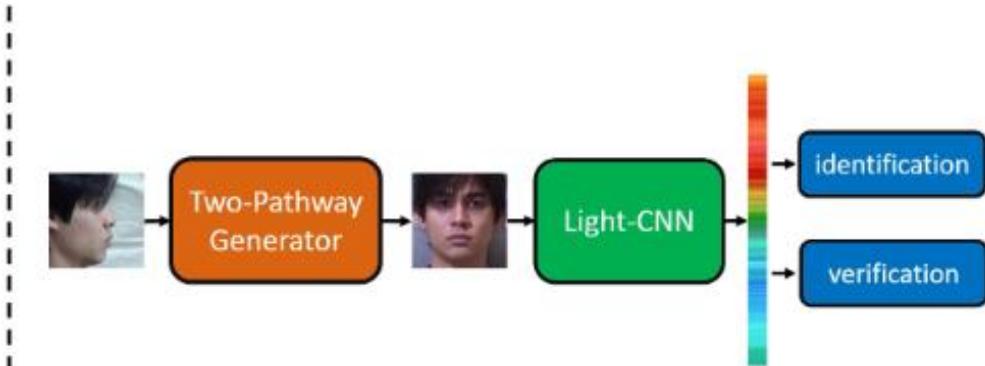
DR-GAN



$$\max_D V_D(D, G) = E_{\mathbf{x}, \mathbf{y} \sim p_d(\mathbf{x}, \mathbf{y})} [\log D_{y^d}^d(\mathbf{x}) + \log D_{y^p}^p(\mathbf{x})] + \\ E_{\substack{\mathbf{x}, \mathbf{y} \sim p_d(\mathbf{x}, \mathbf{y}), \\ \mathbf{z} \sim p_z(\mathbf{z}), \mathbf{c} \sim p_c(\mathbf{c})}} [\log(D_{N^d+1}^d(G(\mathbf{x}, \mathbf{c}, \mathbf{z})))], \quad (4)$$

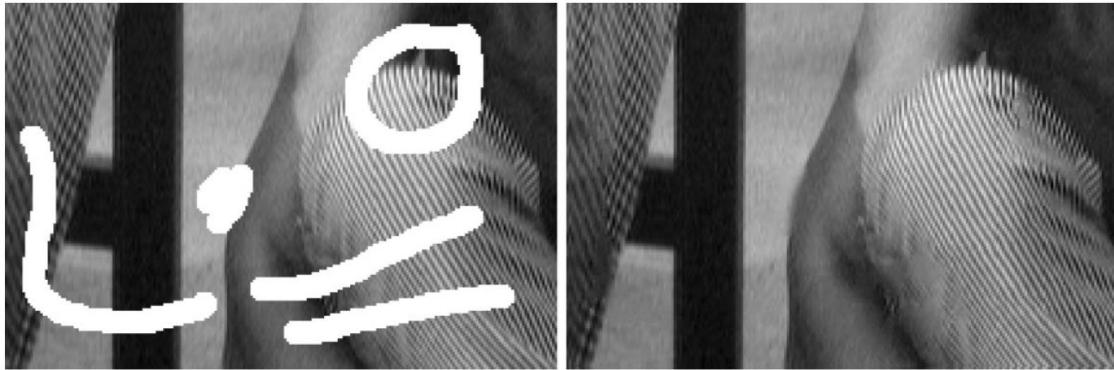
$$\max_G V_G(D, G) = E_{\substack{\mathbf{x}, \mathbf{y} \sim p_d(\mathbf{x}, \mathbf{y}), \\ \mathbf{z} \sim p_z(\mathbf{z}), \mathbf{c} \sim p_c(\mathbf{c})}} [\log(D_{y^d}^d(G(\mathbf{x}, \mathbf{c}, \mathbf{z}))) + \\ \log(D_{y^t}^p(G(\mathbf{x}, \mathbf{c}, \mathbf{z})))]. \quad (5)$$



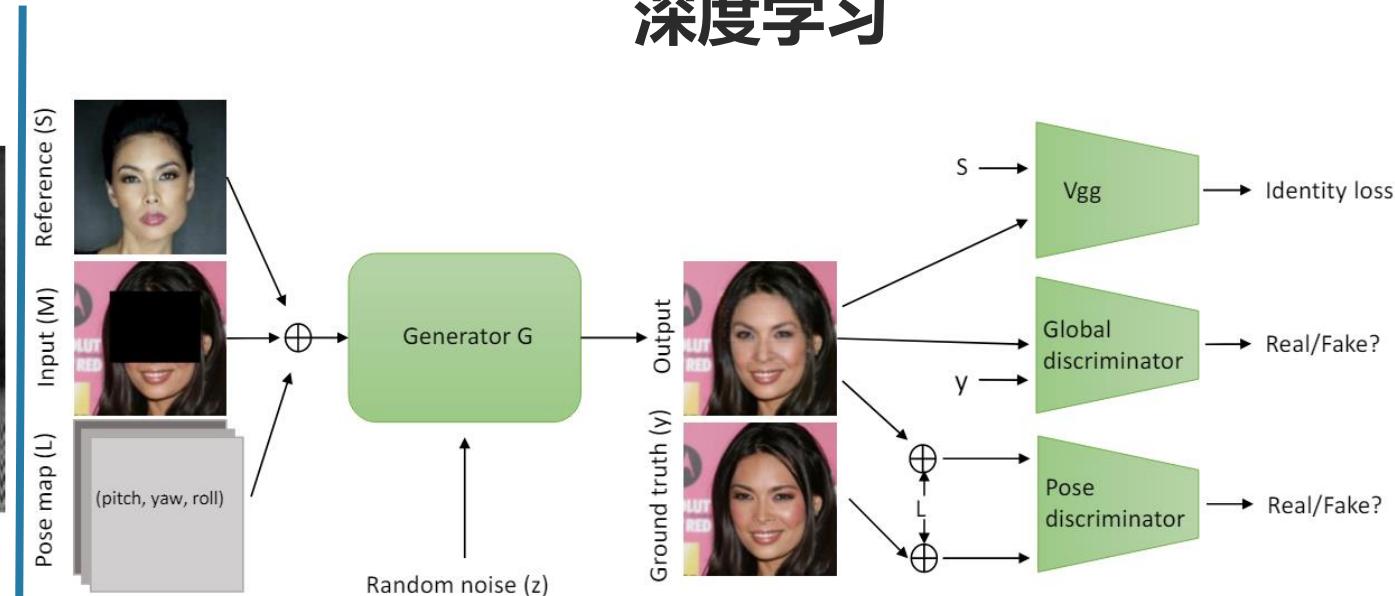
*Two-pathway Generator Network**Discriminator Network**Recognition via Generation*

目标函数:
$$\hat{\theta}_G = \frac{1}{N} \operatorname{argmin}_{\theta_G} \sum_{n=1}^N \{ L_{syn}(G_{\theta_G}(I_n^P), I_n^F) + \alpha L_{cross_entropy}(G_{\theta_E}(I_n^P), y_n) \} \quad (1)$$

传统方法



基于纹理以及结构进行补全图像^[30]



深度学习

利用生成对抗网络来补全人像^[41]

[30] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," in Proc. CVPR, 2003, pp. 1–6.

[41] Y. Zhao et al., "Identity preserving face completion for large ocular region occlusion," in Proc. BMVC, 2018.



人脸补全：深度学习

Li et al. [38] 提出了利用GAN来生成人脸缺失的部分，
其中包含全局与局部鉴别器。



忽视了人脸的身份信息

Yeh et al. [69] 搜索在潜在空间中与图像“最接近”的
损坏图像的编码。然后使用编码让生成器重建图像。

Zhao et al. [41] 将身份损失引入生成器的训练中，能够
够在补全图像的同时保留一定的身份信息。



忽视了人脸的姿态

[38] Y. Li, S. Liu, J. Yang, and M.-H. Yang, "Generative face completion," in Proc. CVPR, 2017, pp. 3911–3919.

[41] Y. Zhao et al., "Identity preserving face completion for large ocular region occlusion," in Proc. BMVC, 2018.

[69] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, "Semantic image inpainting with deep generative models," in Proc. CVPR, 2017, pp. 5485–5493.



Part
02

实现原理



BoostGAN的直观效果

45°



60°

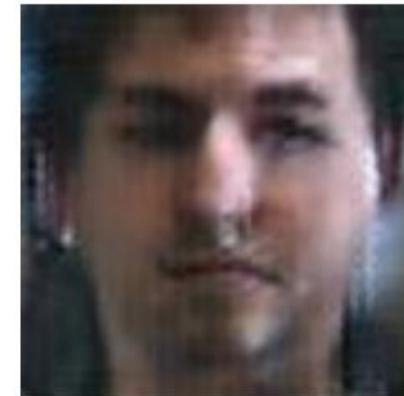


(a)

受遮挡的侧面人脸



BoostGAN



DR-GAN



TP-GAN



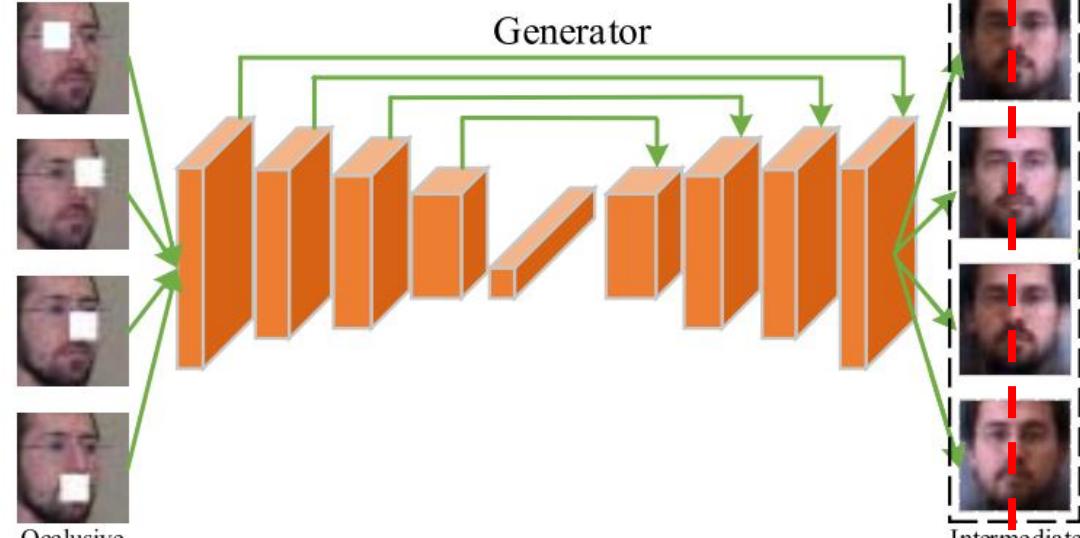
真实的正面人脸

BoostGAN的整体架构

Boosting Generator

Multi-occlusion Frontal View

Generator



Multi-input Boosting Network

细生成器 G^f

生成器 G

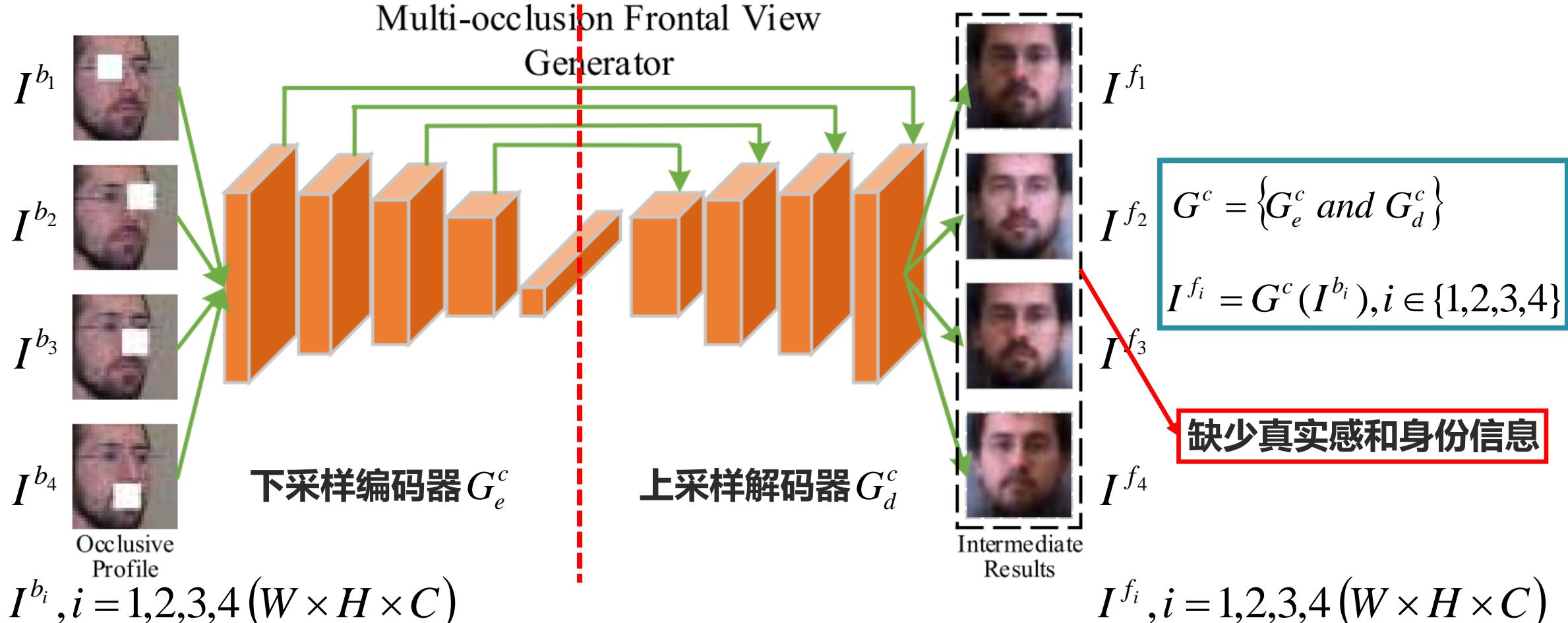
Discriminator

Real/Fake

判别器 D

Multiocclusion Frontal View Generator

Multiocclusion Frontal View Generator用 G^c 表示，其功能是通过输入的4张受遮挡的侧面人脸去分别对应生成4张粗糙的正面人脸，结构与TP-GAN^[24]中的全局生成器相同。



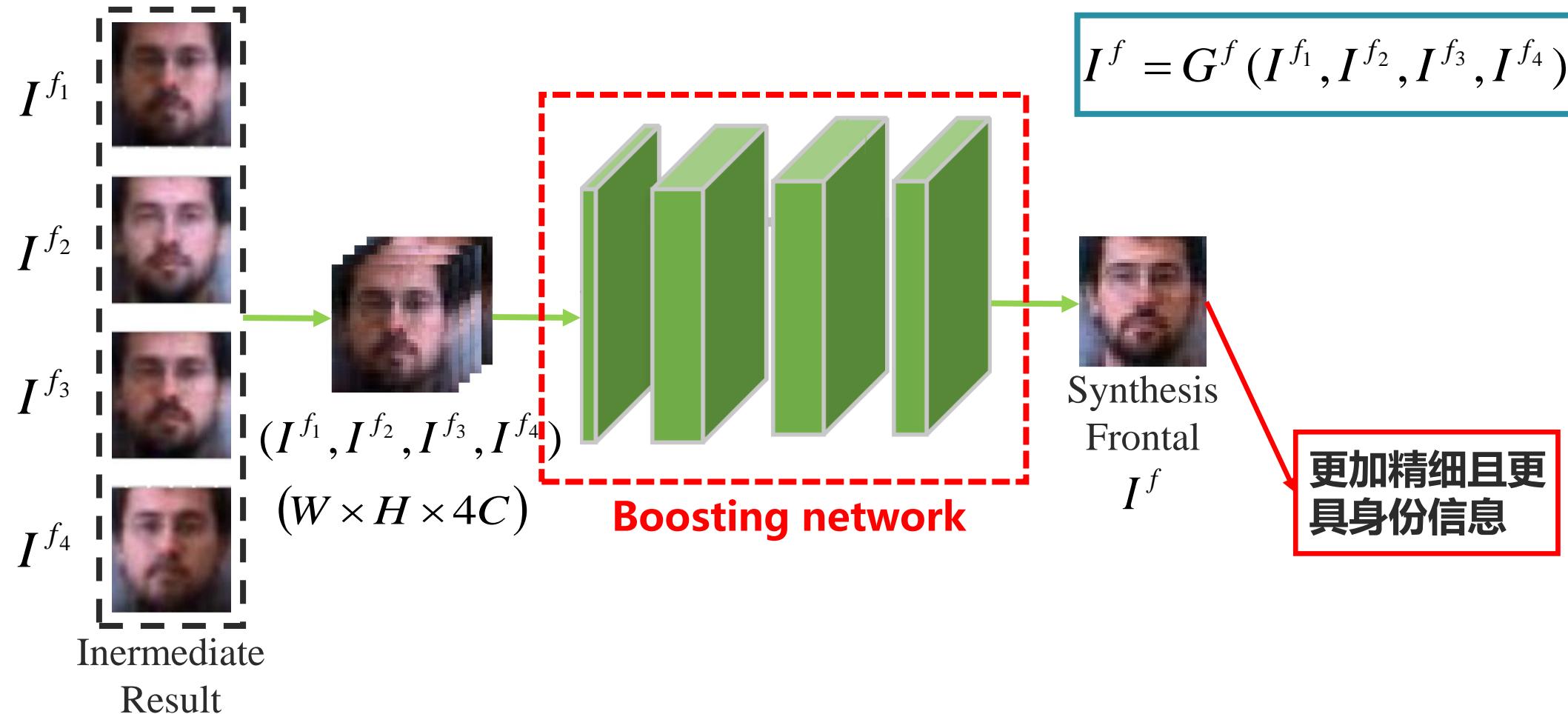
[24] R. Huang, S. Zhang, T. Li, and R. He, "Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis," in Proc. ICCV, 2017, pp. 2439–2448.



Multi-Input Boosting Network

Multi-Input Boosting Network用 G^f 表示，其功能是通过 G^c 生成的4张粗糙正面人脸去生成1张精细的正面人脸。

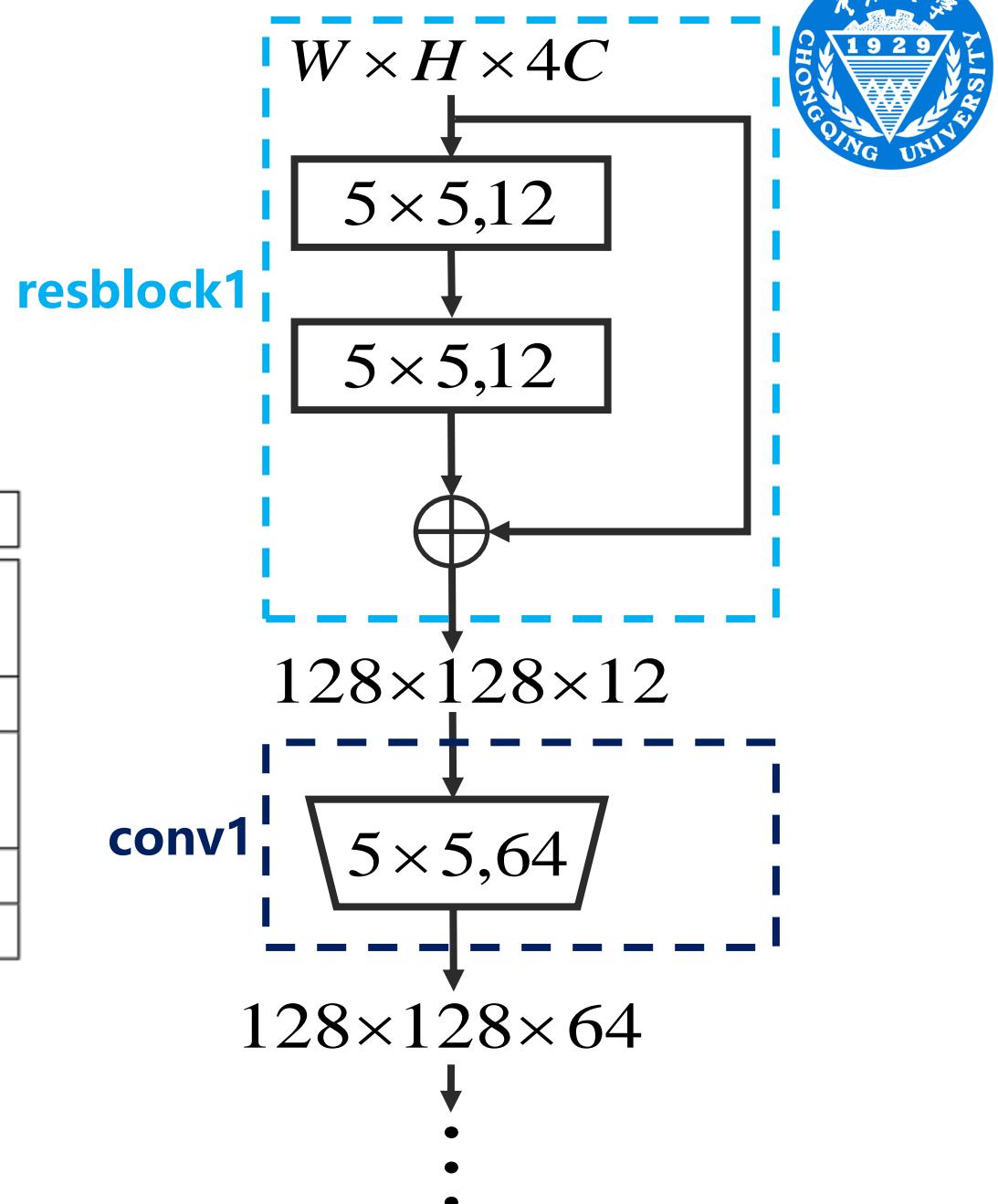
Boosting：一种基于集成策略的元算法，它可以通过模型融合将一组弱学习器转化为一组强学习器。



Boosting Network的具体结构

CONFIGURATION OF THE BOOSTING NETWORK. THE RESBLOCK DENOTES THE RESIDUAL CONNECTION

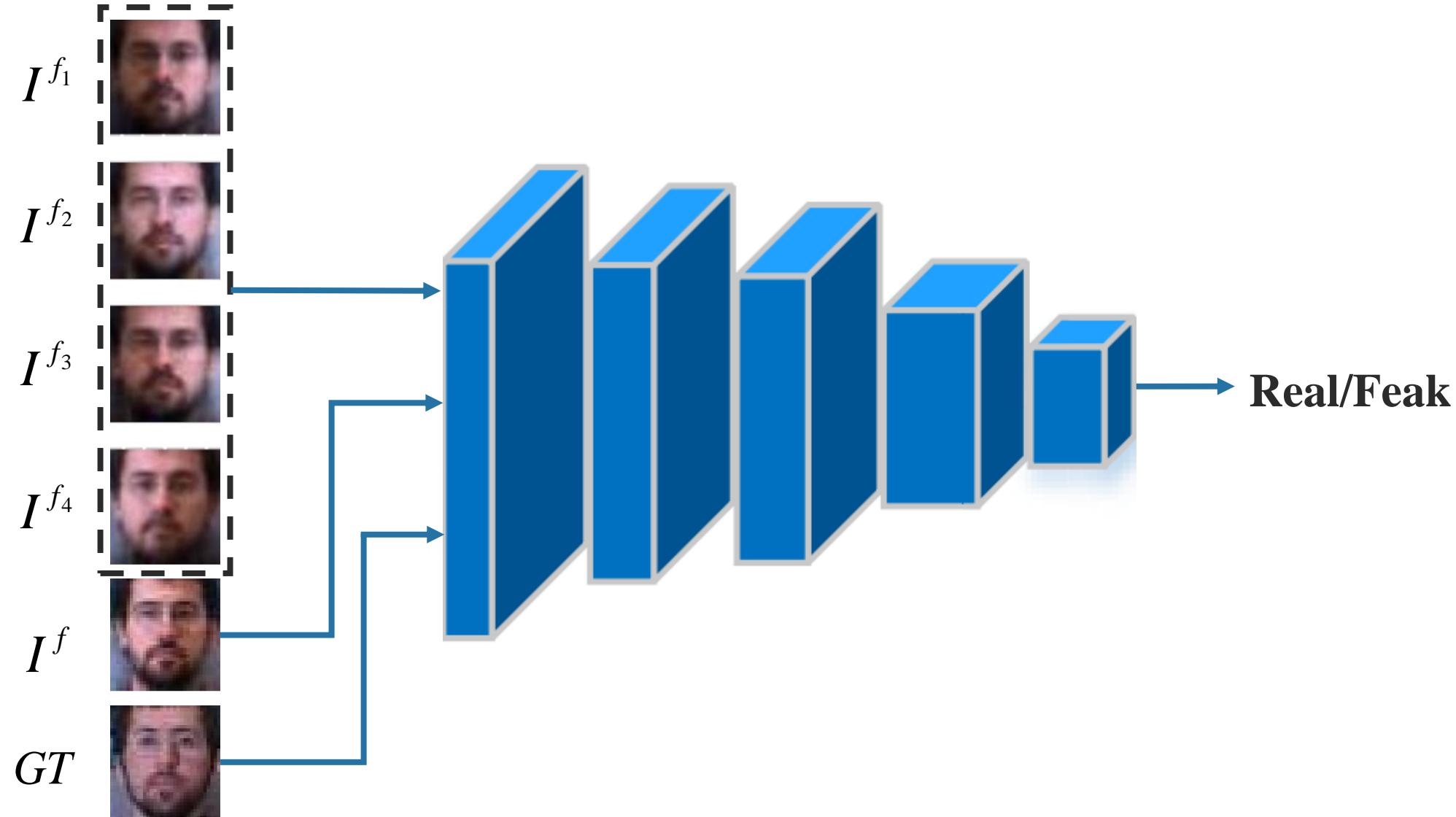
Layer	Input	Filter Size	Output Size
resblock1	Images	$5 \times 5, 12$ $5 \times 5, 12$ $\times 1$	$128 \times 128 \times 12$
conv1	resblock1	$5 \times 5, 64$	$128 \times 128 \times 64$
resblock2	conv1	$3 \times 3, 64$ $3 \times 3, 64$ $\times 1$	$128 \times 128 \times 64$
conv2	resblock2	$3 \times 3, 32$	$128 \times 128 \times 32$
conv3	conv2	$3 \times 3, 3$	$128 \times 128 \times 3$





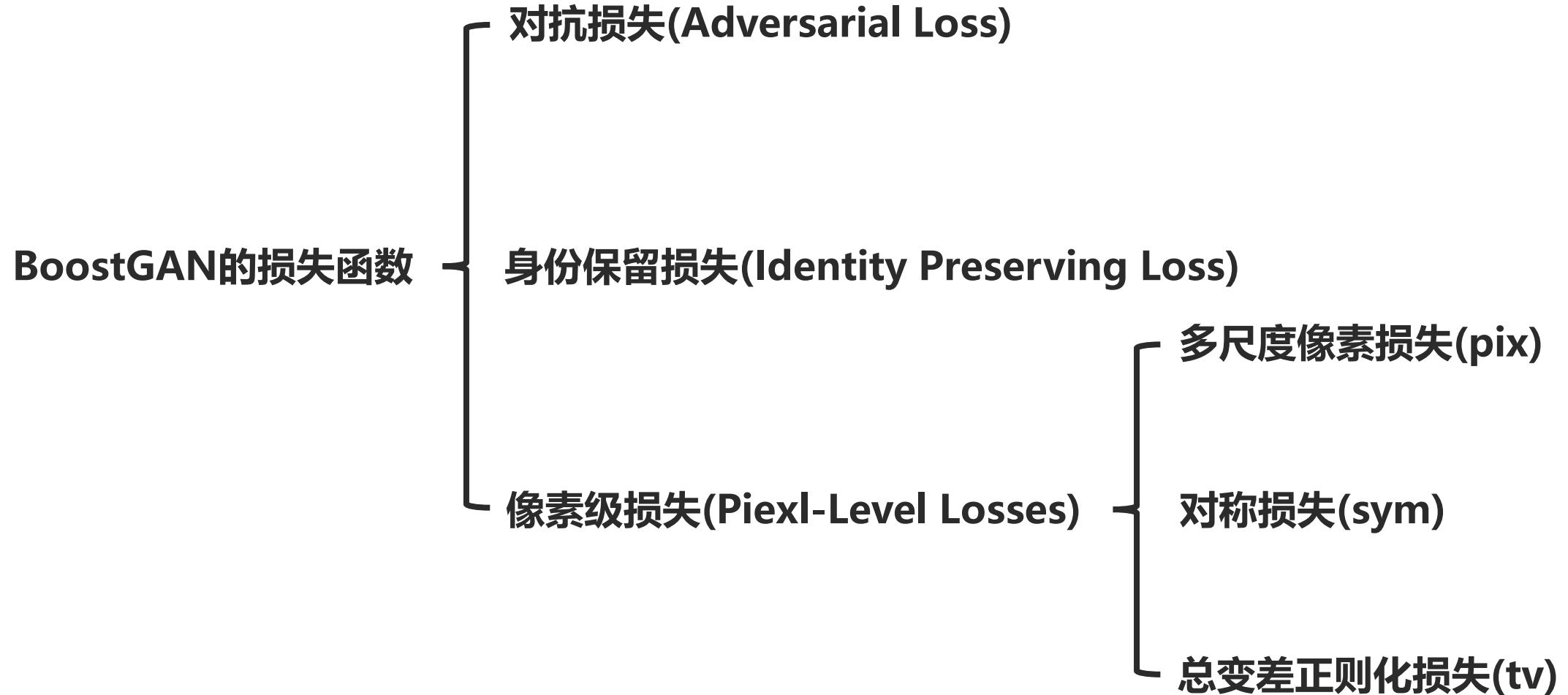
Multi-Input Discriminator

Multi-Input Discriminator用 D 表示，其功能是判断多种输入的去遮挡正面人脸的真假。





损失函数



BoostGAN的训练目标就是最小化上述5种损失的加权和。



对抗损失

在一个batch中， G^c 、 G^f 与 D 的完整博弈过程可以表示为：

$$\min_{G^c, G^f} \max_D V(D, G) = \frac{1}{N} \sum_{n=1}^N \left\{ \log D(I_n^{gt}) + \frac{1}{5} \left(\sum_{i=1}^4 \log(1 - D(G^c(I_n^{b_i}))) + \log(1 - D(G^f(I_n^{f_1}, I_n^{f_2}, I_n^{f_3}, I_n^{f_4}))) \right) \right\} \quad (3)$$

只与判别器 D 有关的项，
让判别器判别真图像

让判别器判别 G^c 生成的4张假图像

让判别器判别 G^f 生成的1张假图像

对应 G^c 的项

对应 G^f 的项

在实践中，通过交替优化 D 和 $\{G^c, G^f\}$ 来完成训练，即有：

固定 $\{G^c, G^f\}$ 的参数， D 的训练目标：

$$L_D = \max_D V(D, G) = \frac{1}{N} \sum_{n=1}^N \left\{ \log D(I_n^{gt}) + \frac{1}{5} \left(\sum_{i=1}^4 \log(1 - D(G^c(I_n^{b_i}))) + \log(1 - D(G^f(I_n^{f_1}, I_n^{f_2}, I_n^{f_3}, I_n^{f_4}))) \right) \right\} \quad (4)$$



对抗损失

固定 D 的参数, G^c 和 G^f 的训练目标:

$$L_{adv} = \max_{G^c, G^f} V(D, G) = \frac{1}{5N} \sum_{n=1}^N \left\{ \sum_{i=1}^4 \log D(G^c(I_n^{b_i})) + \log D(G^f(I_n^{f_1}, I_n^{f_2}, I_n^{f_3}, I_n^{f_4})) \right\} \quad (5)$$

$$\min_{G^c, G^f} \max_D V(D, G) = \frac{1}{N} \sum_{n=1}^N \left\{ \log D(I_n^{gt}) + \frac{1}{5} \left(\sum_{i=1}^4 \log(1 - D(G^c(I_n^{b_i}))) + \log(1 - D(G^f(I_n^{f_1}, I_n^{f_2}, I_n^{f_3}, I_n^{f_4}))) \right) \right\} \quad (3)$$



身份保留损失

用于确保最终输出的人脸在生成的过程中能保留原始人脸的身份信息。身份保留损失通常是在生成的人脸和真实的人脸上应用预训练的人脸识别模型来得到的。

这里使用到的预训练人脸识别模型是Light CNN。

$$L_{ip} = \sum_{i=1}^5 \left| F^{po}(I^{gt}) - F^{po}(\hat{I}^i) \right| + \left| F^{fc}(I^{gt}) - F^{fc}(\hat{I}^i) \right| \quad (6)$$

在式(6)中：

$F^{po}(I^{gt})$ 表示真实人脸在Light CNN最后的池化层的输出

$F^{fc}(I^{gt})$ 表示真实人脸在Light CNN最后的全连接层的输出

$F^{po}(\hat{I}^i)$ 表示所有的生成人脸在Light CNN最后的池化层的输出

$F^{fc}(\hat{I}^i)$ 表示所有的生成人脸在Light CNN最后的全连接层的输出



像素级损失

像素级损失用于在像素级别上量化生成人脸与真实人脸的差异，并约束生成的人脸有一定的视觉外观。

(1) 多尺度像素损失(pix)

$$L_{pix} = \sum_{i=1}^5 \frac{1}{K} \sum_{k=1}^K \frac{1}{W_k H_k C} \sum_{w,h,c=1}^{W_k, H_k, C} |\hat{I}_{k,w,h,c}^i - I_{k,w,h,c}^{gt}| \quad (7)$$

在式(7)中， K 表示人脸图像可以划分出的尺度数量， W_k 和 H_k 分别表示第 k 个尺度上的宽和高。共测试了3种尺度，分别为 32×32 、 64×64 和 128×128 。

(2) 对称损失(sym)

$$L_{sym} = \sum_{i=1}^5 \frac{1}{W/2 \times H} \sum_{w=1}^{W/2} \sum_{h=1}^H |\hat{I}_{w,h}^i - \hat{I}_{w_s,h}^i| \quad (8)$$

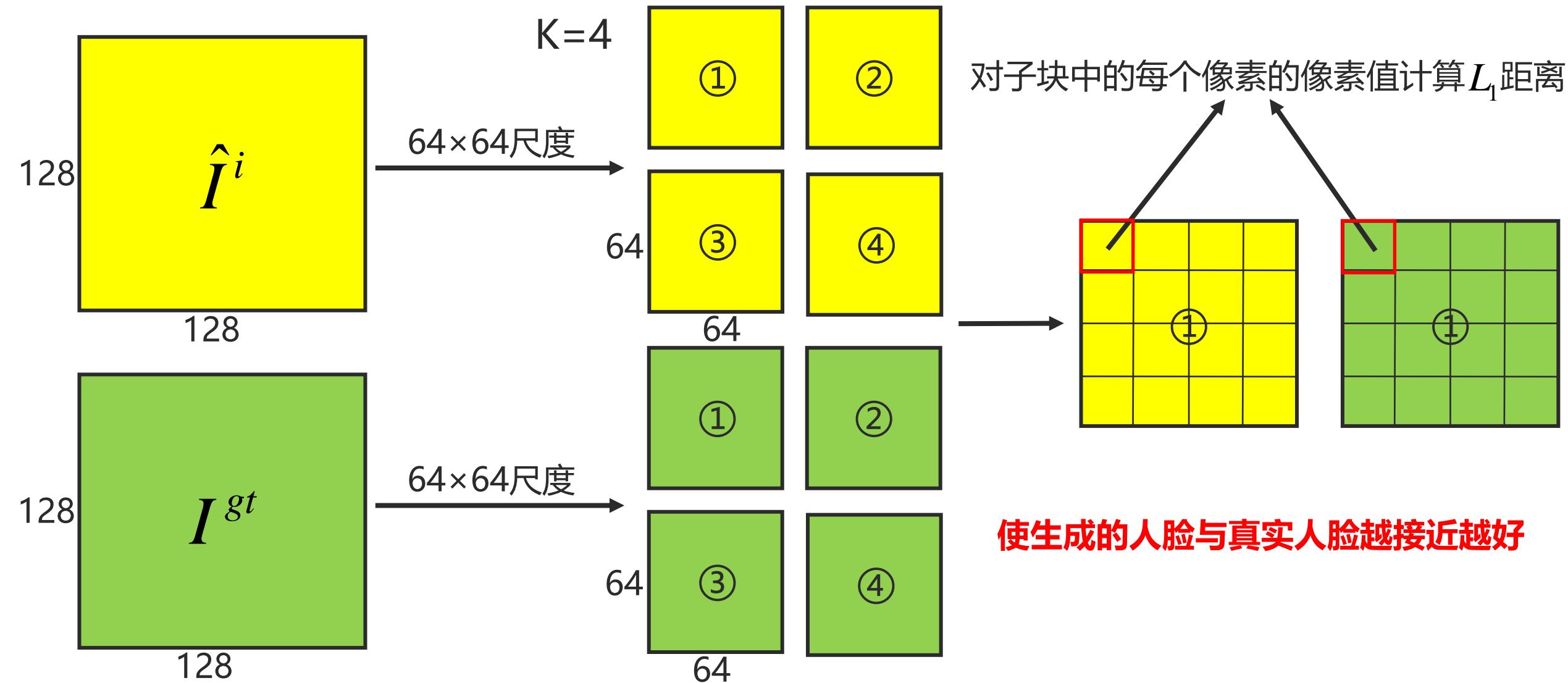
在式(8)中， $w_s = W - (w - 1)$ 表示生成的人脸图像中，横坐标为 w 的像素的对称像素的横坐标。

(3) 总变差正则化损失(tv)

$$L_{tv} = \sum_{i=1}^5 \sum_{c=1}^C \sum_{w,h=1}^{W,H} |\hat{I}_{w+1,h,c}^i - \hat{I}_{w,h,c}^i| + |\hat{I}_{w,h+1,c}^i - \hat{I}_{w,h,c}^i| \quad (9)$$

对多尺度像素损失的理解

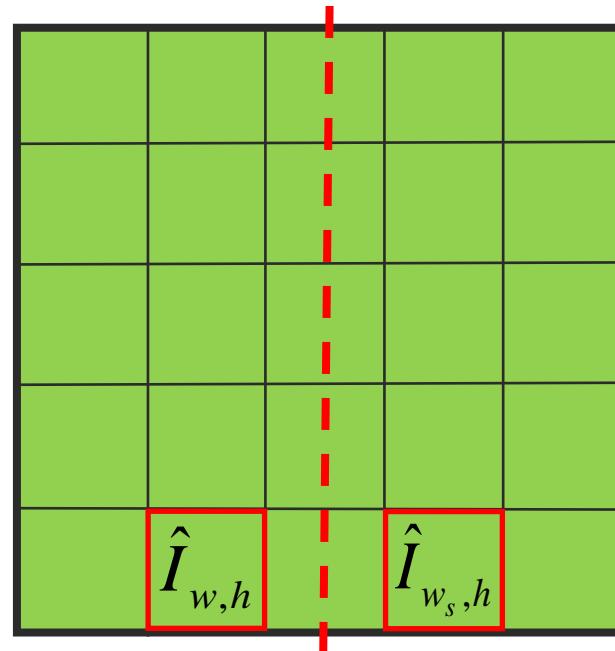
$$L_{pix} = \sum_{i=1}^5 \frac{1}{K} \sum_{k=1}^K \frac{1}{W_k H_k C} \sum_{w,h,c=1}^{W_k, H_k, C} |\hat{I}_{k,w,h,c}^i - I_{k,w,h,c}^{gt}|$$



对称损失和总变差正则化损失的理解

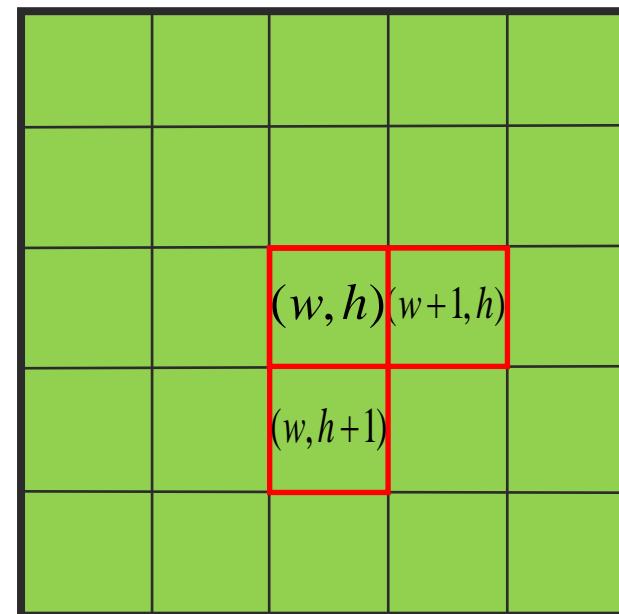
$$L_{sym} = \sum_{i=1}^5 \frac{1}{W/2 \times H} \sum_{w=1}^{W/2} \sum_{h=1}^H \left| \hat{I}_{w,h}^i - \hat{I}_{w_s,h}^i \right|$$

$$w_s = W - (w-1)$$



鼓励生成的人脸具有对称的外观

$$L_{tv} = \sum_{i=1}^5 \sum_{c=1}^C \sum_{w,h=1}^{W,H} \left| \hat{I}_{w+1,h,c}^i - \hat{I}_{w,h,c}^i \right| + \left| \hat{I}_{w,h+1,c}^i - \hat{I}_{w,h,c}^i \right|$$



鼓励生成的人脸具有平滑的外观



生成器的总损失

综上，生成器的总损失就是上述5种损失的加权和。

$$L_G = \lambda_1 L_{adv} + \lambda_2 L_{ip} + \lambda_3 L_{pix} + \lambda_4 L_{sym} + \lambda_5 L_{tv} \quad (10)$$

在式(10)中， $\lambda_1 - \lambda_5$ 是权重参数。

两种不同的测试方案

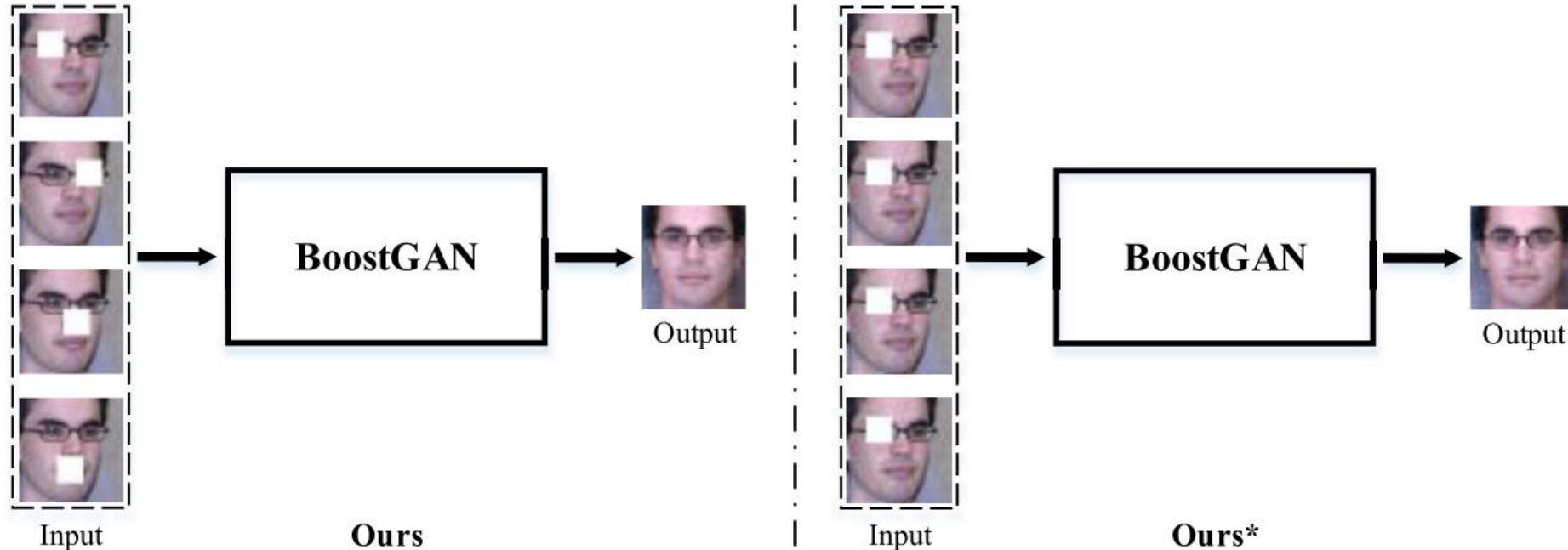


Fig. 3. Two different testing schemes with regard to inputs. Left: four faces with different keypoint occlusions are feeded as inputs, which is called Ours. Right: four faces with the same occlusions are feeded as inputs, which is called Ours*.

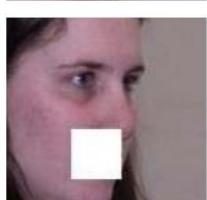


Part
03

实验



实验数据分类和数据集



规则遮挡

不规则遮挡

Multi-PIE数据集：**受控**数据集，记录了来自337名受试者的超过750,000张图片。使用所有表情中性、姿势在 ± 75 度内的337名受试者，前200个受试者组成训练集，其余137个受试者作为测试集。

LFW (Labeled Faces in the Wild) 数据集：**非受控**数据集，更贴合实际自然，包含5749个受试者的13233张图片，只用作测试。



规则遮挡：数据预处理和训练参数设置

使用MTCNN 检测并将这两个数据集对齐到 128×128 的规范视图。

添加规则遮挡，包括关键点遮挡和随机遮挡，每个遮挡掩模的大小设为 32×32 。

设置 $\lambda_1 = 2e1$ 、 $\lambda_2 = 4e1$ 、 $\lambda_3 = 1$ 、 $\lambda_4 = 3e-1$ 、 $\lambda_5 = 1e-3$ 。

使用 Adam 优化器训练我们的网络。学习率固定在 0.0001。

Batch 设置为 4，训练在 4-5 个 epoch 后停止。

模型在 NVIDIA 1080Ti GPU 上进行训练，整个模型训练过程大约需要 6 ~ 8 天。

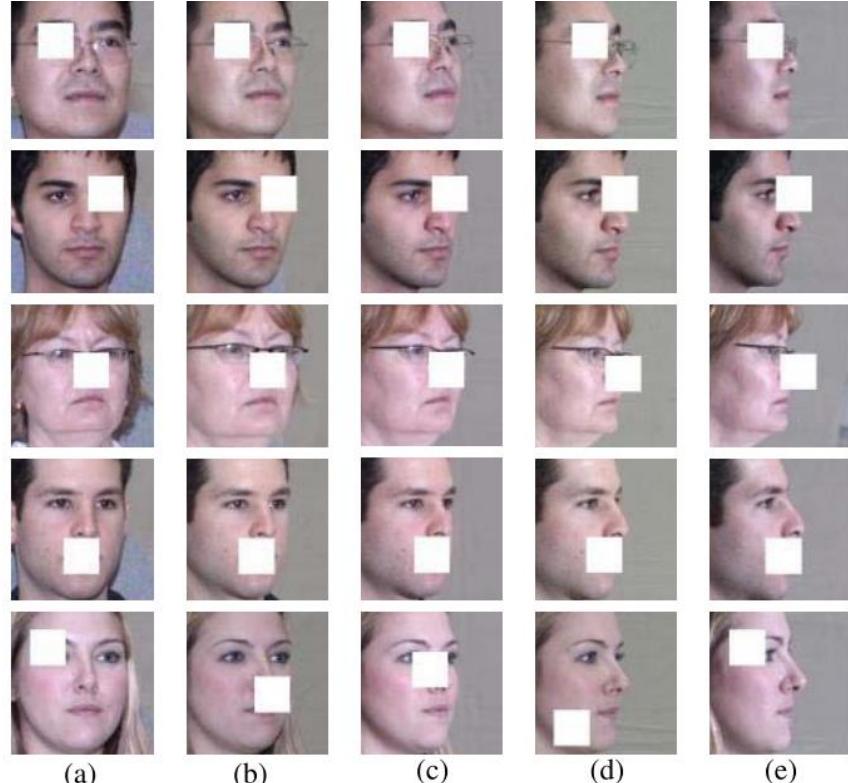


Fig. 4. Examples on regular block occluded Multi-PIE. From left to right, the poses are (a) 15° , (b) 30° , (c) 45° , (d) 60° , and (e) 75° . From top to bottom, the occluded masks are located on left eye, right eye, the tip of nose, the center of mouth, and random block, respectively.

规则遮挡：Multi-PIE关键点遮挡实验结果

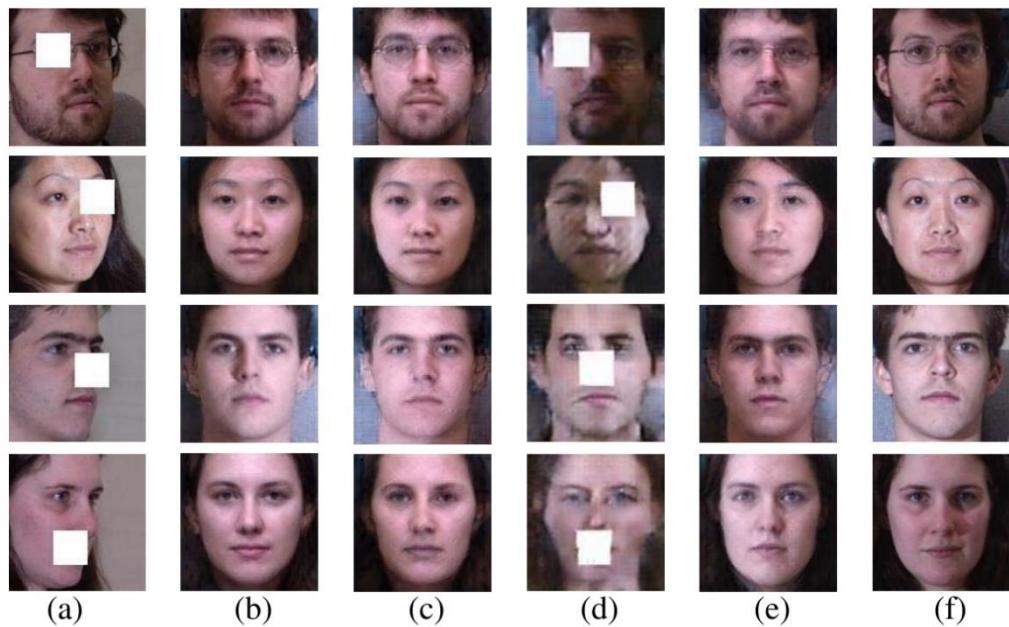


Fig. 5. Synthesis results on keypoint region occluded Multi-PIE data set. From top to bottom, the poses are 15° , 30° , 45° , 60° . The ground-truth frontal faces are provided at column (f). (a) Profile. (b) Ours. (c) Ours*. (d) [23]. (e) [24]*. (f) GT.

Ours的效果明显优于其他方法；即使用四个相同的输入（Ours*），仍然优于DR-GAN和TP-GAN*。

RANK-1 RECOGNITION RATE (%) COMPARISON ON KEYPOINT REGION OCCLUDED MULTI-PIE. BLACK: RANKS THE FIRST;
RED: RANKS THE SECOND; BLUE: RANKS THE THIRD

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$
DR-GAN [23] (k1)	67.38	60.68	55.83	47.25	39.34
DR-GAN [23] (k2)	73.24	65.37	59.90	51.18	42.24
DR-GAN [23] (k3)	66.93	60.60	56.54	49.70	39.77
DR-GAN [23] (k4)	71.33	63.72	57.59	50.10	40.87
DR-GAN [23] (mean)	69.72	62.59	57.47	49.56	40.55
DR-GAN [23] (k3_DP)	54.9	54.9	53.3	50.7	-
TP-GAN [24]* (k1)	98.17	95.46	86.60	65.91	39.51
TP-GAN [24]* (k2)	99.27	97.25	88.37	66.03	40.82
TP-GAN [24]* (k3)	95.04	90.95	82.72	62.40	38.67
TP-GAN [24]* (k4)	97.80	93.66	83.84	62.27	36.76
TP-GAN [24]* (mean)	97.57	94.33	85.38	64.15	38.94
TP-GAN [24]* (k3_DP)	63.44	54.93	51.82	44.69	25.72
Ours*(k1)	99.03	96.21	86.66	64.45	39.74
Ours*(k2)	99.12	96.33	85.41	63.49	39.63
Ours*(k3)	96.03	92.06	83.10	63.45	39.92
Ours*(k4)	98.13	94.88	84.24	62.17	39.45
Ours*(mean)	98.08	94.87	84.85	63.39	39.69
Ours	99.48	97.75	91.55	72.76	48.43

规则遮挡：Multi-PIE随机遮挡实验结果

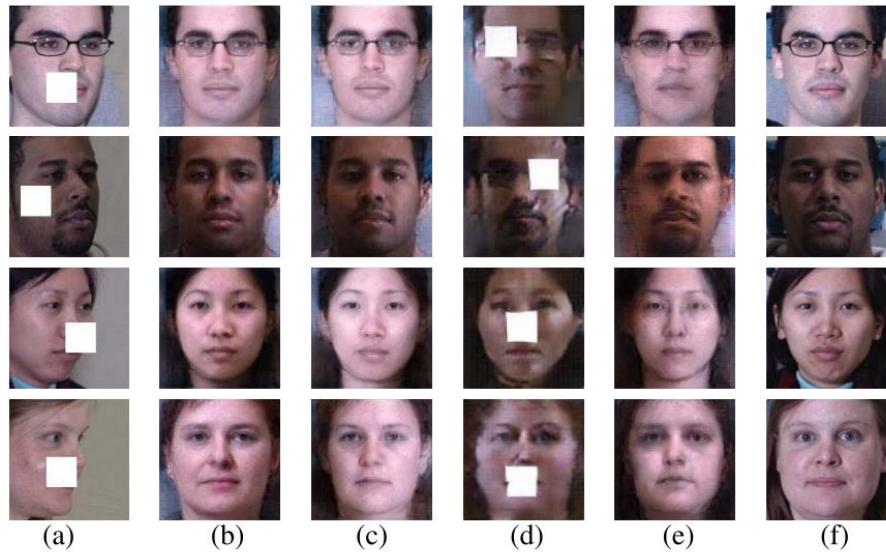


Fig. 6. Frontalization results on random block occluded Multi-PIE. The poses are 15° , 30° , 45° , and 60° from top to bottom. The GT frontal faces are provided at (f). Notably, all the models are trained solely on keypoint position occluded Multi-PIE data set. (a) Profile. (b) Ours. (c) Ours*. (d) [23]. (e) [24]*. (f) GT.

DR-GAN和TP-GAN*随机遮挡对比关键点
遮挡识别性能下降明显，而BoostGAN性能几乎
没有下降。

RANK-1 RECOGNITION RATE (%) COMPARISON ON RANDOM BLOCK OCCLUDED MULTI-PIE. BLACK: RANKS THE FIRST; RED: RANKS THE SECOND; BLUE: RANKS THE THIRD

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$	$\pm 15^\circ$
DR-GAN [23] (r1)	47.64	38.93	33.21	25.38	18.92	67.38
DR-GAN [23] (r2)	65.75	55.15	46.52	38.33	29.00	73.24
DR-GAN [23] (r3)	56.01	46.27	39.13	29.11	23.01	66.93
DR-GAN [23] (r4)	59.10	47.92	39.97	33.69	25.20	71.33
DR-GAN [23] (mean)	57.13	47.07	39.71	31.63	24.03	69.72
TP-GAN [24]* (r1)	89.81	83.88	74.94	54.83	31.34	98.17
TP-GAN [24]* (r2)	77.98	71.68	60.52	42.68	23.92	99.27
TP-GAN [24]* (r3)	79.12	72.45	60.00	41.37	24.11	95.04
TP-GAN [24]* (r4)	86.13	77.76	64.84	45.08	25.15	97.80
TP-GAN [24]* (mean)	83.26	76.44	65.08	45.99	26.13	97.57
Ours*(r1)	98.16	95.07	86.67	66.47	43.06	99.03
Ours*(r2)	98.10	94.97	86.51	66.56	42.96	99.12
Ours*(r3)	98.02	94.99	86.78	66.25	42.69	96.03
Ours*(r4)	98.12	95.12	86.69	65.78	42.67	98.13
Ours*(mean)	98.10	95.04	86.66	66.27	42.85	98.08
Ours	99.45	97.50	91.11	72.12	48.53	99.48

关键点遮挡 $\pm 15^\circ$ 实验结果

规则遮挡：LFW关键点遮挡实验结果

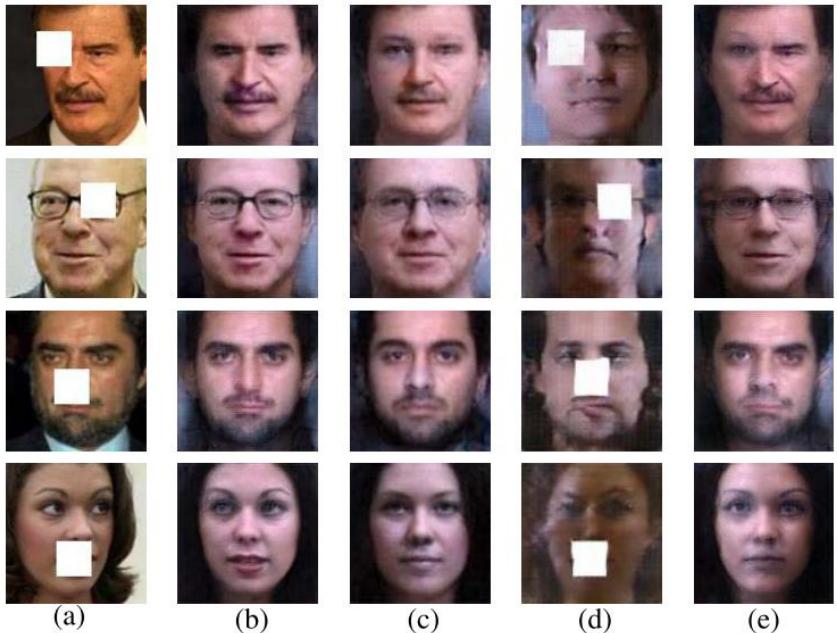


Fig. 7. Frontalization results on keypoint region occluded LFW data set in the wild. Notably the ground-truth frontal images for this data set are unavailable. The models are trained based solely on keypoint position occluded Multi-PIE data set. (a) Profile. (b) Ours. (c) Ours*. (d) [23]. (e) [24]*.

LFW真值图不可用，面部验证性能通过识别准确率(ACC)和特征曲线下面积(AUC)评估。我们的两个模型仍然优于DR-GAN和TP-GAN*的性能。

FACE VERIFICATION ACCURACY (ACC) AND AUC RESULTS ON KEYPOINT REGION OCCLUDED LFW

Method	ACC(%)	AUC(%)
DR-GAN [23] (k1)	67.60	73.65
DR-GAN [23] (k2)	67.28	72.94
DR-GAN [23] (k3)	58.43	59.19
DR-GAN [23] (k4)	69.50	76.05
DR-GAN [23] (mean)	65.71	70.46
TP-GAN [24]* (k1)	86.52	92.81
TP-GAN [24]* (k2)	87.83	93.96
TP-GAN [24]* (k3)	85.17	91.63
TP-GAN [24]* (k4)	87.78	93.97
TP-GAN [24]* (mean)	86.83	93.09
Ours*(k1)	88.47	94.07
Ours*(k2)	88.13	93.85
Ours*(k3)	86.62	93.66
Ours*(k4)	87.88	93.92
Ours*(mean)	87.78	93.88
Ours	89.57	94.90

规则遮挡：LFW随机遮挡实验结果

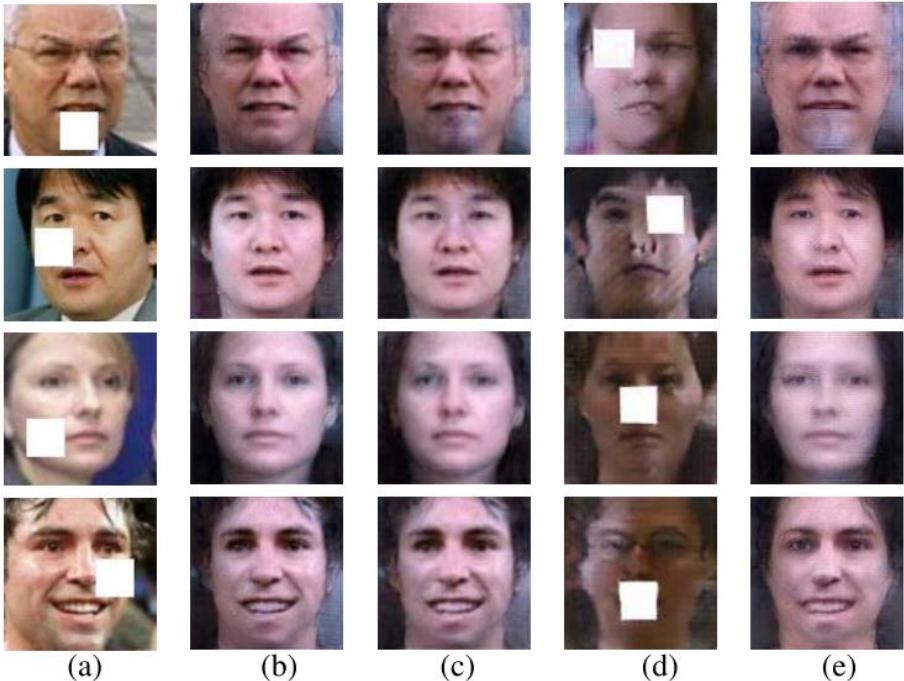


Fig. 8. Frontalization results on random block occluded LFW data set. Notably, all the models are trained solely on keypoint position occluded Multi-PIE data set, without retraining on randomly blocked data sets. (a) Profile. (b) Ours. (c) Ours*. (d) [23]. (e) [24]*.

模型在受控和非受控场景中都对规则遮挡展示了良好的泛化能力和鲁棒性。

FACE VERIFICATION ACCURACY (ACC) AND AUC RESULTS ON RANDOM BLOCK OCCLUDED LFW

Method	ACC(%)	AUC(%)	ACC(%)
DR-GAN [23] (r1)	63.28	67.20	67.60
DR-GAN [23] (r2)	65.53	71.79	67.28
DR-GAN [23] (r3)	57.15	57.76	58.43
DR-GAN [23] (r4)	64.82	70.35	69.50
DR-GAN [23] (mean)	62.70	66.78	65.71
TP-GAN [24]* (r1)	82.75	89.86	86.52
TP-GAN [24]* (r2)	77.65	84.63	87.83
TP-GAN [24]* (r3)	81.07	88.24	85.17
TP-GAN [24]* (r4)	83.25	90.15	87.78
TP-GAN [24]* (mean)	81.18	88.22	86.83
Ours*(r1)	87.92	93.46	88.47
Ours*(r2)	86.78	92.81	88.13
Ours*(r3)	87.28	93.15	86.62
Ours*(r4)	87.18	93.26	87.88
Ours*(mean)	87.29	93.17	87.78
Ours	89.58	94.75	89.57

关键点遮挡ACC实验结果



不规则遮挡：数据预处理

使用MTCNN 检测并将这两个数据集对齐到 128×128 的规范视图

添加不规则遮挡。为遮挡提供了四种不同类型的不规则遮挡掩模M。不规则掩模M的尺寸为 128×128 ，与输入图像相同。

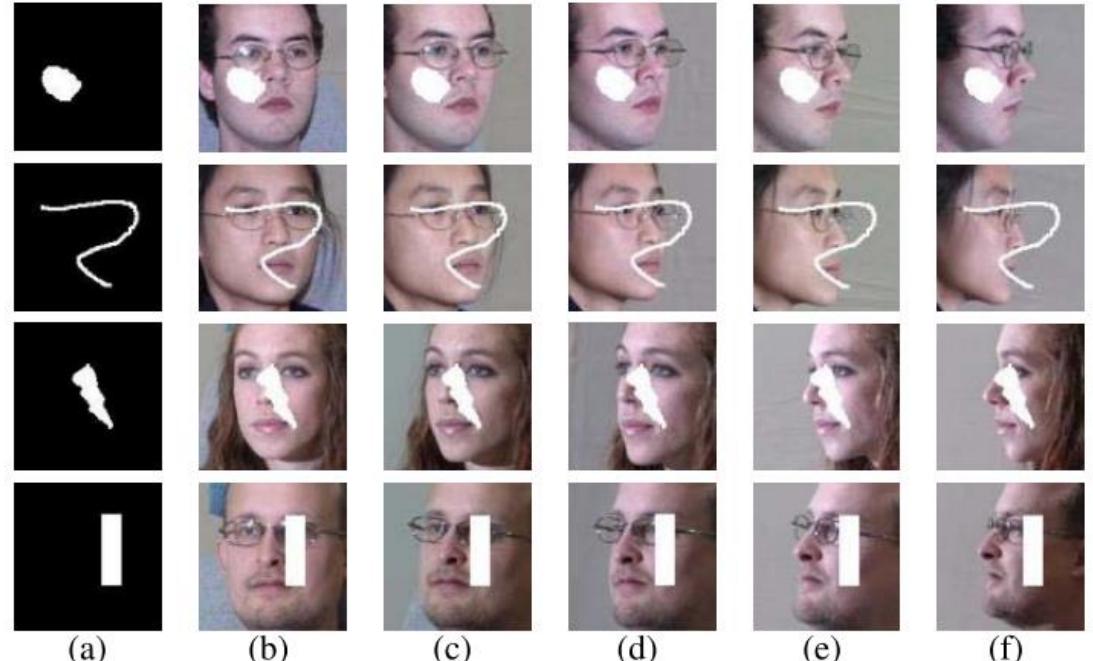


Fig. 9. Examples on irregular block occluded Multi-PIE. (a) Four kinds of irregular masks deployed in this article. (b)–(f) Irregular block occluded face samples under the poses of $\pm 15^\circ$, $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$, and $\pm 75^\circ$ based on the four masks. (a) Mask. (b) $\pm 15^\circ$. (c) $\pm 30^\circ$. (d) $\pm 45^\circ$. (e) $\pm 60^\circ$. (f) $\pm 75^\circ$.

不规则遮挡：Multi-PIE实验结果

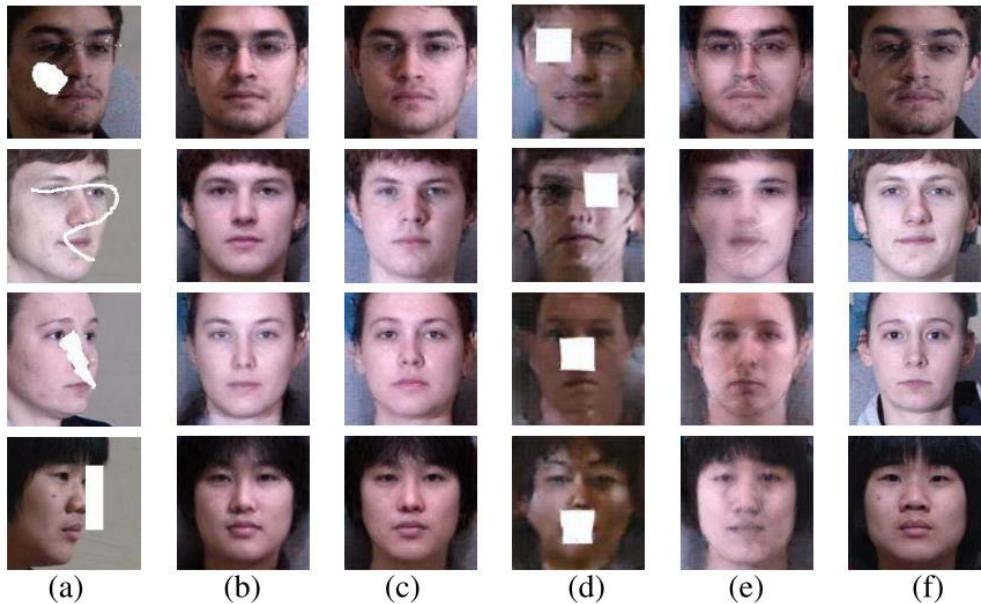


Fig. 10. Frontalization results on irregular block occluded Multi-PIE. The poses are 15° , 30° , 45° , and 60° from top to bottom. The GT frontal images are provided at (f). Notably, all the models are trained solely on keypoint position occluded Multi-PIE data set. (a) Profile. (b) Ours. (c) Ours*. (d) [23] (e) [24]*. (f) GT.

DR-GAN和TP-GAN*显示出明显的人工痕迹和鬼影效果。

COMPARISON OF RANK-1 RECOGNITION RATE (%) ON IRREGULAR BLOCK OCCLUDED MULTI-PIE. BLACK: RANKS THE FIRST;
RED: RANKS THE SECOND; BLUE: RANKS THE THIRD

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$
DR-GAN [23] (m1)	60.51	52.68	44.88	35.69	23.59
DR-GAN [23] (m2)	73.40	63.59	50.74	40.10	29.13
DR-GAN [23] (m3)	73.38	65.44	55.63	44.16	33.56
DR-GAN [23] (m4)	63.70	43.62	35.74	32.17	27.34
DR-GAN [23] (mean)	67.75	56.33	46.75	38.03	28.41
TP-GAN [24]* (m1)	94.11	89.98	81.32	57.23	28.93
TP-GAN [24]* (m2)	61.60	48.65	39.19	24.47	13.98
TP-GAN [24]* (m3)	86.21	79.40	67.69	47.69	24.34
TP-GAN [24]* (m4)	85.84	76.71	64.97	44.35	26.18
TP-GAN [24]* (mean)	81.94	73.69	63.29	43.44	23.36
Ours*(m1)	99.12	96.84	90.20	70.36	44.29
Ours*(m2)	93.27	86.13	70.92	46.08	27.65
Ours*(m3)	93.14	88.17	76.53	56.67	34.52
Ours*(m4)	98.56	95.21	85.99	65.59	42.11
Ours*(mean)	96.02	91.59	80.91	59.68	37.14
Ours	98.81	96.10	88.14	67.23	44.03

不规则遮挡：LFW实验结果

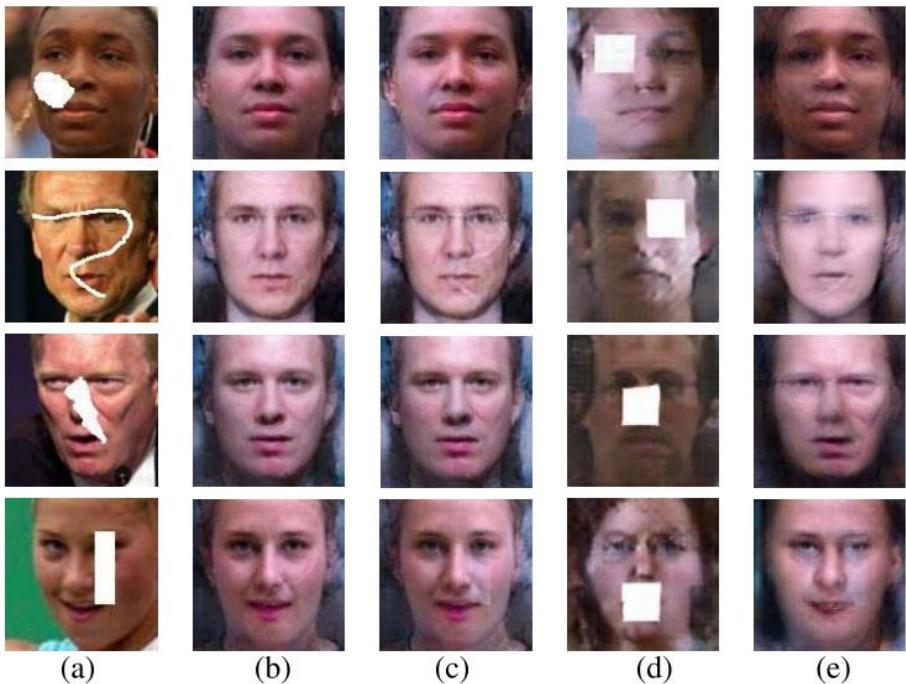


Fig. 11. Frontalization results on irregular block occluded LFW data set. Notably, all the models are trained solely on keypoint position occluded Multi-PIE data set, without retraining on irregular occluded data sets. (a) Profile. (b) Ours. (c) Ours*. (d) [23]. (e) [24]*.

BoostGAN对于不同的遮挡在受控和非受控场景中均不敏感，反映了其良好的泛化能力和鲁棒性。

FACE VERIFICATION ACCURACY (ACC) AND AUC RESULTS
IRREGULAR BLOCK OCCLUDED LFW

Method	ACC(%)	AUC(%)
DR-GAN [23] (m1)	66.27	71.85
DR-GAN [23] (m2)	66.93	73.11
DR-GAN [23] (m3)	58.22	59.20
DR-GAN [23] (m4)	66.82	72.58
DR-GAN [23] (mean)	64.56	69.19
TP-GAN [24]* (m1)	84.17	91.22
TP-GAN [24]* (m2)	77.12	83.88
TP-GAN [24]* (m3)	82.80	89.87
TP-GAN [24]* (m4)	84.32	91.03
TP-GAN [24]* (mean)	82.10	89.00
Ours*(m1)	88.45	94.09
Ours*(m2)	84.98	91.59
Ours*(m3)	86.23	92.83
Ours*(m4)	86.98	93.06
Ours*(mean)	86.66	92.89
Ours	88.52	93.91



两步法去遮挡实验：动机

通用的**两步法**解决带遮挡的侧面人脸正面化和识别问题：

- (1) 去遮挡（图像补全）
- (2) 面部正面化

图像补全方法面临的问题：

- (1) 图像补全去除遮挡后将丢失纹理细节和身份信息。
- (2) 图像修复旨在恢复精细的细节和语义面部结构，而不是精确地回归真实情况。
- (3) 面部补全方法旨在恢复正面或近正面图像，而不是侧面图像。

综上，两步法可能不适合解决遮挡下的面部正面化和识别问题。



两步法去遮挡实验：实验结果

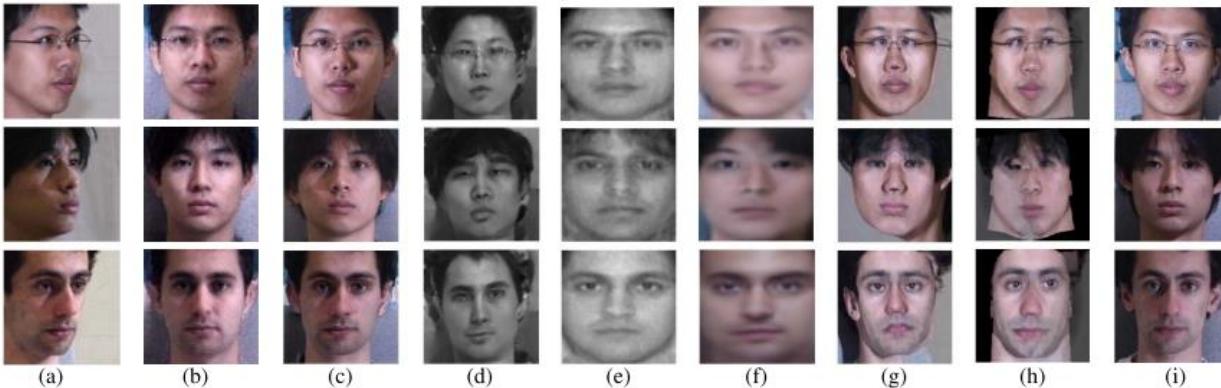


Fig. 12. Comparison with SOTA synthesis models under the pose variation of 45° (the first two rows) and 30° (the third row). Our BoostGAN model is trained from scratch on the nonoccluded Multi-PIE data set. (a) Profile. (b) Ours. (c) [24]. (d) [23]. (e) [76]. (f) [77]. (g) [60]. (h) [21]. (i) GT.

RANK-1 RECOGNITION RATE (%) COMPARISON ON PROFILE
MULTI-PIE WITH OCCLUSION REMOVED FIRST

采用无遮挡的图像进行实验。Light CNN的结果作为基准。BoostGAN在识别清晰侧面图像方面优于所有其他方法。

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	mean
FIP+LDA [78]	90.7	80.7	64.1	45.9	70.35
MVP+LDA [79]	92.8	83.7	72.9	60.1	77.38
CPF [76]	95.0	88.5	79.9	61.9	81.33
DR-GAN [23]	94.0	90.1	86.2	83.2	88.38
DR-GAN _{AM} [29]	95.0	91.3	88.0	85.8	90.03
FF-GAN [62]	94.6	92.5	89.7	85.2	90.50
TP-GAN [24]	98.68	98.06	95.38	87.72	94.96
Light CNN [70]	98.59	97.38	92.13	62.09	87.55
Ours	99.88	99.19	96.84	87.52	95.86



生成器的消融分析

COMPARISON WITH DIFFERENT BOOSTING ACROSS COARSE AND FINE
GENERATOR ON KEYPOINT REGION OCCLUDED MULTI-PIE
FOR BETTER INSIGHT OF THE BOOSTING EFFECT

Boosting	±15°	±30°	±45°	±60°	±75°	
Only 1 fine	95.52	86.20	69.29	46.99	32.18	← 体现粗生成器 G^c 的作用
1 coarse + 1 fine	98.47	93.13	78.16	53.93	34.70	
2 coarse + 1 fine	98.95	95.53	86.09	61.62	36.97	体现中间结果的作用
3 coarse + 1 fine	99.39	96.23	85.20	60.67	39.91	
Only 4 coarse	97.63	90.78	76.14	51.38	33.18	← 体现细生成器 G^f 的作用
4 coarse + 1 fine	99.48	97.75	91.55	72.76	48.43	

总结



这篇文章提出了一个BoostGAN模型。它用于具有**姿态和遮挡**的面部正面化和识别，是一个端到端的框架，配备了一个粗到细的面部去遮挡和正面化网络集合。

粗生成器用于在多重遮挡和大姿态变化中实现粗略的正面化和去遮挡。增强网络旨在通过增强多个输入的补充信息，生成干净、正面、逼真且保留身份的面部。

在规则和不规则遮挡下进行实验，实验结果表明 BoostGAN确实优于其他最先进的GAN模型。



谢谢！
请老师和同学批评指正

| 汇报人：赵思桓、刘睿龙、曾强文

2024.5.15