vision@OUC

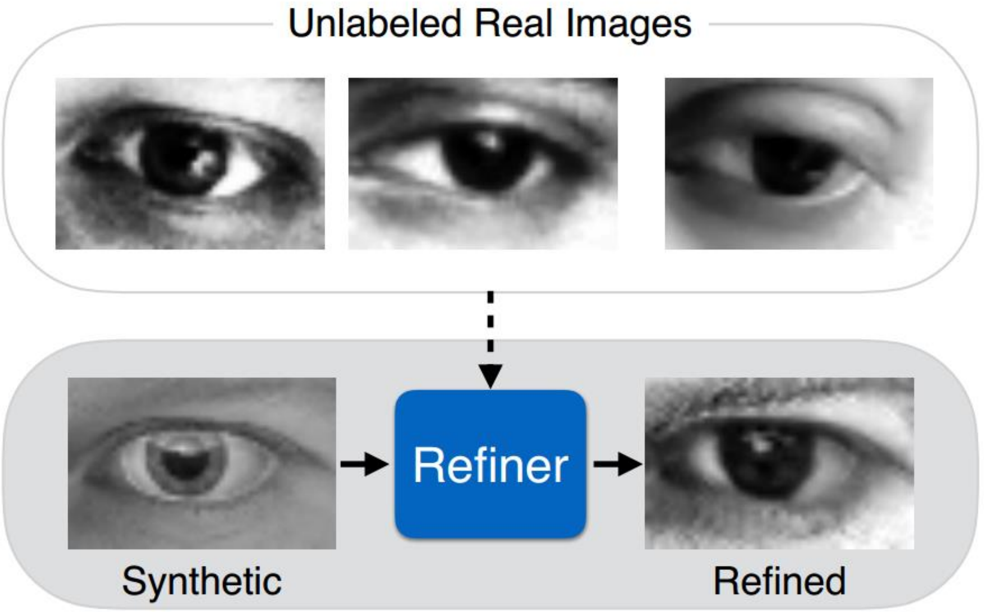
GAN in CVPR2017

Wang Chao, Group of DL

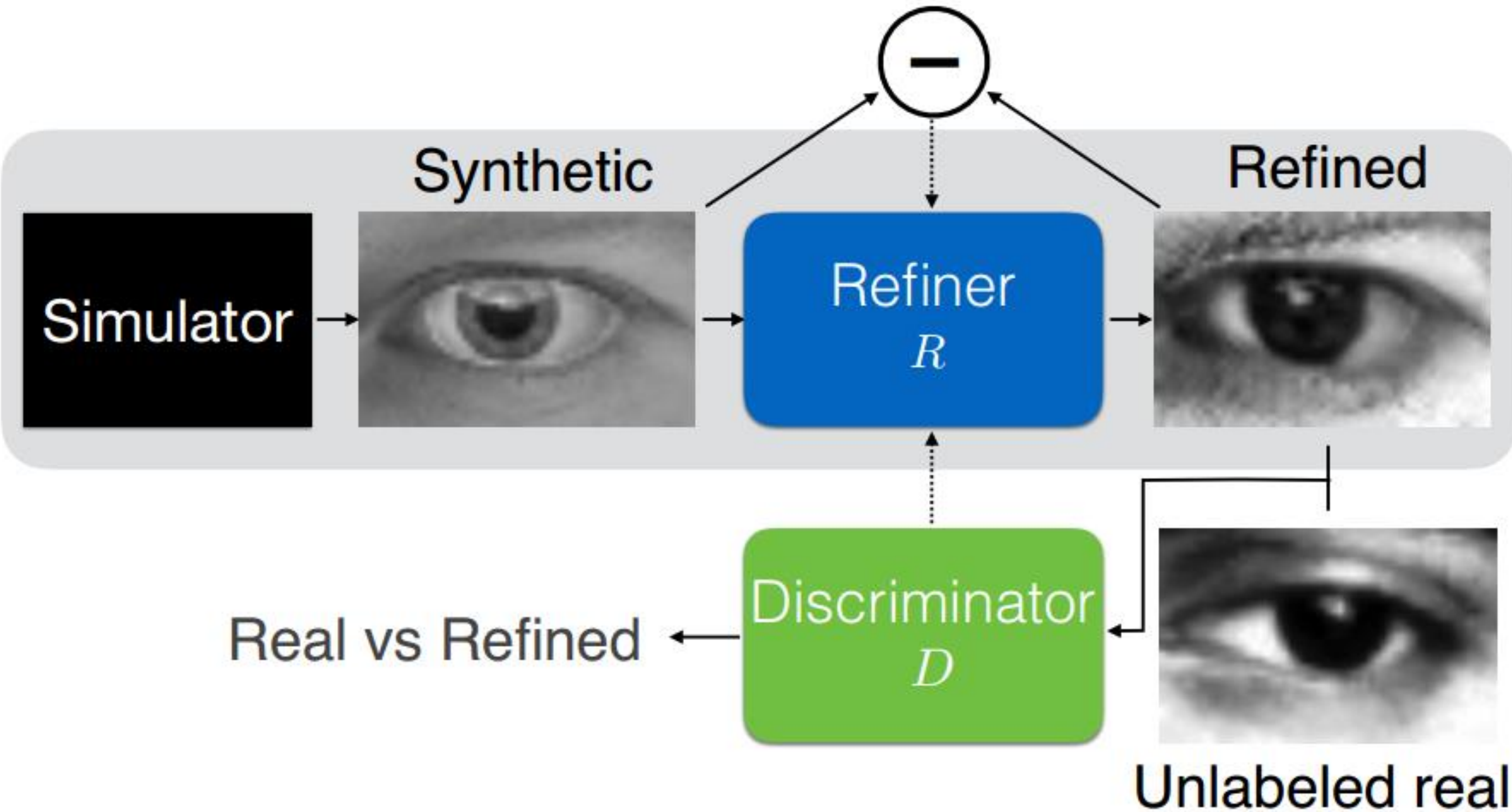
2017-7-22

SimGAN

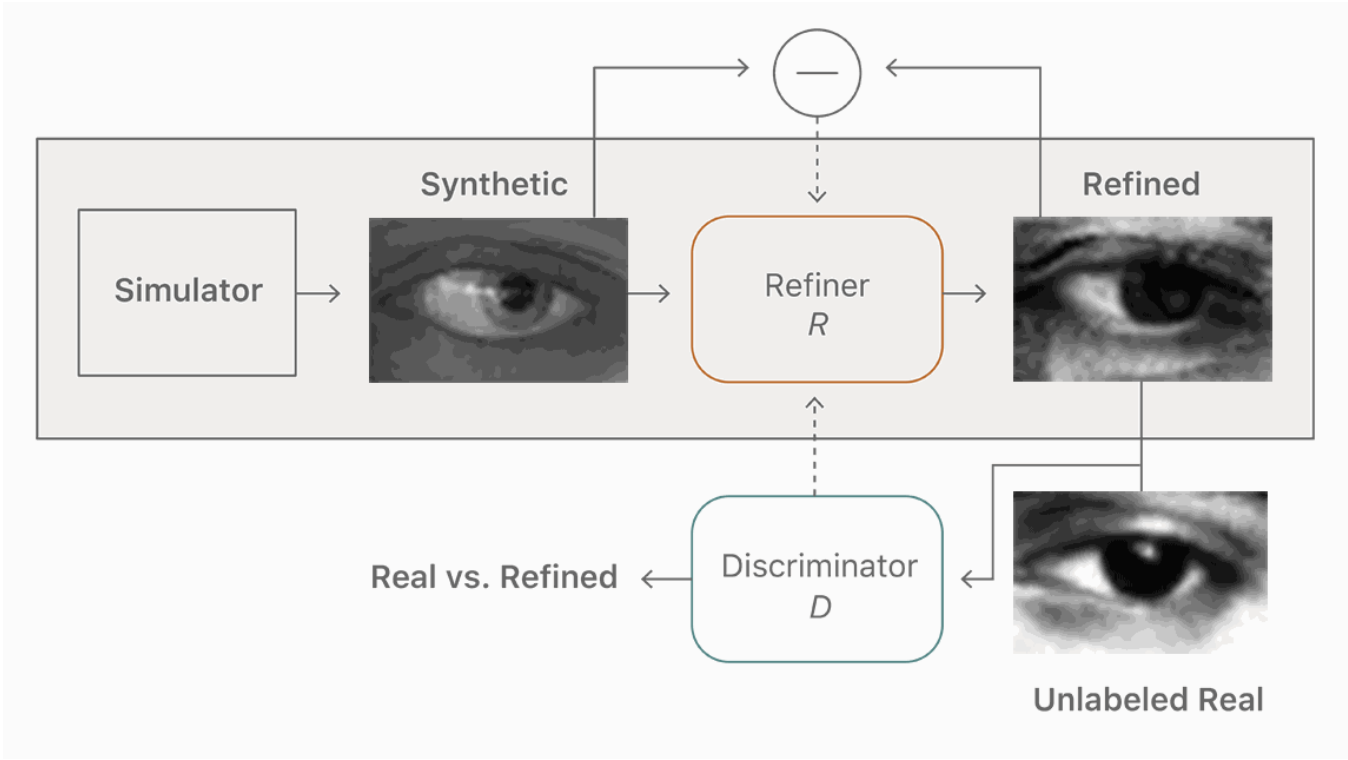
- A 'self-regularization' term
- A local adversarial loss
- Updating the discriminator using a history of refined images



SimGAN

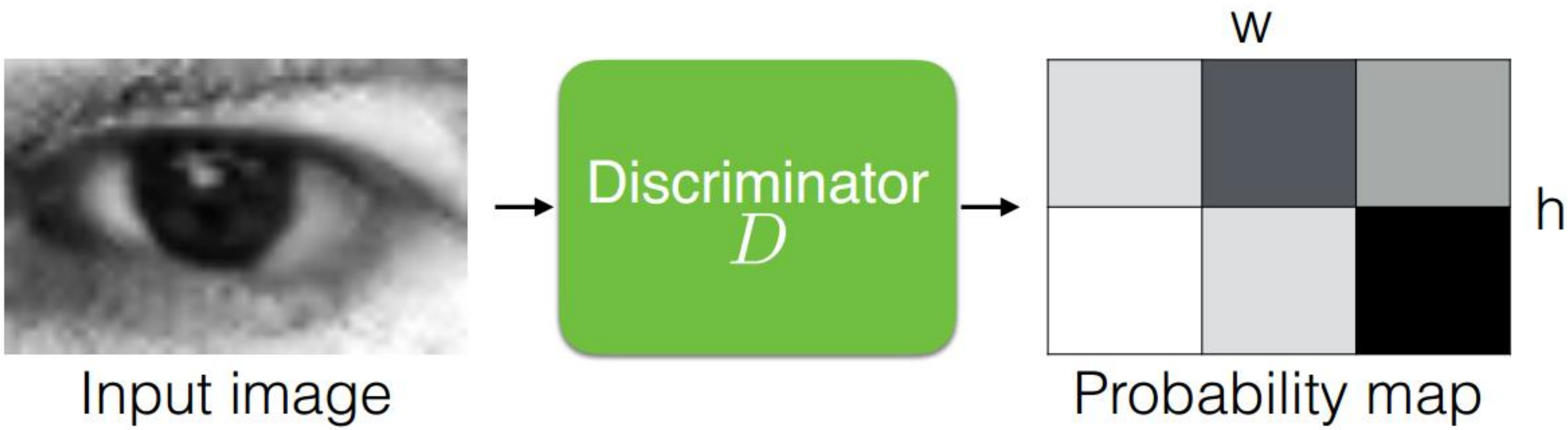

Synthetic
Refined
Simulator
Refiner
R
Discriminator
D
Real vs Refined
Unlabeled real

Adversarial Loss with Self-Regularization

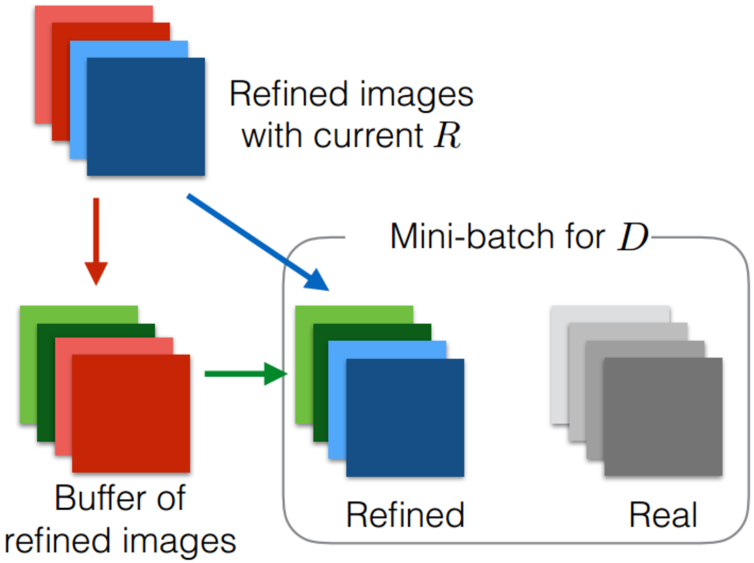


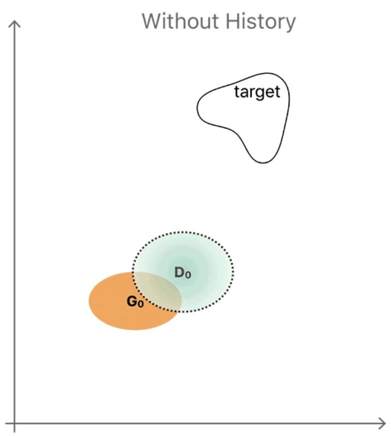
$$L_R(\theta) = -\sum_i \log(1 - D_\phi(R_\theta(x_i))) + \lambda \| \psi(R_\theta(x_i)) \|_1$$

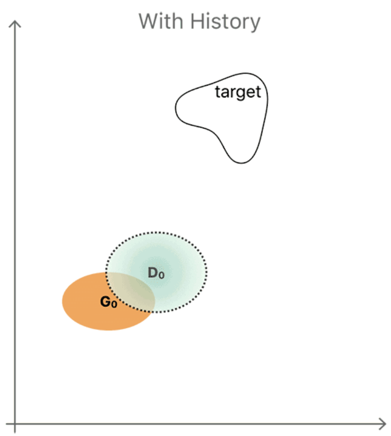
Local Adversarial Loss



Updating the Discriminator using History

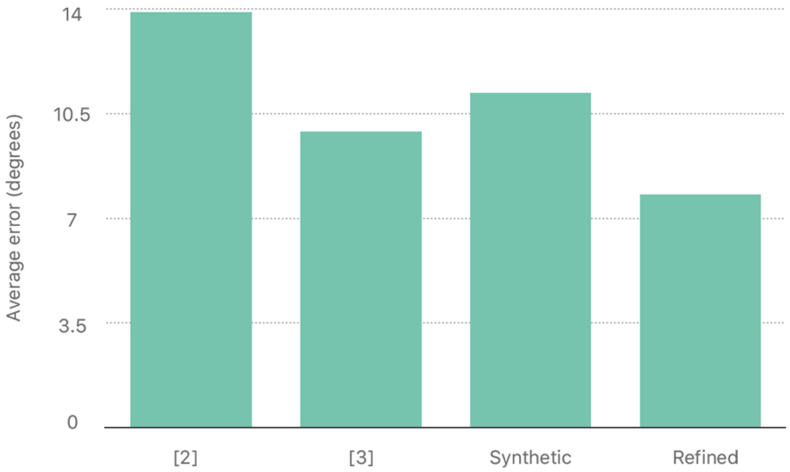






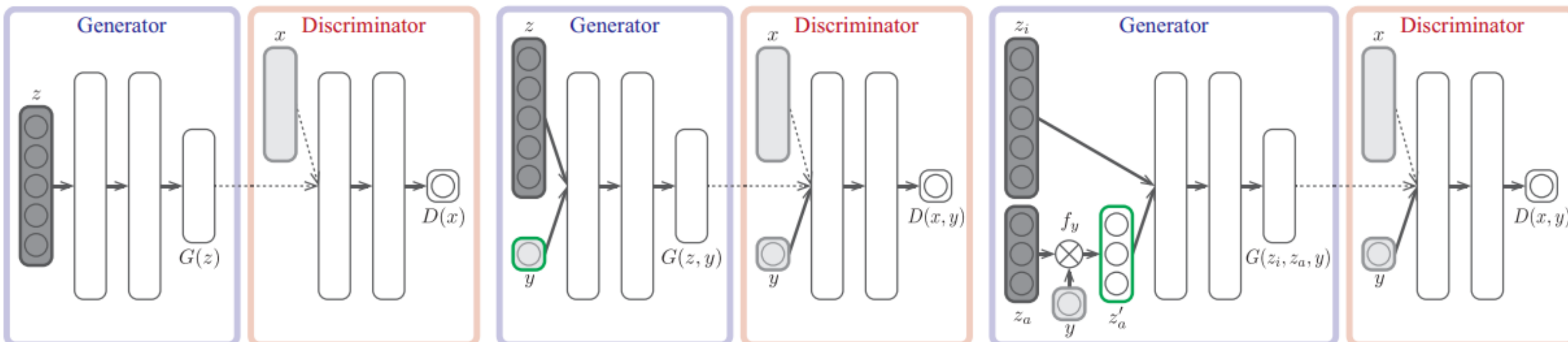
SimGAN

Synthetic | Refined (with history) | Refined (without history)



Training data	% of images within d
Synthetic Data	62.3
Synthetic Data 4x	64.9
Refined Synthetic Data	69.4
Refined Synthetic Data 4x	**87.2**

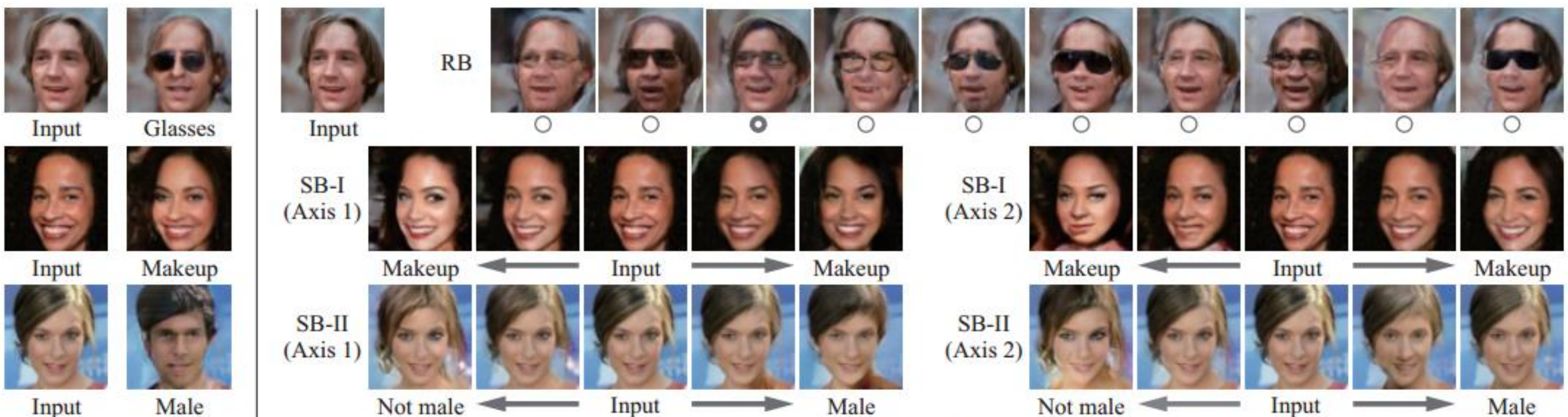
Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks



(a) GAN (b) CGAN (c) CFGAN

$$\min_{\mathbf{G}} \max_{\mathbf{D}} V(D,G) = E_{x\sim p_{data}(x,y)}[\log D(x,y)] + E_{z_i \sim p_{z_i}(z_i), z_a \sim p_{z_a}(z_a), y \sim p_y(y)}$$
$$[\log(1 - D(G(z_i, z_a, y), y)]$$

Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks

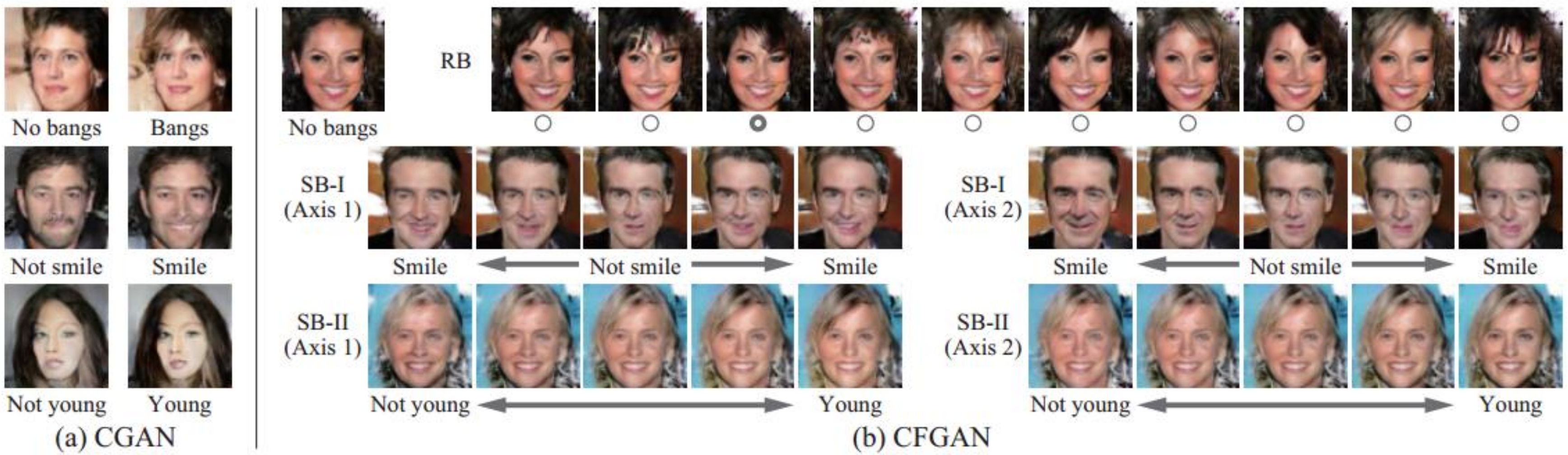


(a) CGAN (b) CFGAN

$$f_y(z_a) = \begin{cases} z_a, & (y = 1) \\ 0, & (y = 0) \end{cases} \quad [z_a \sim Cat(K = k, p = \frac{1}{k})]$$

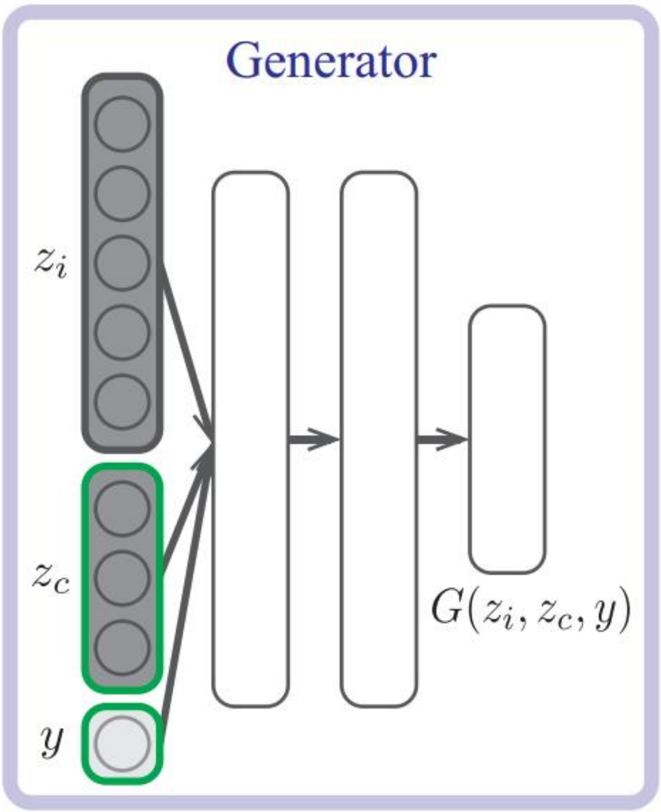
Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks". CVPR 2017.

Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks

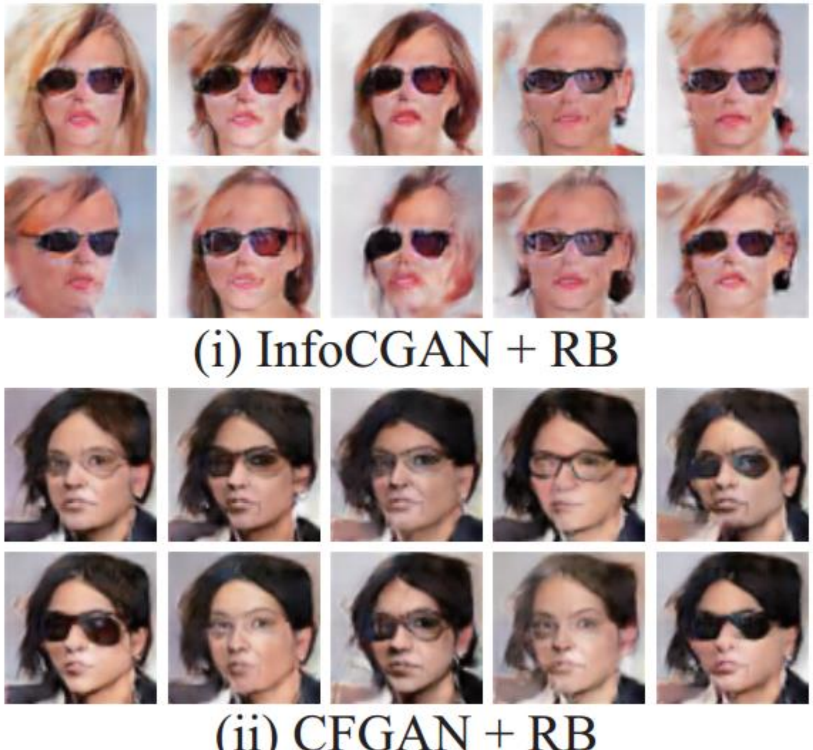


$$f_y(z_a) = \begin{cases} z_a, & (y = 1) \\ 0, & (y = 0) \end{cases} \quad [z_a \sim Unif(-1,1)]$$

Generative Attribute Controller With Conditional Filtered Generative Adversarial Networks



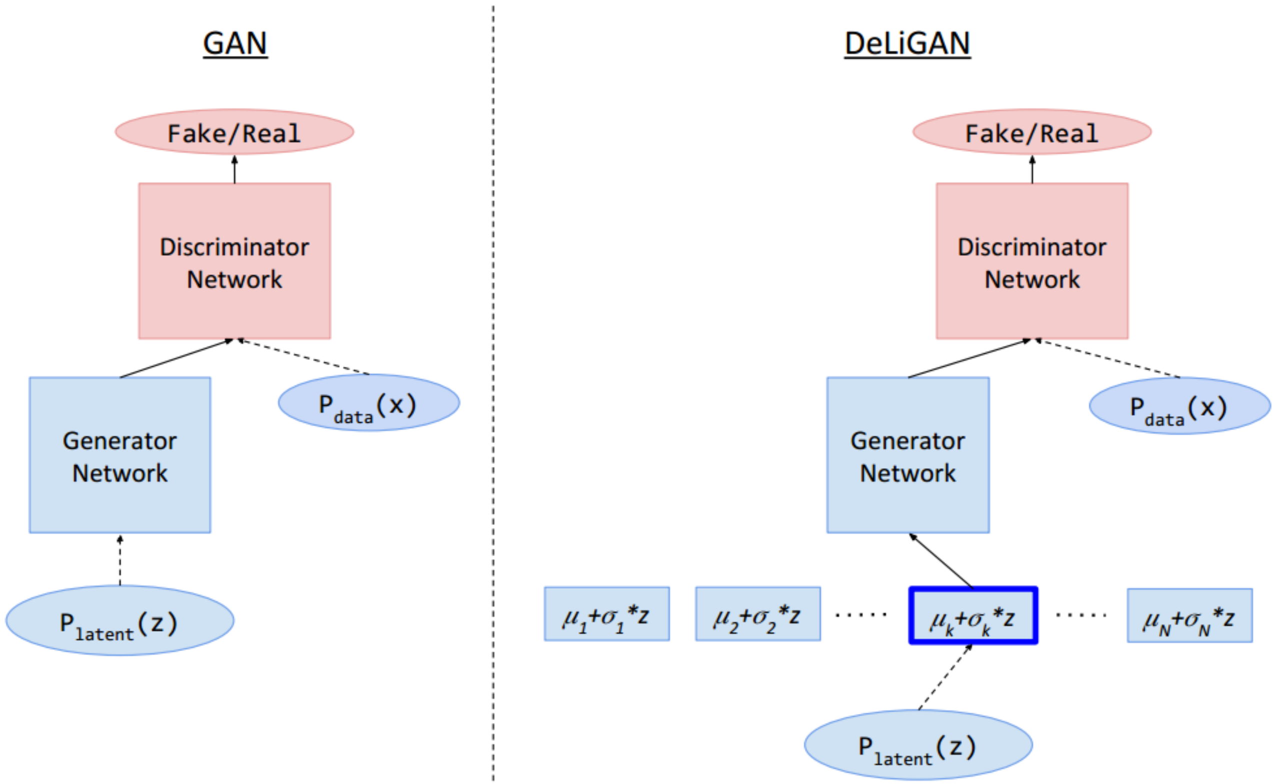
(a) InfoCGAN



Attribute y: Glasses

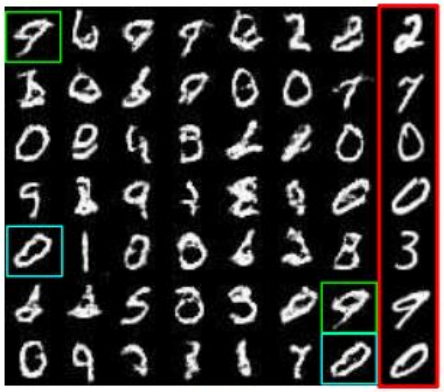
(b) Example results

DeliGAN



$$z = \mu_i + \sigma_i \varepsilon, where, \varepsilon \sim N(0,1)$$

DeliGAN



a) GAN

b) DeliGAN

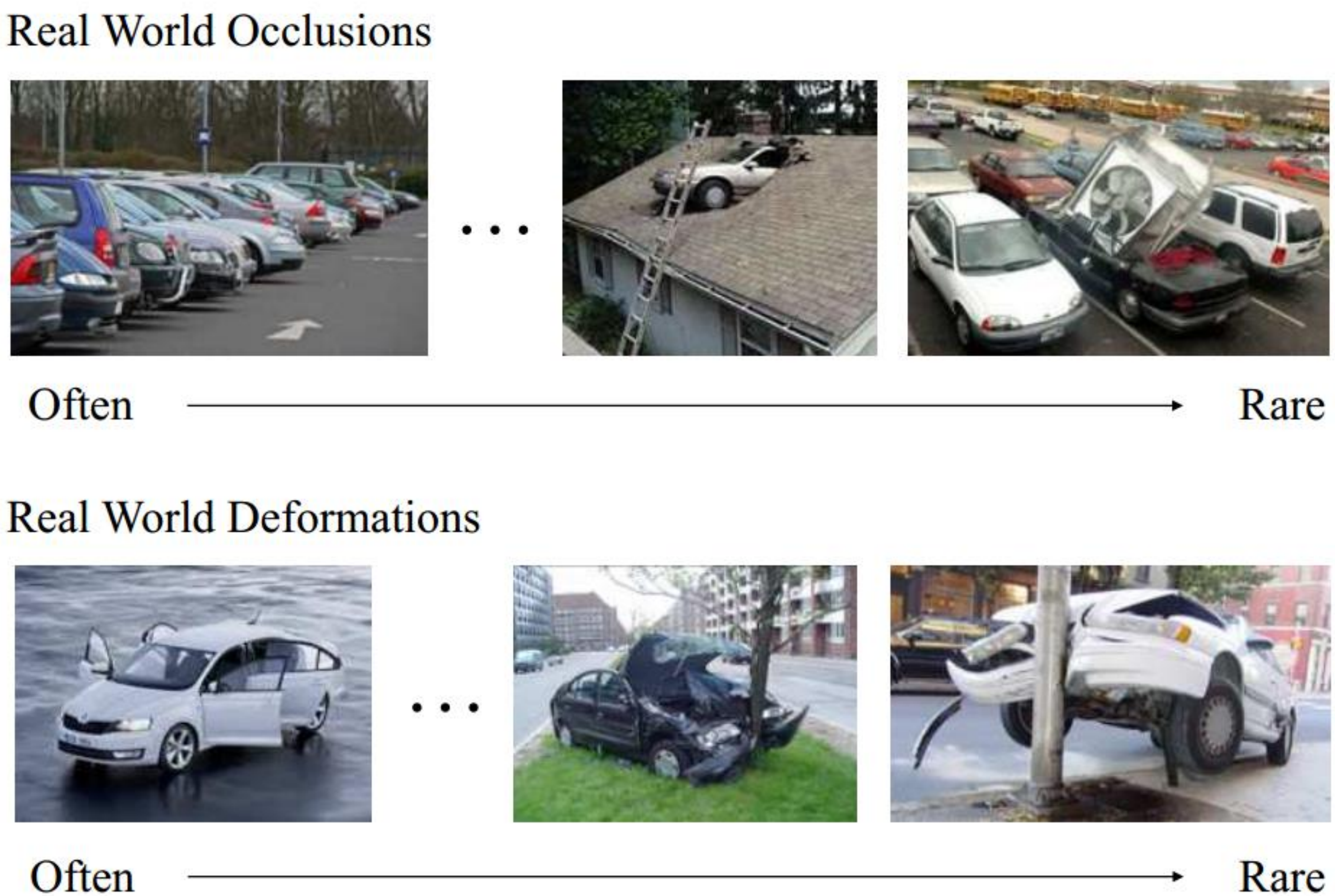
a) GAN

b) DeLiGAN

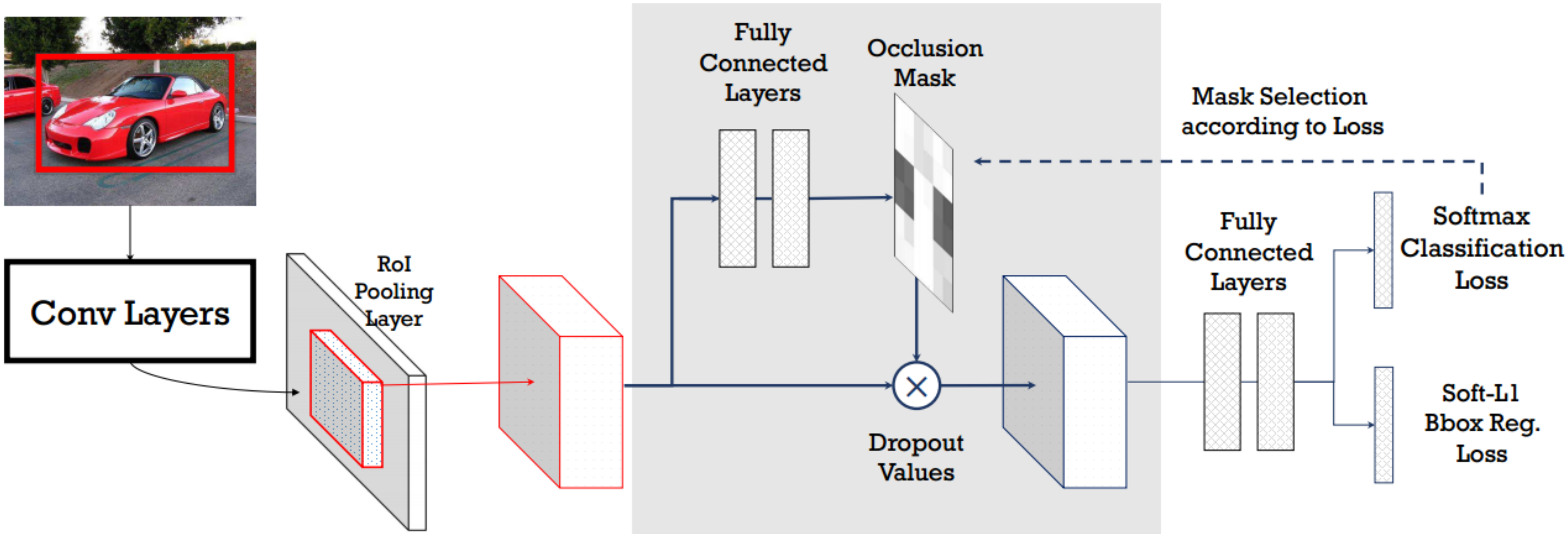
GAN++	2.44 ∓ 0.06	1.73 ∓ 0.04	1.68 ∓ 0.05	2.27 ∓ 0.06	2.23 ∓ 0.04	1.73 ∓ 0.03	1.26 ∓ 0.02	1.21 ∓ 0.04	1.25 ∓ 0.02	1.23 ∓ 0.02	1.76 ∓ 0.40
MoE-GAN	2.69 ∓ 0.08	2.08 ∓ 0.05	2.01 ∓ 0.06	2.19 ∓ 0.04	2.16 ∓ 0.03	1.85 ∓ 0.09	1.84 ∓ 0.07	2.14 ∓ 0.08	1.60 ∓ 0.04	1.85 ∓ 0.05	2.04 ∓ 0.28
DeLiGAN	**2.78** ∓ 0.02	**2.36** ∓ 0.06	**2.44** ∓ 0.07	2.17 ∓ 0.04	**2.31** ∓ 0.02	1.27 ∓ 0.01	**2.31** ∓ 0.02	**3.63** ∓ 0.14	**1.21** ∓ 0.03	2.00 ∓ 0.05	**2.28** ∓ 0.62
GAN	2.72 ∓ 0.20	2.02 ∓ 0.18	2.21 ∓ 0.44	**2.43** ∓ 0.19	2.06 ∓ 0.09	**2.22** ∓ 0.23	1.82 ∓ 0.08	2.12 ∓ 0.55	1.19 ∓ 0.19	**2.16** ∓ 0.15	2.15 ∓ 0.25
	Plane	Car	Bird	Cat	Deer	Dog	Frog	Horse	Ship	Truck	Overall

Inception score

A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection



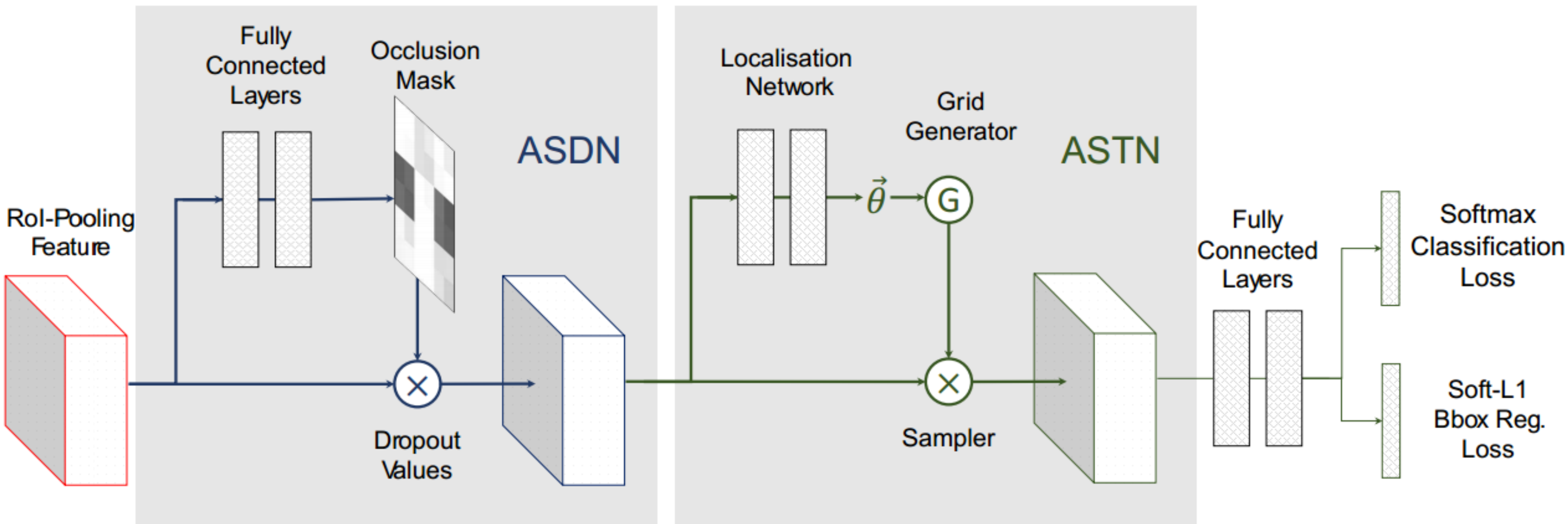
A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection



(a) Pre-training via Searching

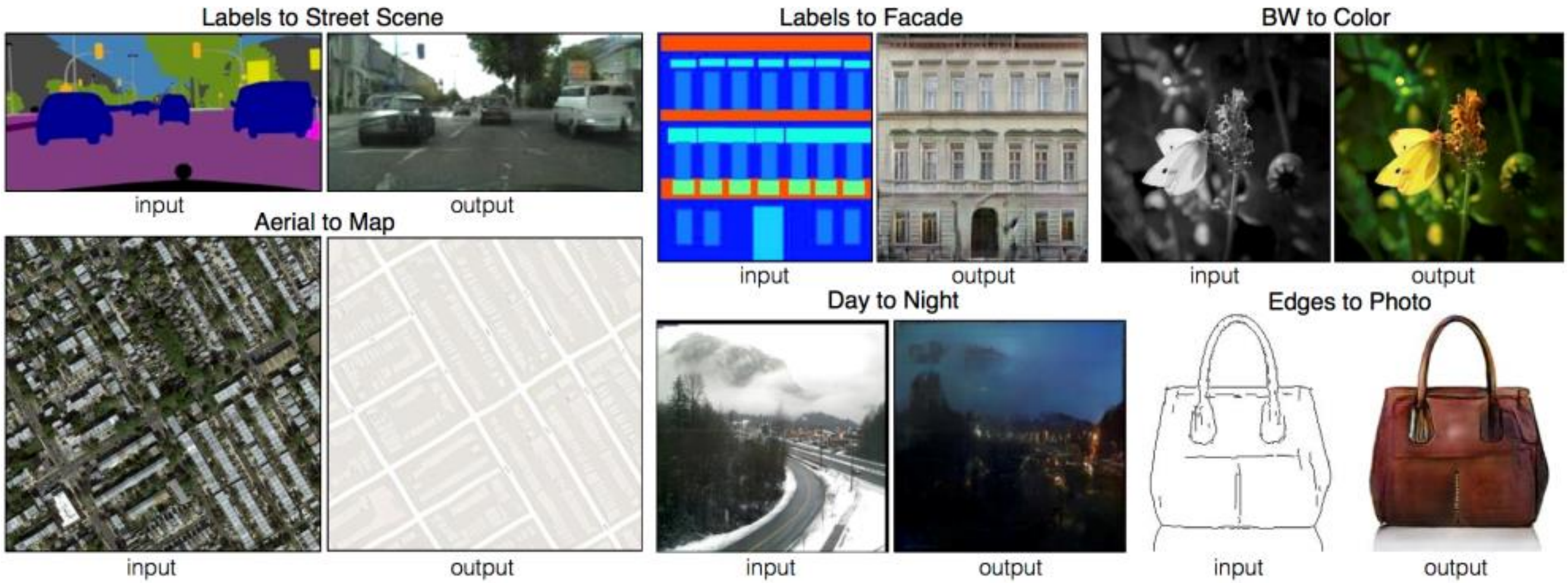
(b) Generated Masks

A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection



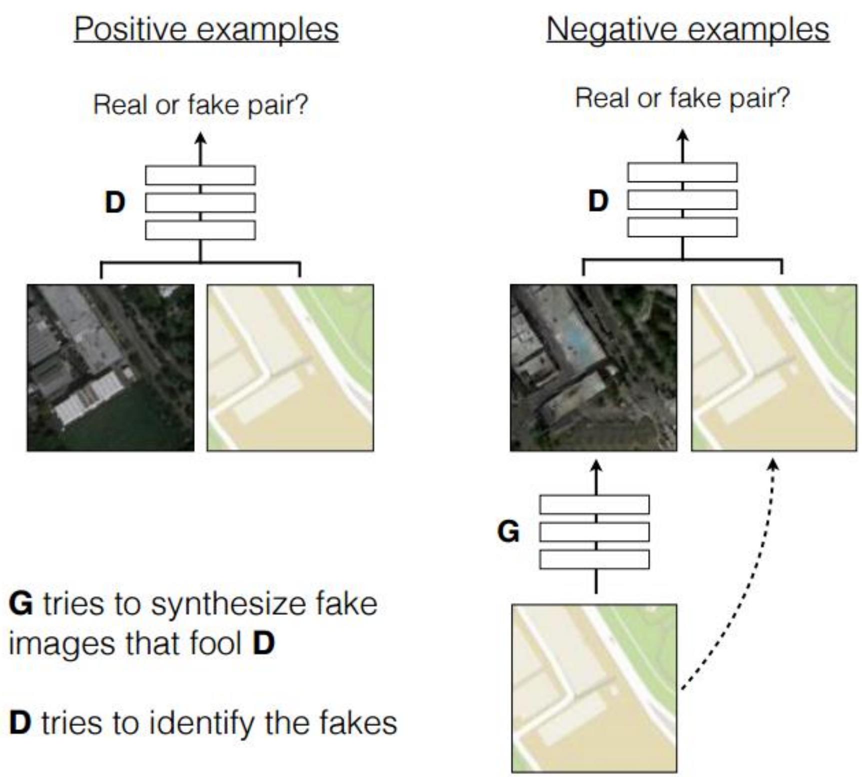
A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection CVPR 2017.

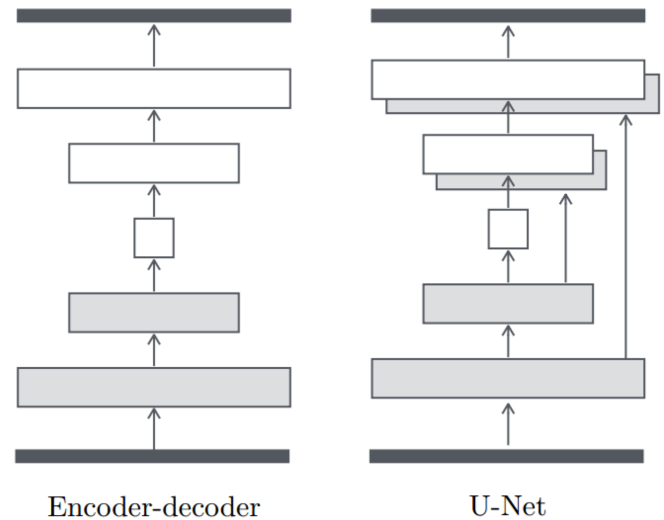
Pix2pix



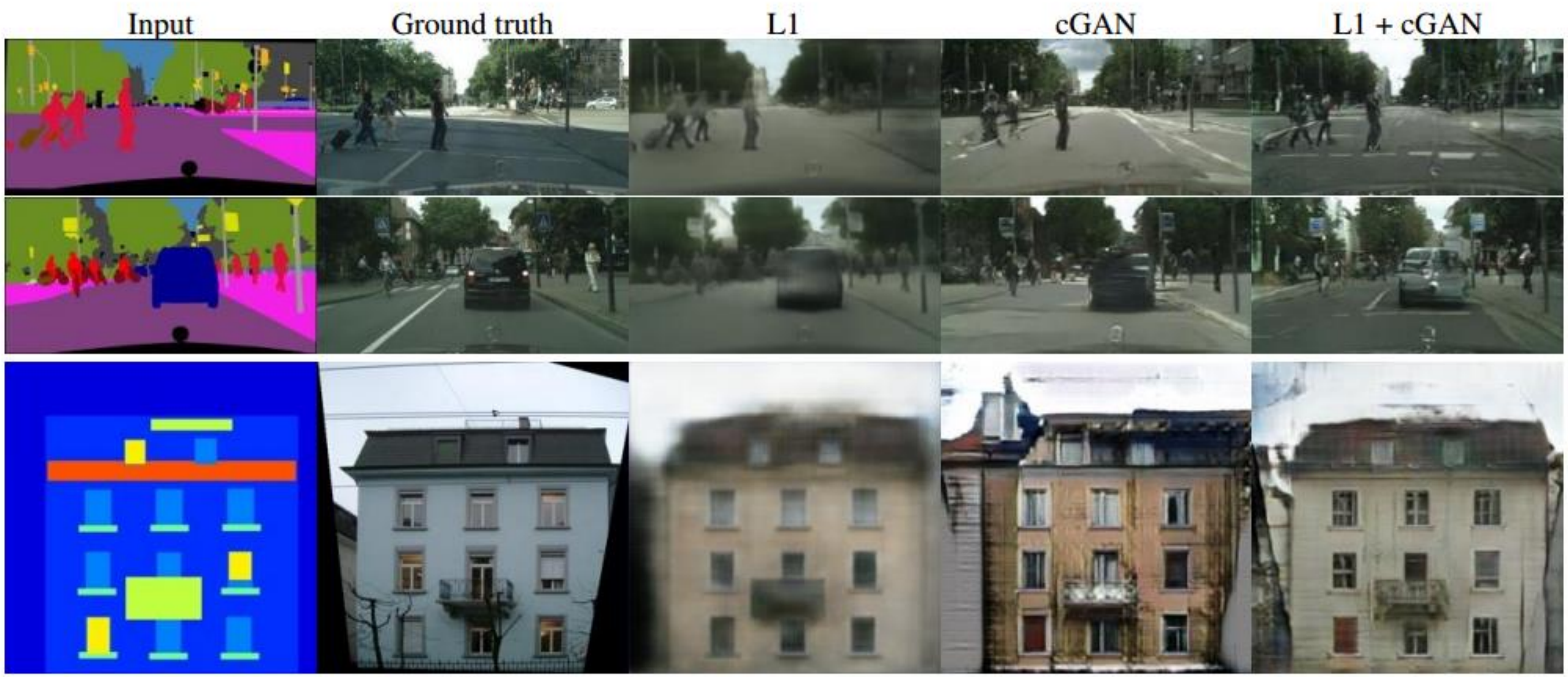
Isola *et al.* "Image-to-Image Translation with Conditional Adversarial Networks". CVPR 2017.

Pix2pix


Positive examples
Real or fake pair?
D
Negative examples
Real or fake pair?
D
G
G tries to synthesize fake images that fool **D**
D tries to identify the fakes

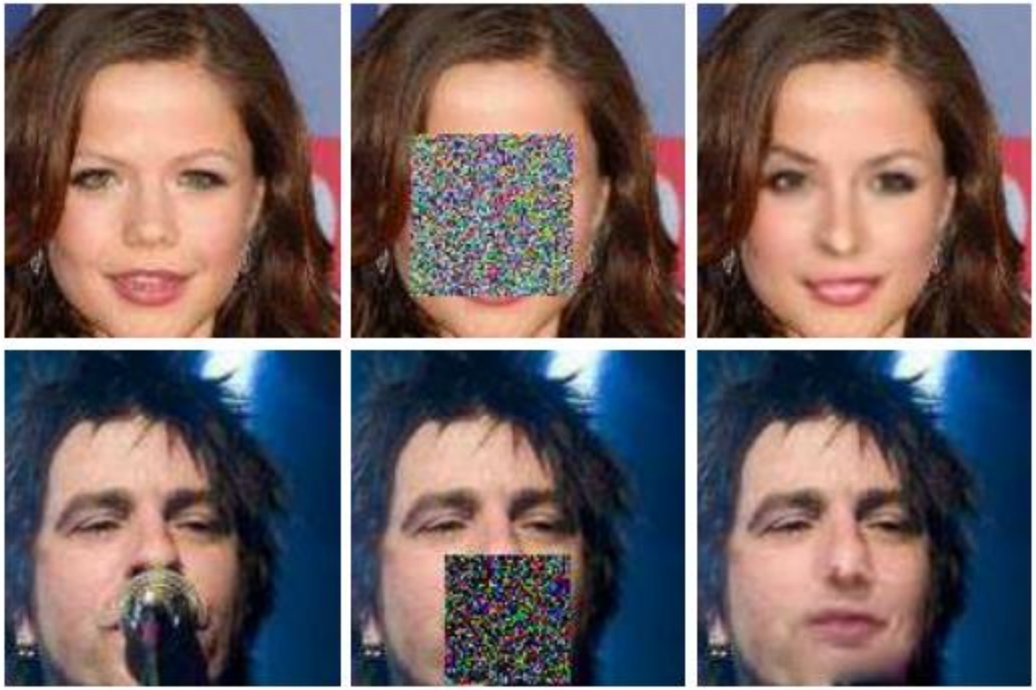

Encoder-decoder
U-Net

Pix2pix



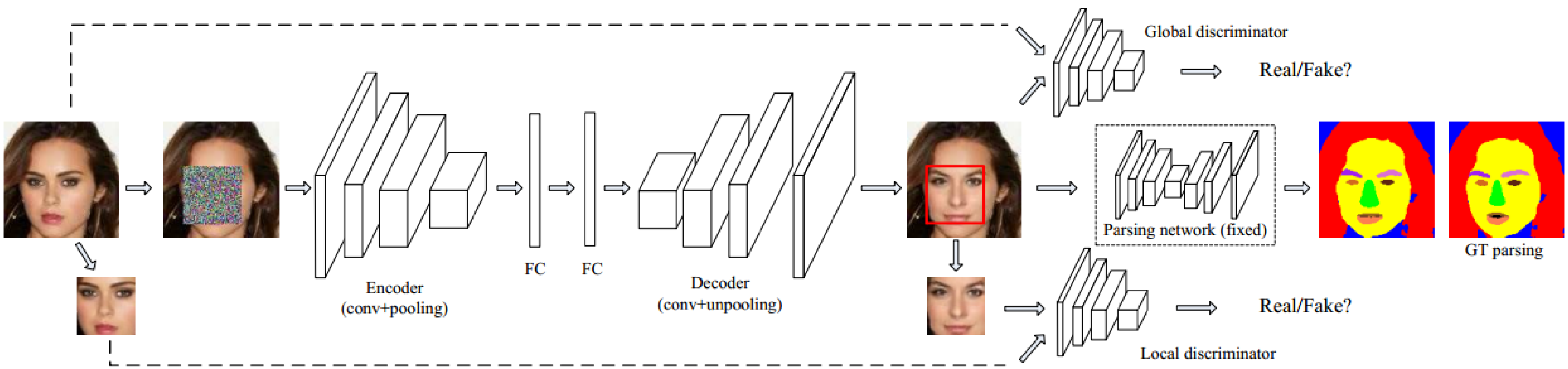
Isola *et al.* “Image-to-Image Translation with Conditional Adversarial Networks”. CVPR 2017.

Generative Face Completion



(a) (b) (c)

Generative Face Completion



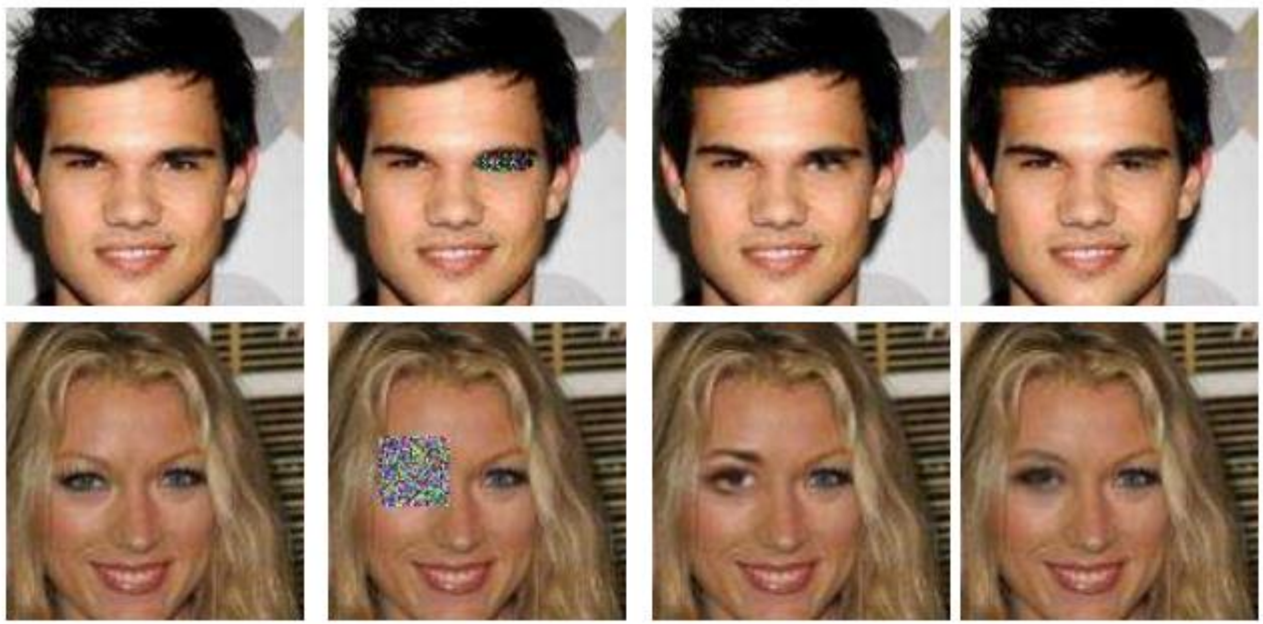
$$L = l_r + \lambda_1 l_{a_1} + \lambda_2 l_{a_2} + \lambda_3 l_p$$

Generative Face Completion

(a) Original image (b) Masked input (c) M1 (d) M2 (e) M3 (f) M4 (g) M5

Figure 3. Completion results under different settings of our model. (c) M1: L_r. (d) M2: $L_r + L_{a_1}$. (e) M3: $L_r + L_{a_1} + L_{a_2}$. (f) M4: $L_r + L_{a_1} + L_{a_2} + L_p$. The result in (f) shows the most realistic and plausible completed content. It can be further improved through post-processing techniques such as (g) M5: M4 + Poisson blending [18] to eliminate subtle color difference along mask boundaries.

Generative Face Completion



(a) original (b) masked input (c) w/o parsing (d) w/ parsing

Q&A