



Introduction to Generative Adversarial Networks

IOTG / VMC / ICV / Algo R&D

Vladislav Sovrasov

2020-07-21

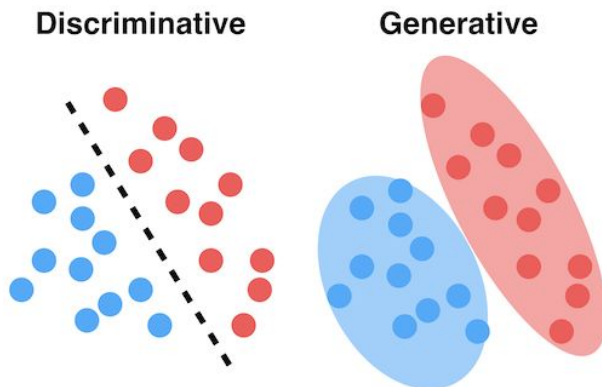
Internet of Things Group

Agenda

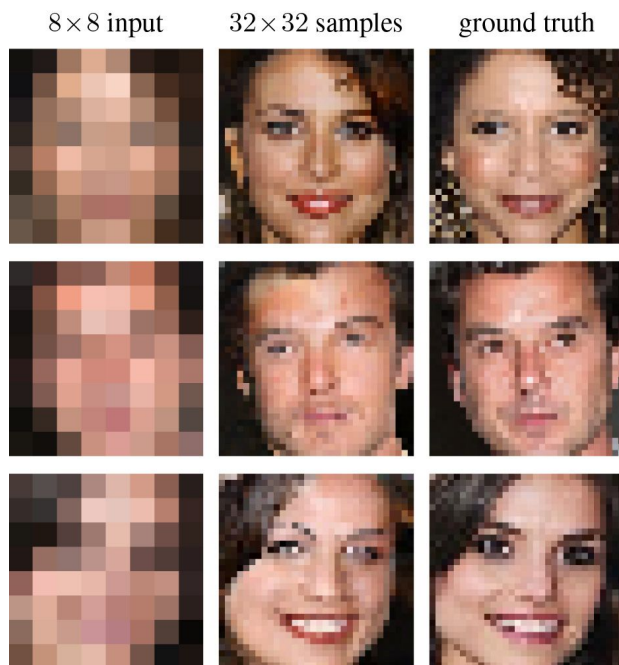
- What is a generative model?
- When do we need to generate new data?
- Pioneering works
- Overview of the GAN learning framework
- How to evaluate GANs
- Problems and limitations of the approach
- Examples of practical application

Discriminative and generative learning

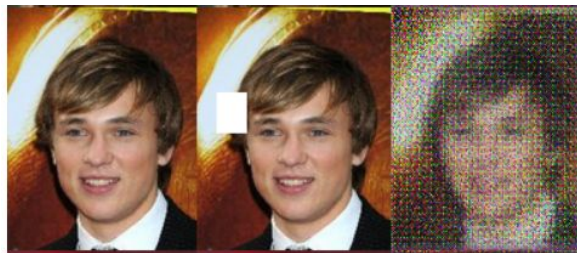
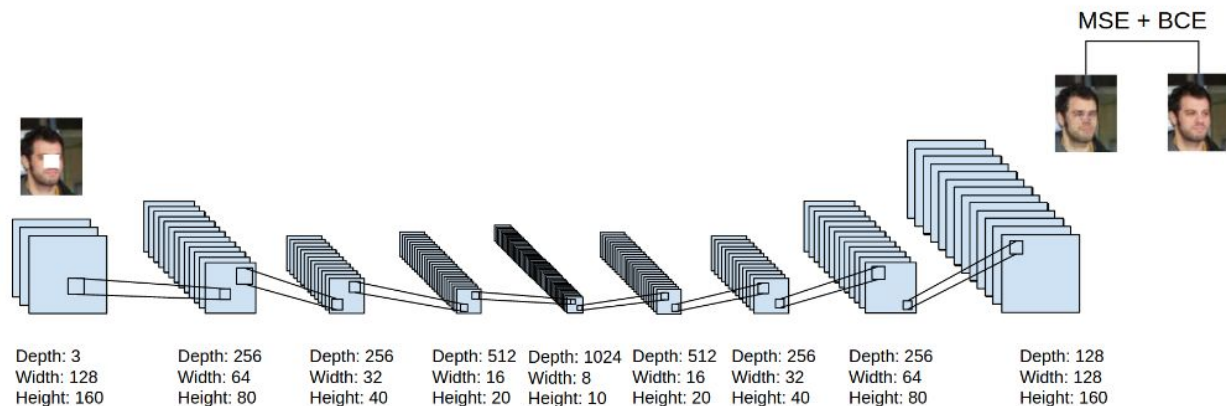
- Discriminative model tries to establish a boundary between classes. It learns a conditional distribution $p(x|y)$.
- Generative models can generate new data instances. They learn the joint distribution $p(x,y)$ or $p(x)$ if there are no labels.



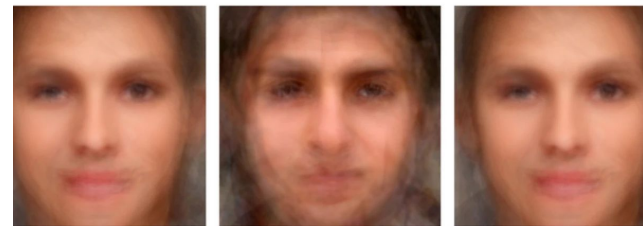
When do we need to generate new data?



Why not L2 reconstruction?



THE AVERAGE FACE



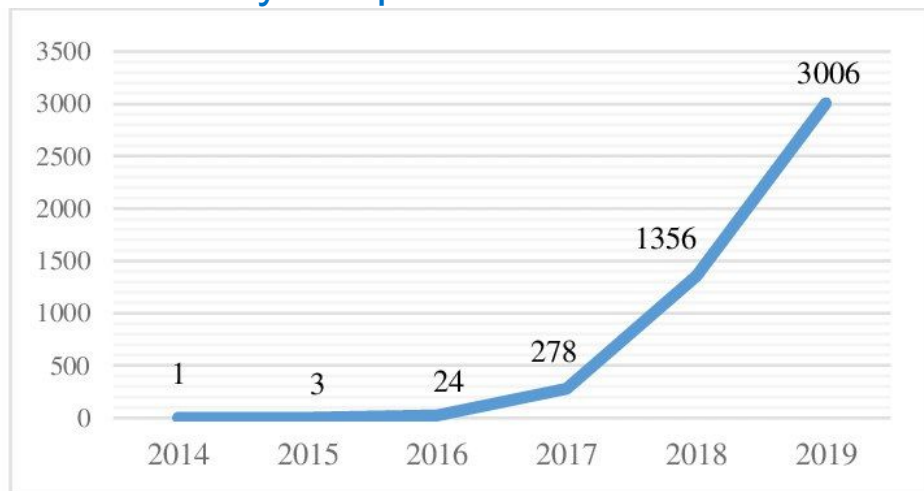
Pioneering works

- Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio - Generative Adversarial Networks (2014)
- Mehdi Mirza, Simon Osindero - Conditional Generative Adversarial Nets (2014)
- Alec Radford, Luke Metz, Soumith Chintala - Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (2015)

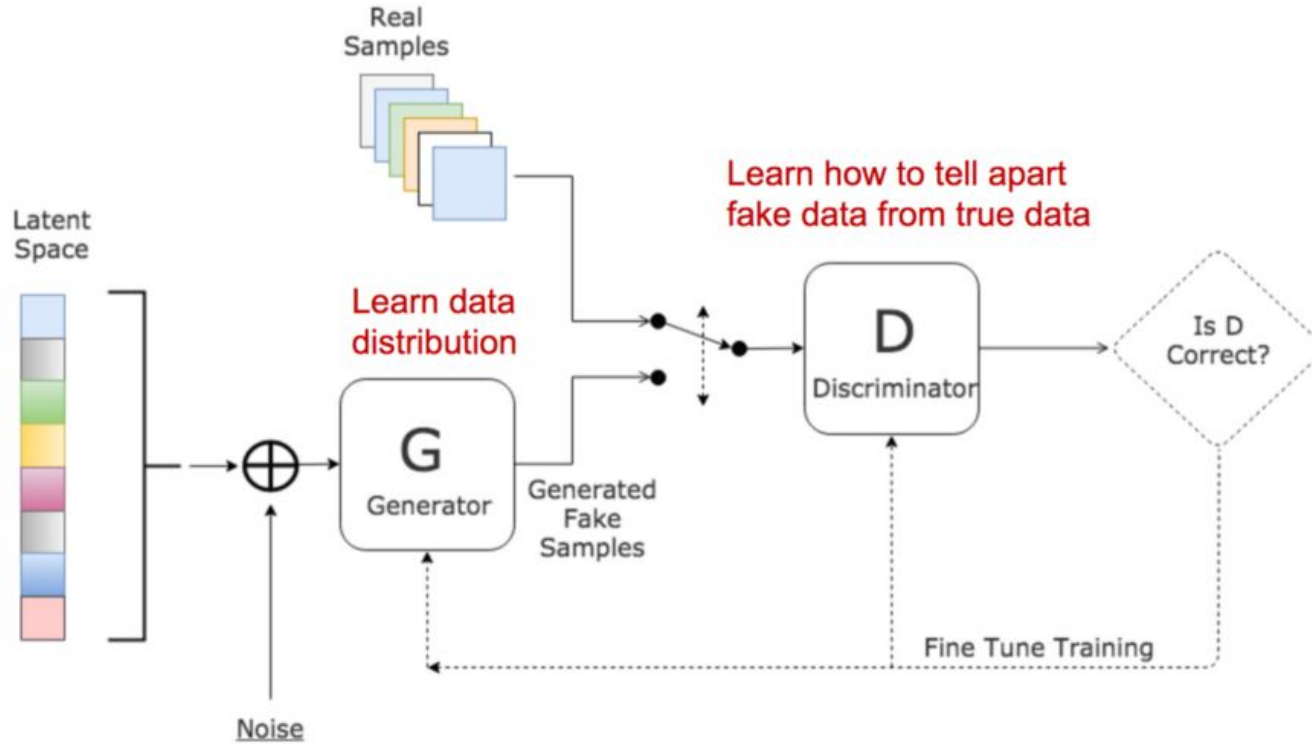
Dynamics of research about GANs

“Generative Adversarial Networks is the most interesting idea in the last 10 years in Machine Learning.” — Yann LeCun, Chief AI scientist at Facebook

Number of articles indexed by Scopus on GANs from 2014 to 2019:



Overview of the GAN learning framework



Adversarial learning technique

x - real data

z - latent variable

D - discriminator

G - generatr

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

How to evaluate GANs?



How to evaluate GANs?

- Manual Image Inspection
- Birthday paradox test
- Nearest Neighbors (find the closest image from a set of real images)
- Inception Score

Collect output of the InceptionV3 model on fake images: $p(y|x)$.

Compute the marginal distribution of labels $p(y)$.

Compute the average KL-divergence between $p(y|x)$ and $p(y)$.

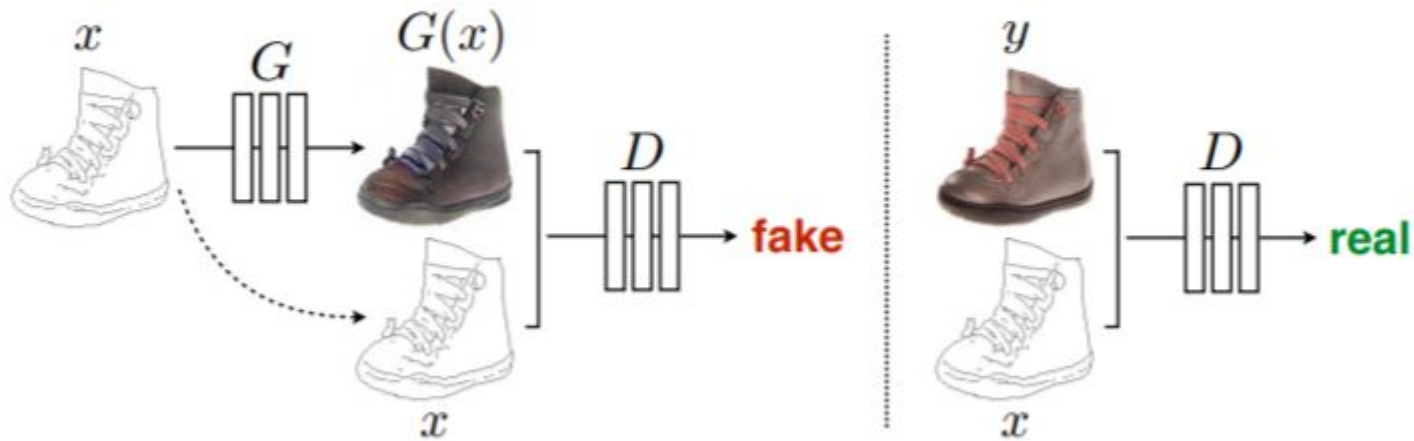
$$D_{KL}(p(y|x)||p(y)) = \sum_{y=1}^C p(y|x) \log \left(\frac{p(y|x)}{p(y)} \right)$$
$$IS = \exp [\mathbb{E}_x D_{KL}(p(y|x)||p(y))]$$

Problems and limitations of GANs

- Optimization process is not always stable and hard to tune
- Mode collapse
- Inconsistent background
- High-frequency artifacts
- Restricted capacity of the generator and discriminator
- Inconsistency between the learned and actual distribution

Pix2Pix

Conditional GAN trained on image pairs.



Note: the generator has no additional random input

Pix2Pix

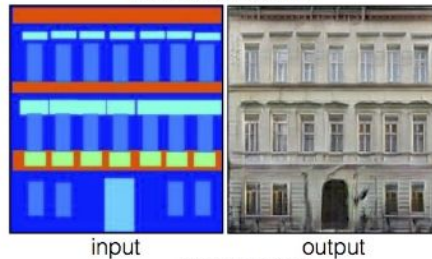
Labels to Street Scene



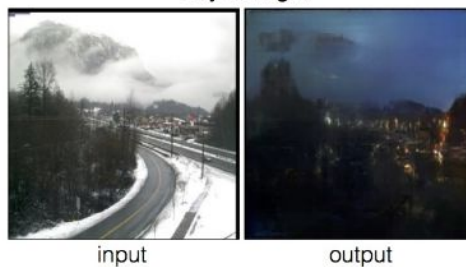
Aerial to Map



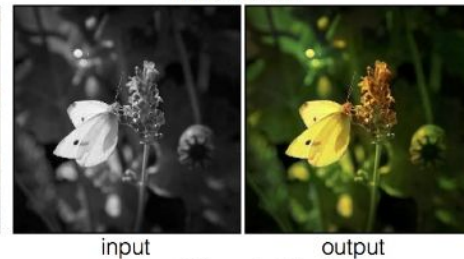
Labels to Facade



Day to Night



BW to Color

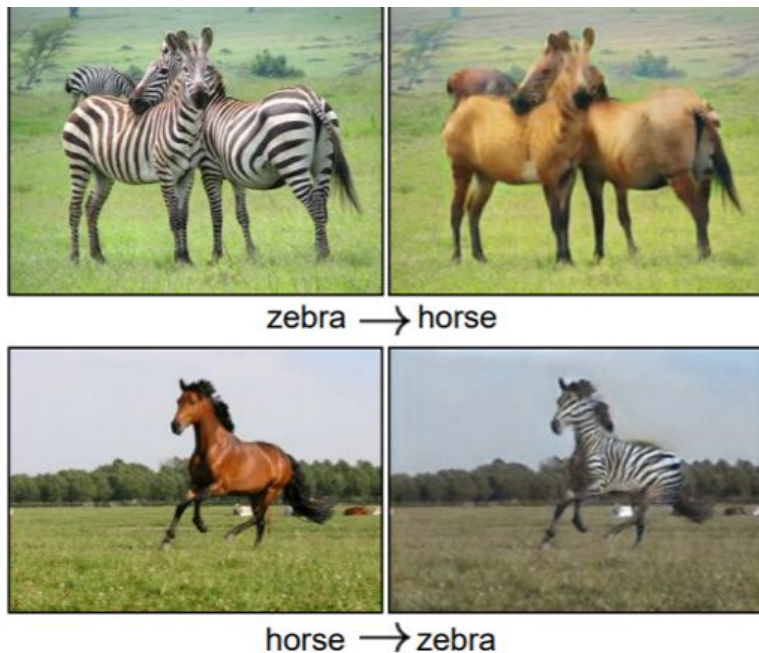


Edges to Photo



Cycle GAN

Problem formulation: let we need to learn to transfer images from one domain to another without implicit pairwise annotation



Cycle GAN

Input



Monet



Van Gogh



Cezanne



Ukiyo-e



Input



Output



Input



Output



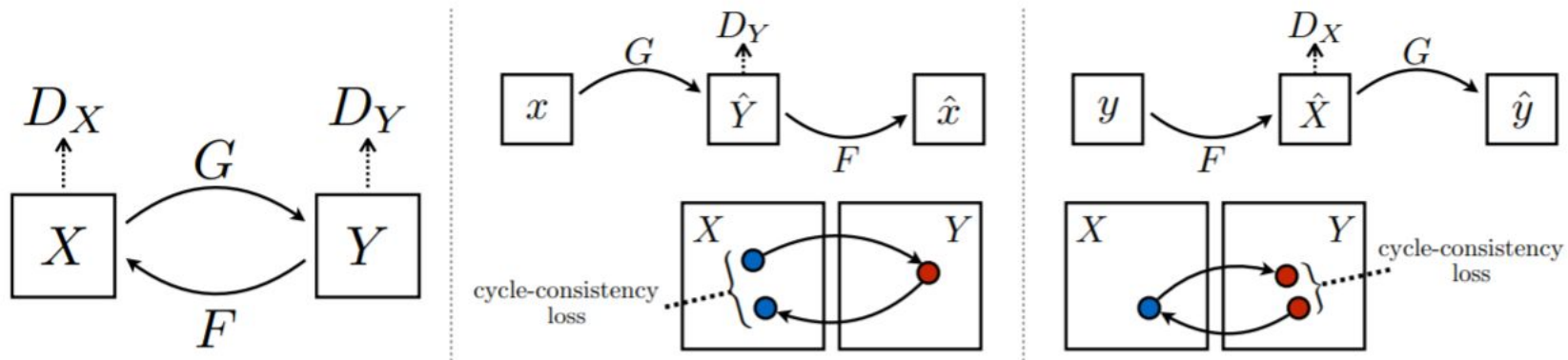
Input



Output



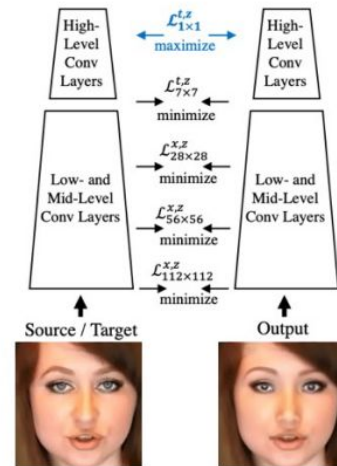
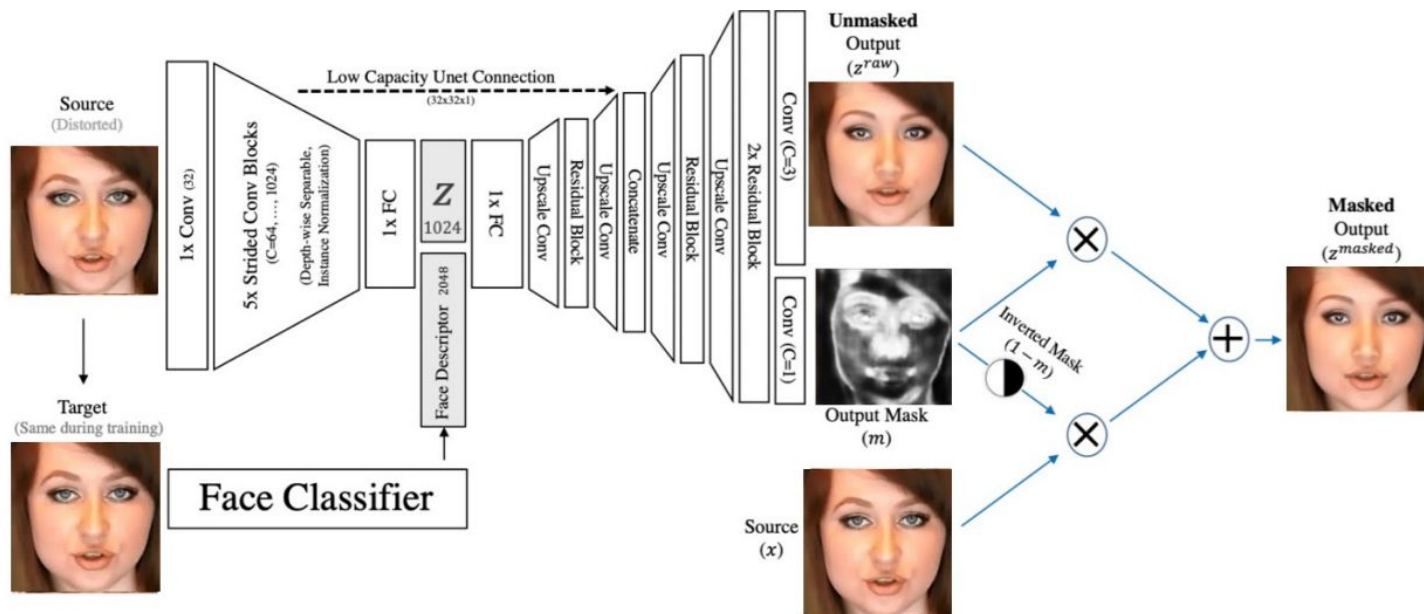
Cycle GAN



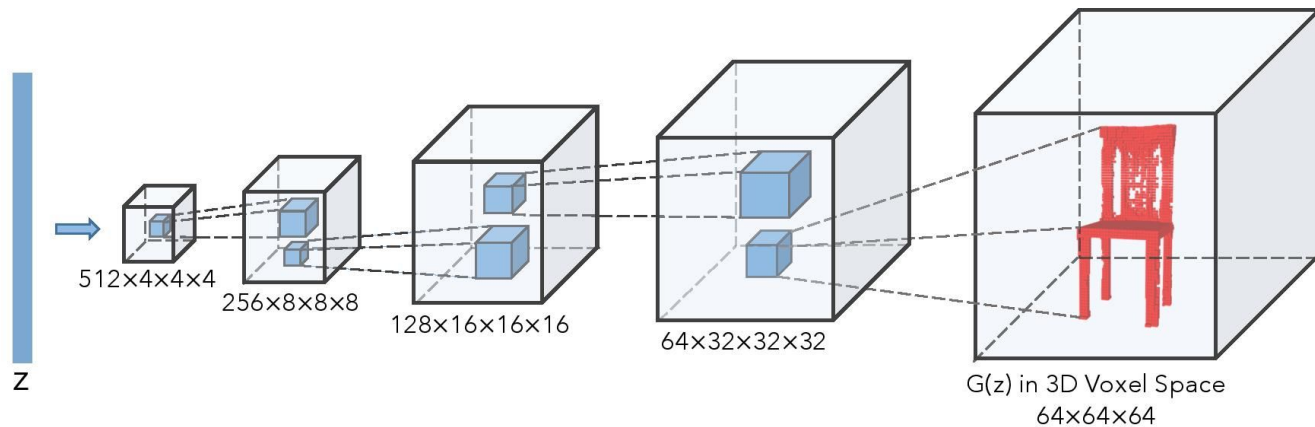
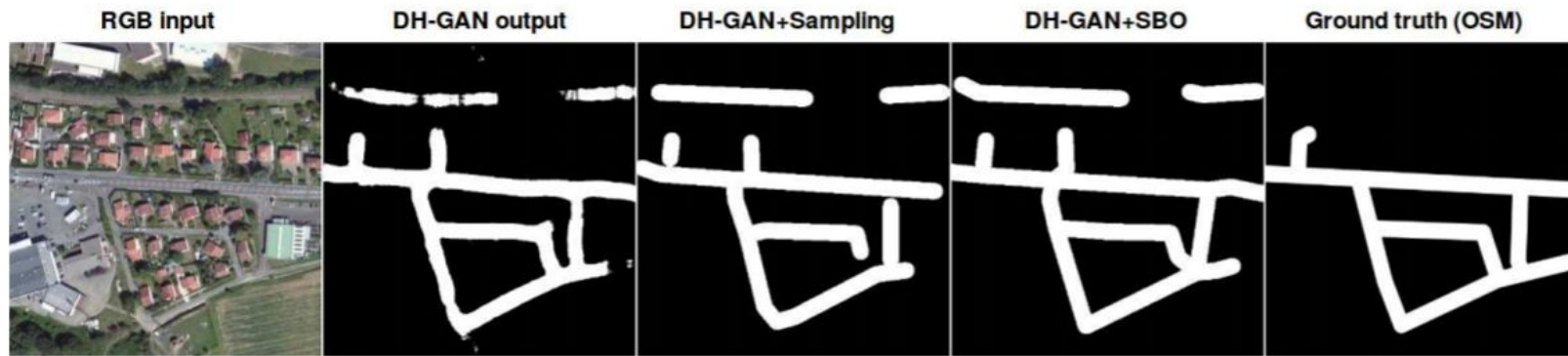
Cycle GAN



Live Face De-Identification in Video



Other examples of application



Practice

https://github.com/dkurt/opencvino_practice/tree/master/modules/9_gan

