

Contribution

- **Two generators** are designed to learn the two-way conditional distributions between the two domains
- **Two discriminators** implicitly define a ternary discriminative function

Model

- Original GAN: learn data distribution in *one* domain

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

- Triangle GAN = Cond. GAN + ALI/BiGAN

$$\begin{aligned} \min_{G_x, G_y} \max_{D_1, D_2} V(G_x, G_y, D_1, D_2) &= \mathbb{E}_{(\mathbf{x}, \mathbf{y}) \sim p(\mathbf{x}, \mathbf{y})} [\log D_1(\mathbf{x}, \mathbf{y})] \\ &+ \mathbb{E}_{\mathbf{y} \sim p(\mathbf{y}), \tilde{\mathbf{x}} \sim p_x(\mathbf{x}|\mathbf{y})} \left[\log \left((1 - D_1(\tilde{\mathbf{x}}, \mathbf{y})) \cdot D_2(\tilde{\mathbf{x}}, \mathbf{y}) \right) \right] \\ &+ \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x}), \tilde{\mathbf{y}} \sim p_y(\mathbf{y}|\mathbf{x})} \left[\log \left((1 - D_1(\mathbf{x}, \tilde{\mathbf{y}})) \cdot (1 - D_2(\mathbf{x}, \tilde{\mathbf{y}})) \right) \right] \end{aligned}$$

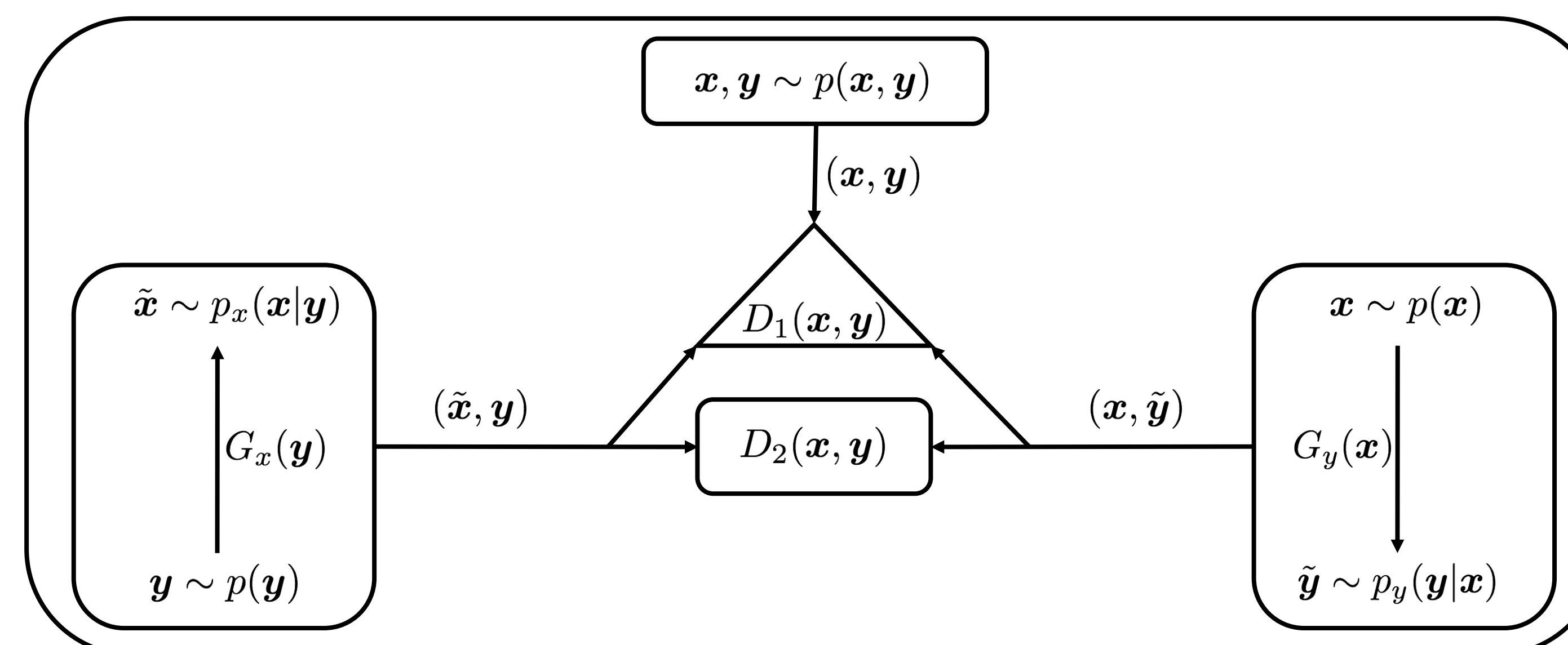
- Blue part accounts for *supervised* learning
- Red part accounts for *unsupervised* learning

Theorem

The equilibrium of $V(G_x, G_y, D_1, D_2)$ is achieved if and only if $p(\mathbf{x}, \mathbf{y}) = p_x(\mathbf{x}, \mathbf{y}) = p_y(\mathbf{x}, \mathbf{y})$ with $D_1^*(\mathbf{x}, \mathbf{y}) = \frac{1}{3}$ and $D_2^*(\mathbf{x}, \mathbf{y}) = \frac{1}{2}$, and the optimum value is $-3 \log 3$.

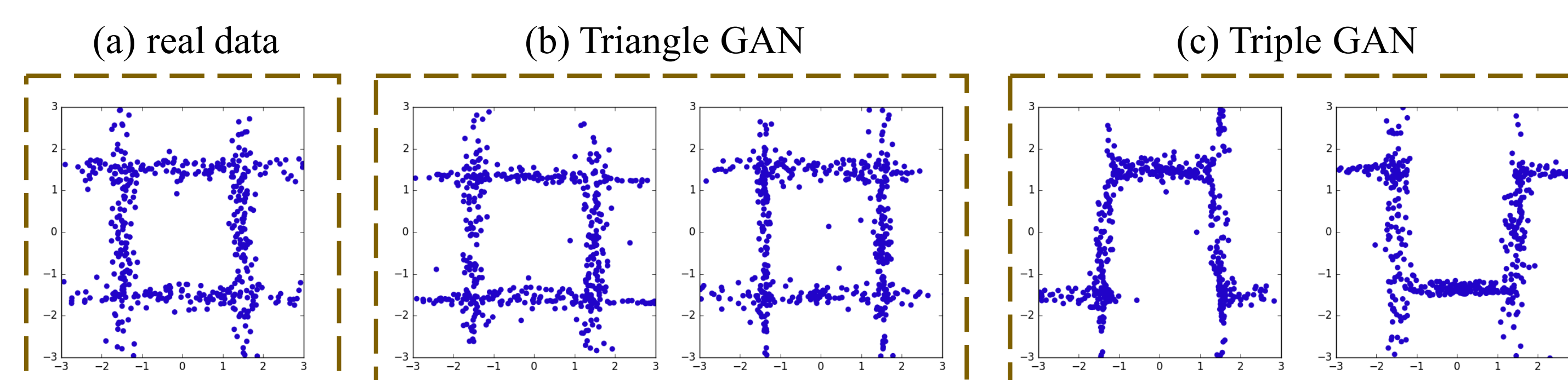
$$\begin{aligned} C(G_x, G_y) &= \max_{D_1, D_2} V(G_x, G_y, D_1, D_2) \\ &= -3 \log 3 + 3 \cdot JSD(p(\mathbf{x}, \mathbf{y}), p_x(\mathbf{x}, \mathbf{y}), p_y(\mathbf{x}, \mathbf{y})) \end{aligned}$$

Framework



Experiment

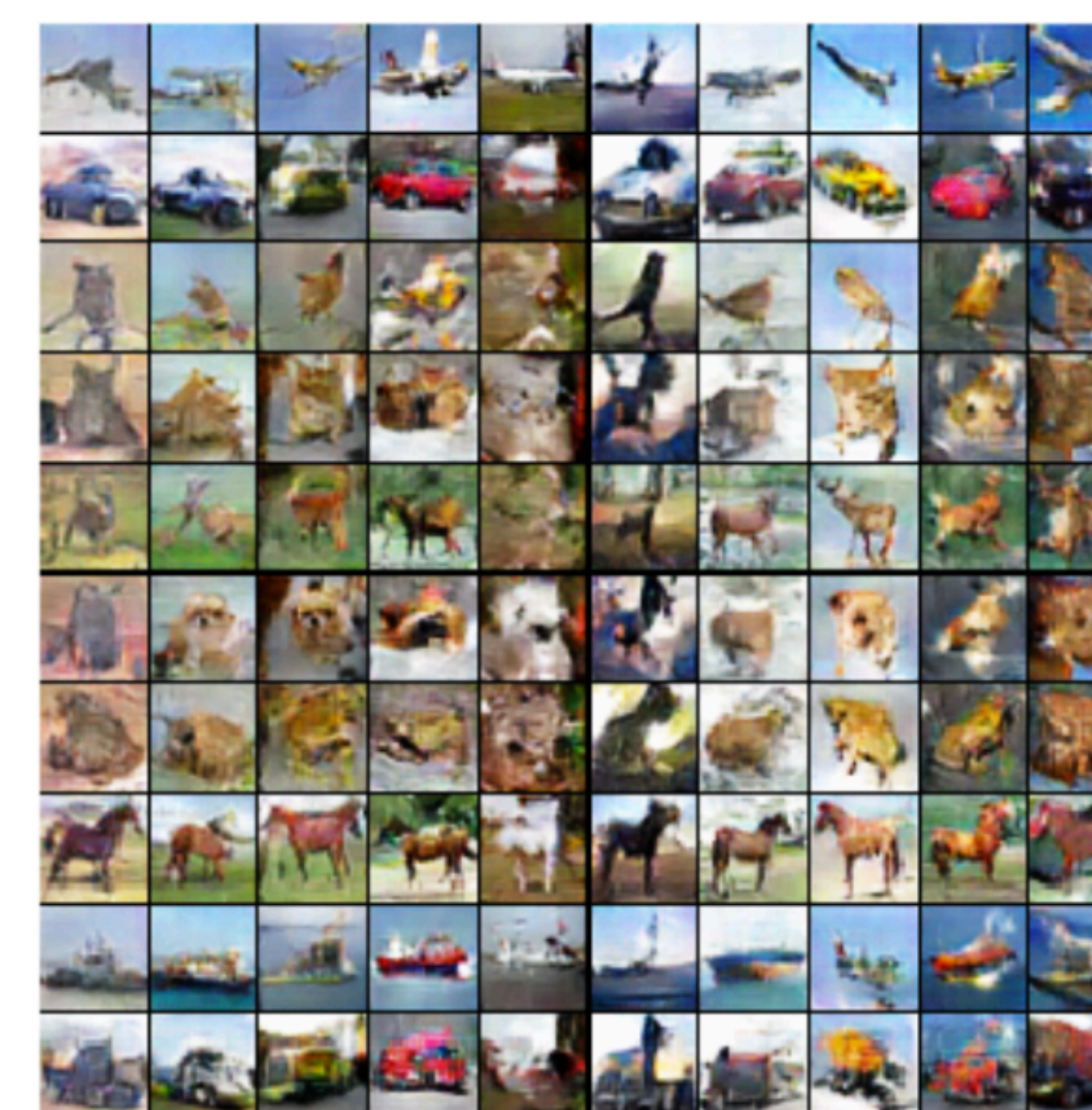
- Toy data



- Image & label pair

Table 1: Error rates (%) on the partially labeled CIFAR10 dataset.

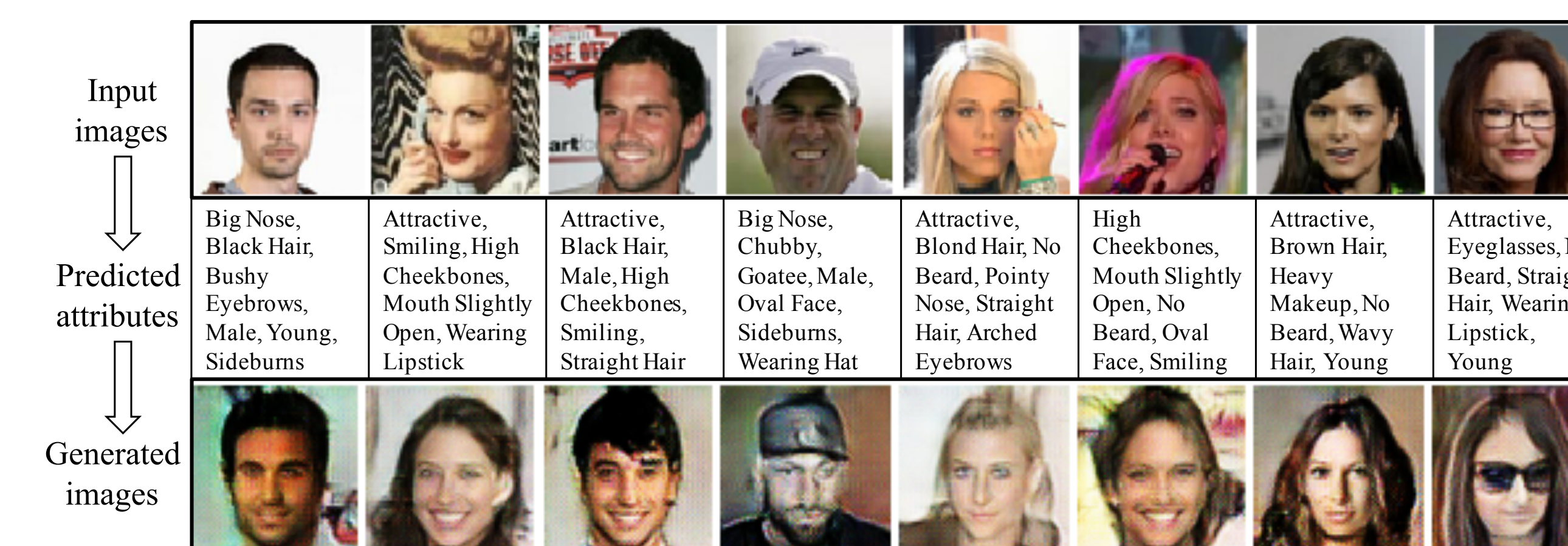
Algorithm	$n = 4000$
CatGAN [36]	19.58 ± 0.58
Improved GAN [21]	18.63 ± 2.32
ALI [10]	17.99 ± 1.62
Triple GAN [12]	16.99 ± 0.36
Δ -GAN (ours)	16.80 ± 0.42



- Image & image pair: 15% vs. 93% Accuracy

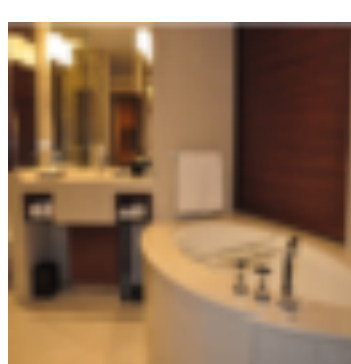
	DiscoGAN				Δ -GAN			
Input:	4	5	0	7	7	5	9	5
Output:	3	2	8	2	1	5	5	1
Input:	6	6	0	1	8	1	0	1
Output:	4	2	9	0	8	1	0	1

- Image & attribute pair



1st row+pale skin/mouth slightly open/eyeglasses/wearing hat=2nd row



Input	Predicted attributes	Generated images
	baseball, standing, next, player, man, group, person, field, sport, ball, outdoor, game, grass, crowd	
	sink, shower, indoor, tub, restroom, bathroom, small, standing, room, tile, white, stall, tiled, black, bath	