Generative Adversarial Networks (GANs)

Objectives

Objective

Describe basic concepts and architecture for GANs

Objective

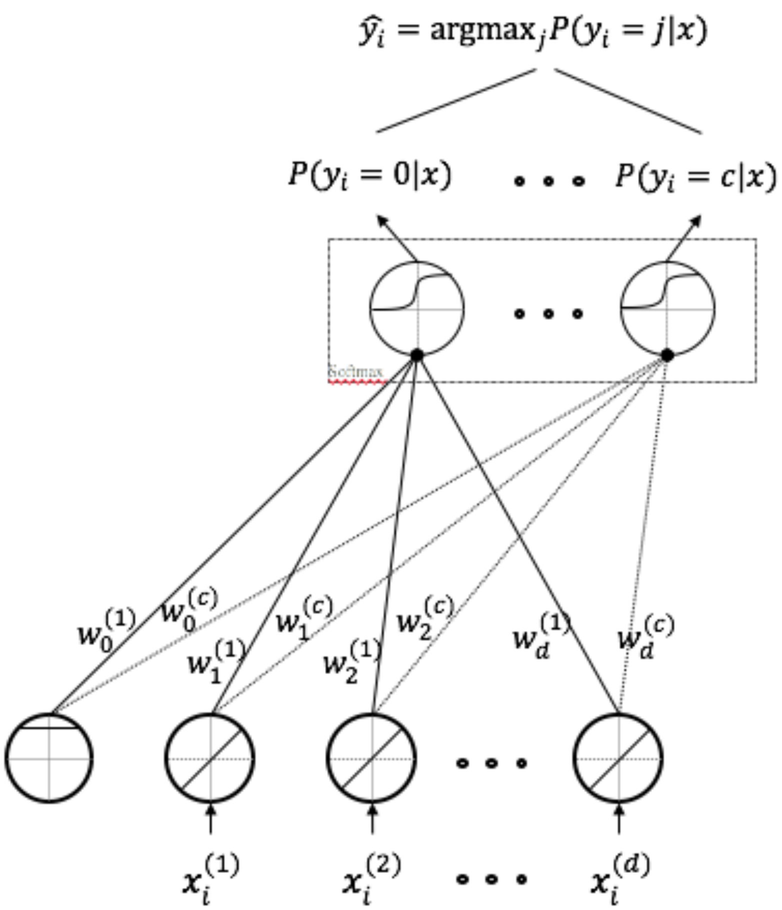
Illustrate variants of GANs and their applications

Generative Adversarial Networks (GANs)

- Proposed in 2014 by Goodfellow *et al.*
- An architecture with two neural networks gaming against each other.
  - One attempting to learn a *generative model*
- Many variants have been proposed since the initial model.

Discriminative vs Generative Models - 1/4

- Discriminative models: E.g., the familiar MLP
  - Given $\{(x_i, y_i)\}$, to learn $P(y_i | x)$



- More generally, we try to learn a *posterior distribution* of y given x, $p(y|x)$
  - Usually reduced to posterior probabilities for classification problems

➔ See also earlier discussion on Naïve Bayes vs Logistic Regression.

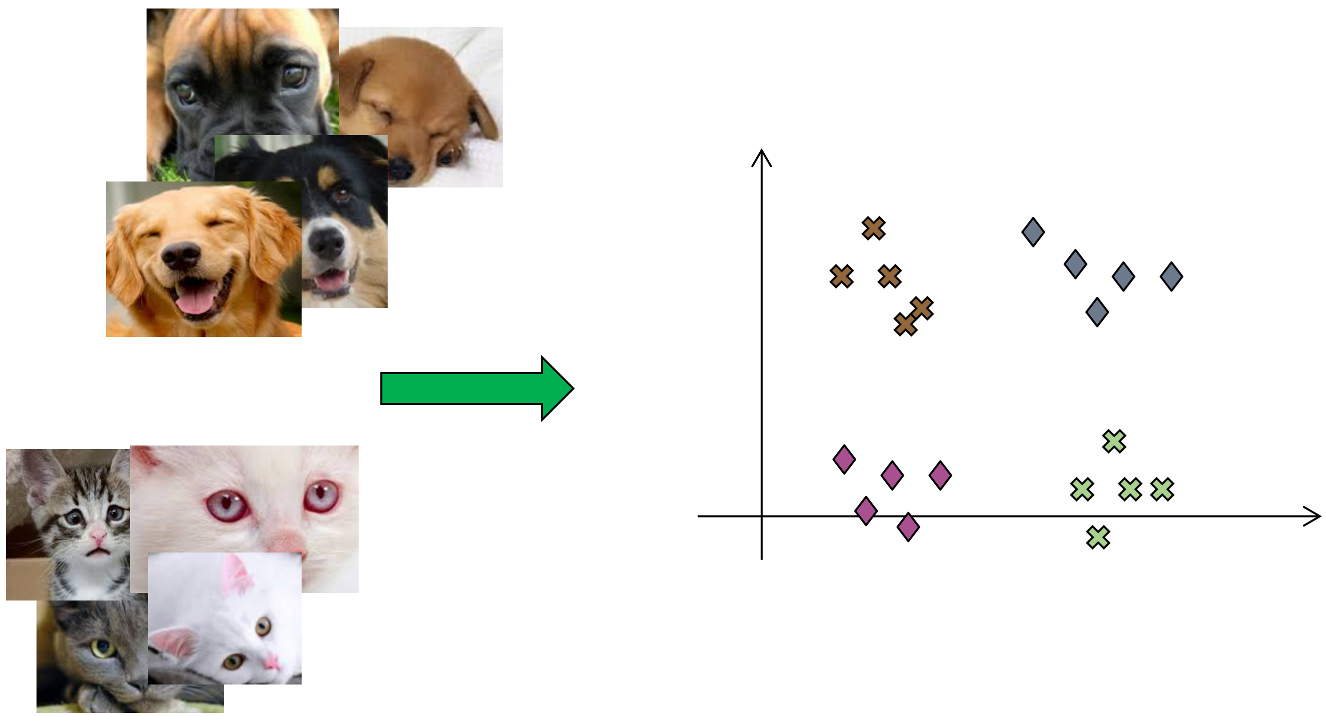
Discriminative vs Generative Models - 2/4

- Generative models think the other direction: how to generate x given y
  - E.g., x_i=? if y_i = 2?

- More generally, we try to learn a *conditional distribution* of x given y, $p(x|y)$

Discriminative vs Generative Models - 3/4

Illustrating the ideas

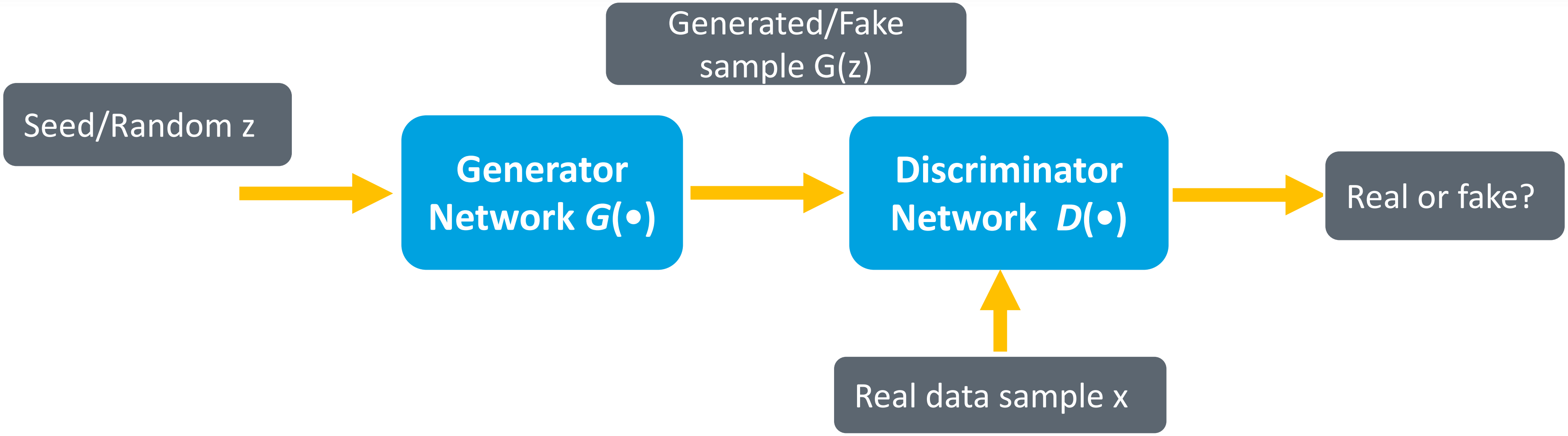


Discriminative vs Generative Models - 4/4

- Estimating $p(x|y)$ (or, in general any $p(x)$, if we drop y by assuming it is given)
  - Explicit density estimation: assuming some parametric or non-parametric models.
  - Implicit density estimation: learn (essentially equivalent) models that may create good samples (as if from the "true" model), without explicitly defining the true model.

→ **GAN is such an approach**

Basic GAN Architecture



- Objective of the Discriminator Network:

  making $D(x) \rightarrow 1$, $D(G(z)) \rightarrow 0$

- Objective of the Generator Network:

  making $D(G(z)) \rightarrow 1$

Basic GAN Training Algorithm

for number of training iterations **do**

  for k steps **do**

    Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise distribution $p_g(\boldsymbol{z})$

    Sample minibatch of m examples $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}\}$ from data distribution $p_{data}(\boldsymbol{x})$

    Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)\right]$$

  end for

    Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise distribution $p_g(\boldsymbol{z})$

    Update the generator by descending its stochastic gradient

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)$$

end for

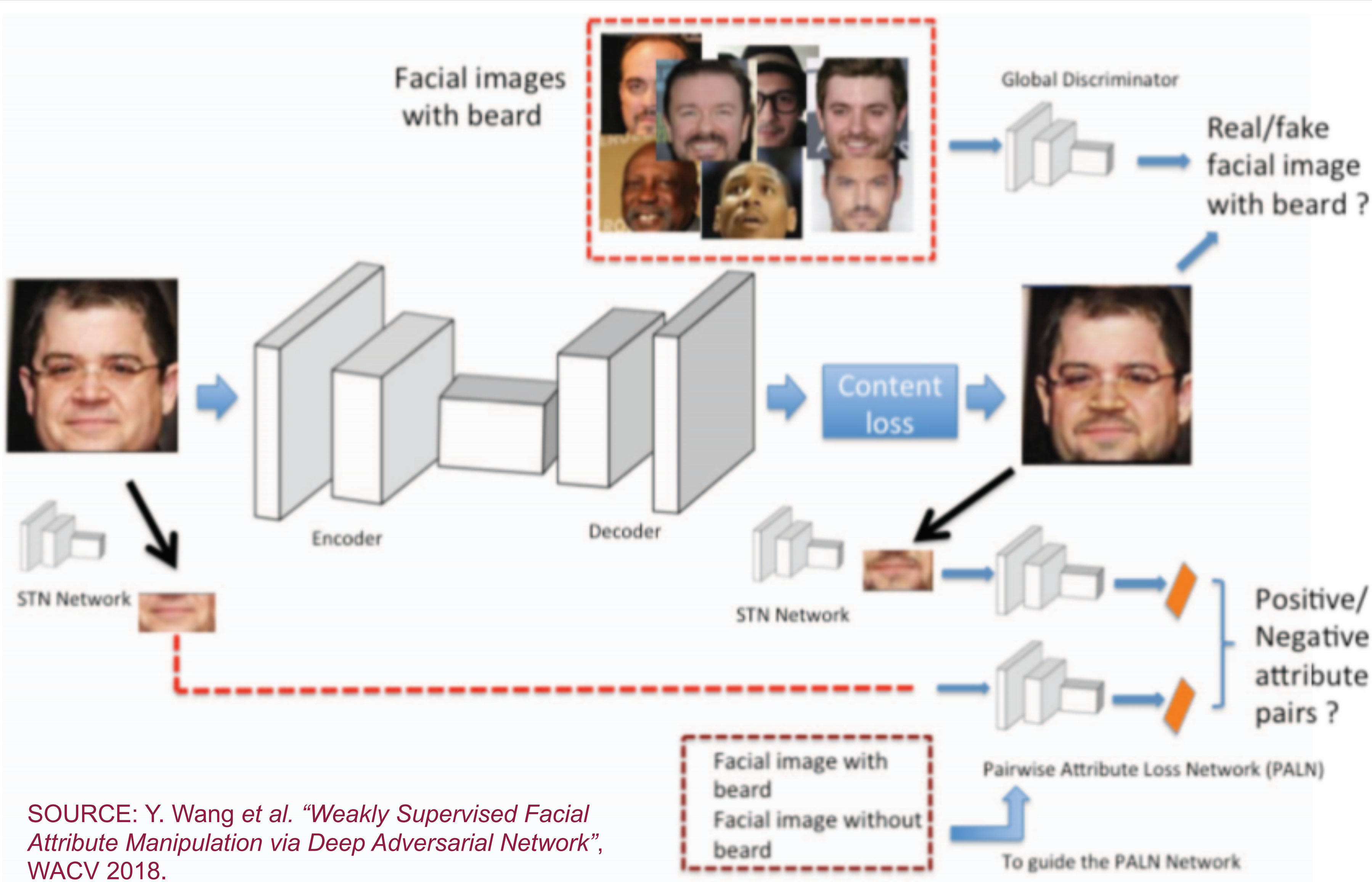
θ_d and θ_g are the parameters of the discriminator and generator respectively.

SOURCE: Goodfellow *et al.* https://arxiv.org/pdf/1406.2661.pdf

Applications of GAN

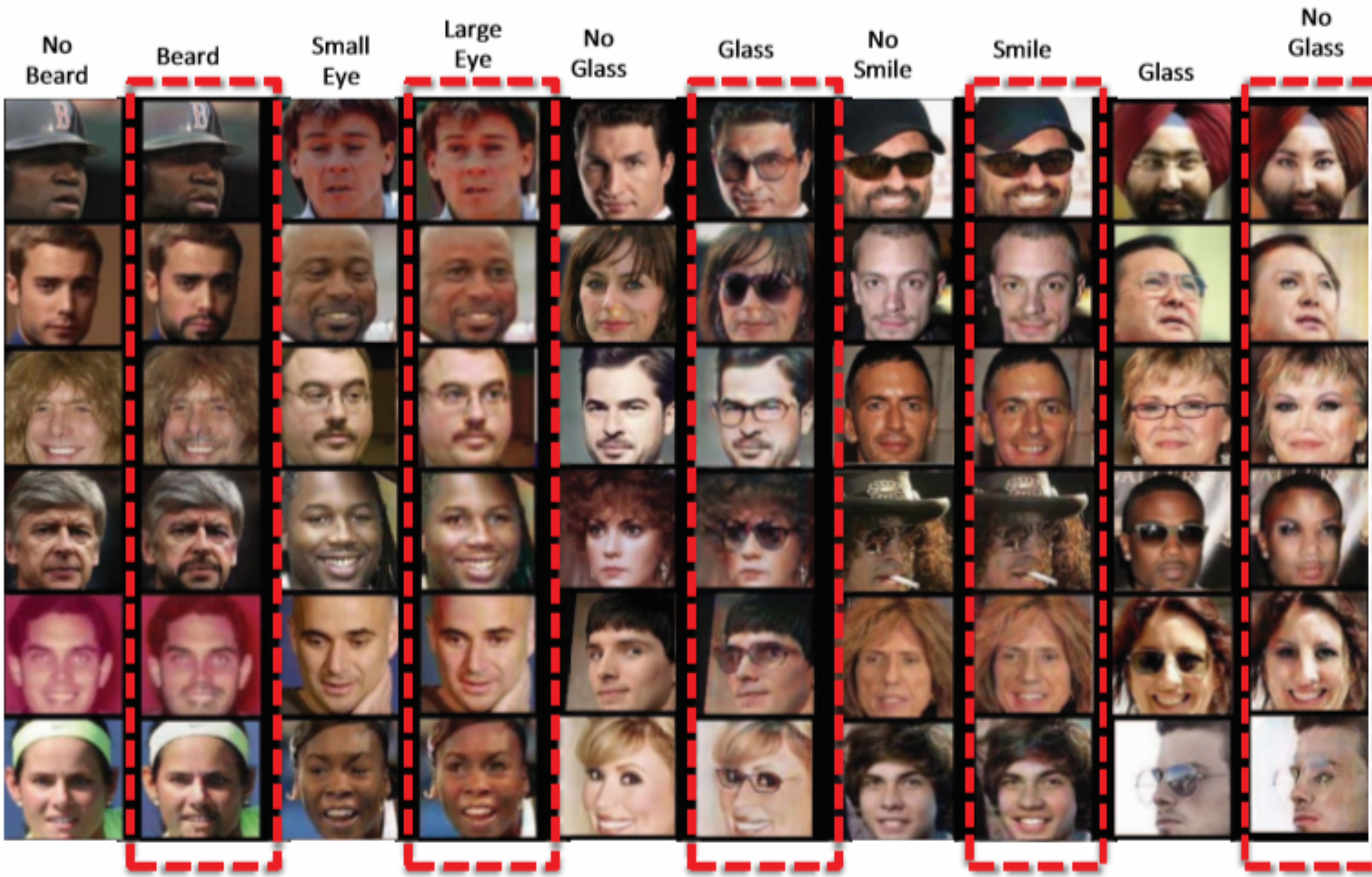
- GAN enabled many novel/interesting/fun applications.
- Many GAN-based models have been proposed, following the initial paper.
- Consider one example: Facial attribute manipulation
  - Y. Wang *et al.* *"Weakly Supervised Facial Attribute Manipulation via Deep Adversarial Network"*, WACV 2018.

Facial Attribute Manipulation - 1/2



SOURCE: Y. Wang *et al. "Weakly Supervised Facial Attribute Manipulation via Deep Adversarial Network"*, WACV 2018.

Facial Attribute Manipulation - 2/2



SOURCE: Y. Wang *et al.* *"Weakly Supervised Facial Attribute Manipulation via Deep Adversarial Network"*, WACV 2018.