Daniele Perlo – Dipartimento di Informatica UniTo

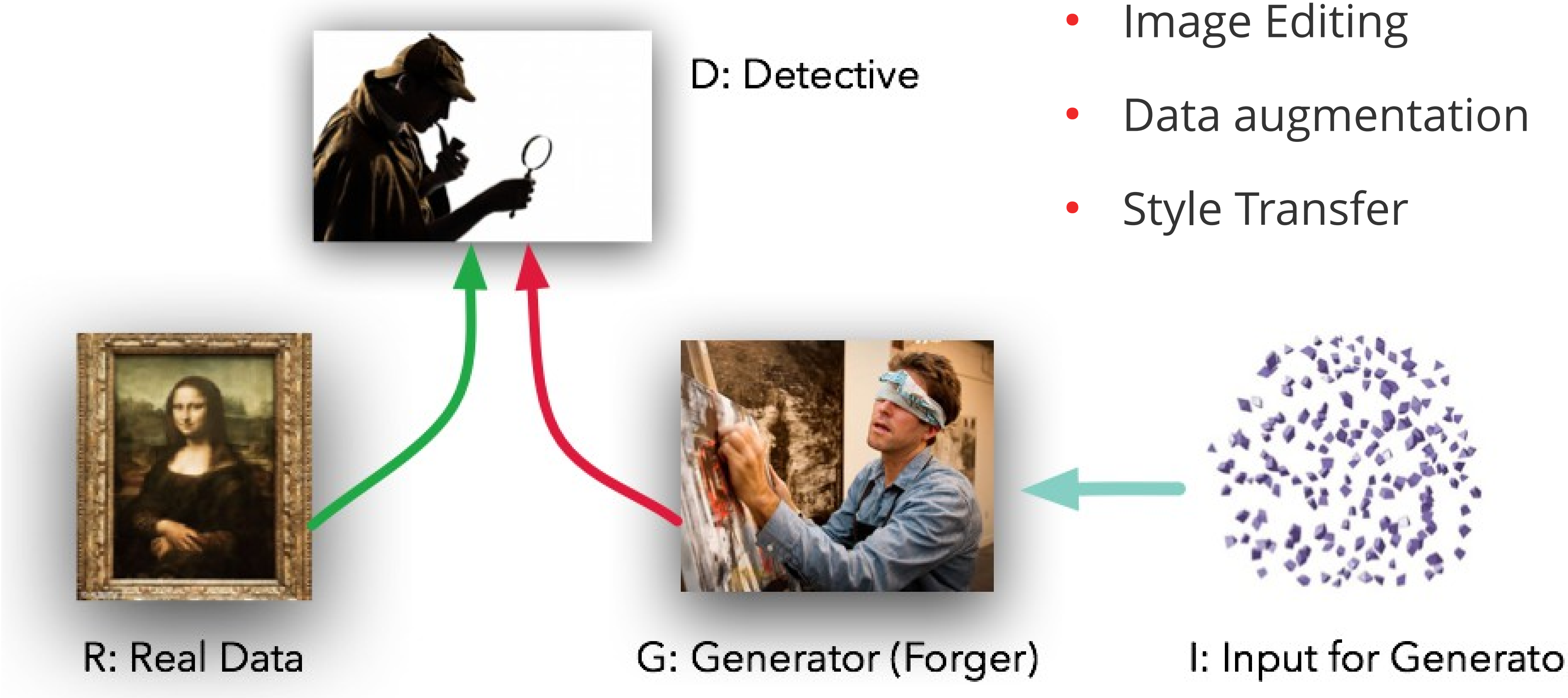
Generative Adversarial Networks: An Overview

Ref] http://ieeexplore.ieee.org/document/8253599/
Generative Adversarial Networks: An Overview

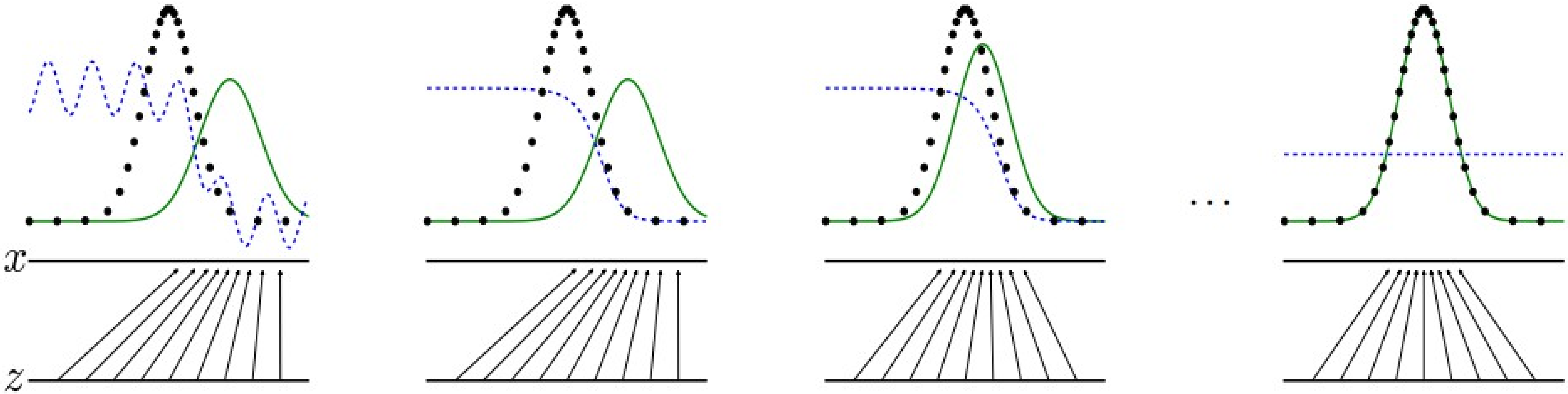
GAN Framework

Creazione di data samples sintetici:

- Image Editing
- Data augmentation
- Style Transfer



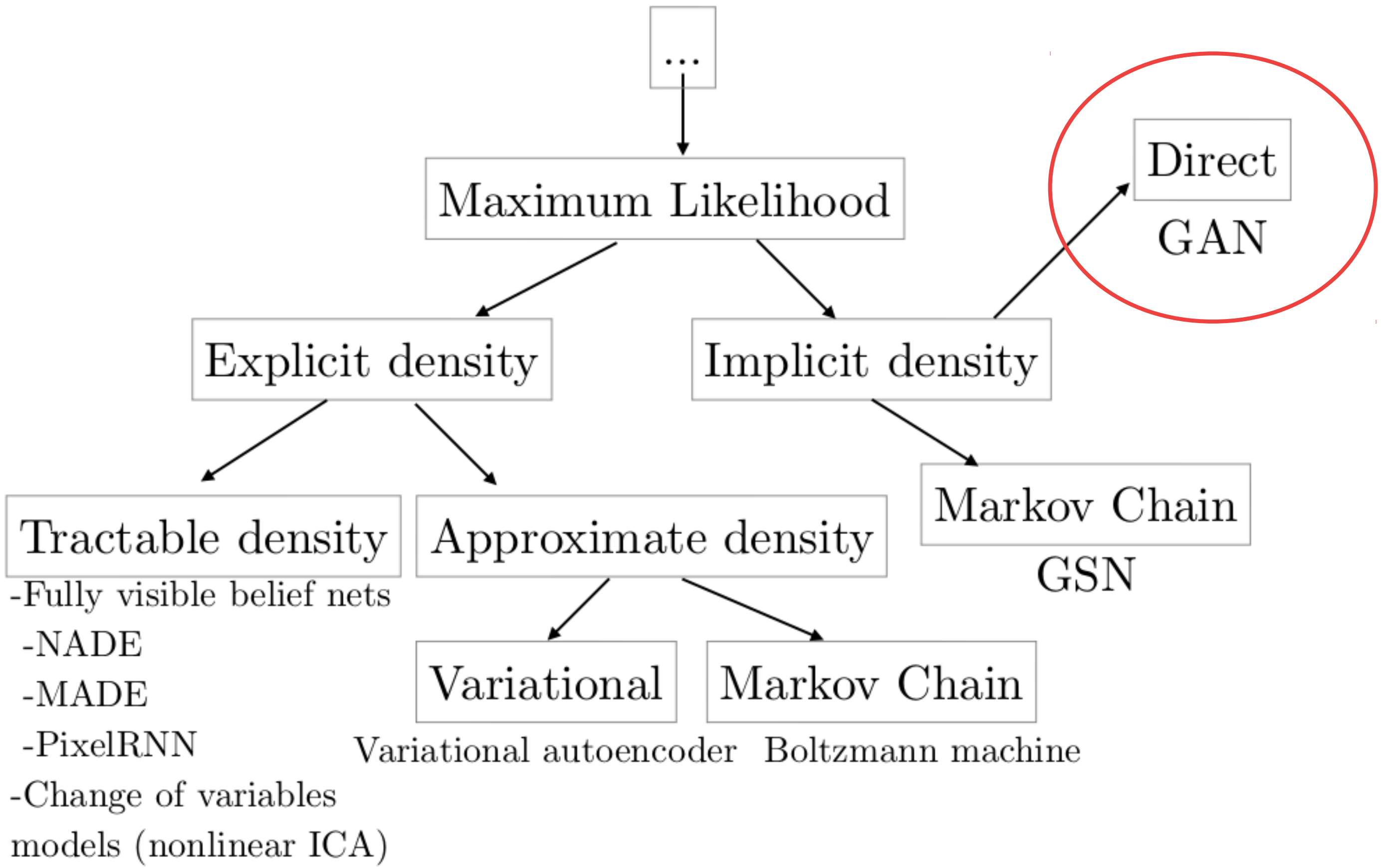
GAN Framework



$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \prod_{i=1}^{m} p_{\text{model}}\left(\boldsymbol{x}^{(i)}; \boldsymbol{\theta}\right)$$

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} D_{\text{KL}}\left(p_{\text{data}}(\boldsymbol{x}) \| p_{\text{model}}(\boldsymbol{x}; \boldsymbol{\theta})\right).$$

Generative Models Family



Ref] https://arxiv.org/abs/1701.00160
NIPS 2016 Tutorial: Generative Adversarial Networks, Goodfellow

Comparing GAN

- “They can generate samples in parallel, instead of using runtime proportional to the dimensionality of x. This is an advantage relative to FVBNs.

- The design of the generator function has very few restrictions. This is an advantage relative to Boltzmann machines, for which few probability distributions admit tractable Markov chain sampling, and relative to non-linear ICA, for which the generator must be invertible and the latent code z must have the same dimension as the samples x.

- No Markov chains are needed. This is an advantage relative to Boltzmann machines and GSNs.

- No variational bound is needed, and specific model families usable within the GAN framework are already known to be universal approximators, so GANs are already known to be asymptotically consistent. Some VAEs are conjectured to be asymptotically consistent, but this is not yet proven.

- GANs are subjectively regarded as producing better samples than other methods.”

Ref] https://arxiv.org/abs/1701.00160
NIPS 2016 Tutorial: Generative Adversarial Networks, Goodfellow

How do GANs Work

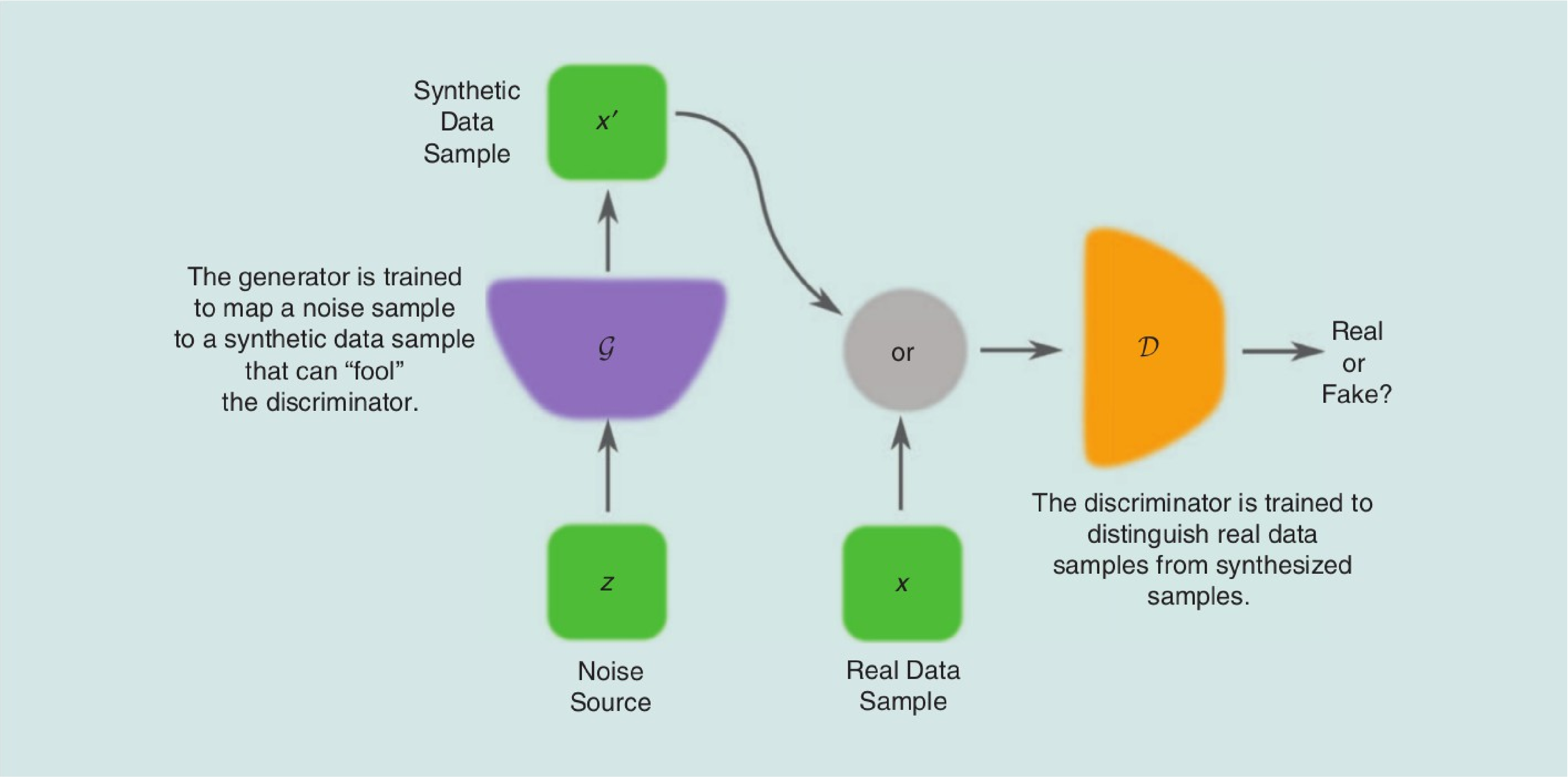


FIGURE 1. The two models that are learned during the training process for a GAN are the discriminator ($\mathcal{D}$) and the generator ($\mathcal{G}$). These are typically implemented with neural networks, but they could be implemented by any form of differentiable system that maps data from one space to another; see article text for details.

How do GANs Work

- Cost Functions

D: Cross Entropy

$$J^{(D)}(\boldsymbol{\theta}^{(D)}, \boldsymbol{\theta}^{(G)}) = -\frac{1}{2}\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}} \log D(\boldsymbol{x}) - \frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log\left(1 - D\left(G(z)\right)\right)$$

$$J^{(G)} = -J^{(D)}$$

G: Minimax -

$$J^{(G)} = -\frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log D(G(\boldsymbol{z}))$$

Non-Saturating H

$$J^{(G)} = -\frac{1}{2}\mathbb{E}_{z} \exp\left(\sigma^{-1}\left(D(G(\boldsymbol{z}))\right)\right)$$

MaxLikehood -

GANs Architectures

- **Fully Connected GANs**

  D & G composte da soli livelli FC

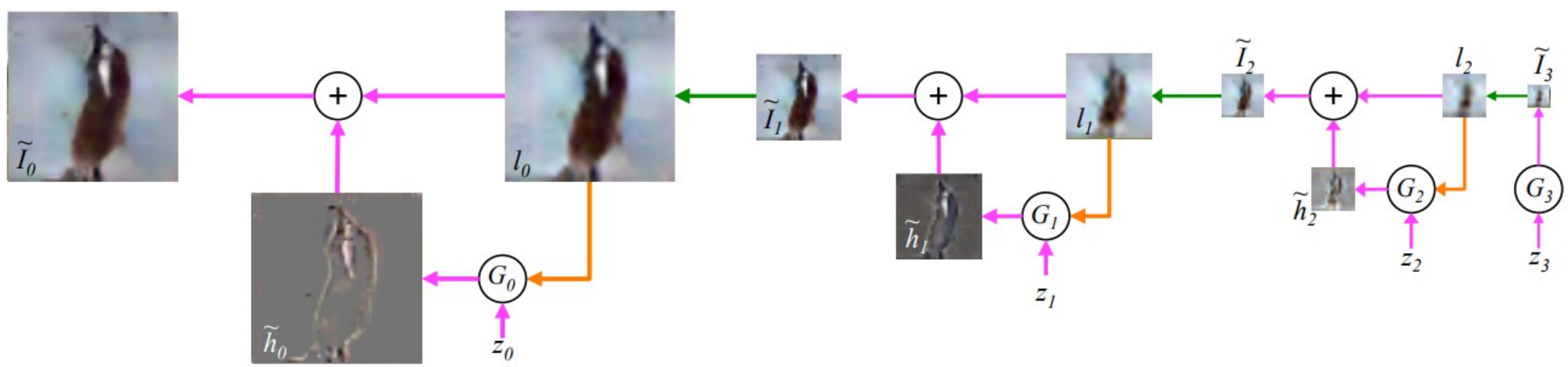
  → Utilizzo su Dataset semplici.

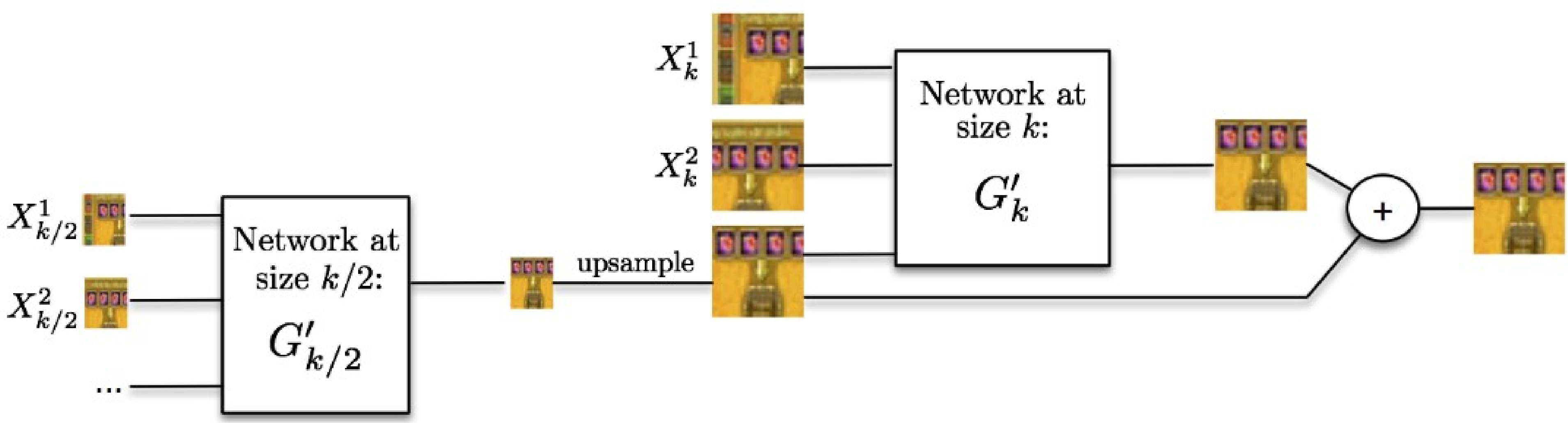
- **Convolutional GANs** → image data

  LAPGAN: NN multiscale, Laplacian pyramid

  DCGAN: Operatori di up&down scaling non sono definiti a priori

LAPGAN



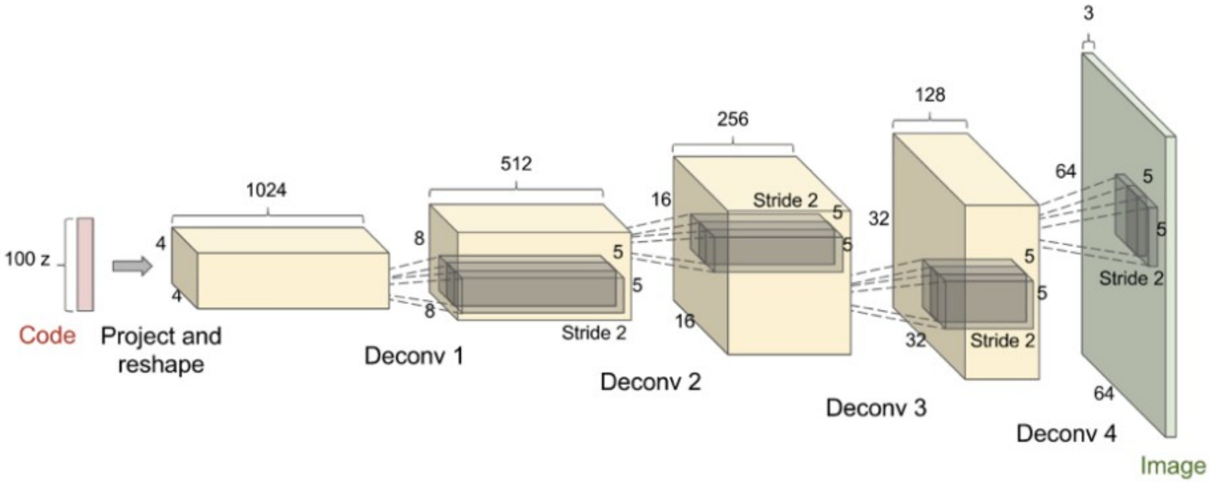


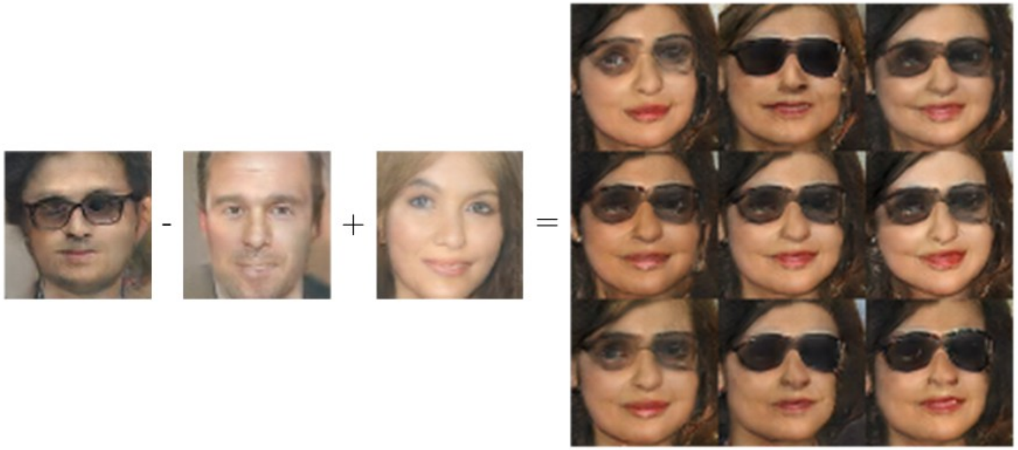
Ref] https://arxiv.org/abs/1511.05440
Deep multi-scale video prediction beyond mean square error

Ref] https://arxiv.org/abs/1506.0575

Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks

DCGAN



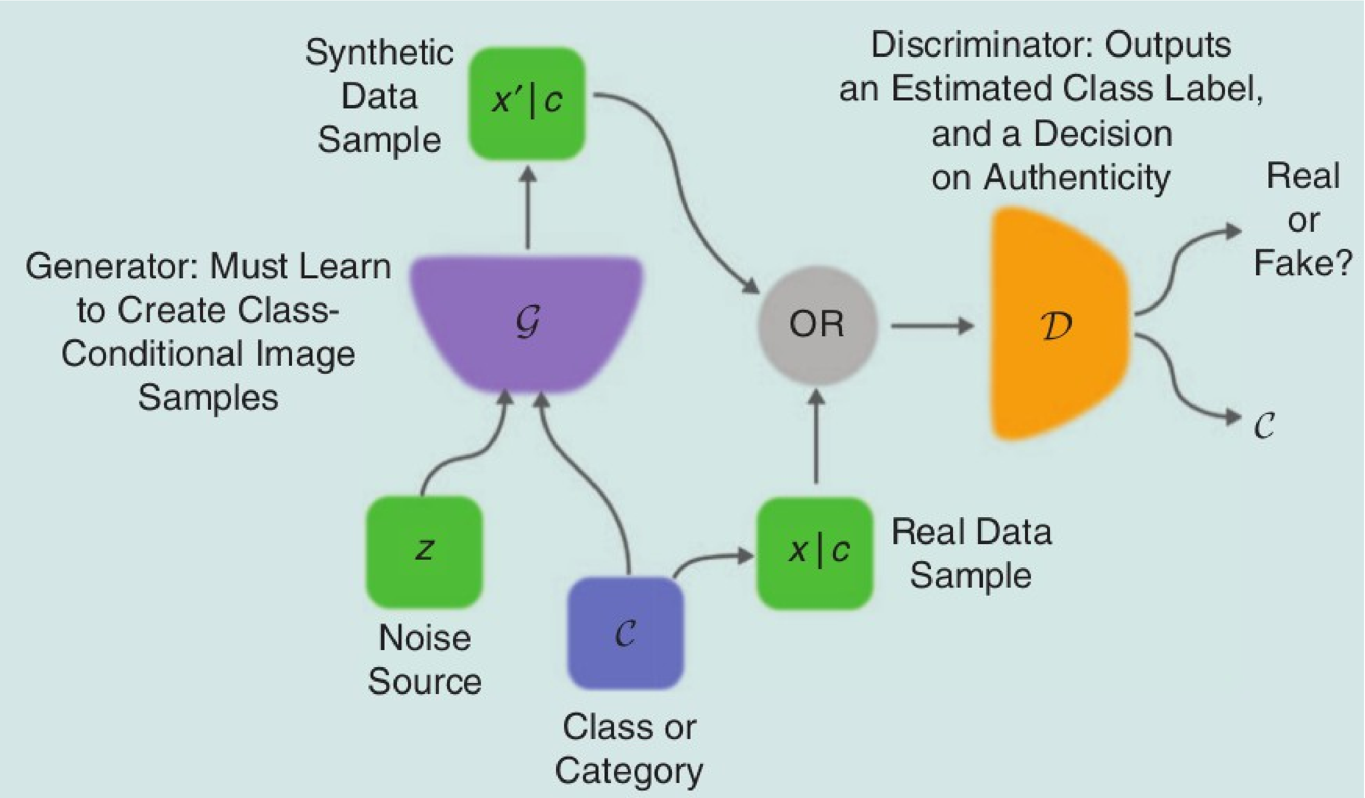


Ref] https://arxiv.org/abs/1511.06434
Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks

GANs Architectures

- **Conditional GANs**
  D & G diventano class-conditional

  evoluzione
  InfoGAN



Ref] https://arxiv.org/abs/1411.1784
Conditional generative adversarial nets

Ref] https://arxiv.org/abs/1606.03657
Infogan: Interpretable representation learning by information maximizing generative adversarial nets

GANs Architectures

- **GANs with Inference Models**
  map da Input space in “Latent space”

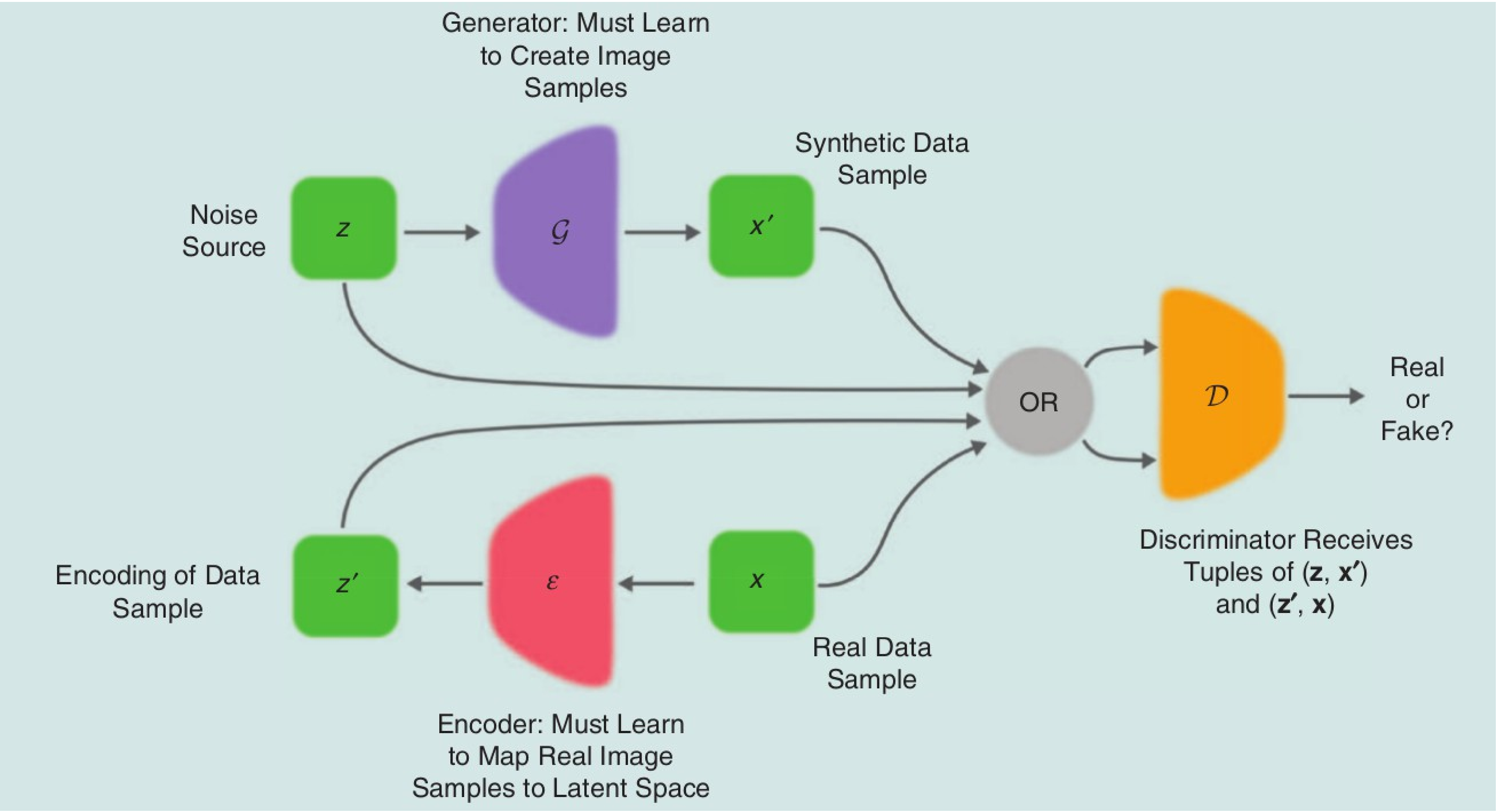
  → catturare encoding e decoding transformations (ricostruzione input)

  Gan in cui *G è divisa in encoder e decoder,*

  La rete discriminativa impara a distinguere le coppie (codifica, input)

  Bidirectional GANs

ALI/BIGANs

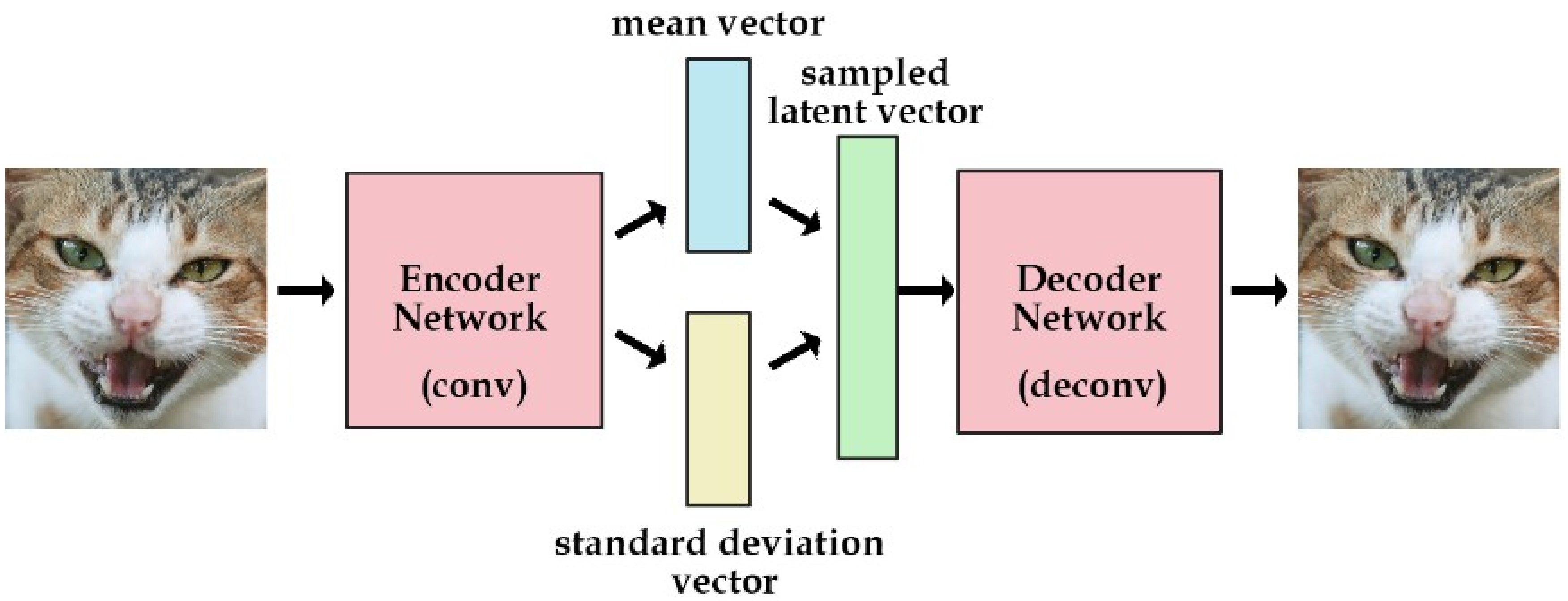


Ref] https://arxiv.org/abs/1411.1784
Adversarially learned inference

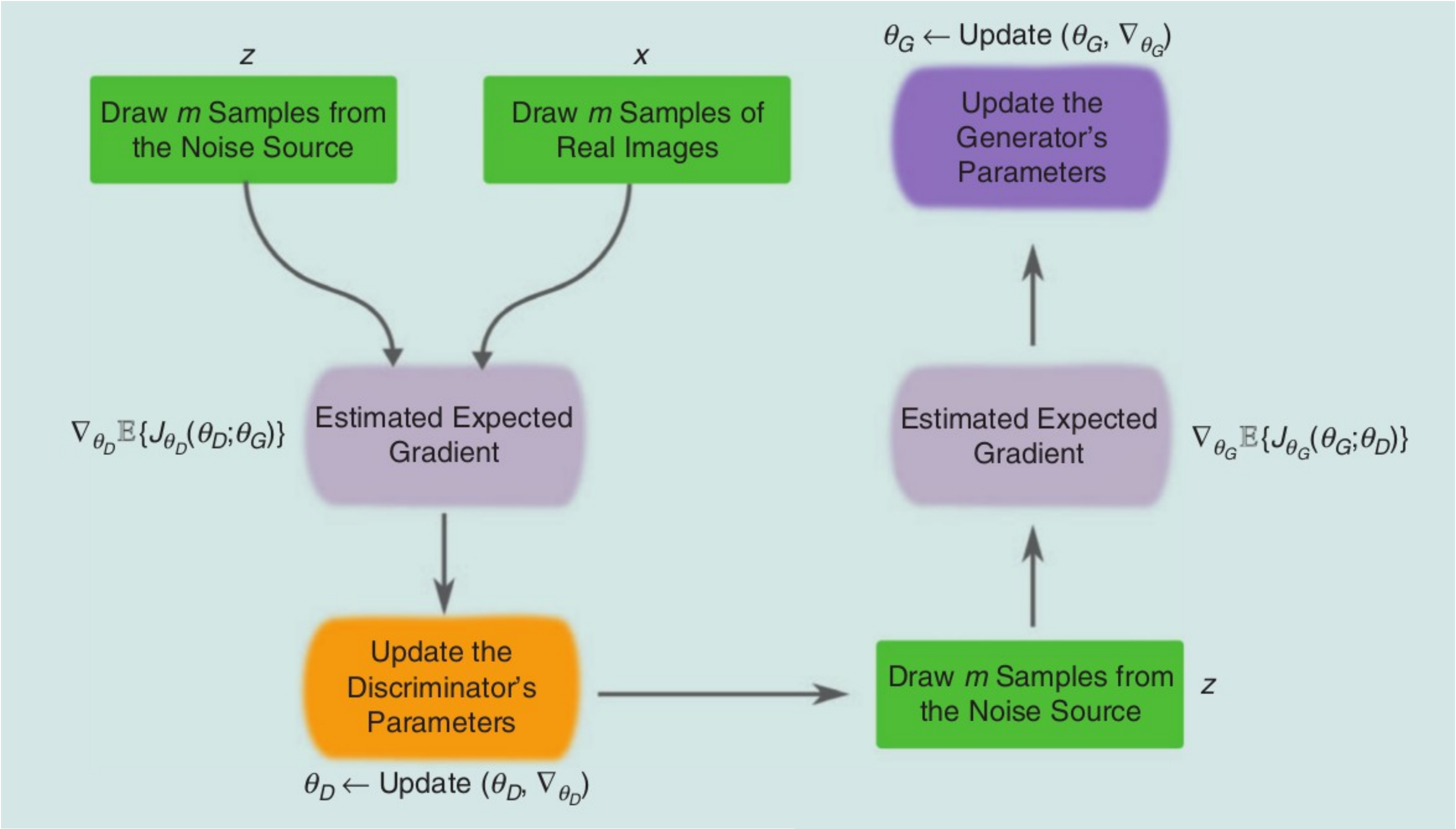
Ref] https://arxiv.org/abs/1606.03657

GANs Architectures

- **Adversarial Variational Autoencoders**



Training GANs



$$\mathcal{D}^*(\mathbf{x}) = p_{\text{data}}(\mathbf{x}) / (p_{\text{data}}(\mathbf{x}) + p_g(\mathbf{x}))$$

Training GANs

Training Instabile:

- Difficoltà nel far convergere entrambe le reti, a causa di molteplici fattori (hyper-parametri, tempo di training, dataset)

- Collasso della rete generativa verso la produzione di un singolo “modello”

- Gradient Vanishing → la rete discriminativa converge più velocemente della generativa

Training Tricks

- Use ReLU (D & G)
- Feature Matching (G)
- Minibatch discrimination (D)
- Heuristic Averaging (G)
- Virtual batch normalization (D & G)
- One-side label smoothing (D)
- Gaussian Noise
- Different Cost Functions (es. WGAN)

Usages

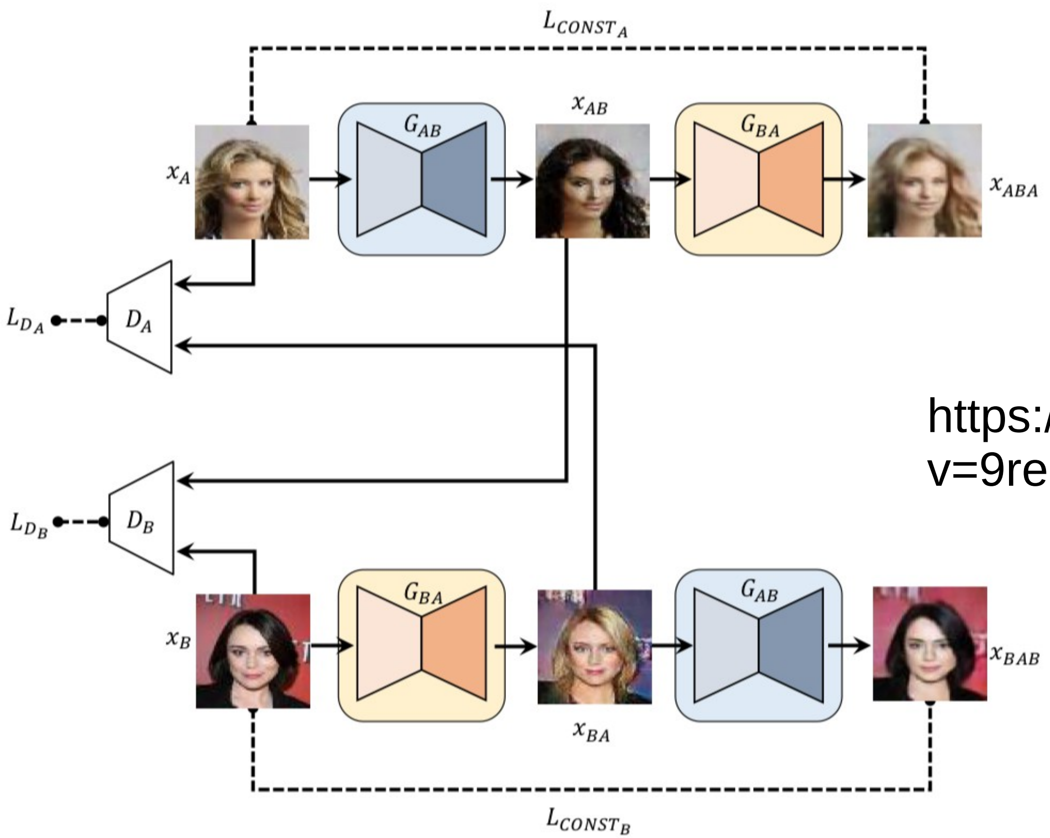
- Classification and regression → DCGAN / ALI
- Image Synthesis → LAPGAN + Conditional
- Image to Image Translation → pix2pix / CycleGAN
- Super Resolution – SRGAN

Open Questions

- Mode Collapse
- Saddle Point
- Evaluating Generative models

MIND BREAKER

Cycle/Disco GAN



https://www.youtube.com/watch?v=9reHvktowLY

Ref] https://arxiv.org/abs/1703.05192
Learning to Discover Cross-Domain Relations with Generative Adversarial Networks

PROGRESSIVE GROWING GANs

Figure 5: 1024 × 1024 images generated using the CELEBA-HQ datase

Our (256 × 256)

Ref] http://research.nvidia.com/sites/default/files/pubs/2017-10_Progressive-Growing-of/karras2018iclr-paper.pdf