Pose-Guided Photorealistic Face Rotation

Yibo Hu[1,2], Xiang Wu[1], Bing Yu[3], Ran He[1,2], Zhenan Sun[1,2]

[1]CRIPAC & NLPR & CEBSIT, CASIA

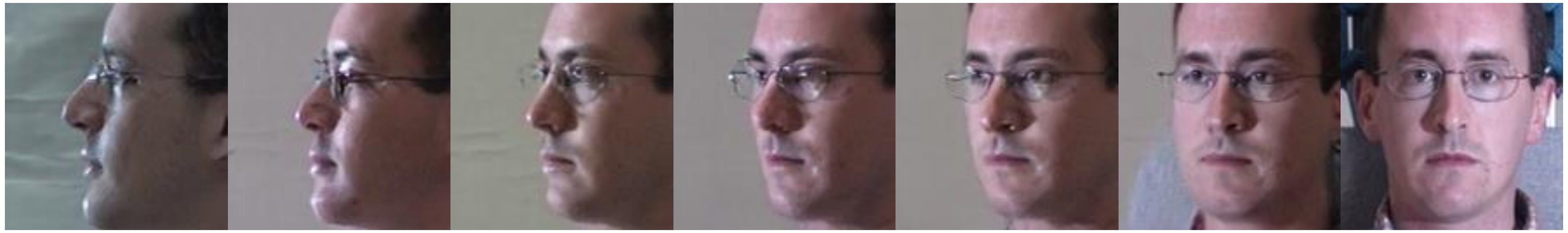
[2]University of Chinese Academy of Sciences

[3]Noah’s Ark Laboratory, Huawei Technologies Co., Ltd.

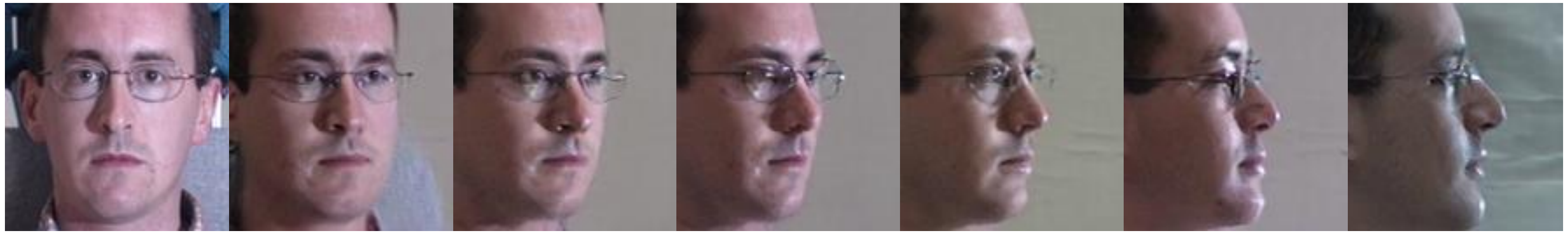
Background

Goal: Rotating a normalized face to arbitrary poses, where only yaw is considered.

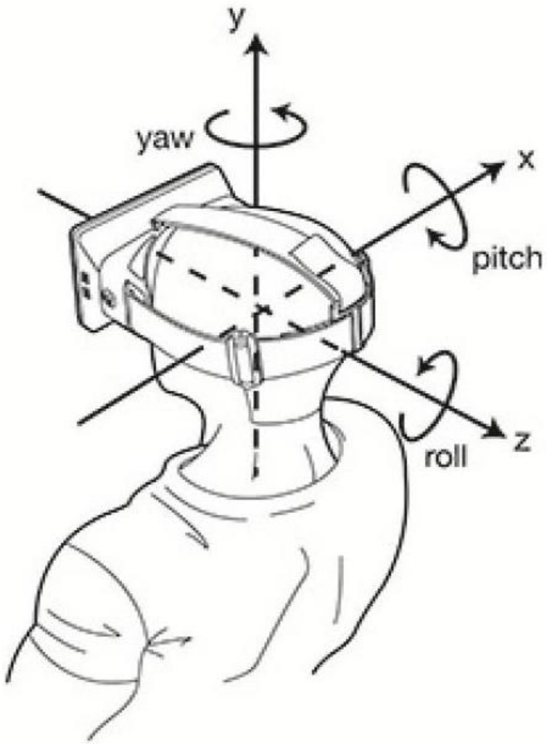
Application: Face rotation provides a cheap but effective way for data augmentation and representation learning of face recognition.



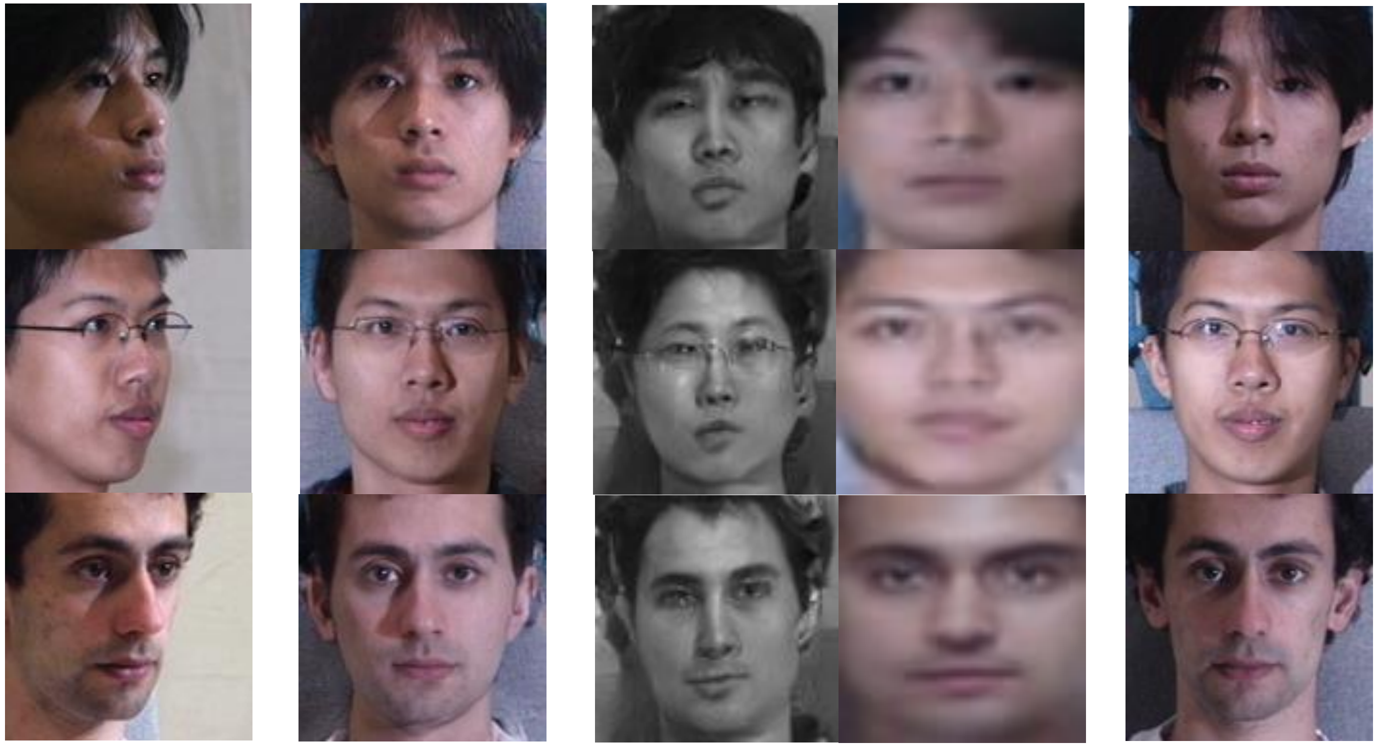
Profile → Frontal



Frontal → Profile



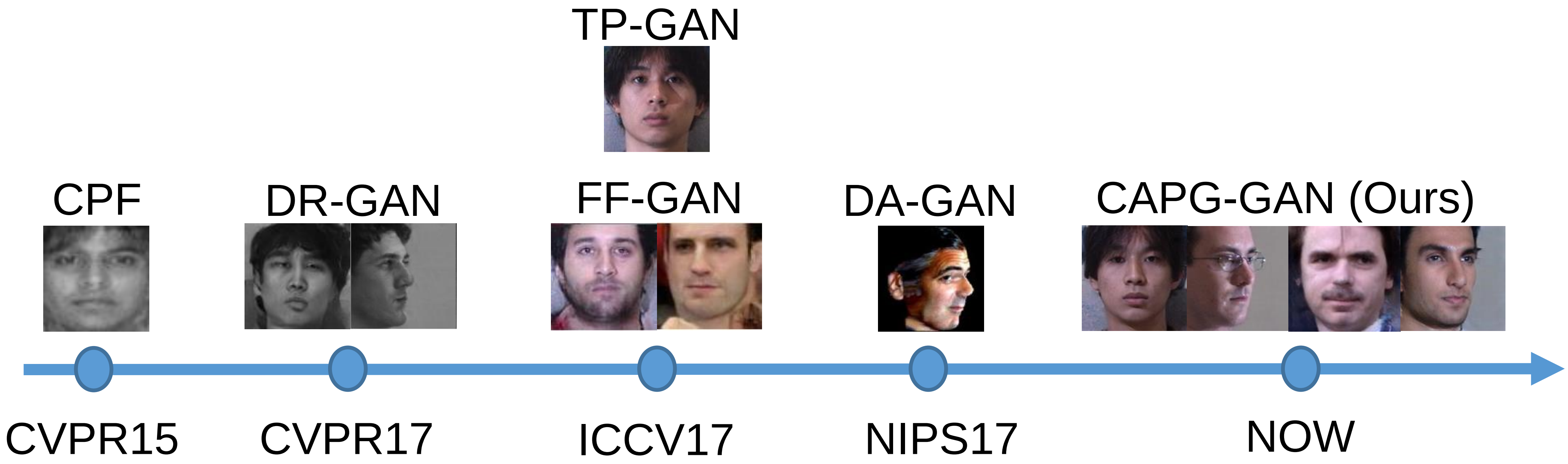
Background



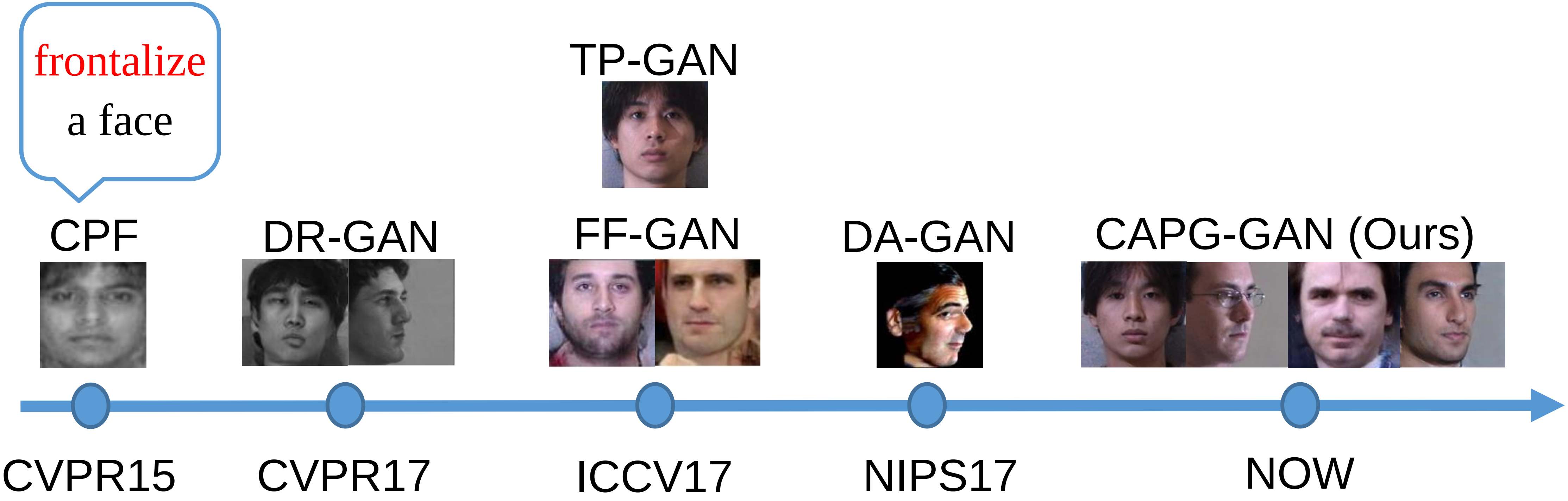
Input Ours Others GT

- Photo-realistic
- High-resolution
- Identity preserving
- Ill-posed problem

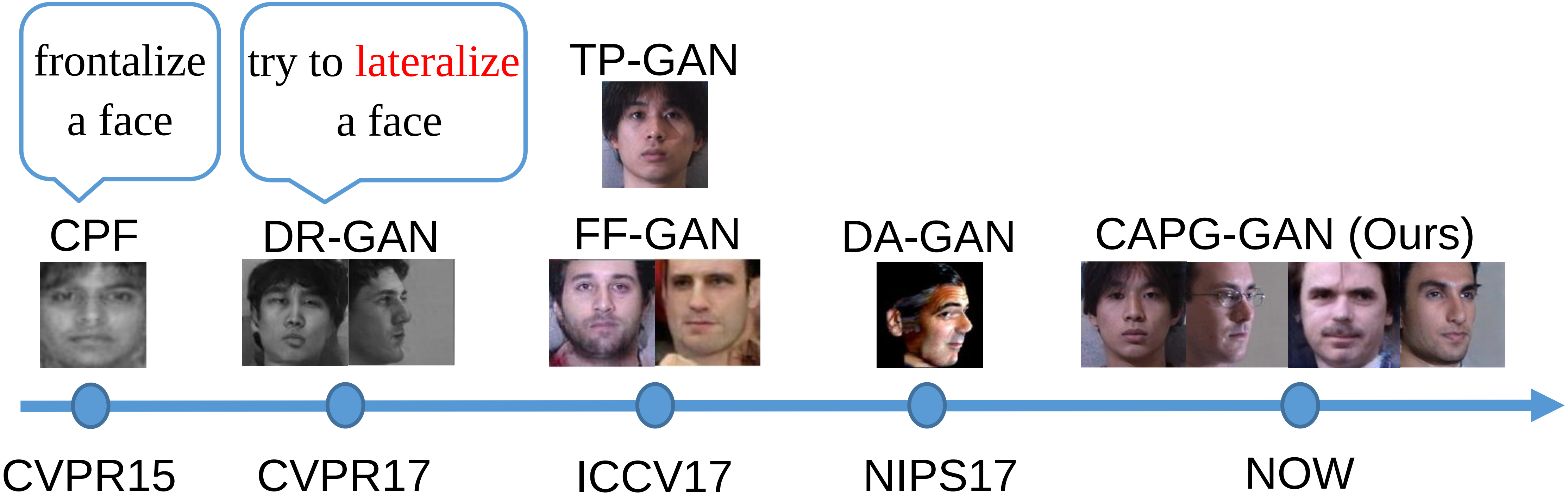
Related Work



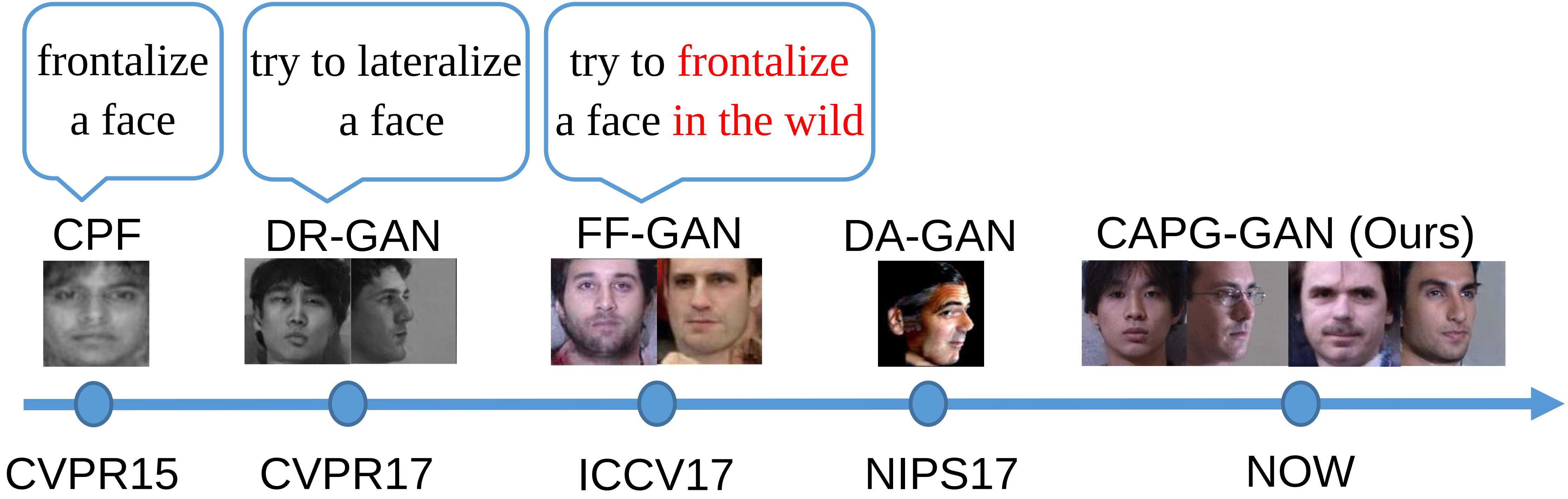
Related Work



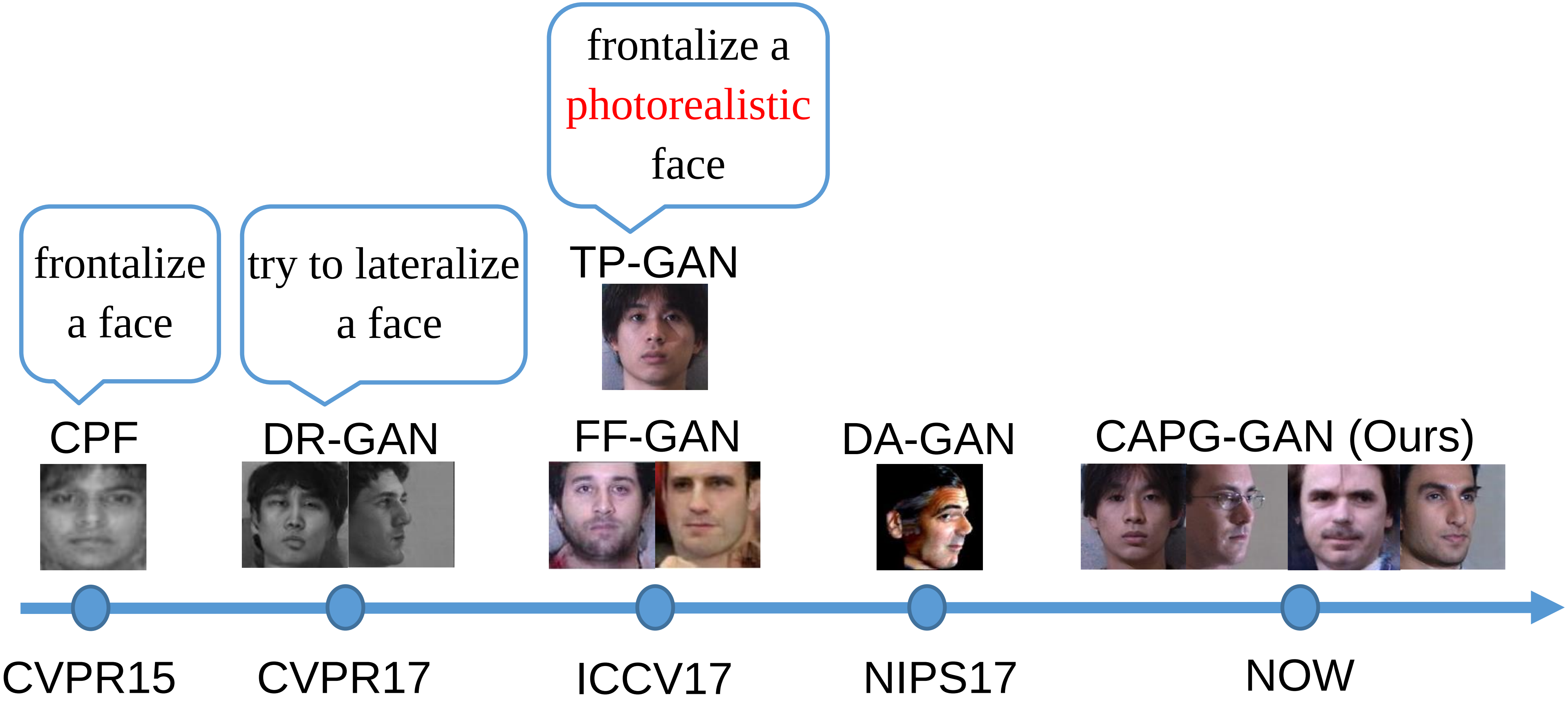
Related Work



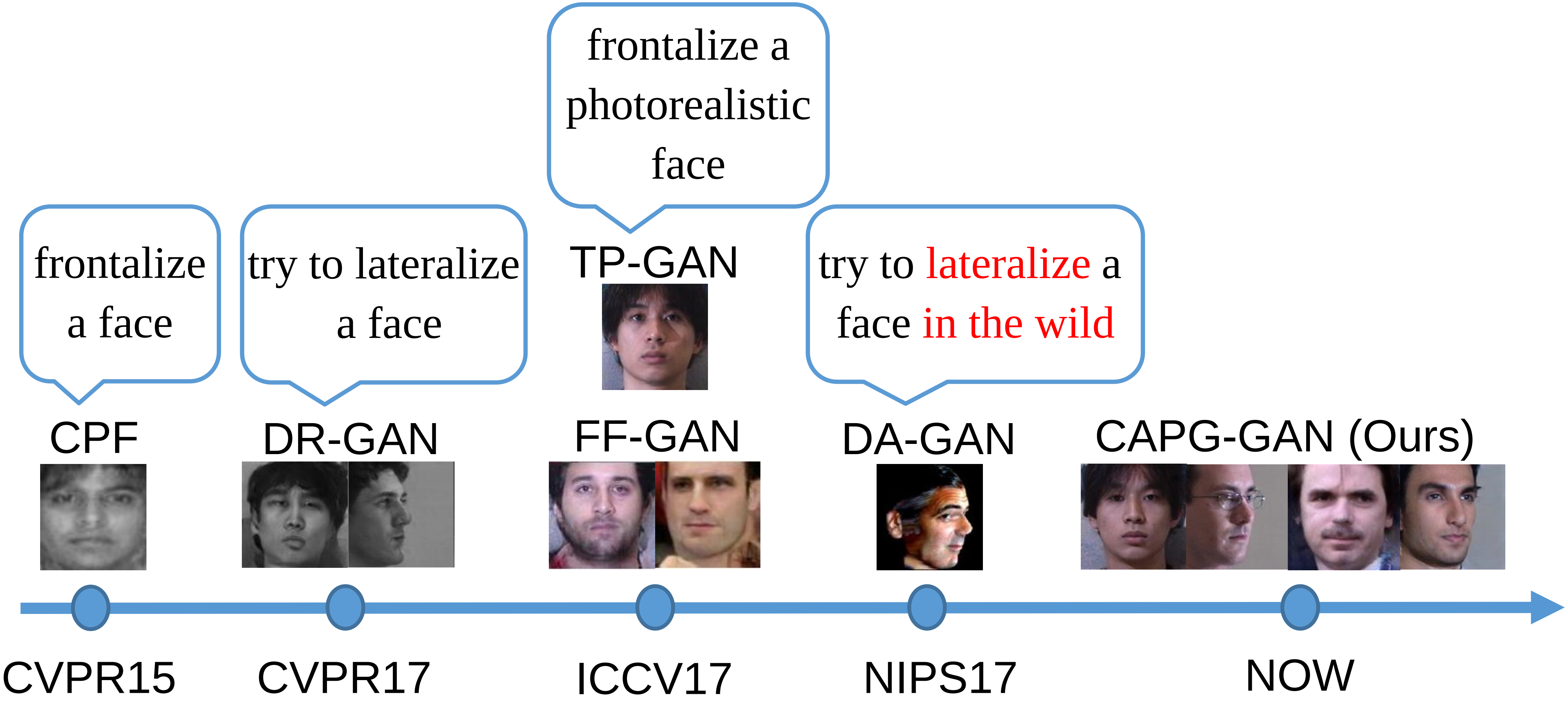
Related Work



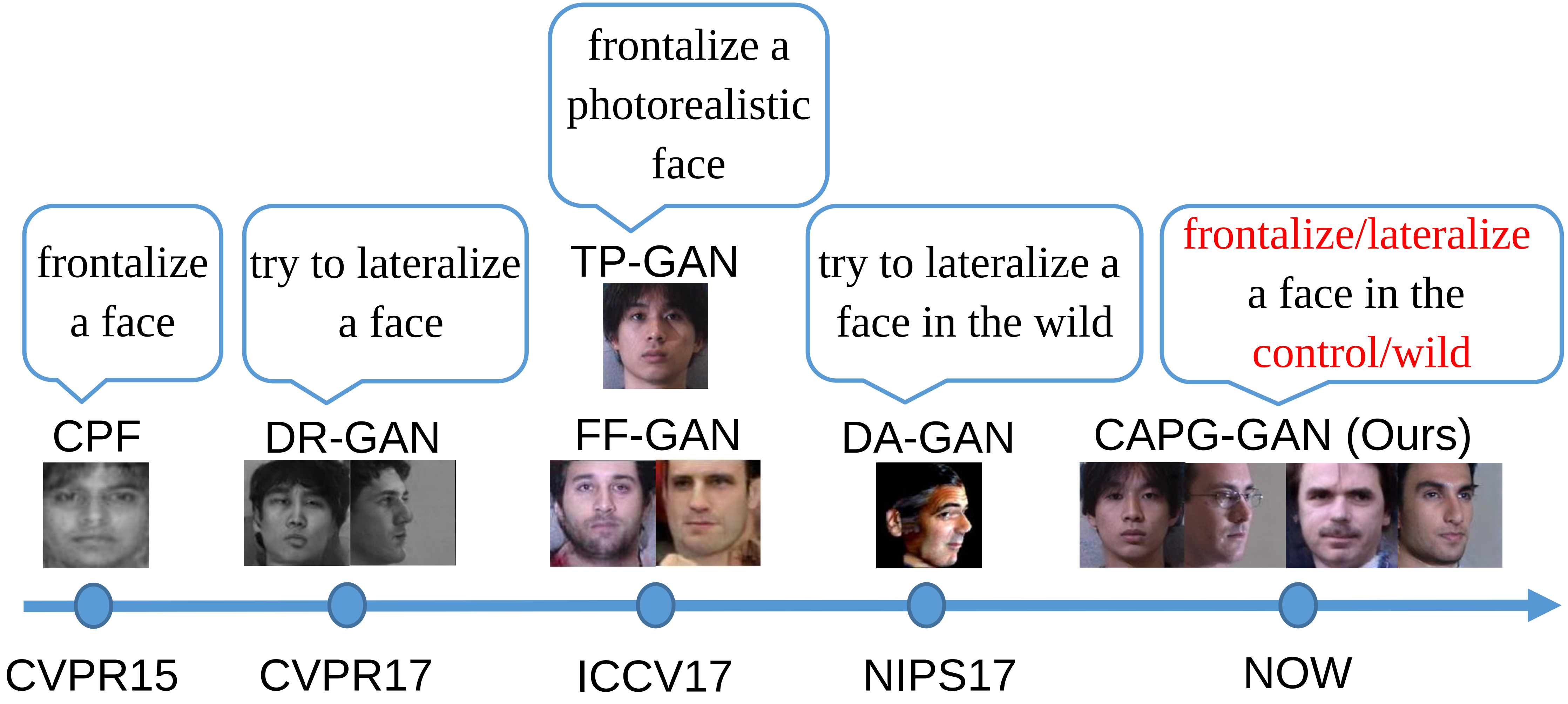
Related Work



Related Work



Related Work



Framework — Couple-Agent Pose-Guided GAN

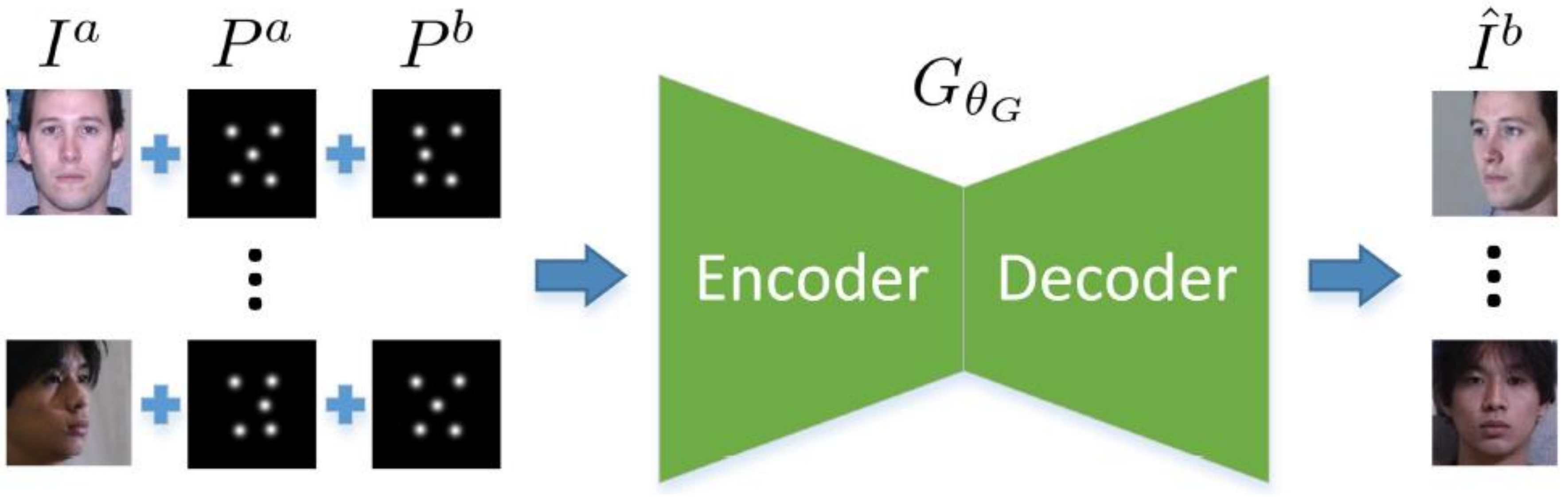
Contributions:

- We propose **Couple-Agent Pose-Guided GAN** (CAPG-GAN) for face rotation in 2D space.

Framework — Couple-Agent Pose-Guided GAN

Contributions:

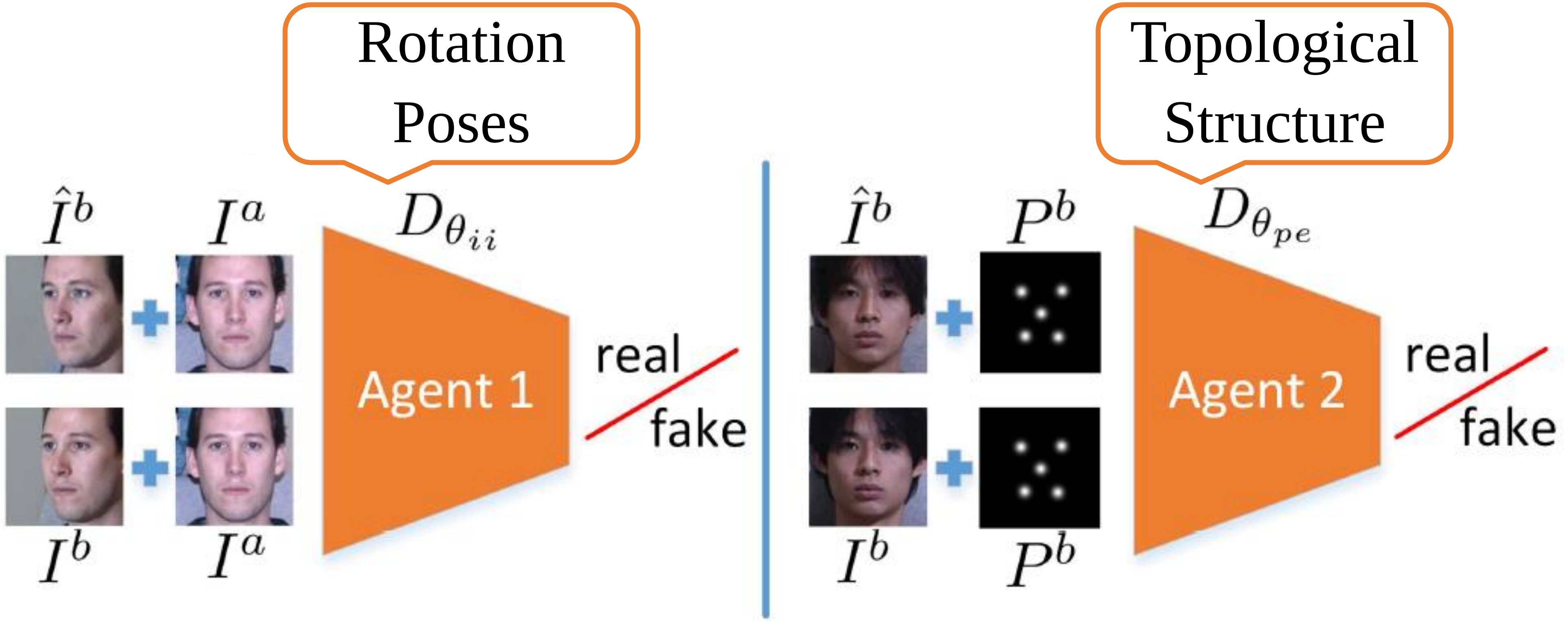
- The **Pose-guided generator** uses landmark heatmaps as controllable signals to synthesize arbitrary poses.



Framework — Couple-Agent Pose-Guided GAN

Contributions:

- The **Couple-agent discriminator** combines prior domain knowledge of poses and topological structure of faces to reinforce the realism.



Losses

- Conditional Adversarial Loss (Coupel-Agent Discriminator)

$$L_{adv}^{ii} = E_{I^b \sim P(I^b)} \left[\log D_{\theta_{ii}} \left(I^b, I^a\right)\right] + E_{\hat{I}^b \sim P(\hat{I}^b)} \left[\log \left(1 - D_{\theta_{ii}} \left(\hat{I}^b, I^a\right)\right)\right]$$

$$L_{adv}^{pe} = E_{I^b \sim P(I^b)} \left[\log D_{\theta_{pe}} \left(I^b, P^b\right)\right] + E_{\hat{I}^b \sim P(\hat{I}^b)} \left[\log \left(1 - D_{\theta_{pe}} \left(\hat{I}^b, P^b\right)\right)\right]$$

- Multi-Scale Pixel-Wise Loss

$$L_{pix} = \frac{1}{S} \sum_{s=1}^{S} \frac{1}{W_s H_s C} \sum_{w,h,c=1}^{W_s,H_s,C} \left| \hat{I}_{s,w,h,c}^{b} - I_{s,w,h,c}^{b} \right|$$

- Identity Preserving Loss

$$L_{ip} = \left\| D_{ip}^{p}(\hat{I}^b) - D_{ip}^{p}(I^b) \right\|_F^2 + \left\| D_{ip}^{fc}(\hat{I}^b) - D_{ip}^{fc}(I^b) \right\|_2^2$$

- Total Variation Regularization

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

Results — Multi-PIE Frontalization

Results — Multi-PIE Frontalization

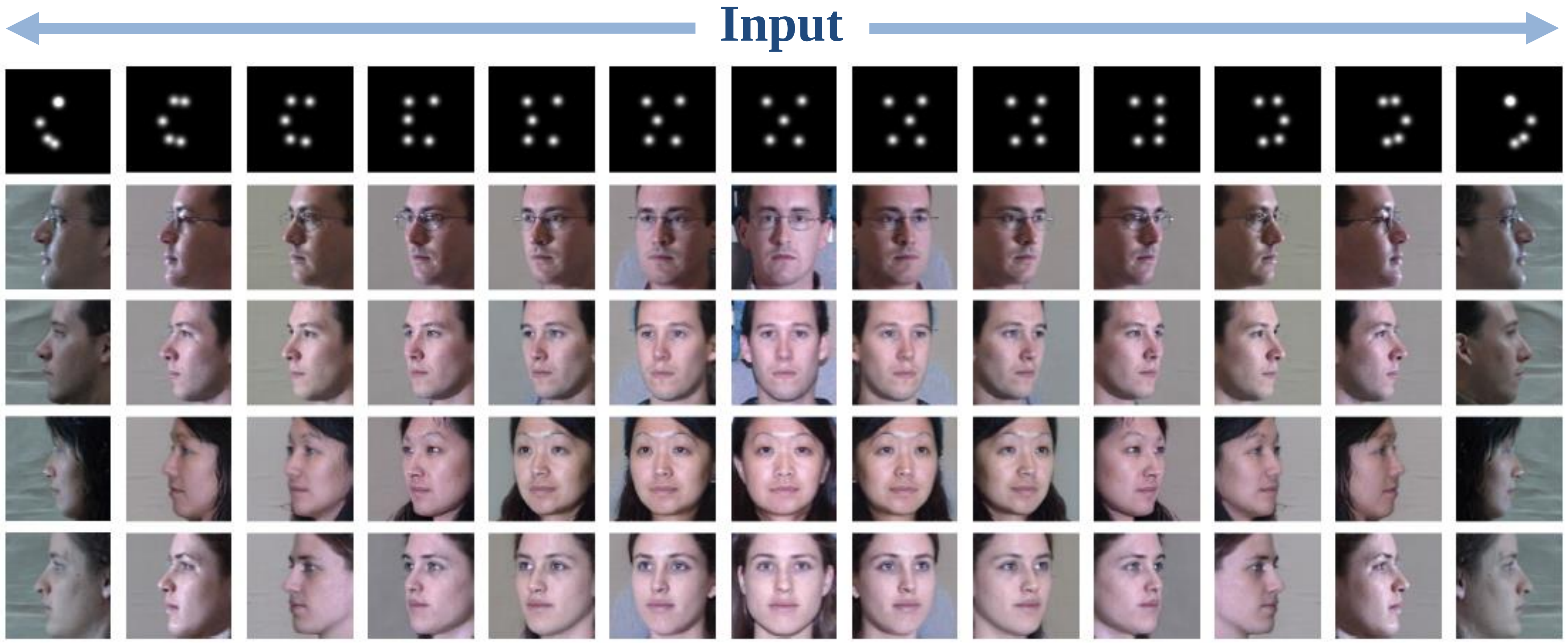
(a) Profile (b) Ours (c) [14] (d) [35] (e) Frontal

Results — Multi-PIE Frontalization

Table 2. Rank-1 recognition rates (%) across views, illuminations and sessions under Setting 2.

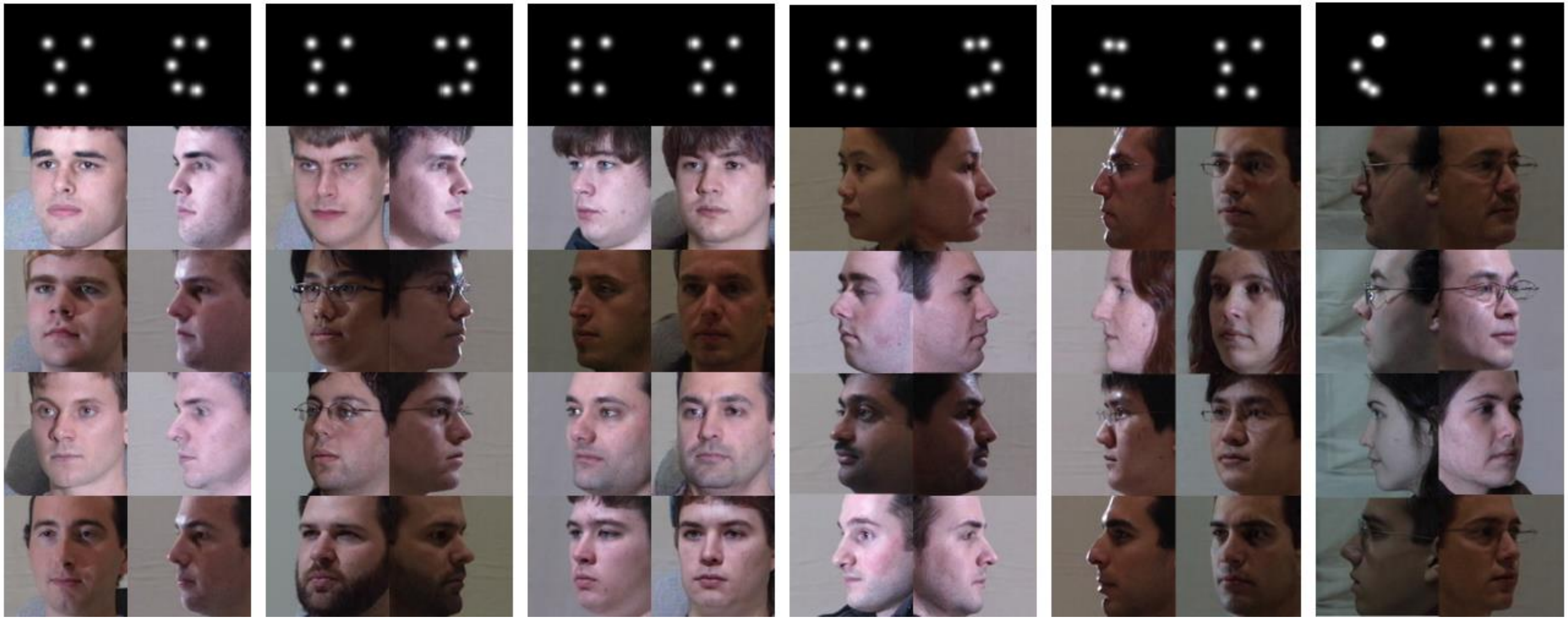
Method	$\pm 90^\circ$	$\pm 75^\circ$	$\pm 60^\circ$	$\pm 45^\circ$	$\pm 30^\circ$	$\pm 15^\circ$
FIP+LDA[40]	-	-	45.9	64.1	80.7	90.7
MVP+LDA[41]	-	-	60.1	72.9	83.7	92.8
CPF[34]	-	-	61.9	79.9	88.5	95.0
DR-GAN[28]	-	-	83.2	86.2	90.1	94.0
FF-GAN[35]	61.2	77.2	85.2	89.7	92.5	94.6
TP-GAN[14]	64.64	77.43	87.72	95.38	98.06	98.68
Light CNN[29]	5.51	24.18	62.09	92.13	97.38	98.59
CAPG-GAN	**66.05**	**83.05**	**90.63**	**97.33**	**99.56**	**99.82**
	↑60.54	**↑58.87**	**↑28.54**	**↑5.2**	**↑2.18**	**↑1.23**

Results — Multi-PIE Rotation



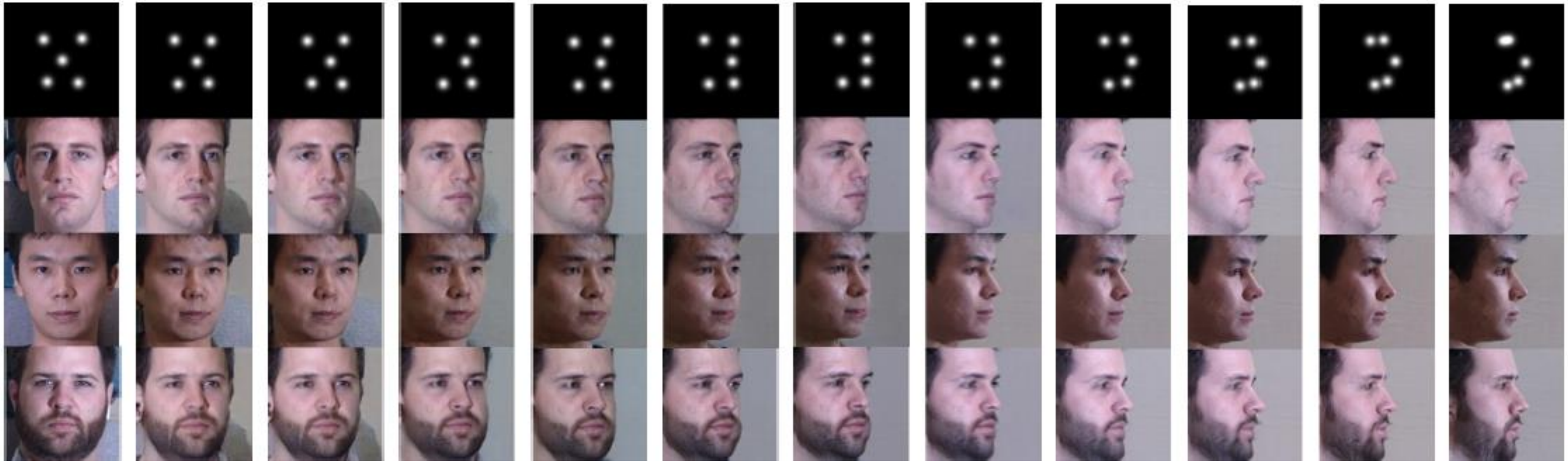
(a) +90° (b) +75° (c) +60° (d) +45° (e) +30° (f) +15° (g) 0° (h) −15° (i) −30° (j) −45° (k) −60° (l) −75° (m) −90°

Results — Multi-PIE Rotation

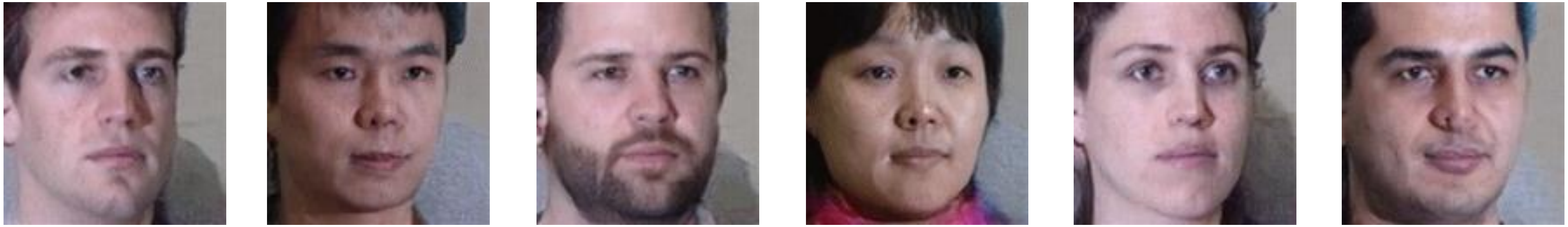


(a) $+15^\circ \rightarrow +60^\circ$ (b) $+30^\circ \rightarrow -60^\circ$ (c) $+45^\circ \rightarrow +15^\circ$ (d) $+60^\circ \rightarrow -75^\circ$ (e) $+75^\circ \rightarrow +30^\circ$ (f) $+90^\circ \rightarrow -45^\circ$

Results — Multi-PIE Rotation



(a) Source (b) 7.5° (c) 15° (d) 22.5° (e) 30° (f) 37.5° (g) 45° (h) 52.5° (i) 60° (j) 67.5° (k) 75° (l) 82.5°



Results — LFW