

# Evaluation of Generative Models

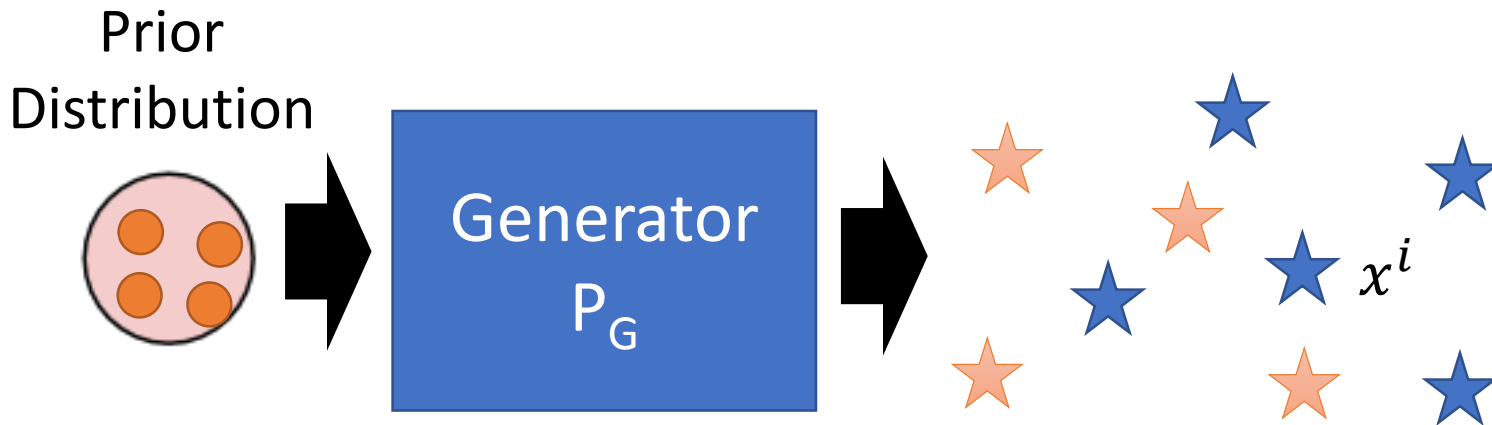
# Reference

- Lucas Theis, Aäron van den Oord, Matthias Bethge, “A note on the evaluation of generative models”, arXiv preprint, 2015

# Likelihood

★ : real data (not observed during training)

★ : generated data



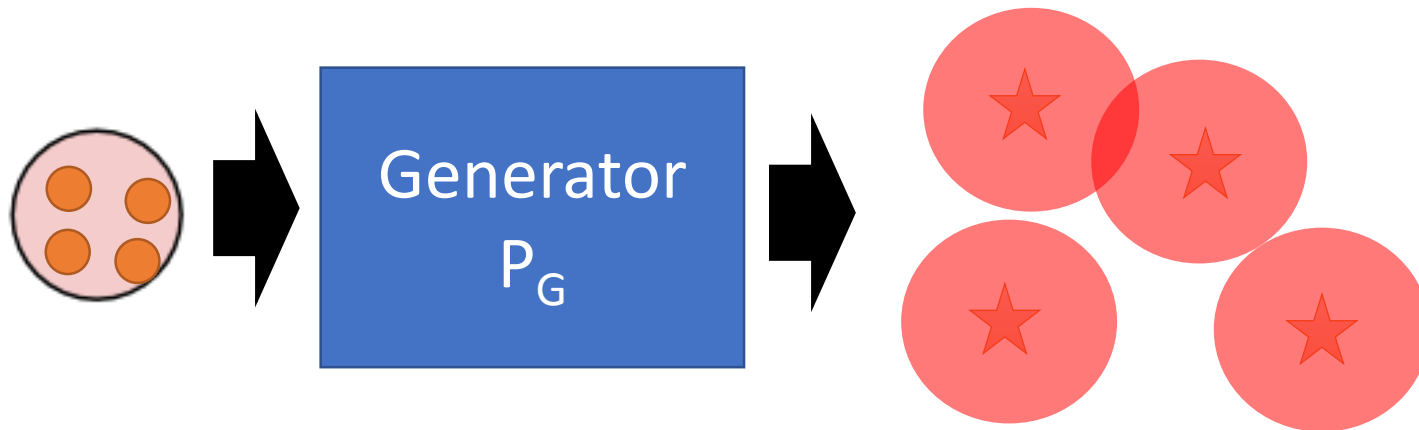
$$\text{Log Likelihood: } L = \frac{1}{N} \sum_i \log P_G(x^i)$$

We cannot compute  $P_G(x^i)$ . We can only sample from  $P_G$ .

# Likelihood

## - Kernel Density Estimation

- Estimate the distribution of  $P_G(x)$  from sampling



Each sample is the mean of a Gaussian with the same covariance.

Now we have an approximation of  $P_G$ , so we can compute  $P_G(x^i)$  for each real data  $x^i$

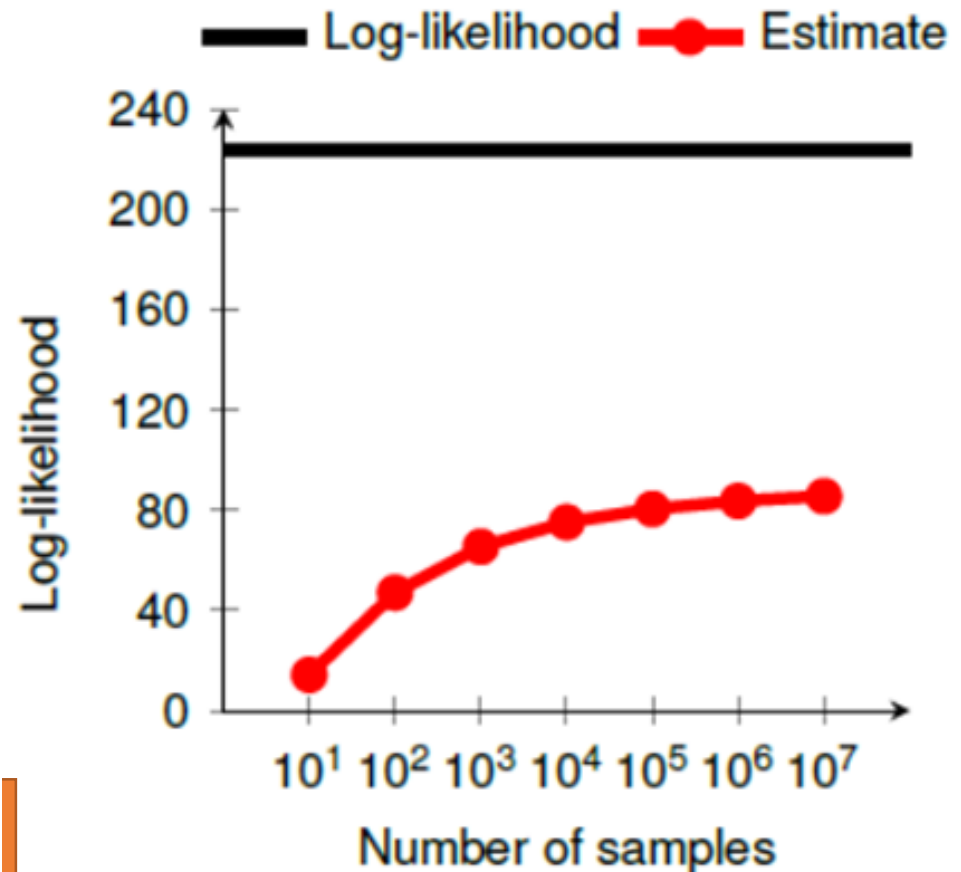
Then we can compute the likelihood.

# Likelihood

## - Kernel Density Estimation

- How many samples?
- Weird results?

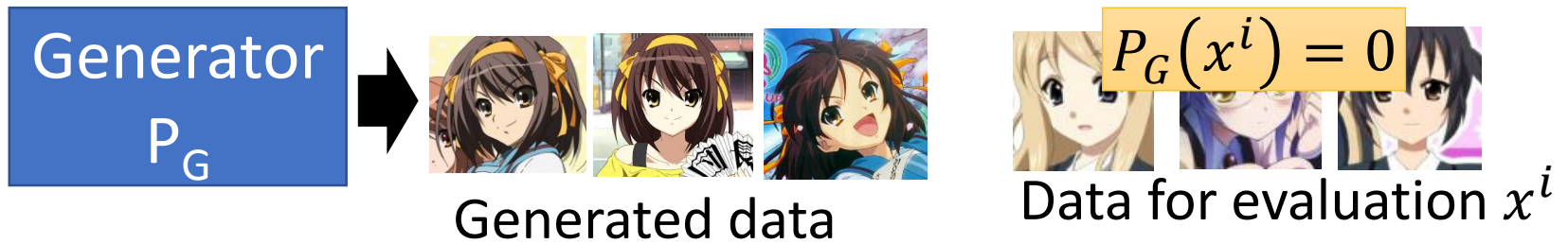
Model	Likelihood
DBN	138
GAN	225



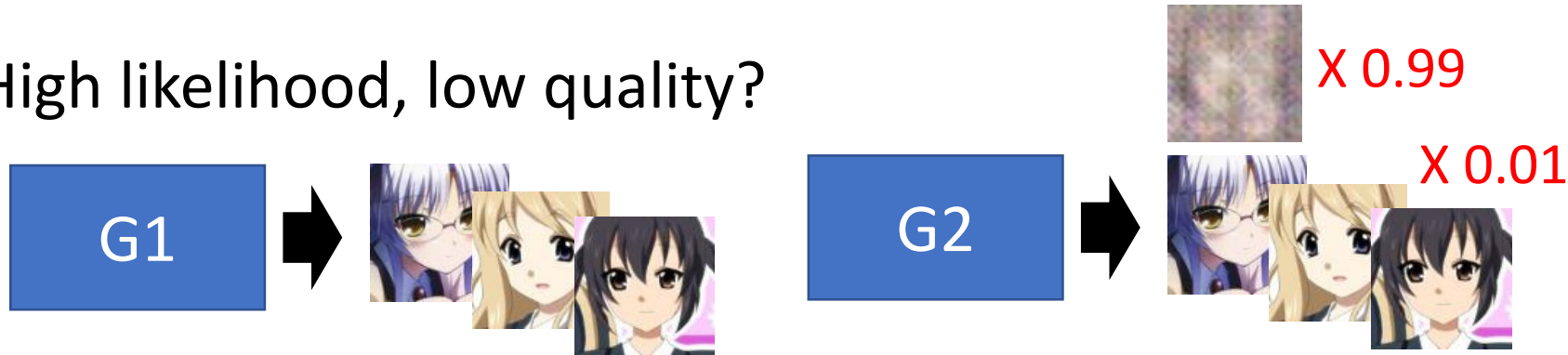
# Likelihood v.s. Quality

- Low likelihood, high quality?

Considering a model generating good images (small variance)



- High likelihood, low quality?



$$L = \frac{1}{N} \sum_i \log \frac{P_G(x^i)}{100} = -\log 100 + \frac{1}{N} \sum_i \log P_G(x^i)$$

$4.6$

# Evaluate by Other Networks



Lower entropy means  
higher visual quality



⋮

$$\frac{1}{N} \sum_n P(y^n|x^n)$$

Higher entropy means  
higher variety

# Evaluate by Other Networks

## - Inception Score

Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, "Improved Techniques for Training GANs", arXiv preprint, 2016

$$\exp(\mathbb{E}_x \text{KL}(p(y|x) || p(\bar{y})))$$

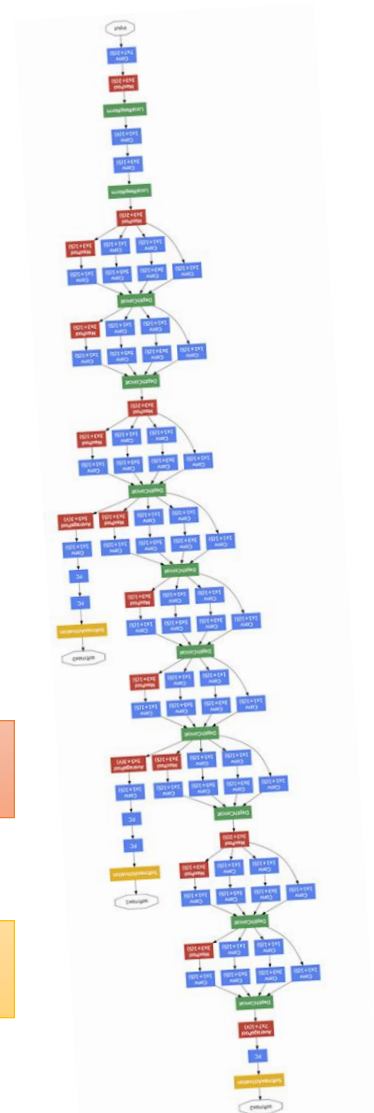
$$= \sum_x \sum_y P(y|x) \log \frac{P(y|x)}{P(y)}$$

$$= \sum_x \sum_y P(y|x) \log P(y|x)$$

Negative entropy

$$- \sum_x \sum_y P(y|x) \log P(y)$$

Cross entropy

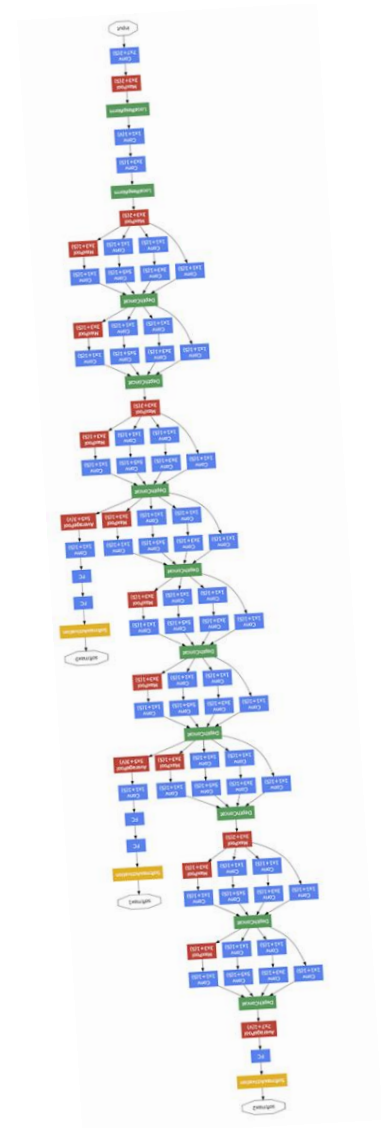
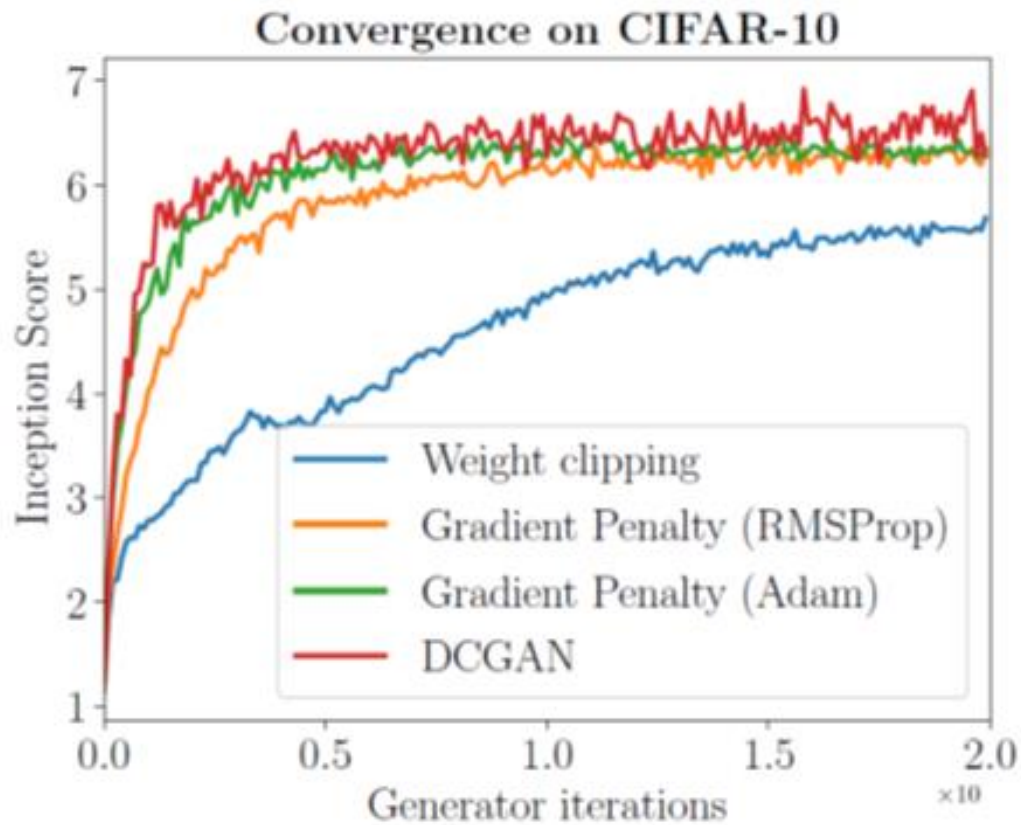




# Evaluate by Other Networks

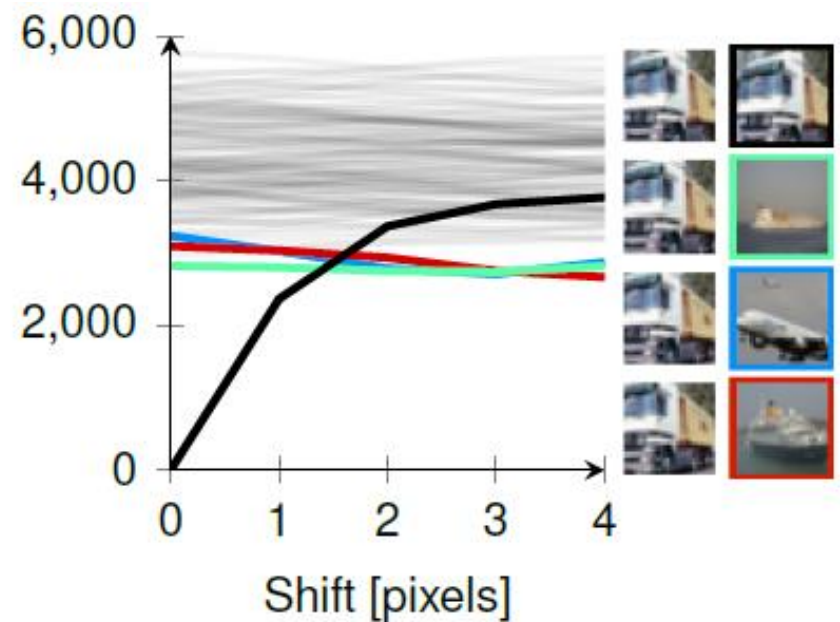
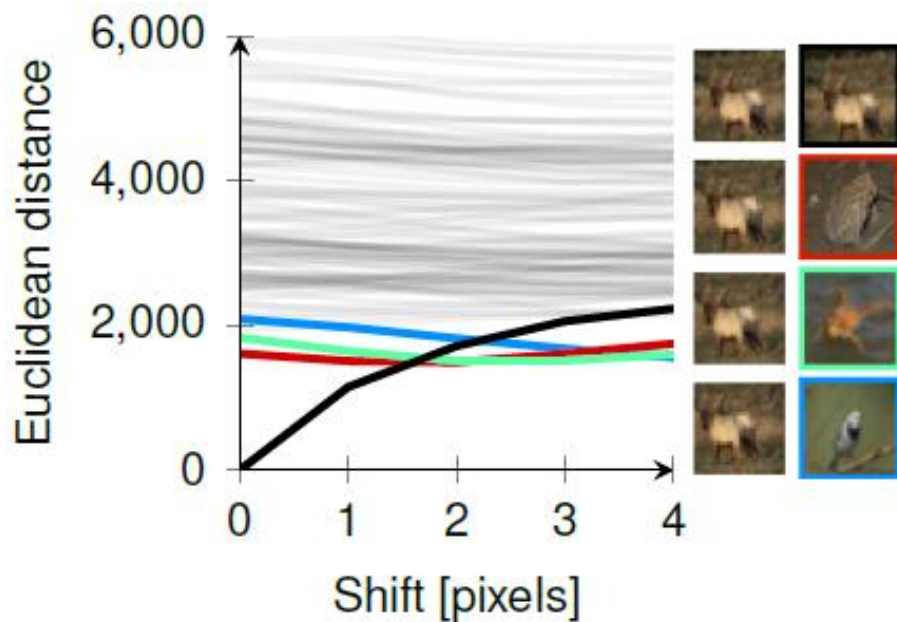
## - Inception Score

- Improved W-GAN



# K-Nearest Neighbor

- Using k-nearest neighbor to check whether the generator generates new objects

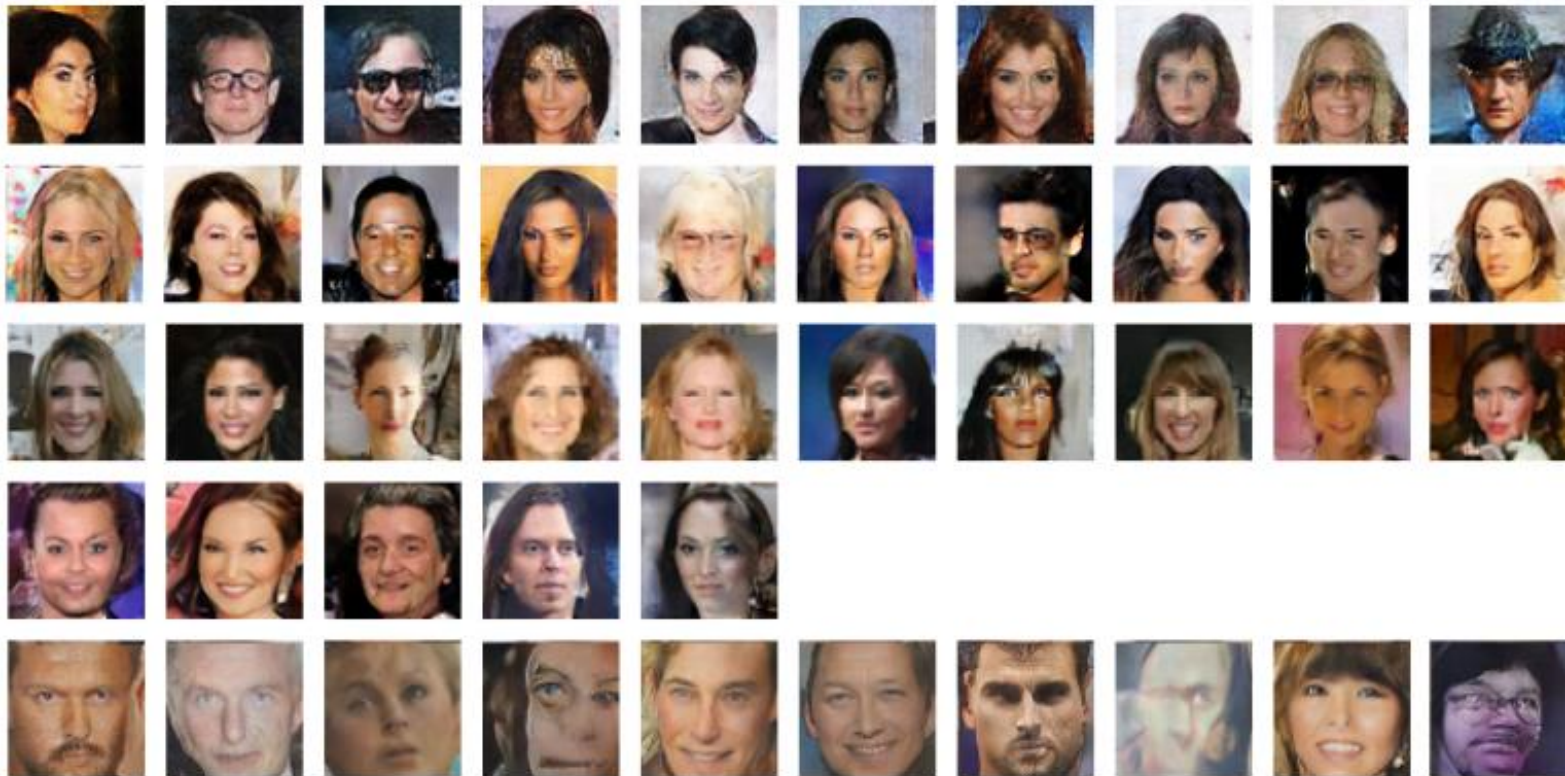


# Mode Collapse



<https://devblogs.nvidia.com/parallelforall/photo-editing-generative-adversarial-networks-2/>

# Missing Mode



Missing anything?



Discriminator always knows it is real with high confidence