



# Simultaneous Face Completion and Frontalization via Mask Guided Two-Stage GAN

指导老师： 张磊 汇报人： 陈江垚、潘建臣、刘瑞欣

# 目录

## CONTENTS



研究背景与  
动机



研究方法与  
内容



研究结果与  
讨论



结论与展望



# 研究背景与动机



# 研究背景与动机

## 计算机视觉中的面部识别和合成

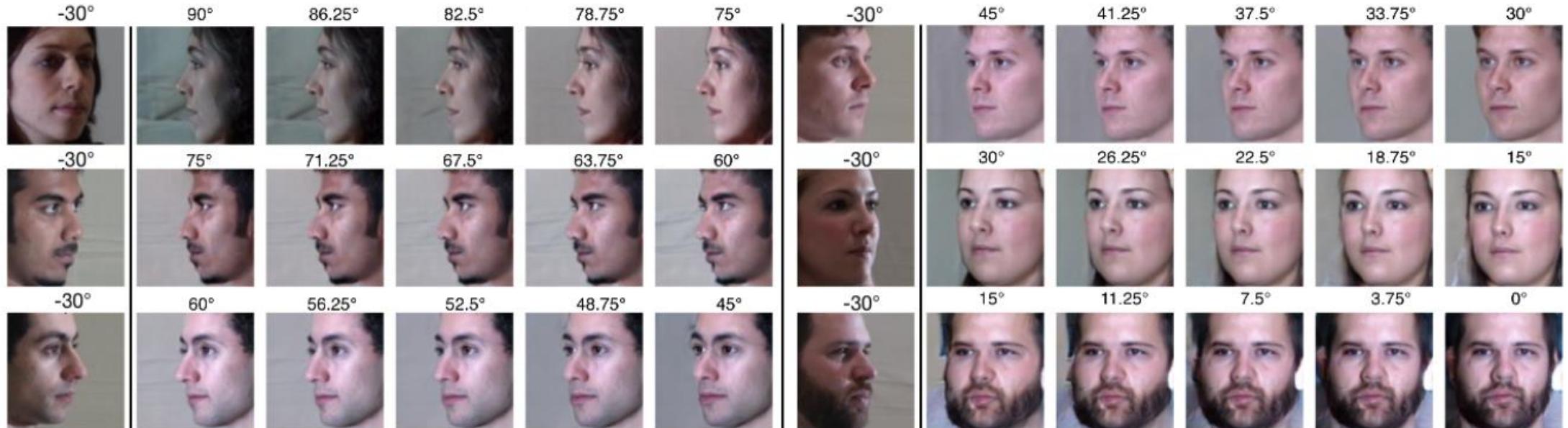


面部识别与合成是计算机视觉领域的基础课题，在许多领域都有广泛的应用。得益于海量的已标记人脸数据和深度学习的快速发展，虽然取得了长足的进步，但是也存在着许多制约算法性能进一步提高的因素。



# 研究背景与动机

## 制约算法性能提高的因素



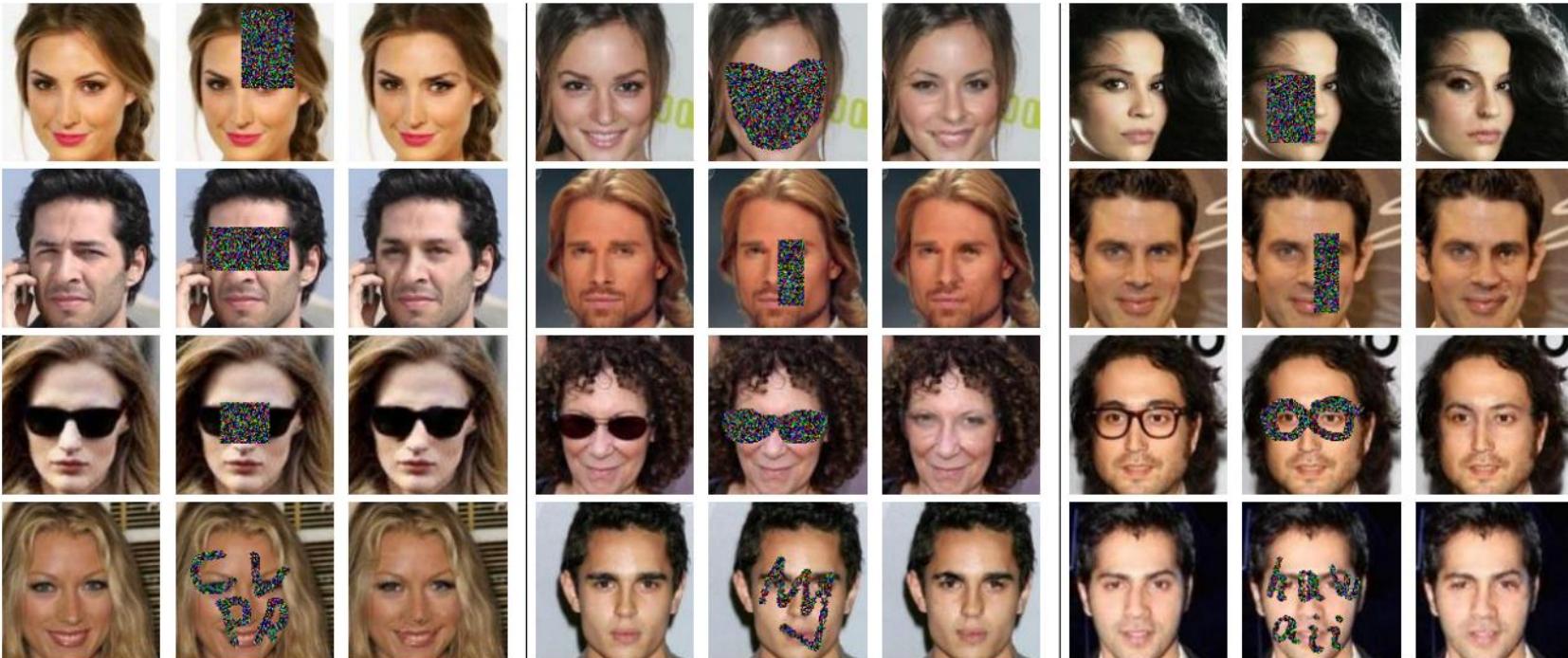
姿态变化：最近的一项研究<sup>[1]</sup>观察到姿态变化可能导致10%的准确度下降

[1] Frontal to profile face verification in the wild, (WACV 2016)



# 研究背景与动机

## 制约算法性能提高的因素

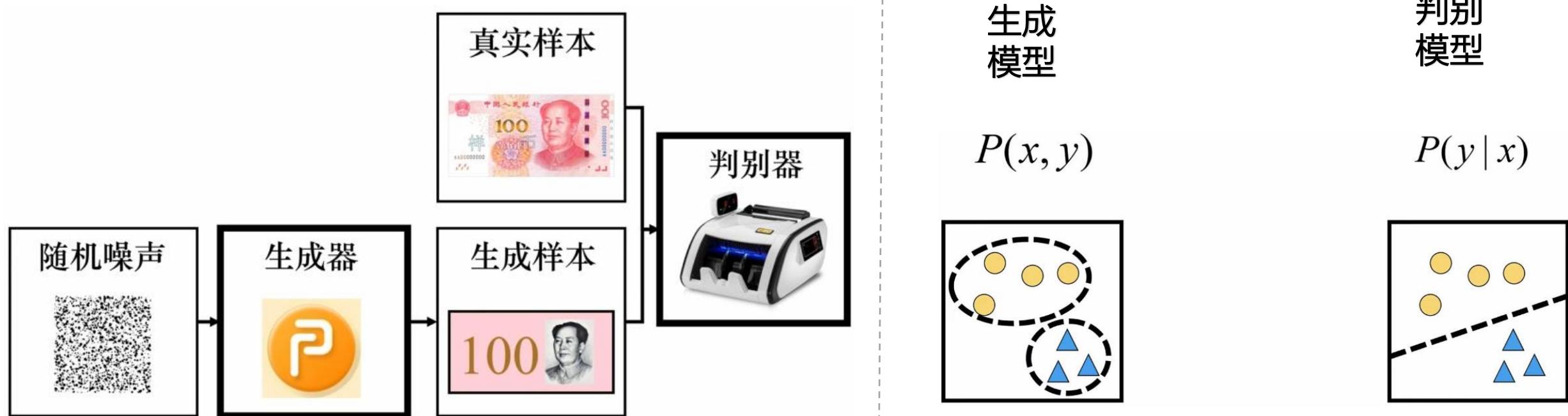


面部遮挡：遮挡也是严重影响人脸识别性能的一个关键因素。其原因很明显，  
人脸上某些重要的局部区域的“缺失”会导致人脸特征的识别损失。



# 相关工作

## Generative adversarial network(GAN)





# 相关工作

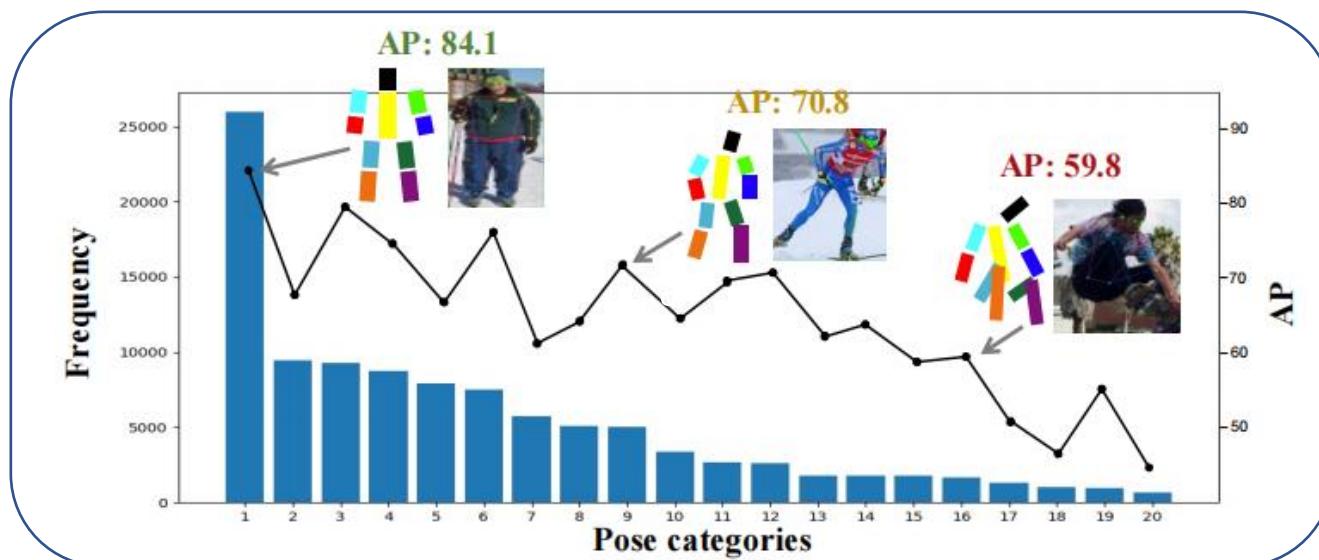
## 姿态变化相关工作

相关工作: direct pose-invariant feature based approaches 和 indirect identity-preserved synthesis based approaches

➤ direct pose-invariant feature based approaches: 直接基于姿态不变特征的方法

① 通过提出不同的监督损失函数来增强所提取特征的区分度 (三重损失<sup>[2]</sup>、中心损失、大边缘余弦损失<sup>[3]</sup>)

**弊端:** 海量人脸图像的采集、标注和处理是耗时且昂贵的以及存在长尾分布的问题



[2] FaceNet: A unified embedding for face recognition and clustering(CVPR 2015)

[3] A discriminative feature learning approach for deep face recognition(ECCV 2016)



# 相关工作

## 姿态变化相关工作

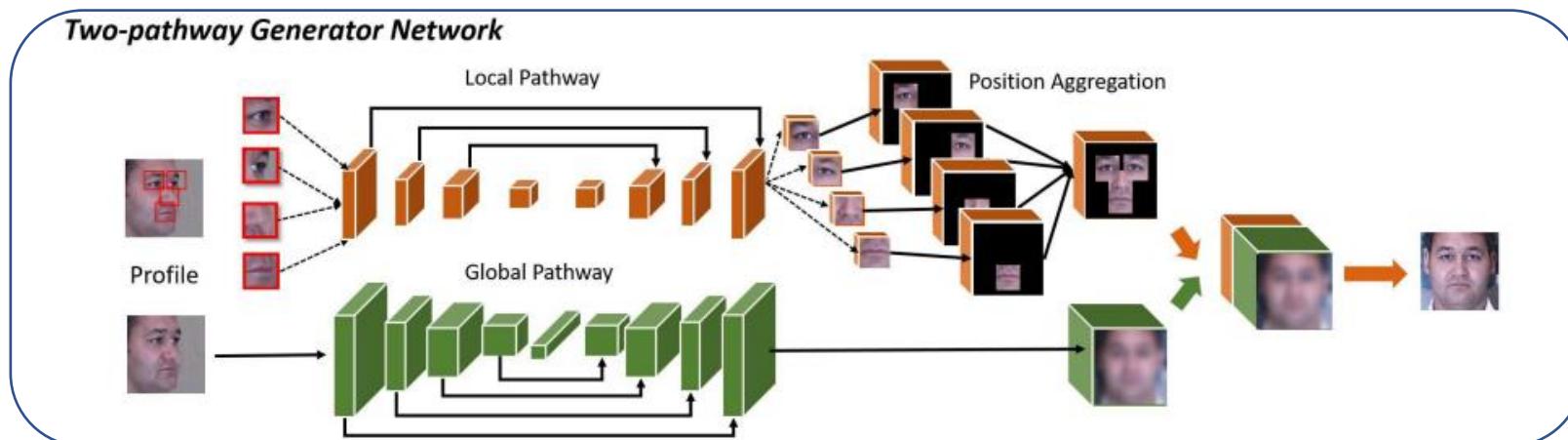
➤ indirect identity-preserved synthesis based approaches: 间接基于身份保留合成的方法

② 利用3D变形模型 (3DMM) 拟合和渲染面部图像来完成面部合成任务

**弊端：**拟合和绘制的高计算量与高精度，增加了基于3D的间接合成方法的实现难度

③ 以生成对抗网络 (GAN) 为基础的各个代表性工作 (DR-GAN<sup>[4]</sup>、TP-GAN<sup>[5]</sup>)

**弊端：**仅与姿势变化相关联，而不考虑实际生活中经常发生的面部遮挡



[4] FaceNet: A unified embedding for face recognition and clustering(CVPR 2015)

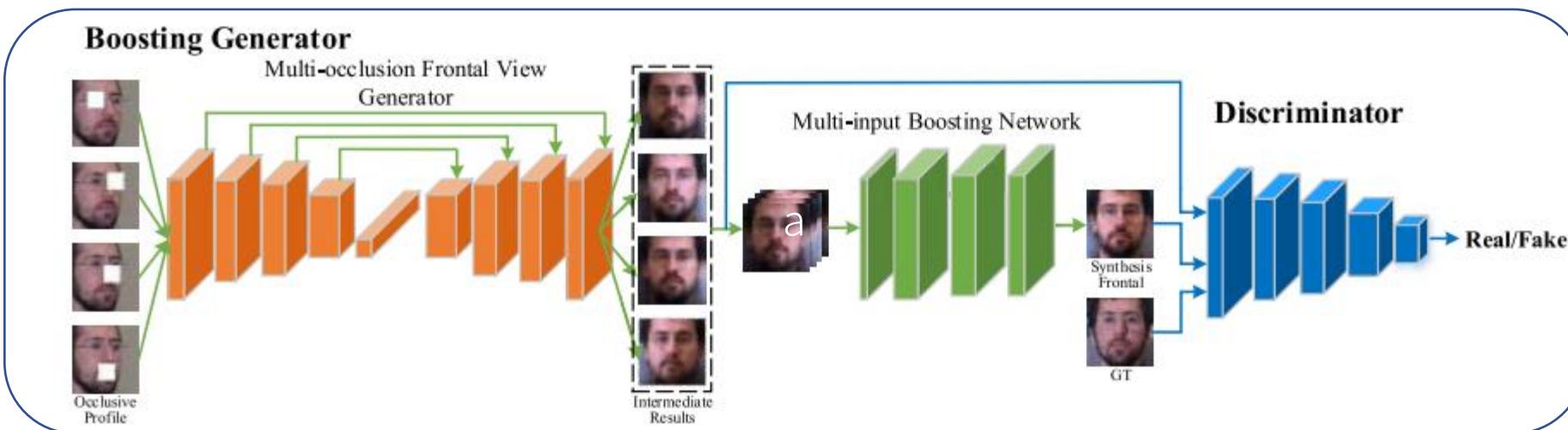
[5] Beyond Face Rotation: Global and Local Perception GAN for Photorealistic and Identity Preserving Frontal View Synthesis(ICCV2017)



# 相关工作

## 面部遮挡+姿态变换相关工作

相关工作：Q. Duan and L. Zhang提出BoostGAN<sup>[6]</sup>



**问题1：**被遮挡轮廓图像的噪声先验信息作为重要的先验信息，没有被充分利用，使得图像恢复过程无方向性，影响了图像的真实感。

**问题2：**由于姿态变化和面部遮挡的共存，这人脸恢复和人脸正面化需要分别完成。BoostGAN作为一种单阶段方法，不能很好地处理这两个任务，并且去遮挡过程是隐式和不可解释的。这些缺陷限制了它的可扩展性。

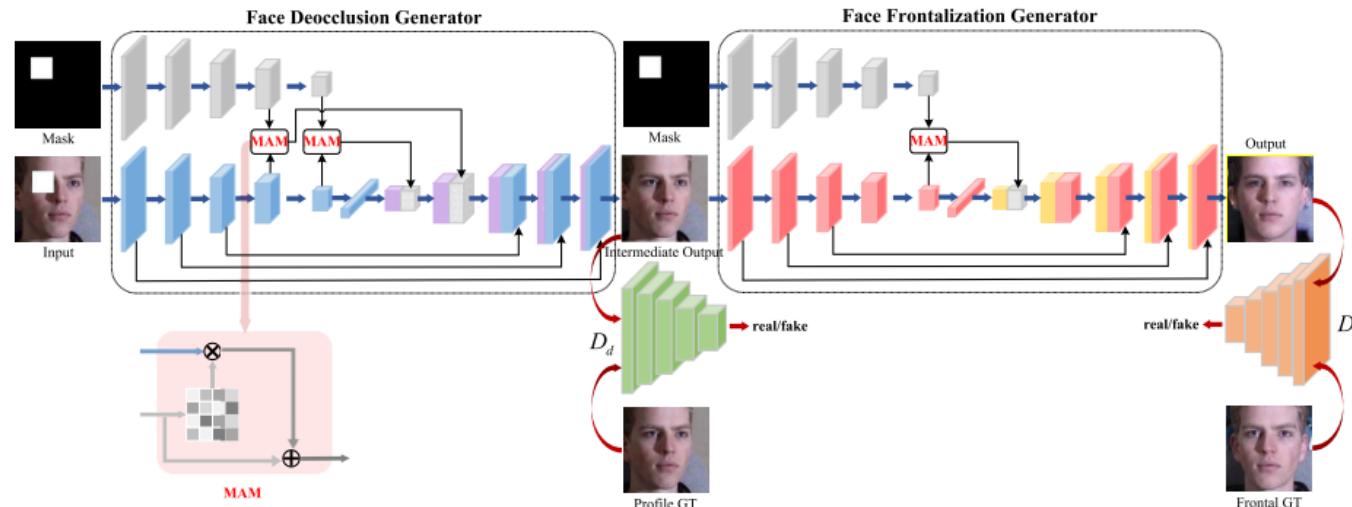
[6] Look more into occlusion: Realistic face frontalization and recognition with BoostGAN( IEEE Trans. Netw.Learn. Syst., vol. 32, no. 1, pp. 214–228, Jan. 2021.)



# 动机

## 问题与解决方法

### 解决问题



分别完成人脸  
恢复和正面化

- ① **两阶段生成对抗网络 (TSGAN)**，人脸恢复模块和人脸正面化模块。
- ②设计了一个带**U型连接的掩模关注模块 (MAM)**，采用遮挡掩模作为注意力，充分利用遮挡的先验知识

**新的损失函数 双三重损失：**通过最小化类内差异和最大化类间差异来提高身份信息的区分度



# 研究方法

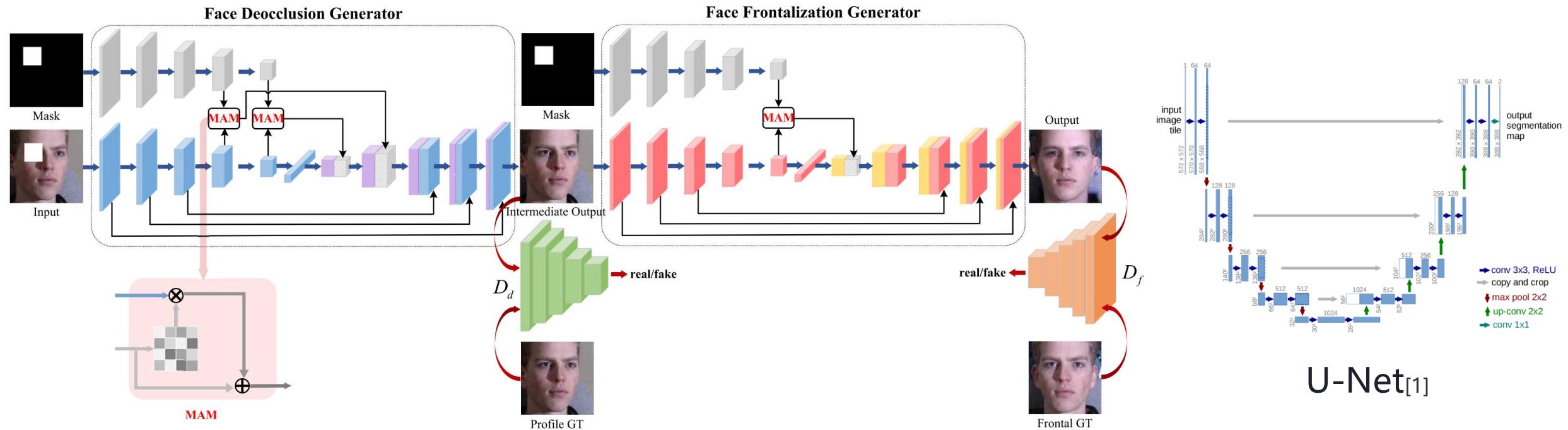


# 研究方法

## A. Face Deocclusion Module and Face Frontalization Module

TSGAN网络由两个模块组成：

人脸去遮挡(Deocclusion)模块和人脸正向化(Frontalization)模块

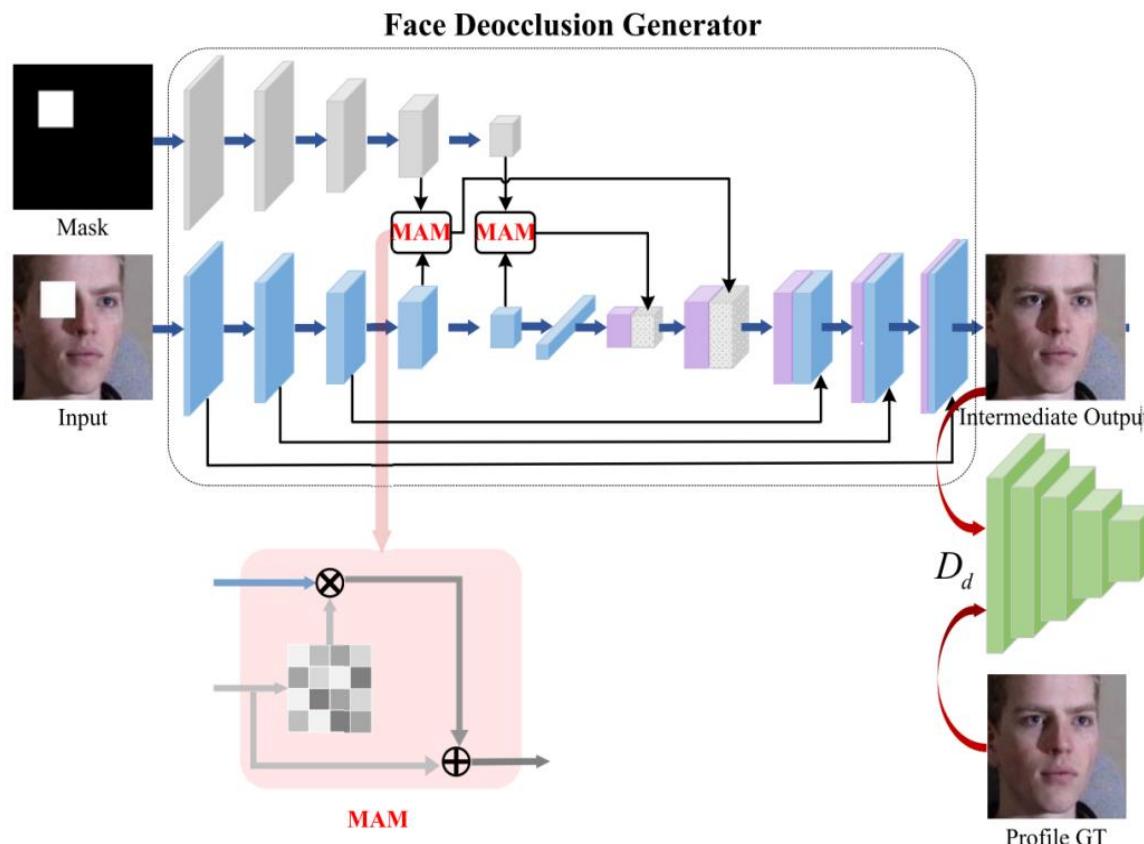




# 研究方法

## A. Face Deocclusion Module and Face Frontalization Module

### 人脸去遮挡(Deocclusion)模块



- 人脸去遮挡模块目的：通过恢复被遮挡的侧面图像来获得一个清晰的侧面图像
- 人脸去遮挡模块组成：包括一个人脸去遮挡生成器  $G_d$  和一个人脸去遮挡鉴别器  $D_d$
- 人脸去遮挡生成器  $G_d$  组成：包含人脸去遮挡编码器网络  $G_d^{enc}$  和人脸去遮挡解码器网络  $G_d^{dnc}$ 。

人脸去遮挡的处理可以如下表述：

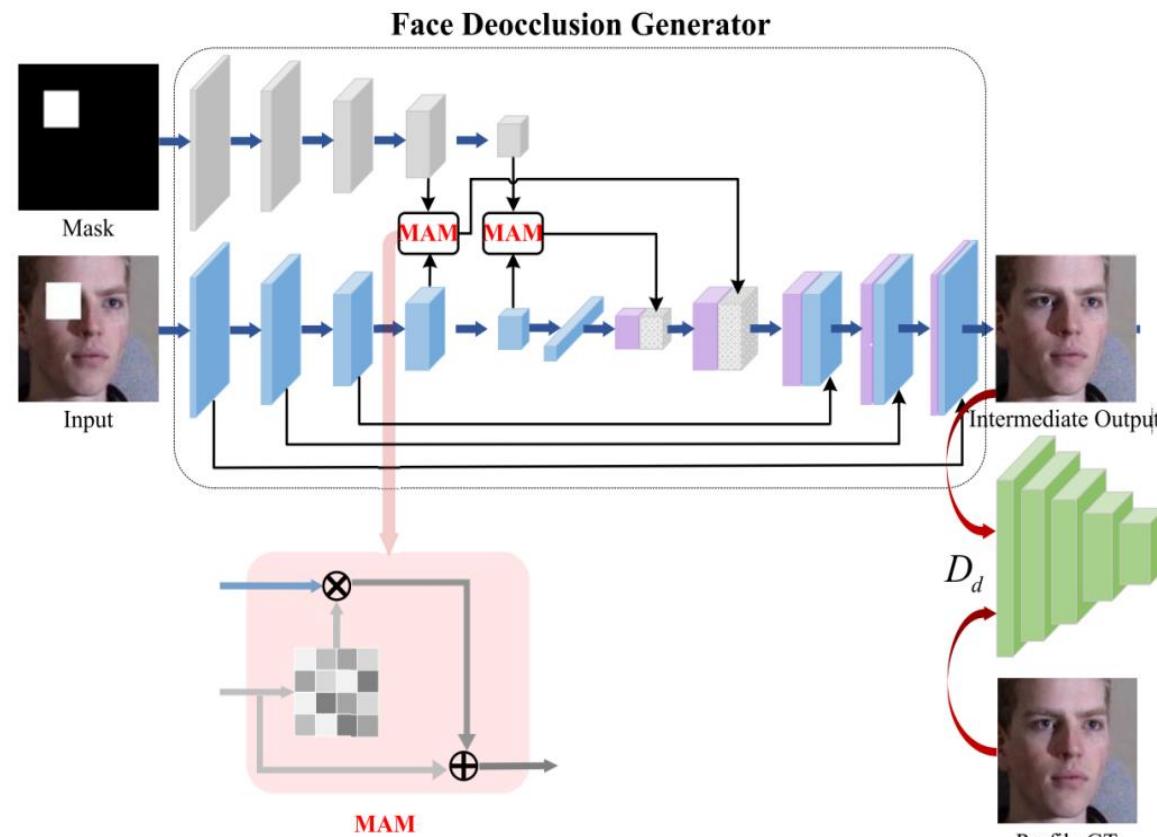
$$I_r = G_d(I, M) = G_d^{dec}(G_d^{enc}(I, M)), \quad (1)$$



# 研究方法

## A. Face Deocclusion Module and Face Frontalization Module

### 人脸去遮挡(Deocclusion)模块



➤ 人脸去遮挡处理表述如下：

$$I_r = G_d(I, M) = G_d^{dec}(G_d^{enc}(I, M)), \quad (1)$$

➤ 人脸去遮挡生成器  $G_d$  和一个人脸去遮挡鉴别器  $D_d$  之间的博弈可以表述如下：

$$\begin{aligned} \min_{G_d} \max_{D_d} V_d(G_d, D_d) = & \frac{1}{N} \sum_{n=1}^N \{\log D_d(\hat{I}_p^n) \\ & + \log(1 - D_d(G_d(I^n, M^n)))\}, \end{aligned} \quad (2)$$

➤ 在训练中， $G_d$  和  $D_d$  是通过交替求解以下优化问题得到的：

$$\begin{aligned} L_{adv}^{G_d} &= \min_{G_d} V_d(G_d, D_d) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log(1 - D_d(G_d(I^n, M^n)))\}. \end{aligned} \quad (3)$$

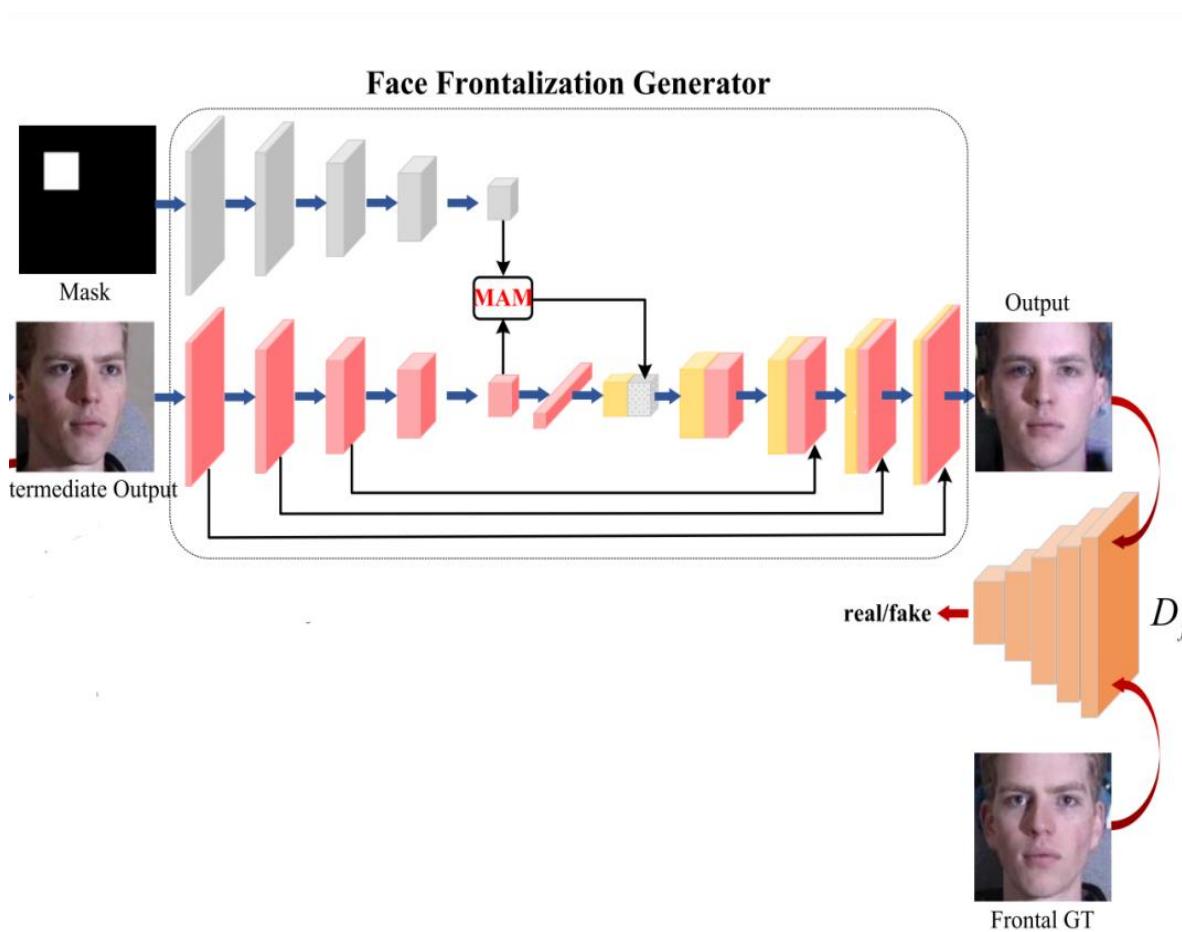
$$\begin{aligned} L_{D_d} &= \max_{D_d} V_d(G_d, D_d) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log D_d(\hat{I}_p^n) \\ &+ \log(1 - D_d(G_d(I^n, M^n)))\}. \end{aligned} \quad (4)$$



# 研究方法

## A. Face Deocclusion Module and Face Frontalization Module

### 人脸正向化(Frontalization)模块



- 人脸正向化生成器的骨干网络称为 $G_f$ 。 $G_f$ 的编码器网络和解码器网络分别表示为 $G_f^{enc}$ 和 $G_f^{dec}$

$$I_f = G_f(I_r, M) = G_f^{dec}(G_f^{enc}(I_r, M)), \quad (5)$$

- 相应地，鉴别器 $D_f$ 也被设计用于对抗过程。表述如下：

$$\begin{aligned} \min_{G_f} \max_{D_f} V_f(G_f, D_f) = & \frac{1}{N} \sum_{n=1}^N \{\log D_f(\hat{I}_f^n) \\ & + \log(1 - D_f(G_f(I_r^n, M^n)))\}, \end{aligned} \quad (6)$$

- 在训练中， $G_f$ 和 $D_f$ 是通过交替求解以下优化问题得到的：

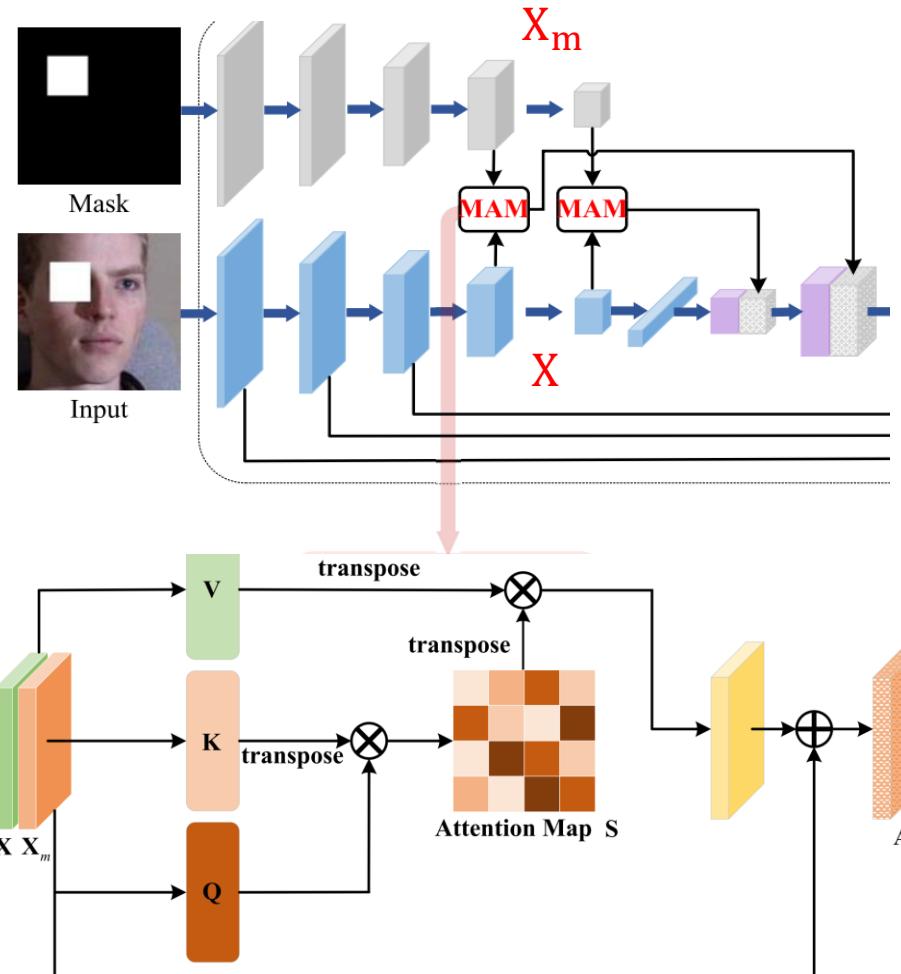
$$\begin{aligned} L_{adv}^{G_f} = \min_{G_f} V_f(G_f, D_f) \\ = \frac{1}{N} \sum_{n=1}^N \{\log(1 - D_f(G_f(I_r^n, M^n)))\}. \end{aligned} \quad (7)$$

$$\begin{aligned} L_{D_f} = \max_{D_f} V_f(G_f, D_f) \\ = \frac{1}{N} \sum_{n=1}^N \{\log D_f(\hat{I}_f^n) \\ + \log(1 - D_f(G_f(I_r^n, M^n)))\}. \end{aligned} \quad (8)$$



# 研究方法

## B. Mask-Attention Module



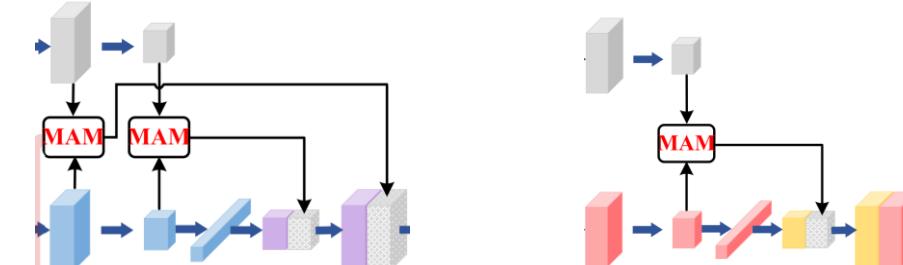
掩模注意力(Mask-Attention)模块

- 掩模注意力模块目的：进一步提高第一阶段去遮挡生成器的人脸还原能力和身份鉴别能力。
- 首先， $X_m$ 分别输入两个不同的一维卷积层，生成两个新的特征图 $K$ 和 $Q$ ，其中  $K, Q \in \mathbb{R}^{B \times C}$
- 其次，对 $K$ 和 $Q$ 的转置进行矩阵乘法，得到一个大小为 $B \times B$ 的注意力图。然后，采用softmax层计算得到的注意力图的相似度：

$$S_{ij} = \frac{\exp(\mathbf{K}_i^T * \mathbf{Q}_j)}{\sum_{b=1}^B \exp(\mathbf{K}_i^T * \mathbf{Q}_b)}, \quad (9)$$

- 最后通过逐元素求和操作将其与 $X_m$ 合并

$$\mathbf{A} = \alpha (\mathbf{V}^T * \mathbf{S}^T) + \mathbf{X}_m, \quad (10)$$

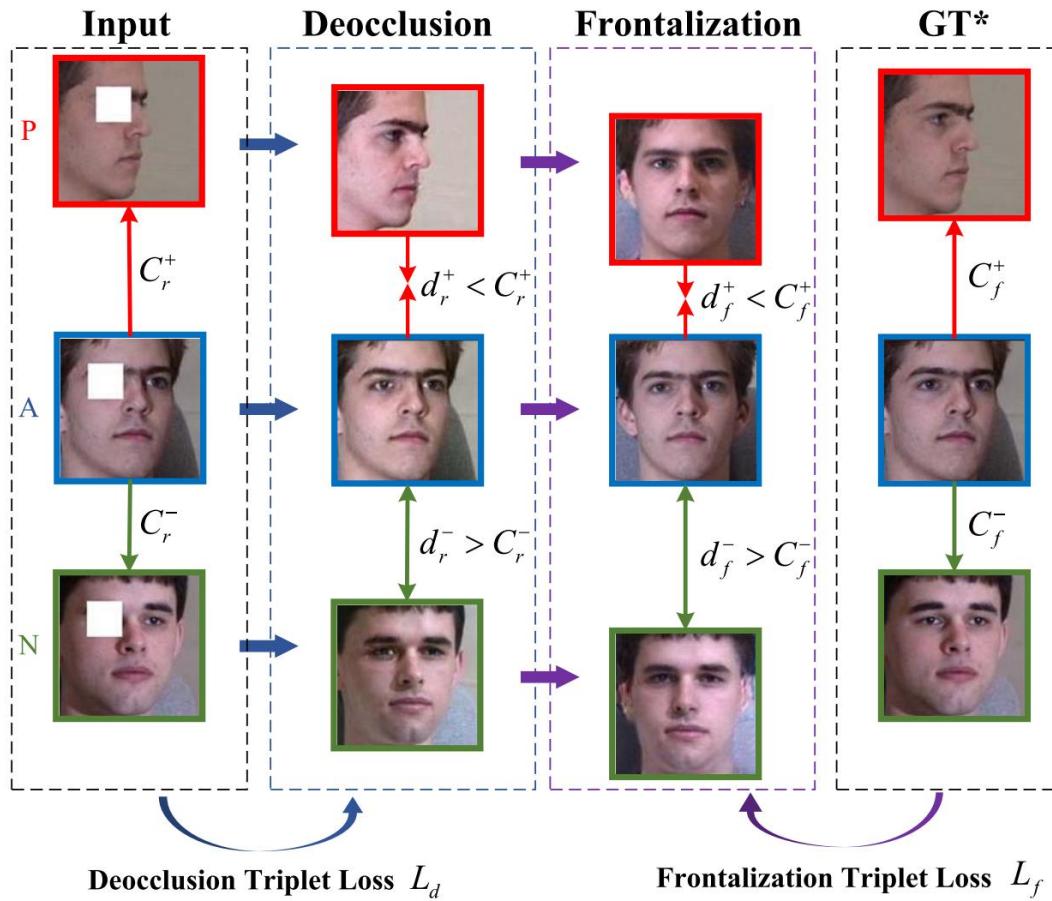




# 研究方法

## C. Dual Triplet Loss

### 双重三元损失



➤ 人脸去遮挡三元损失  $L_d$  和 人脸正向化模块 三元损失  $L_f$  定义为:

$$L_d = \text{Softplus}(d(f(I_r^a), f(I_r^+)) - C_r^+) + \text{Softplus}(C_r^- - d(f(I_r^a), f(I_r^-))), \quad (11)$$

$$L_f = \text{Softplus}(d(f(I_f^a), f(I_f^+)) - C_f^+) + \text{Softplus}(C_f^- - d(f(I_f^a), f(I_f^-))), \quad (12)$$

- $C_r^-, C_r^+$  表示为  $d(f(I^a), f(I^+))$ ,  $d(f(I^a), f(I^-))$
- $d(x, y)$  定义为 X 和 Y 的距离:  $d(x, y) = 0.5 - \mathbf{x}^T \mathbf{y} / (2\|\mathbf{x}\| \|\mathbf{y}\|)$
- Softplus 定义为单调函数:  $\text{Softplus}(\cdot) = \ln(1 + \exp(\cdot))$
- 双重三元损失  $L_{dtri}$  可由  $L_d$  和  $L_f$  构成:

$$L_{dtri} = L_d + L_f. \quad (13)$$



# 研究方法

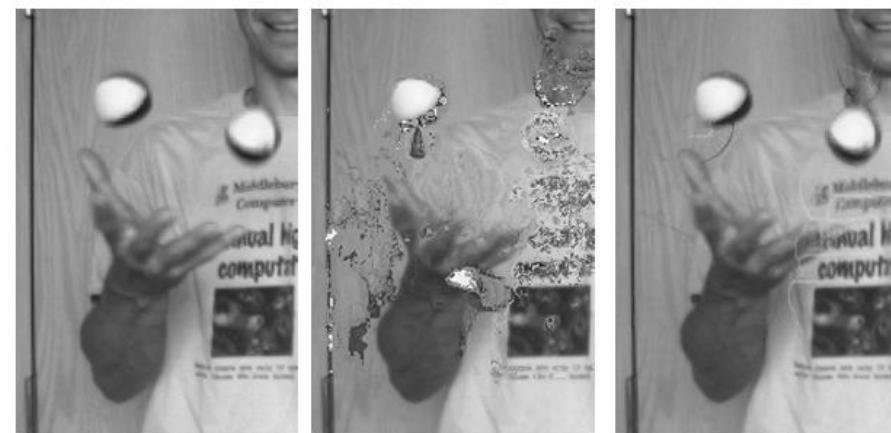
## D. Feature-Level Loss and Pixel-Level Losses

### 特征级损失



$$L_{ip} = |f(\hat{I}_p) - f(I_r)| + |f(\hat{I}_f) - f(I_f)| \\ + |f_{po}(\hat{I}_p) - f_{po}(I_r)| + |f_{po}(\hat{I}_f) - f_{po}(I_f)|, \quad (14)$$

### 像素级损失



$$L_{pix} = \frac{1}{D} \sum_{d=1}^D \frac{1}{W_d H_d C} \sum_{w,h,c=1}^{W_d, H_d, C} (|I_r^{d,w,h,c} - \hat{I}_p^{d,w,h,c}| \\ + |I_f^{d,w,h,c} - \hat{I}_f^{d,w,h,c}|), \quad (15)$$

$$L_{sym} = \frac{1}{W/2 \times H} \sum_{w=1}^{W/2} \sum_{h=1}^H |I_f^{w,h} - I_f^{W-(w-1),h}|, \quad (16)$$

$$L_{tv} = \sum_{c=1}^C \sum_{w,h=1}^{W,H} |I_r^{w+1,h,c} - I_r^{w,h,c}| + |I_r^{w,h+1,c} - I_r^{w,h,c}| \\ + |I_f^{w+1,h,c} - I_f^{w,h,c}| + |I_f^{w,h+1,c} - I_f^{w,h,c}|, \quad (17)$$



# 研究方法

## E. Overall Training Loss Function

### 总体训练损失函数

#### $D_d$ 和 $D_f$ 对抗损失

$$\begin{aligned} L_{D_d} &= \max_{D_d} V_d(G_d, D_d) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log D_d(\hat{I}_p^n) \\ &\quad + \log(1 - D_d(G_d(I^n, M^n)))\}. \end{aligned} \quad (4)$$

$$\begin{aligned} L_{D_f} &= \max_{D_f} V_f(G_f, D_f) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log D_f(\hat{I}_f^n) \\ &\quad + \log(1 - D_f(G_f(I_r^n, M^n)))\}. \end{aligned} \quad (8)$$

$$\left\{ \begin{array}{l} \max_{D_d, D_f} L_D = L_{D_d} + L_{D_f}, \\ \min_{G_d, G_f} L_G = \lambda_1(L_{adv}^{G_d} + L_{adv}^{G_f}) + \lambda_2 L_{dtri} + \lambda_3 L_{ip} \\ \quad + \lambda_4 L_{pix} + \lambda_5 L_{sym} + \lambda_6 L_{tv}, \end{array} \right.$$

#### $G_d$ 和 $G_f$ 对抗损失

$$\begin{aligned} L_{adv}^{G_d} &= \min_{G_d} V_d(G_d, D_d) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log(1 - D_d(G_d(I^n, M^n)))\}. \end{aligned} \quad (3)$$

$$\begin{aligned} L_{adv}^{G_f} &= \min_{G_f} V_f(G_f, D_f) \\ &= \frac{1}{N} \sum_{n=1}^N \{\log(1 - D_f(G_f(I_r^n, M^n)))\}. \end{aligned} \quad (7)$$

#### 双重三元损失

$$L_{dtri} = L_d + L_f$$

$$\begin{aligned} L_d &= Softplus(d(f(I_r^a), f(I_r^+)) - C_r^+) \\ &\quad + Softplus(C_r^- - d(f(I_r^a), f(I_r^-))), \end{aligned} \quad (11)$$

$$\begin{aligned} L_f &= Softplus(d(f(I_f^a), f(I_f^+)) - C_f^+) \\ &\quad + Softplus(C_f^- - d(f(I_f^a), f(I_f^-))), \end{aligned} \quad (12)$$

#### 特征级损失

$$\begin{aligned} L_{ip} &= |f(\hat{I}_p) - f(I_r)| + |f(\hat{I}_f) - f(I_f)| \\ &\quad + |f_{po}(\hat{I}_p) - f_{po}(I_r)| + |f_{po}(\hat{I}_f) - f_{po}(I_f)|, \end{aligned} \quad (14)$$

(18)

#### 像素级损失

$$\begin{aligned} L_{pix} &= \frac{1}{D} \sum_{d=1}^D \frac{1}{W_d H_d C} \sum_{w,h,c=1}^{W_d, H_d, C} (|I_r^{d,w,h,c} - \hat{I}_p^{d,w,h,c}| \\ &\quad + |I_f^{d,w,h,c} - \hat{I}_f^{d,w,h,c}|), \end{aligned} \quad (15)$$

$$L_{sym} = \frac{1}{W/2 \times H} \sum_{w=1}^{W/2} \sum_{h=1}^H |I_f^{w,h} - I_f^{W-(w-1),h}|, \quad (16)$$

$$\begin{aligned} L_{tv} &= \sum_{c=1}^C \sum_{w,h=1}^{W,H} |I_r^{w+1,h,c} - I_r^{w,h,c}| + |I_r^{w,h+1,c} - I_r^{w,h,c}| \\ &\quad + |I_f^{w+1,h,c} - I_f^{w,h,c}| + |I_f^{w,h+1,c} - I_f^{w,h,c}|, \end{aligned} \quad (17)$$

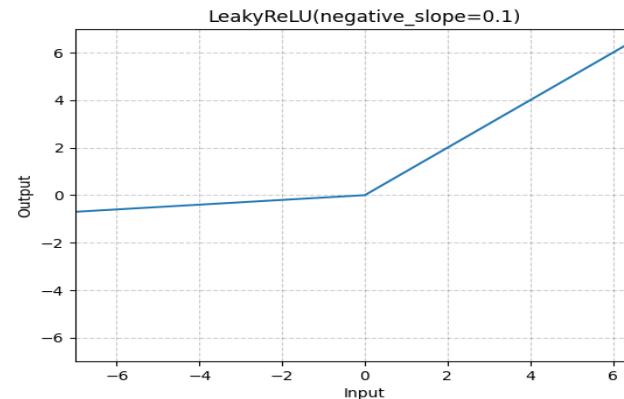
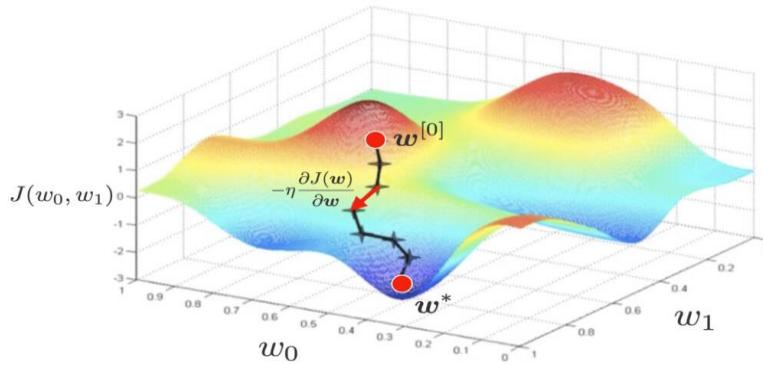
➤  $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$  和  $\lambda_6$ 是权衡参数。



# 研究方法

## F. Optimizer and Implementation

### 优化器与实现



- 使用Adam优化器来训练网络。
- 在模型中使用批归一化和Leaky ReLU来加速收敛。
- 模型在NVIDIA 1080Ti GPU上进行训练，并使用TensorFlow实现。



# 实验与分析



# 实验设置与数据集选用

## 数据集选用

- Multi-PIE: 在**受限**环境下评估姿态、光照和表情变化下人脸识别的最大数据库。是一个由 337 名被试者在不同姿态、光照条件和表情下拍摄的超过750,000张面部图像组成的数据集，有20个照明级别和13个姿势，范围从-90°到90°。
- LFW (Labeled Faces in The Wild): LFW 数据库主要是从互联网上搜集图像 (**非受限**)，它包含来自5749名受试者的13233张面部图像，包含各种姿态、表情和光照。



Multi-PIE



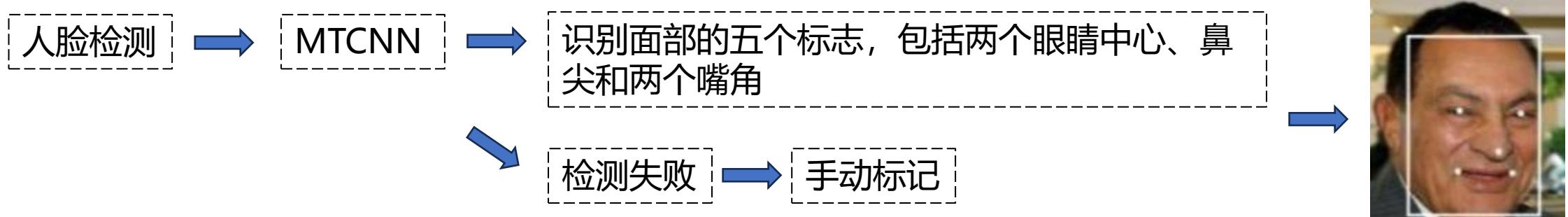
LFW



# 实验设置与数据集选用

## 数据预处理

目的：为了减少不同数据库的分布差异



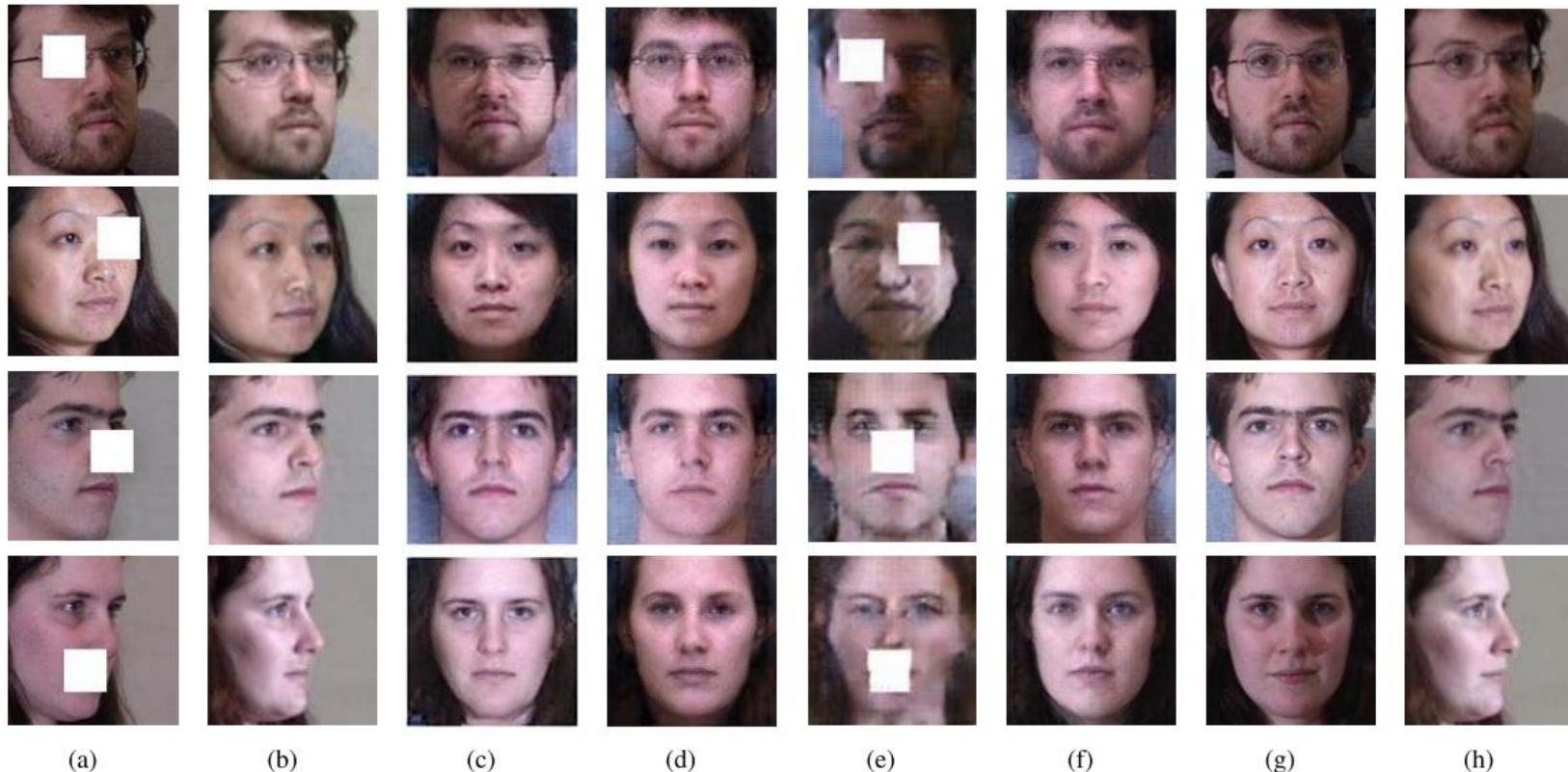
## 配置细节

- 在所有实验中，设置 $\lambda_1 = 2e1$ ,  $\lambda_2 = 4$ ,  $\lambda_3 = 40$ ,  $\lambda_4 = 1$ ,  $\lambda_5 = 3e-1$ ,  $\lambda_6 = 1e-3$ 。
- 学习率为0.0001，在整个训练阶段都是固定的。
- 批处理大小为6，训练在4 ~ 5个epoch后停止。整个模型训练过程大约需要6 ~ 8天。



# 实验——规则遮挡下Multi-PIE人脸合成与识别

## 关键点遮挡



定性分析：

- ① 与 BoostGAN 、 DR-GAN、 TP-GAN相比， TSGAN的合成结果效果更好。
- ② 与原图对比，可以看到，得益于掩码注意模块， TSGAN的去遮挡和正面化的性能得到了提高。

Fig. 4. Synthesis results on **keypoint region** occluded Multi-PIE dataset. From top to bottom, the poses are  $15^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ . The ground truth frontal images are provided at the last column. (a) Profile. (b) Ours\* (synthetic de-occluded profile facial image). (c) Ours. (d) [31]\*. (e) [20]. (f) [14]\*. (g) GT. (h) GT\*.



# 实验——规则遮挡下Multi-PIE人脸合成与识别

## 关键点遮挡

TABLE I

RANK-1 RECOGNITION RATE (%) COMPARISON ON *Keypoint Region*  
OCCLUDED MULTI-PIE. **BLACK**: RANKS THE 1<sup>st</sup>; **RED**: RANKS THE  
2<sup>nd</sup>; **BLUE**: RANKS THE 3<sup>rd</sup>

Method	±15°	±30°	±45°	±60°	±75°
DR-GAN [20] (k1)	67.38	60.68	55.83	47.25	39.34
DR-GAN [20] (k2)	73.24	65.37	59.90	51.18	<b>42.24</b>
DR-GAN [20] (k3)	66.93	60.60	56.54	49.70	39.77
DR-GAN [20] (k4)	71.33	63.72	57.59	50.10	<b>40.87</b>
DR-GAN [20] (mean)	69.72	62.59	57.47	49.56	40.55
TP-GAN [14]* (k1)	98.17	95.46	86.60	65.91	39.51
TP-GAN [14]* (k2)	<b>99.27</b>	<b>97.25</b>	88.37	66.03	40.82
TP-GAN [14]* (k3)	95.04	90.95	82.72	62.40	38.67
TP-GAN [14]* (k4)	97.80	93.66	83.84	62.27	36.76
TP-GAN [14]* (mean)	97.57	94.33	85.38	64.15	38.94
BoostGAN* [31](k1)	99.03	96.21	86.66	64.45	39.74
BoostGAN* [31](k2)	<b>99.12</b>	<b>96.33</b>	<b>85.41</b>	<b>63.49</b>	<b>39.63</b>
BoostGAN* [31](k3)	96.03	92.06	83.10	63.45	39.92
BoostGAN* [31](k4)	98.13	94.88	84.24	62.17	39.45
BoostGAN*(mean)	98.08	94.87	84.85	63.39	39.69
Ours(k1)	<b>99.26</b>	<b>97.68</b>	<b>91.35</b>	<b>70.73</b>	<b>42.01</b>
Ours(k2)	<b>99.23</b>	<b>97.92</b>	<b>91.04</b>	68.77	38.48
Ours(k3)	96.15	94.35	89.27	<b>69.25</b>	40.73
Ours(k4)	97.67	95.60	89.45	68.46	40.53
Ours(mean)	98.08	96.39	<b>90.28</b>	<b>69.30</b>	40.44

遮挡眼部

TSGAN

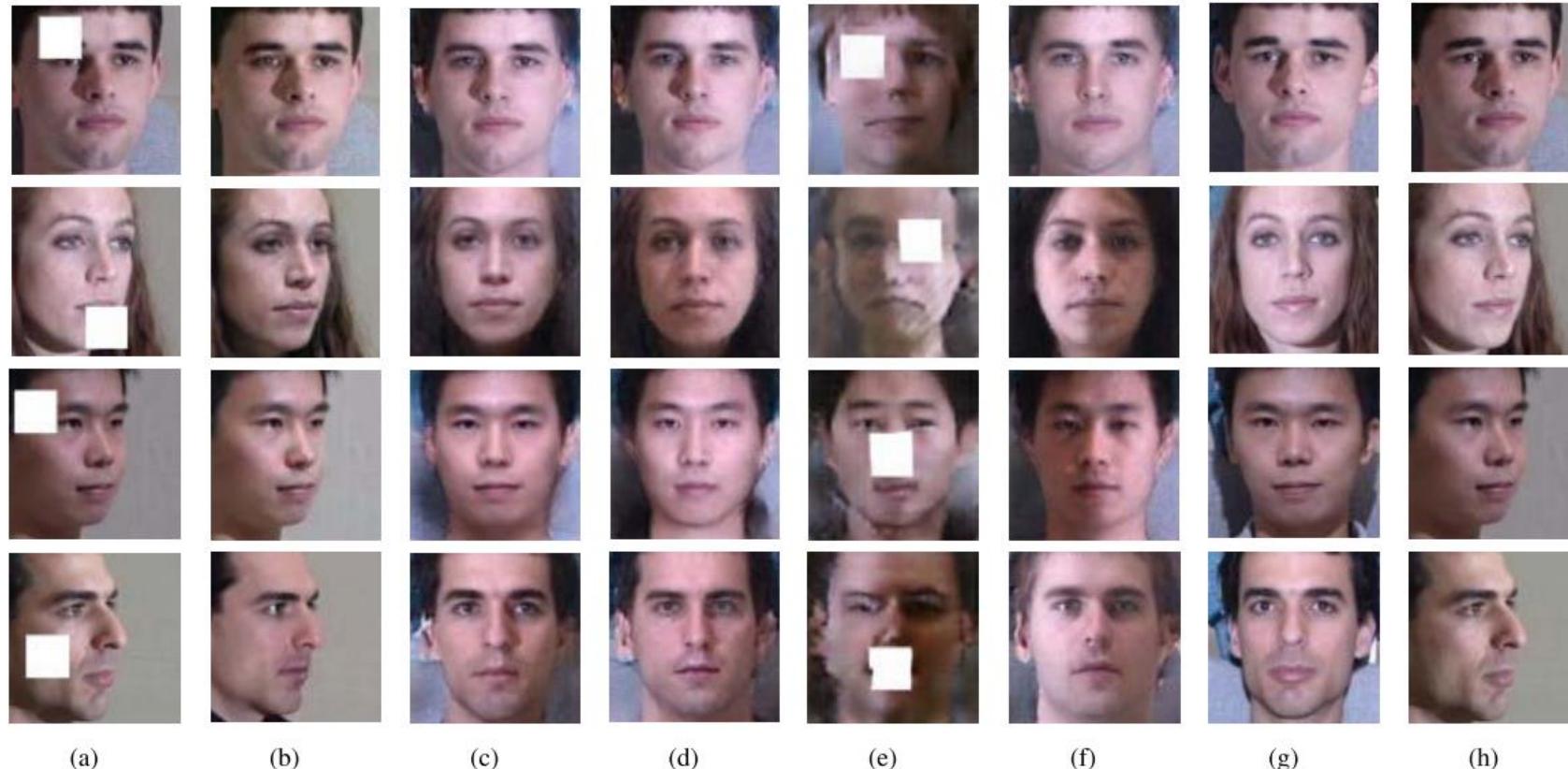
定量分析：

- ① TSGAN性能最优。
- ② 除DR-GAN外，每种方法在左眼遮挡和右眼遮挡上的准确率均高于其他区域遮挡的准确率。



# 实验——规则遮挡下Multi-PIE人脸合成与识别

## 随机块遮挡



### 定性分析：

- ① 测试模型仅基于关键点位置遮挡的Multi-PIE数据集进行训练。
- ② TSGAN的合成效果更好。

Fig. 5. Synthesis results on **random block** occluded Multi-PIE. From top to bottom, the poses are  $15^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ . The ground truth (GT) frontal images are provided at the last column. Note that all the models are trained solely on keypoint occluded Multi-PIE dataset. (a) Profile. (b) Ours\*. (c) Ours. (d) [31]\*. (e) [20]. (f) [14]\*. (g) GT. (h) GT\*.



# 实验——规则遮挡下Multi-PIE人脸合成与识别

## 随机块遮挡

TABLE II

RANK-1 RECOGNITION RATE (%) COMPARISON ON *Random Block Occluded Multi-PIE*. **BLACK**: RANKS THE 1<sup>st</sup>; **RED**: RANKS THE 2<sup>nd</sup>; **BLUE**: RANKS THE 3<sup>rd</sup>

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$
DR-GAN [20] (r1)	47.64	38.93	33.21	25.38	18.92
DR-GAN [20] (r2)	65.75	55.15	46.52	38.33	29.00
DR-GAN [20] (r3)	56.01	46.27	39.13	29.11	23.01
DR-GAN [20] (r4)	59.10	47.92	39.97	33.69	25.20
DR-GAN [20] (mean)	57.13	47.07	39.71	31.63	24.03
TP-GAN [14]* (r1)	89.81	83.88	74.94	54.83	31.34
TP-GAN [14]* (r2)	77.98	71.68	60.52	42.68	23.92
TP-GAN [14]* (r3)	79.12	72.45	60.00	41.37	24.11
TP-GAN [14]* (r4)	86.13	77.76	64.84	45.08	25.15
TP-GAN [14]* (mean)	83.26	76.44	65.08	45.99	26.13
BoostGAN* [31](r1)	98.16	95.07	86.67	66.47	43.06
BoostGAN* [31](r2)	98.10	94.97	86.51	66.56	42.96
BoostGAN* [31](r3)	98.02	94.99	86.78	66.25	42.69
BoostGAN* [31](r4)	98.12	95.12	86.69	65.78	42.67
BoostGAN* [31](mean)	98.10	95.04	86.66	66.27	42.85
Ours(r1)	<b>98.70</b>	<b>97.67</b>	<b>93.78</b>	<b>79.88</b>	<b>44.43</b>
Ours(r2)	<b>98.47</b>	97.01	<b>93.20</b>	<b>80.52</b>	<b>45.22</b>
Ours(r3)	98.33	<b>97.17</b>	<b>93.40</b>	<b>78.80</b>	42.23
Ours(r4)	<b>98.39</b>	96.74	92.21	75.76	<b>44.29</b>
Ours(mean)	<b>98.47</b>	<b>97.15</b>	93.15	78.74	44.04

TABLE I

RANK-1 RECOGNITION RATE (%) COMPARISON ON *Keypoint Region Occluded Multi-PIE*. **BLACK**: RANKS THE 1<sup>st</sup>; **RED**: RANKS THE 2<sup>nd</sup>; **BLUE**: RANKS THE 3<sup>rd</sup>

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$
DR-GAN [20] (k1)	67.38	60.68	55.83	47.25	39.34
DR-GAN [20] (k2)	73.24	65.37	59.90	51.18	<b>42.24</b>
DR-GAN [20] (k3)	66.93	60.60	56.54	49.70	39.77
DR-GAN [20] (k4)	71.33	63.72	57.59	50.10	<b>40.87</b>
DR-GAN [20] (mean)	69.72	62.59	57.47	49.56	40.55
TP-GAN [14]* (k1)	98.17	95.46	86.60	65.91	39.51
TP-GAN [14]* (k2)	<b>99.27</b>	<b>97.25</b>	88.37	66.03	40.82
TP-GAN [14]* (k3)	95.04	90.95	82.72	62.40	38.67
TP-GAN [14]* (k4)	97.80	93.66	83.84	62.27	36.76
TP-GAN [14]* (mean)	97.57	94.33	85.38	64.15	38.94
BoostGAN* [31](k1)	99.03	96.21	86.66	64.45	39.74
BoostGAN* [31](k2)	99.12	96.33	85.41	63.49	39.63
BoostGAN* [31](k3)	96.03	92.06	83.10	63.45	39.92
BoostGAN* [31](k4)	98.13	94.88	84.24	62.17	39.45
BoostGAN*(mean)	98.08	94.87	84.85	63.39	39.69
Ours(k1)	<b>99.26</b>	<b>97.68</b>	<b>91.35</b>	<b>70.73</b>	<b>42.01</b>
Ours(k2)	<b>99.23</b>	<b>97.92</b>	<b>91.04</b>	68.77	38.48
Ours(k3)	96.15	94.35	89.27	<b>69.25</b>	40.73
Ours(k4)	97.67	95.60	89.45	68.46	40.53
Ours(mean)	98.08	96.39	<b>90.28</b>	<b>69.30</b>	40.44

定量分析：

- ① TSGAN 性能最优。
- ② TSGAN 和 BoostGAN 具有更好的鲁棒性。



# 实验二—规则遮挡下LFW的面部正面化和验证

## 关键点遮挡



(a) (b) (c) (d) (e) (f) (g)

Fig. 6. Synthesis results on **keypoint region** occluded LFW dataset in the wild. Note that there are no ground truth frontal images for this dataset. The models are solely trained based on keypoint occluded Multi-PIE dataset.  
(a) Profile. (b) Ours\*. (c) Ours. (d) [31]\*. (e) [20]. (f) [14]\*. (g) GT\*.

定性分析：

- ① 测试模型仅基于关键点位置遮挡的Multi-PIE数据集进行训练。
- ② TSGAN的合成效果更好，泛化能力良好。



# 实验二—规则遮挡下LFW的面部正面化和验证

## 关键点遮挡

TABLE III

FACE VERIFICATION ACCURACY (ACC) AND AREA-UNDER-CURVE (AUC) RESULTS ON *Keypoint Region Occluded LFW*

Method	ACC(%)	AUC(%)
DR-GAN [20] (k1)	67.60	73.65
DR-GAN [20] (k2)	67.28	72.94
DR-GAN [20] (k3)	58.43	59.19
DR-GAN [20] (k4)	69.50	76.05
DR-GAN [20] (mean)	65.71	70.46
TP-GAN [14]* (k1)	86.52	92.81
TP-GAN [14]* (k2)	87.83	93.96
TP-GAN [14]* (k3)	85.17	91.63
TP-GAN [14]* (k4)	87.78	93.97
TP-GAN [14]* (mean)	86.83	93.09
BoostGAN* [31](k1)	88.47	94.07
BoostGAN* [31](k2)	88.13	93.85
BoostGAN* [31](k3)	86.62	93.66
BoostGAN* [31](k4)	87.88	93.92
BoostGAN* [31](mean)	87.78	93.88
Ours(k1)	<b>93.63</b>	<b>97.65</b>
Ours(k2)	<b>93.43</b>	<b>97.64</b>
Ours(k3)	90.32	95.99
Ours(k4)	90.05	95.84
Ours(mean)	<b>91.86</b>	<b>96.78</b>

定量分析:

- ① 使用识别准确率 (recognition accuracy, ACC) 和曲线下面积 (area under curve, AUC) 来评估不同方法的识别能力。
- ② TSGAN 性能最优。



TSGAN



# 实验二—规则遮挡下LFW的面部正面化和验证

## 随机块遮挡

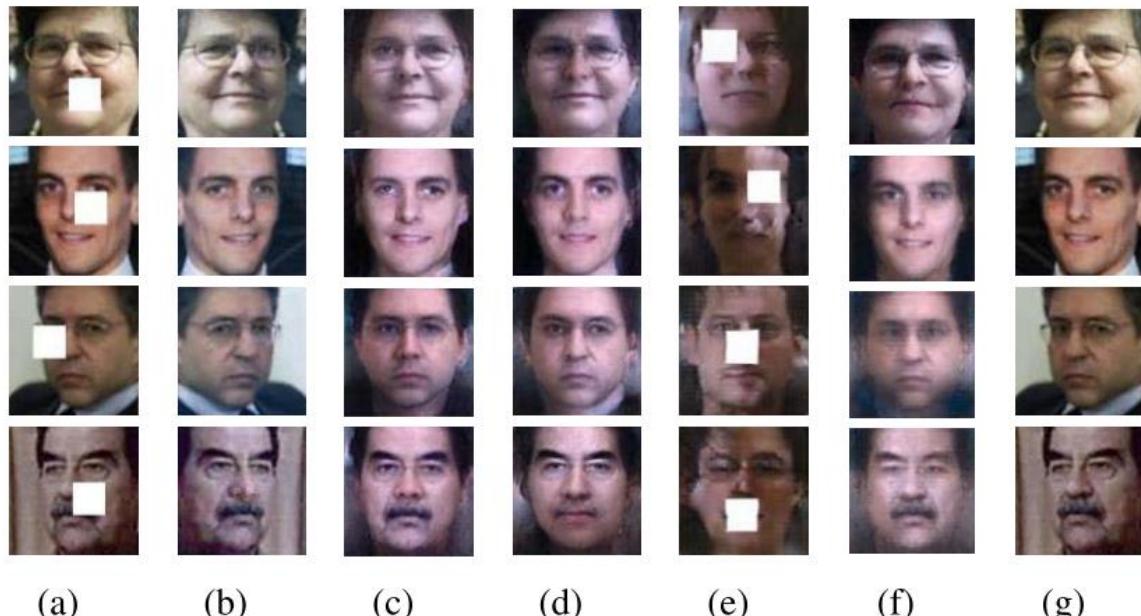


Fig. 7. Synthesis results on **random block** occluded LFW dataset. Note that all the models are trained solely on keypoint occluded Multi-PIE dataset without retraining on randomly blocked datasets. (a) Profile. (b) Ours<sup>\*</sup> (c) Ours. (d) [31]<sup>\*</sup>. (e) [20]. (f) [14]<sup>\*</sup>. (g) GT<sup>\*</sup>.

TABLE IV

FACE VERIFICATION ACCURACY (ACC) AND AREA-UNDER-CURVE (AUC) RESULTS ON *Random Block Occluded LFW*

Method	ACC(%)	AUC(%)
DR-GAN [20] (r1)	63.28	67.20
DR-GAN [20] (r2)	65.53	71.79
DR-GAN [20] (r3)	57.15	57.76
DR-GAN [20] (r4)	64.82	70.35
DR-GAN [20] (mean)	62.70	66.78
TP-GAN [14] <sup>*</sup> (r1)	82.75	89.86
TP-GAN [14] <sup>*</sup> (r2)	77.65	84.63
TP-GAN [14] <sup>*</sup> (r3)	81.07	88.24
TP-GAN [14] <sup>*</sup> (r4)	83.25	90.15
TP-GAN [14] <sup>*</sup> (mean)	81.18	88.22
BoostGAN <sup>*</sup> [31](r1)	87.92	93.46
BoostGAN <sup>*</sup> [31](r2)	86.78	92.81
BoostGAN <sup>*</sup> [31](r3)	87.28	93.15
BoostGAN <sup>*</sup> [31](r4)	87.18	93.26
BoostGAN <sup>*</sup> [31](mean)	87.29	93.17
Ours(r1)	<b>92.65</b>	<b>97.34</b>
Ours(r2)	<b>91.55</b>	<b>96.70</b>
Ours(r3)	88.23	94.76
Ours(r4)	87.38	93.68
Ours(mean)	<b>89.72</b>	<b>95.43</b>

- ① 测试模型仅基于关键点位置遮挡的Multi-PIE数据集进行训练。
- ② TSGAN的合成效果更好，泛化能力良好。



# 实验三—不规则遮挡下Multi-PIE的人脸正面化与识别

不规则遮挡 →

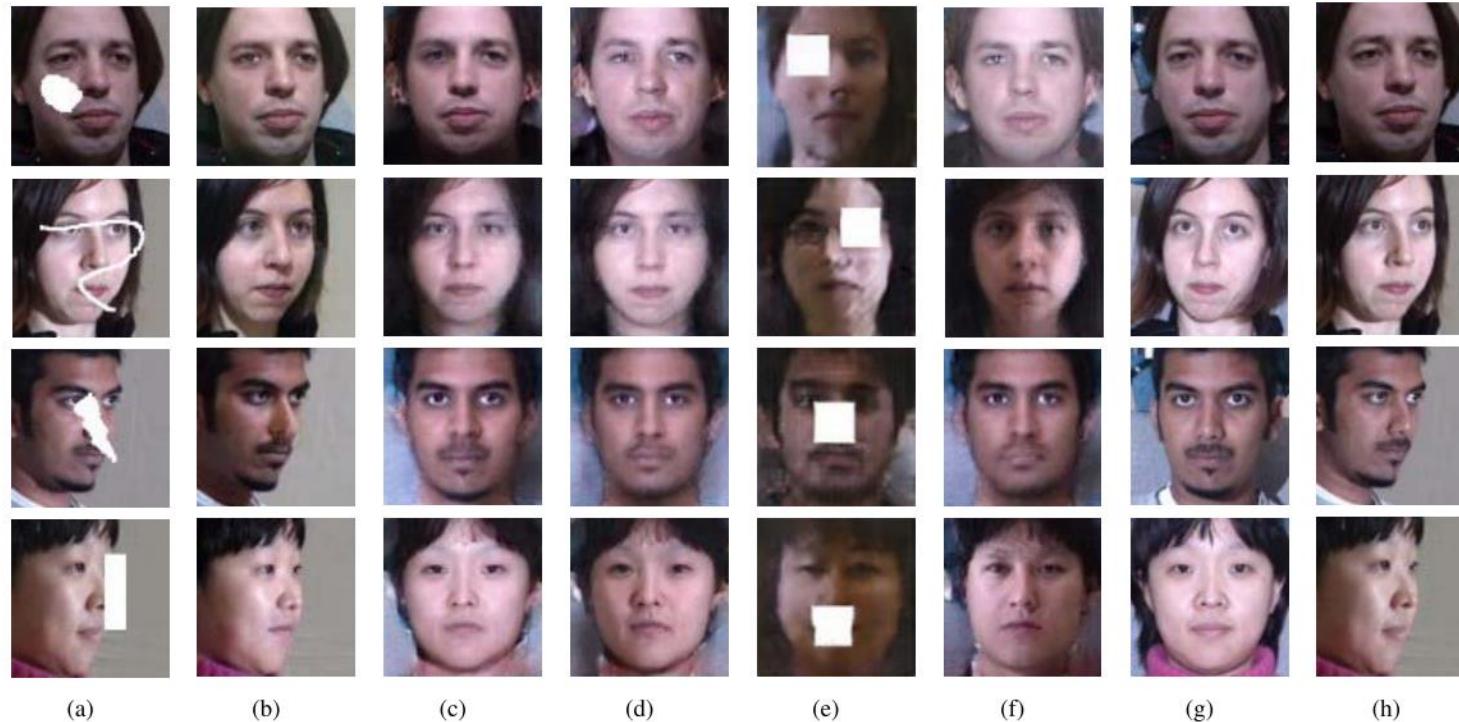
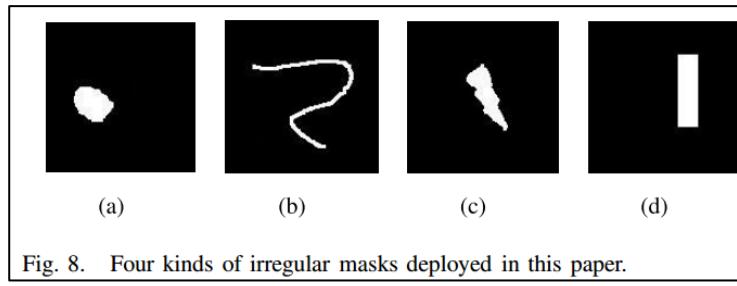


Fig. 9. Examples on **irregular block** occluded Multi-PIE. The first column is irregular mask. From top to bottom, the poses are  $15^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ . Note that the all the models are trained solely on keypoint occluded Multi-PIE dataset. (a) Profile. (b) Ours\*. (c) Ours. (d) [31]\*. (e) [20]. (f) [14]\*. (g) GT. (h) GT\*.

- ① TSGAN在全局结构和局部纹理方面的合成性能优于其他方法。
- ② TSGAN能够有效地保持身份信息。

TABLE V  
RANK-1 RECOGNITION RATE (%) COMPARISON ON *Irregular Block Occluded Multi-PIE*. **BLACK**: RANKS THE 1<sup>st</sup>; **RED**: RANKS THE 2<sup>nd</sup>; **BLUE**: RANKS THE 3<sup>rd</sup>

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	$\pm 75^\circ$
DR-GAN (m1)	60.51	52.68	44.88	35.69	23.59
DR-GAN (m2)	73.40	63.59	50.74	40.10	29.13
DR-GAN (m3)	73.38	65.44	55.63	44.16	33.56
DR-GAN (m4)	63.70	43.62	35.74	32.17	27.34
DR-GAN (mean)	67.75	56.33	46.75	38.03	28.41
TP-GAN* (m1)	94.11	89.98	81.32	57.23	28.93
TP-GAN* (m2)	61.60	48.65	39.19	24.47	13.98
TP-GAN* (m3)	86.21	79.40	67.69	47.69	24.34
TP-GAN* (m4)	85.84	76.71	64.97	44.35	26.18
TP-GAN* (mean)	81.94	73.69	63.29	43.44	23.36
BoostGAN* (m1)	<b>99.12</b>	<b>96.84</b>	<b>90.20</b>	<b>70.36</b>	<b>44.29</b>
BoostGAN* (m2)	93.27	86.13	70.92	46.08	27.65
BoostGAN* (m3)	93.14	88.17	76.53	56.67	34.52
BoostGAN* (m4)	<b>98.56</b>	95.21	85.99	65.59	42.11
BoostGAN* (mean)	96.02	91.59	80.91	59.68	37.14
Ours (m1)	<b>98.73</b>	<b>97.17</b>	<b>93.94</b>	<b>80.48</b>	<b>50.45</b>
Ours (m2)	93.55	88.90	80.02	54.94	27.20
Ours (m3)	97.04	95.12	87.85	67.99	40.47
Ours (m4)	98.14	<b>96.90</b>	<b>92.73</b>	<b>74.85</b>	<b>45.11</b>
Ours (mean)	96.87	94.52	88.64	69.57	40.81



# 实验四—不规则遮挡下LFW的人脸正面化与识别

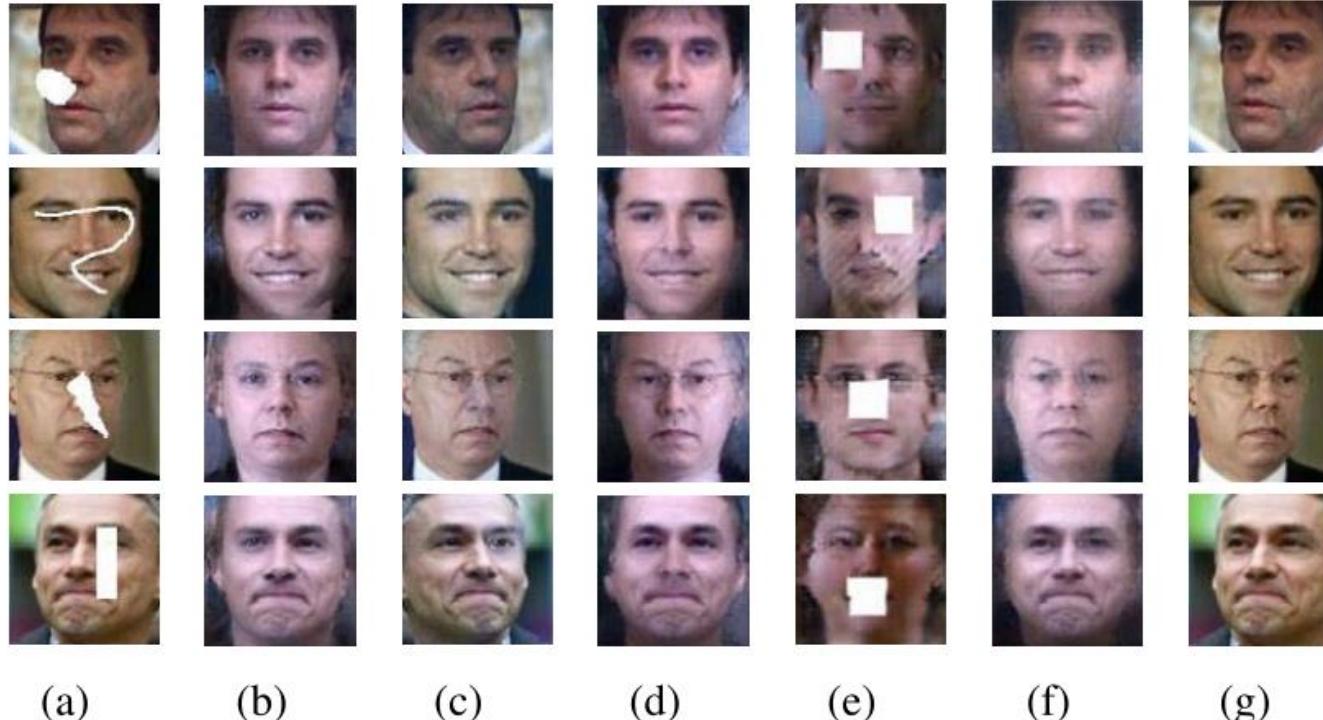


Fig. 10. Synthesis results on **irregular block** occluded LFW dataset. Note that all the models are trained solely on keypoint occluded Multi-PIE dataset, without retraining on randomly blocked datasets. (a) Profile. (b) Ours\*. (c) Ours. (d) [31]\*. (e) [20]. (f) [14]\*. (g) GT\*.

TABLE VI  
FACE VERIFICATION ACCURACY (ACC) AND AREA-UNDER-CURVE (AUC) RESULTS ON *Irregular Block Occluded LFW*

Method	ACC(%)	AUC(%)
DR-GAN [20] (m1)	66.27	71.85
DR-GAN [20] (m2)	66.93	73.11
DR-GAN [20] (m3)	58.22	59.20
DR-GAN [20] (m4)	66.82	72.58
DR-GAN [20] (mean)	64.56	69.19
TP-GAN [14]* (m1)	84.17	91.22
TP-GAN [14]* (m2)	77.12	83.88
TP-GAN [14]* (m3)	82.80	89.87
TP-GAN [14]* (m4)	84.32	91.03
TP-GAN [14]* (mean)	82.10	89.00
BoostGAN* [31](m1)	88.45	94.09
BoostGAN* [31](m2)	84.98	91.59
BoostGAN* [31](m3)	86.23	92.83
BoostGAN* [31](m4)	86.98	93.06
BoostGAN* [31](mean)	86.66	92.89
Ours(m1)	<b>91.27</b>	<b>96.07</b>
Ours(m2)	<b>90.62</b>	<b>95.94</b>
Ours(m3)	86.68	93.19
Ours(m4)	87.08	93.82
Ours(mean)	<b>88.92</b>	<b>94.76</b>

TSGAN展现了在不规则和随机遮挡的非控制数据上的优越性能。



# 分析与讨论—消融分析

## 采用不同的架构和损失函数进行实验

TABLE VII

ABLATION STUDY: RANK-1 RECOGNITION RATE (%) COMPARISON ON  
*Keypoint Region Occluded Multi-PIE*

Method	$\pm 15^\circ$	$\pm 30^\circ$	$\pm 45^\circ$	$\pm 60^\circ$	mean
w/o $L_{ip}$	93.68	90.56	83.27	59.98	30.21
w/o $L_{adv}$	98.72	96.10	86.40	66.01	36.57
w/o MAM	98.24	96.09	89.25	65.74	35.17
w/o DTL	98.82	96.79	87.83	63.71	35.88
w/o $L_{tv}$	98.26	95.93	89.24	68.84	39.23
TSGAN	<b>99.26</b>	<b>97.68</b>	<b>91.35</b>	<b>70.73</b>	<b>42.01</b>

- ① 身份保持损失( $L_{ip}$ )、掩模-注意模块(MAM)、双三元损失(DTL)对识别有很大贡献，特别是对大姿态的识别。
- ② 对抗损失(即 $L_{adv}$ )和总变分正则化(即 $L_{tv}$ )也提高了精度。
- ③ 这些结果证明了TSGAN的有效性。



# 总结



# 总结

## ● 01

- ① 提出了一种同时进行人脸补全和正面化的方法：TSGAN。主要完成：面部去遮挡和面部正脸化，以达到不变形态的面部识别。

## ● 02

- ① 一个设计精良的掩模注意力模块：应用在两个阶段的生成器中，通过U型连接策略采用新型的掩模注意力机制，以改善面部补全的性能。
- ② 提出了一个新的双三元组损失约束，通过最大化类间差异而最小化类内差异，以更有效地保留身份信息，同时监督每个子模块的训练。

## ● 03

- ① 提出的TSGAN在有约束和无约束环境下的定性和定量实验说明了其出色的综合性能，解决了一种部分遮挡但侧面可见的人脸识别问题。
- ② 但是，跨年龄的人脸识别也是一个具有挑战性的问题。如何利用GAN来解决被遮挡、侧面和跨年龄的人脸识别也是一个值得思考的问题。



谢谢！