AUTOMATANTS

Les GAN et leurs mystères (le secret de leur pouvoir)

Samuel Sithakoul

@Rubiksman78

Plan

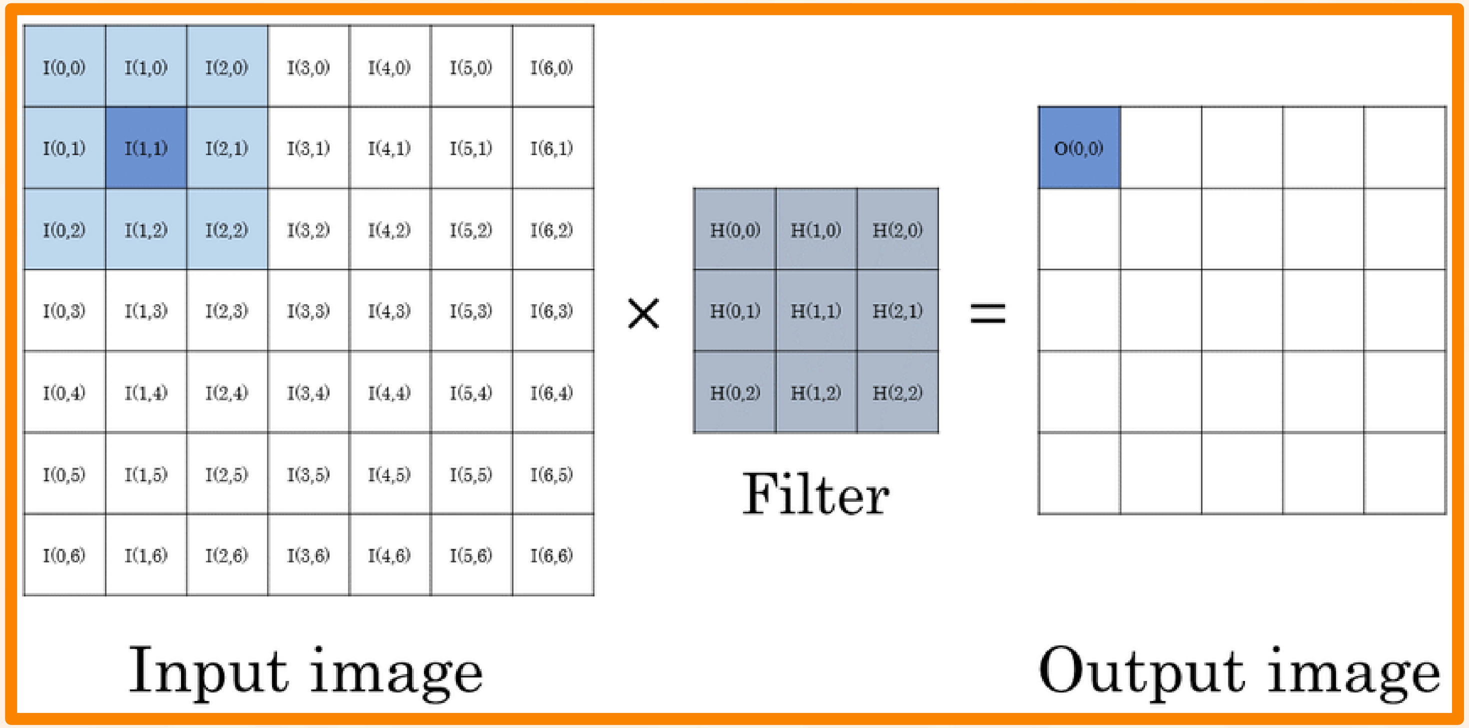
I) Rappels

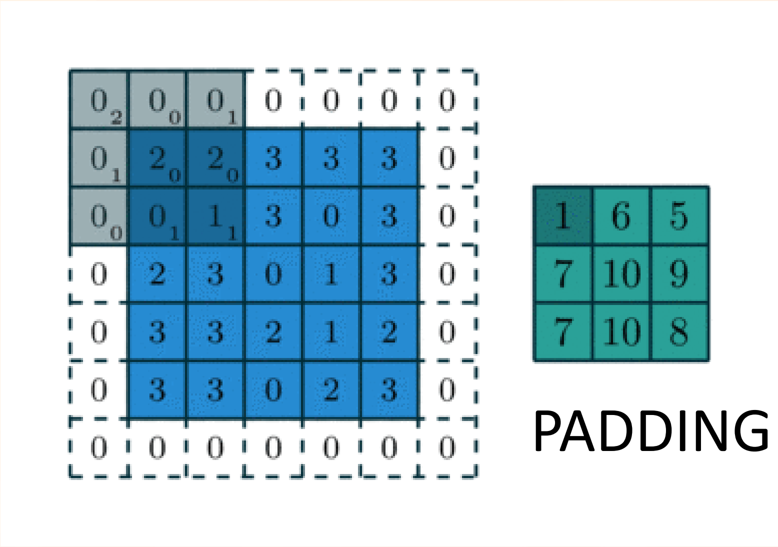
II) Les GAN:

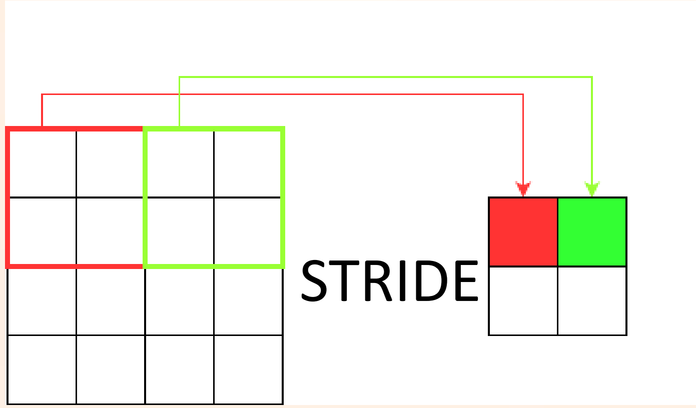
1) Principes de base
2) En pratique
3) Entraînement
4) Interlude: maths
5) Des GAN et des problèmes
6) Y en a encore !

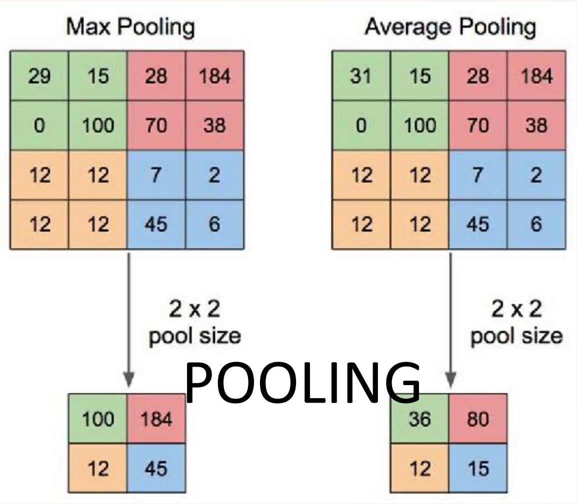
I) Rappels

Convolution

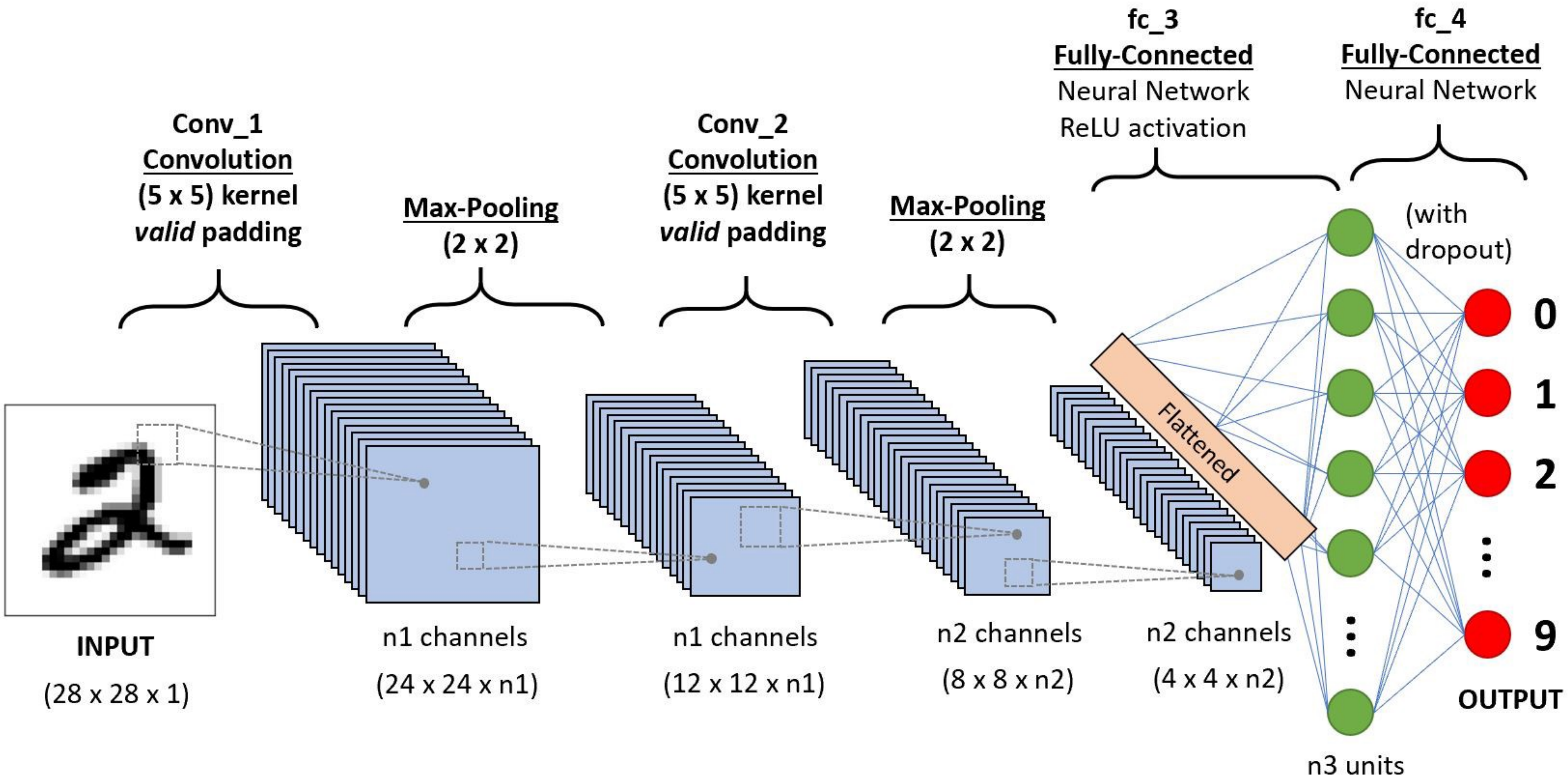








CNN

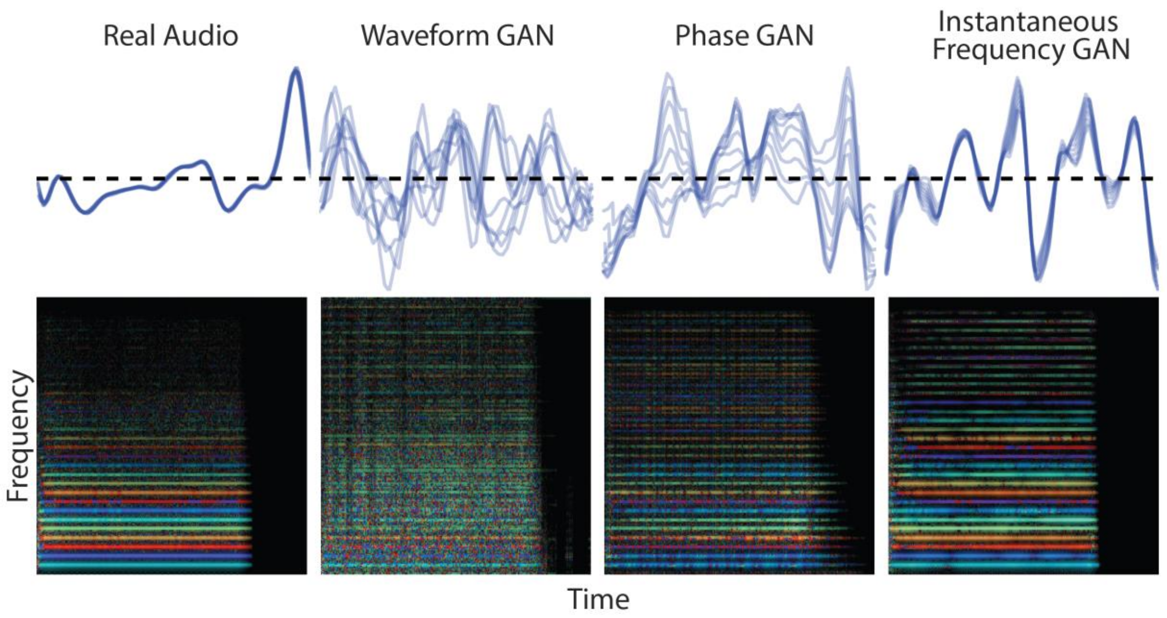


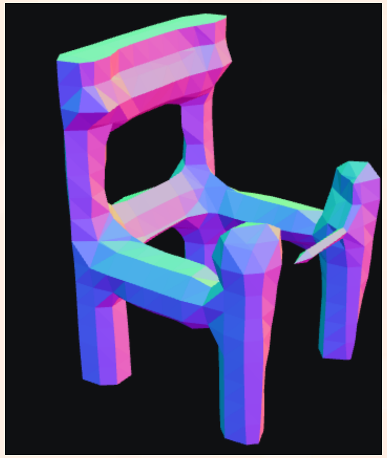
Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
fc_3
Fully-Connected
Neural Network
ReLU activation
fc_4
Fully-Connected
Neural Network
(with dropout)
Flattened
0
1
2
9
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
OUTPUT

Questions?

II) Les GAN

A quoi ça sert un GAN?





IA ou humain?

Questions?

1) Principes de base

Objectif des GAN

Bruit
Loi normale
$z \in \mathbb{R}^{100}$
p_z

GAN

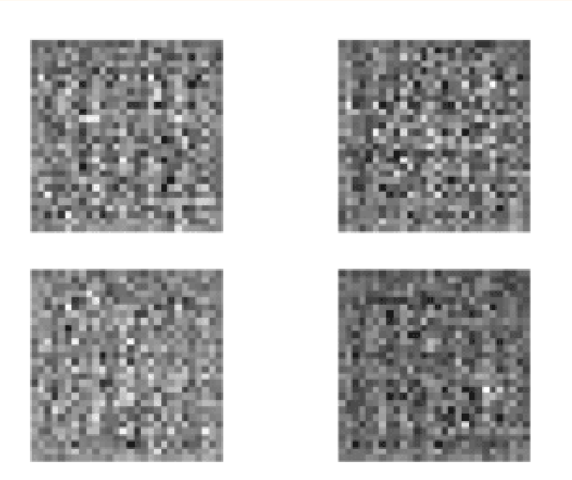
Image générée
$G(z) \in \mathbb{R}^{W \times H}$
p_g

Images réelles
$\mathbb{R}^{W \times H}$
p_x

Objectif: faire coïncider p_g avec p_x

C’est l’heure du-du-du-duel

Générateur
$g: \mathbb{R}^{100} \rightarrow \mathbb{R}^{W \times H}$

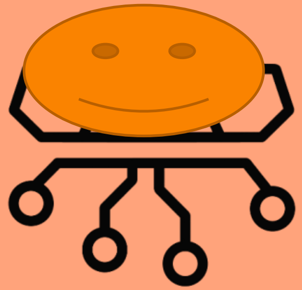


Comment comparer cette image???

Données à approcher

C’est l’heure du-du-du-duel

Le peintre paint



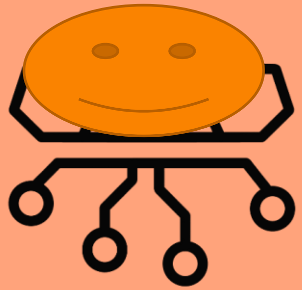
Le faux:

Un vrai dessin:

???

C’est l’heure du-du-du-duel

Le peintre paint



Le faux:

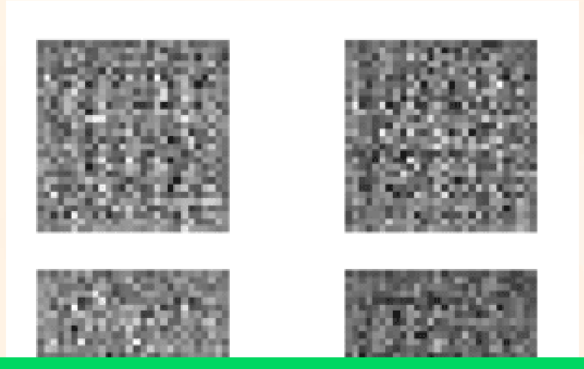
Le critique d’art:
WTF??

Un vrai dessin:

APPREND À DESSINER !!

C’est l’heure du-du-du-duel

Générateur
g: $\mathbb{R}^{100} \rightarrow \mathbb{R}^{W\times H}$



Mais c’est qui ce générateur et discriminateur??
Des réseaux de neurones!

onnées à
pprocher

Discriminateur
d:$\mathbb{R}^{W\times H} \rightarrow [0,1]$

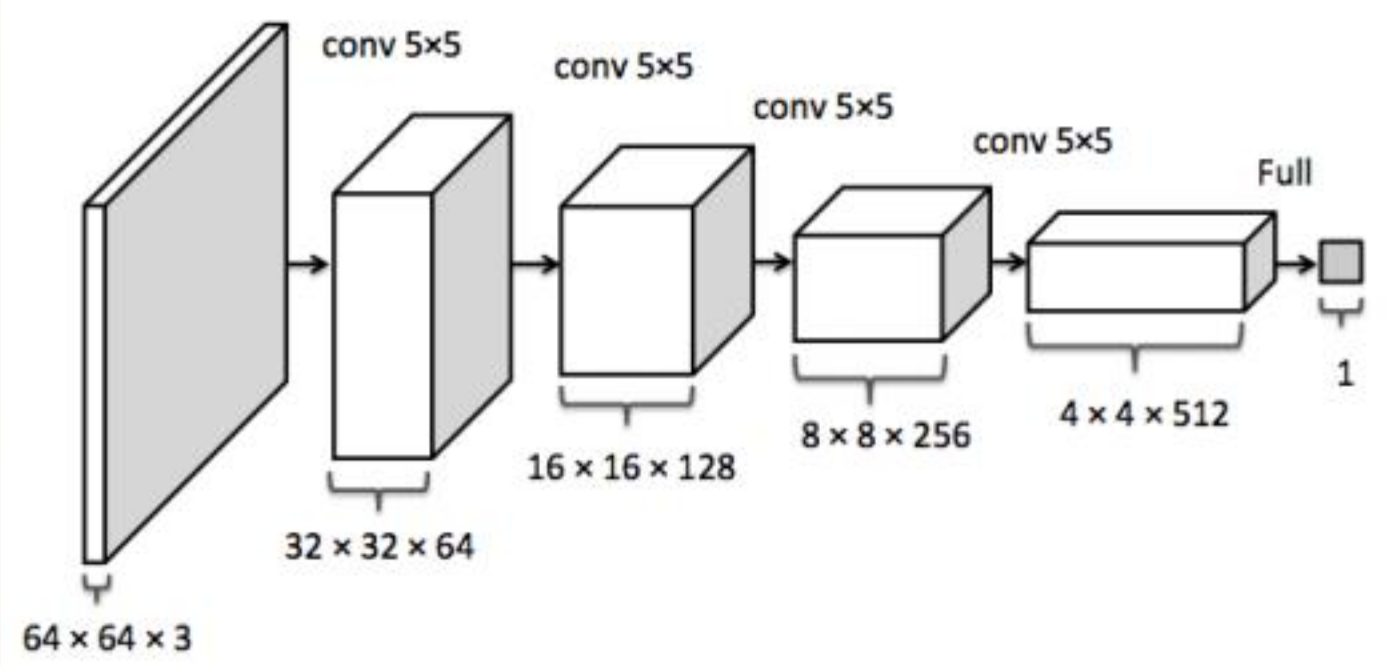
Proba d’être
vraie

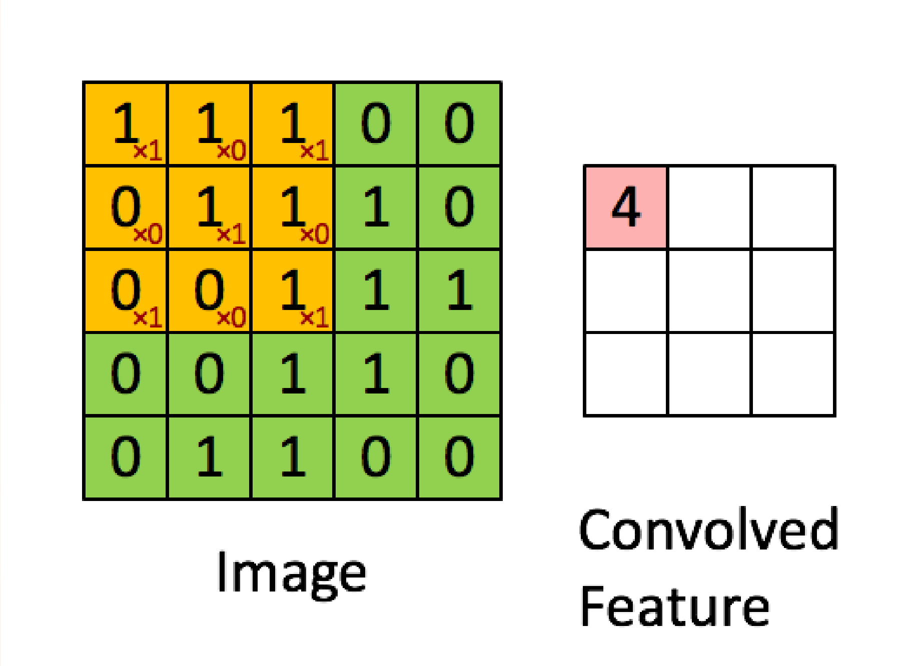
Questions?

2) En pratique

Le discriminateur

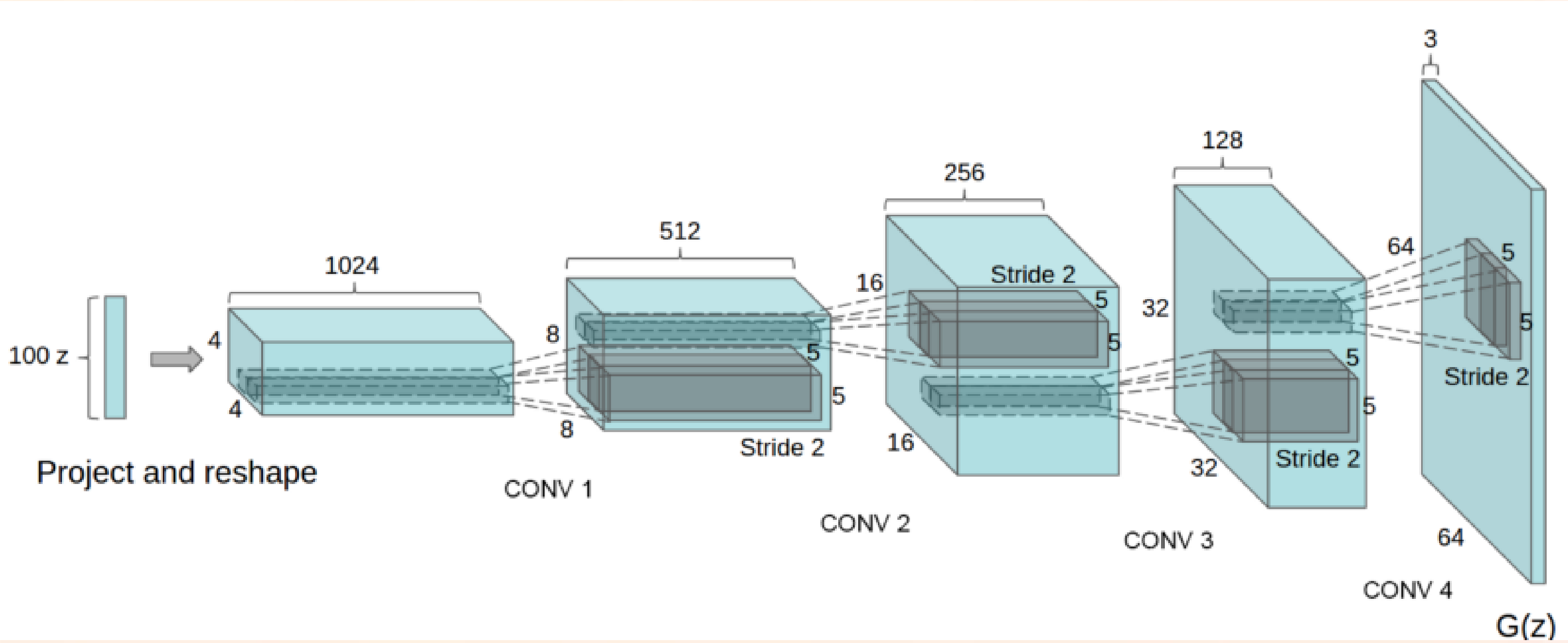
→ Conv2D → Activation →



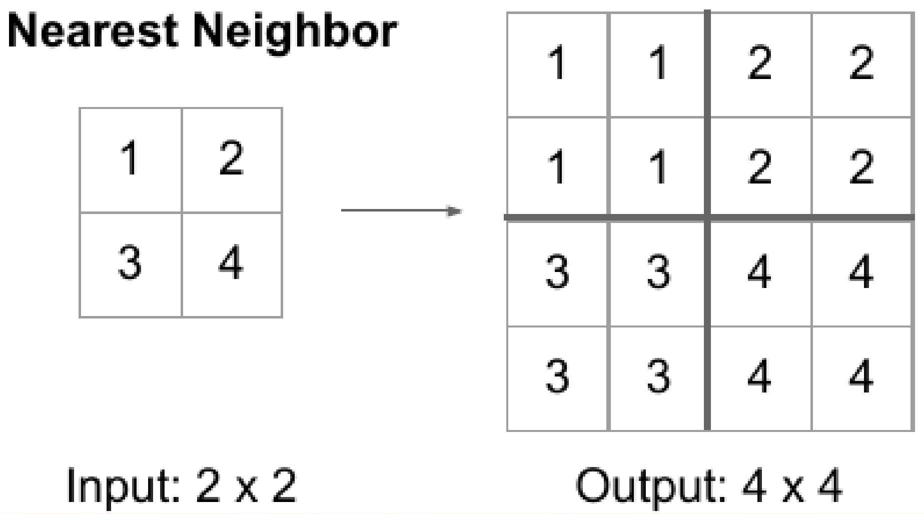


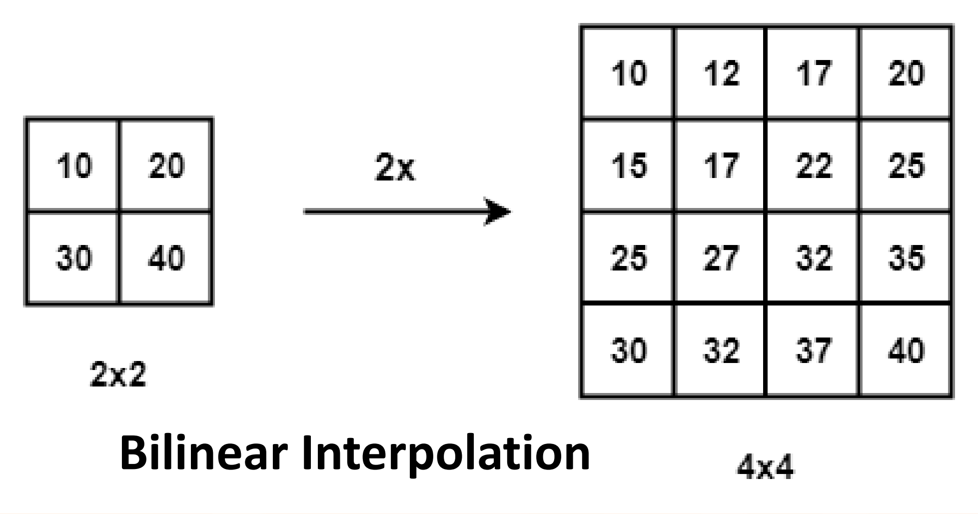
Le générateur

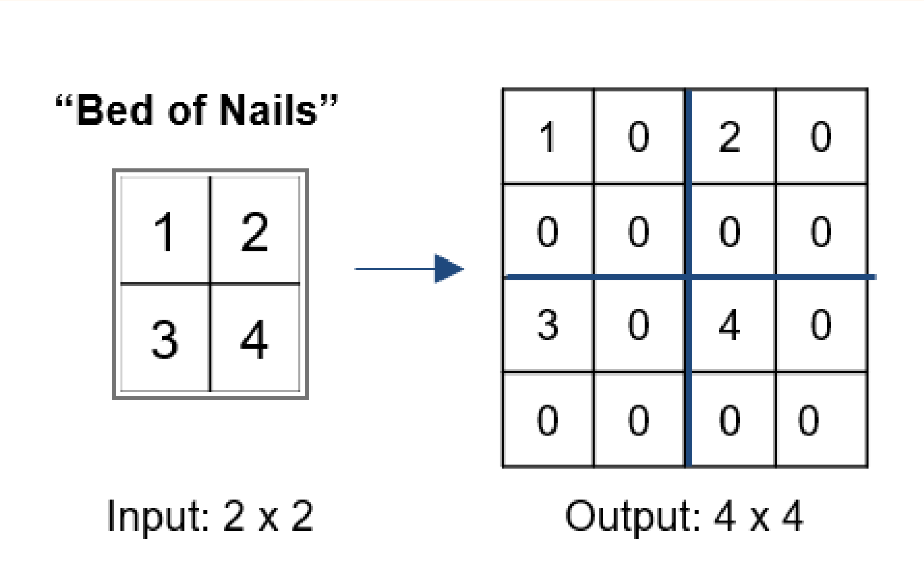
→ Conv2D Transpose → Activation →



Upsampling


Nearest Neighbor
1 2
3 4
1 1 2 2
1 1 2 2
3 3 4 4
3 3 4 4
Input: 2 x 2
Output: 4 x 4


10 20
30 40
2x2
2x
10 12 17 20
15 17 22 25
25 27 32 35
30 32 37 40
Bilinear Interpolation
4x4

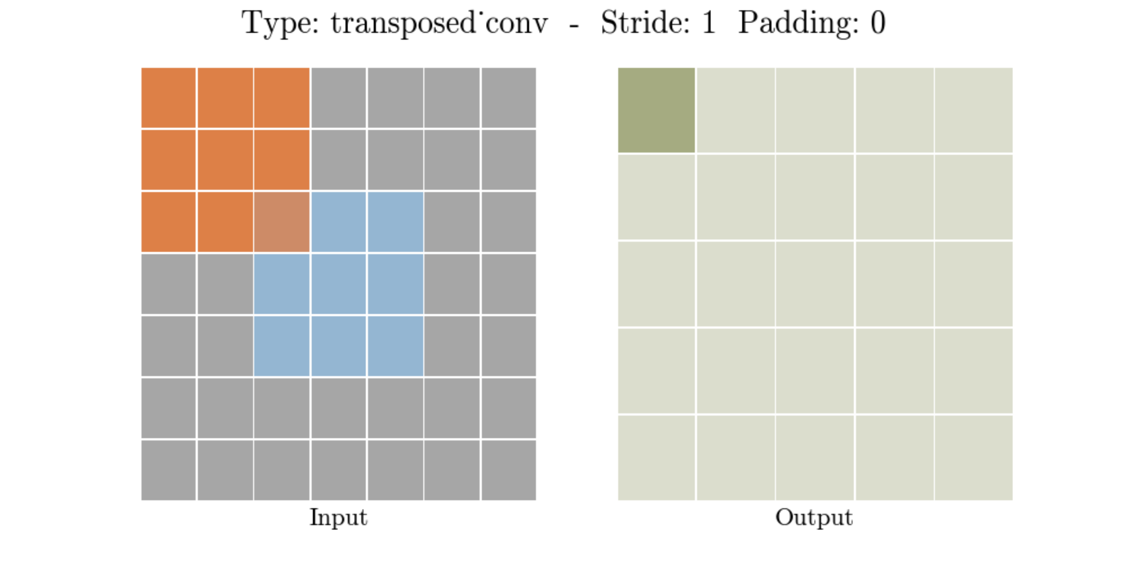

"Bed of Nails"
1 2
3 4
1 0 2 0
0 0 0 0
3 0 4 0
0 0 0 0
Input: 2 x 2
Output: 4 x 4

Les convolutions transposées

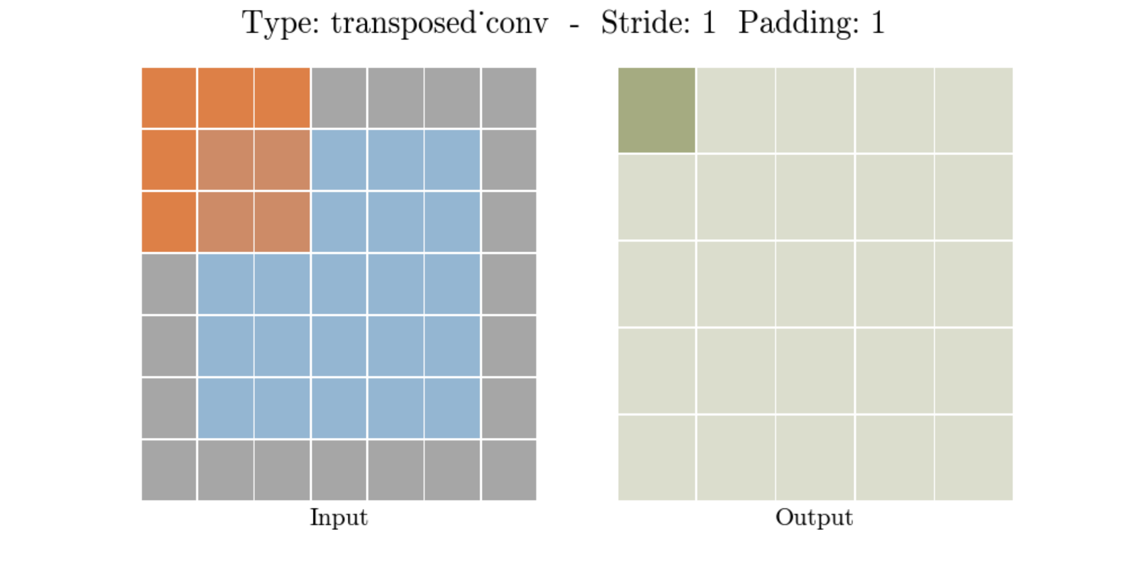
Calcul:
- z = stride – 1
- p’ = kernel_size – padding – 1

Ajouter:
- z zeros entre les pixels
- p’ zeros autour de l’image

Convolution normale



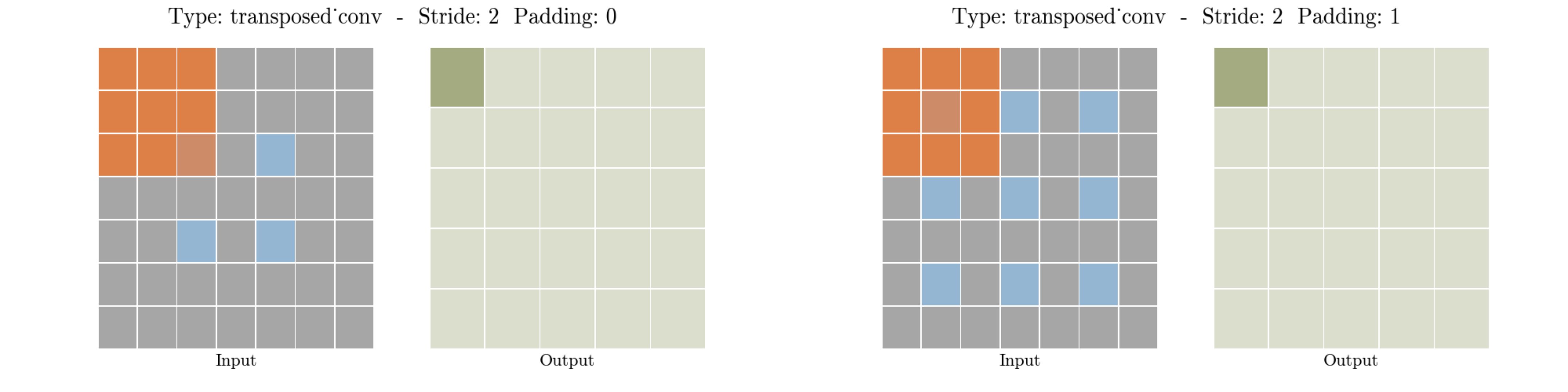
z = 0, p’ = 2



z = 0, p’ = 1

Les convolutions transposées

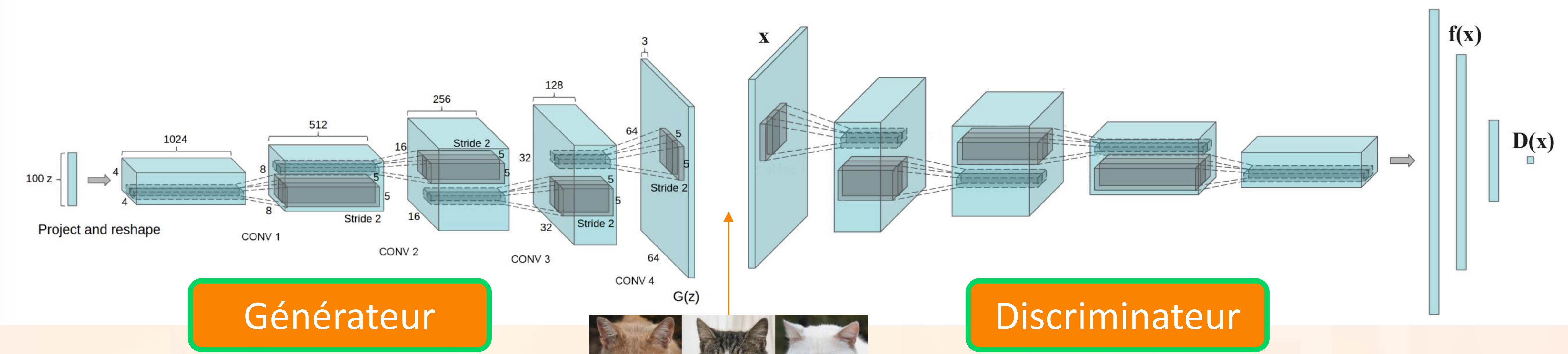
Avec du stride



z = 1, p’=2

z = 1, p’=1

Vue d'ensemble

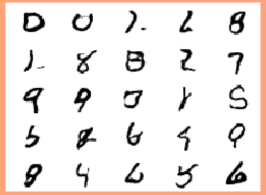


Questions?

3) Entraînement

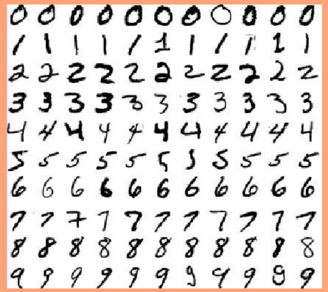
La loss

Images fausses



$z \in \mathbb{R}^{100}$ → Générateur → Discriminateur → Loss

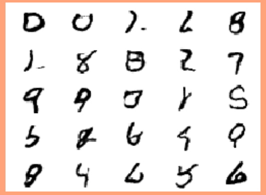
Dataset d'images réelles



Binary Crossentropy (au signe - près):

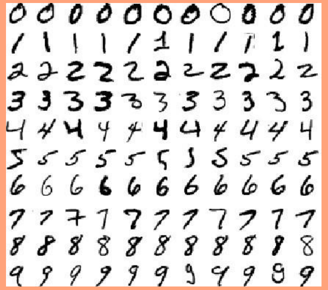
$$\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Le discriminateur apprend

Images fausses 

$z \in \mathbb{R}^{100}$ → Générateur → Discriminateur → Loss: crossentropy

Dataset d'images réelles



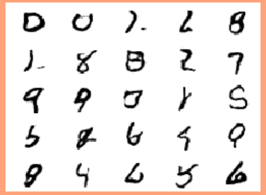
Un vrai est un vrai!
Un faux est un faux!

Backpropagation:
Maximiser
L(D(x),1)+L(D(G(z)),0)

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

Le générateur apprend

Images fausses



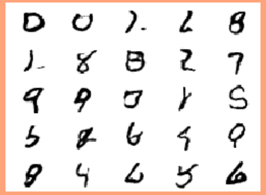
$z \in \mathbb{R}^{100}$ → Générateur → Discriminateur → Loss: crossentropy

Backpropagation:
Minimiser
L(D(G(z)),0)

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

Tu dois te tromper en disant qu'un faux est faux!

Le générateur apprend

Images fausses 

$z \in \mathbb{R}^{100}$ → Générateur → Discriminateur → Loss: crossentropy

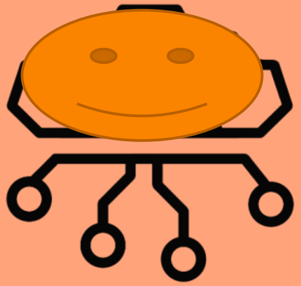
En pratique:
Maximiser
L(D(G(z)),1)

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

Tu dois dire qu'un faux est vrai!

Le combat final

Générateur



Discriminateur

Veut minimiser

Veut maximiser

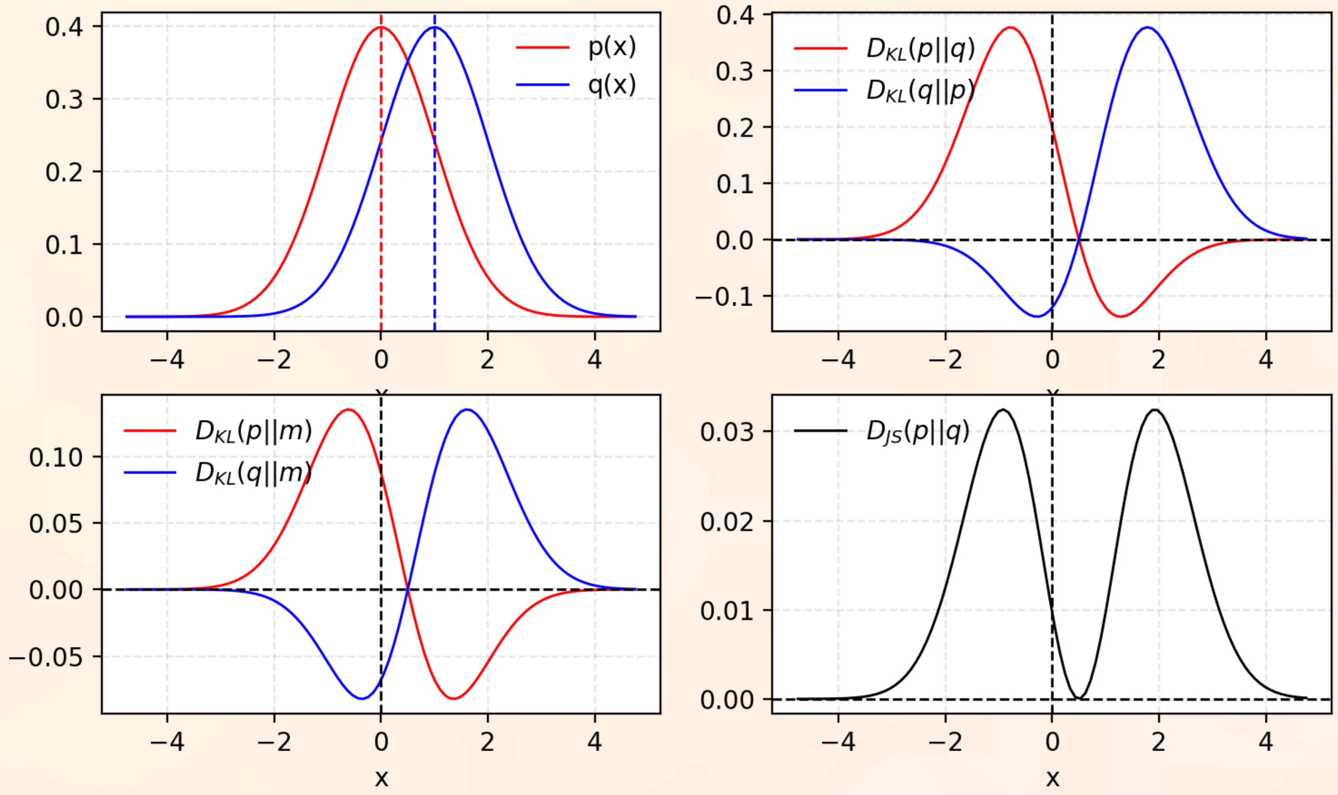
$$\mathrm{L(D,G)} = E_{x\sim p_{data}(x)}[log D(x)] + E_{z\sim p_z(z)}[log(1 - D(G(z)))]$$

Questions?

4) Interlude : maths

KL/JS les bros

Comment mesurer la similarité entre 2 distributions ??



$$D_{KL}(p||q) = \int_x p(x) log \frac{p(x)}{q(x)} dx$$

Kullback-Leibler Divergence

$$D_{JS}(p||q) = \frac{1}{2} D_{KL}(p||\frac{p+q}{2}) + \frac{1}{2} D_{KL}(q||\frac{p+q}{2})$$

Jensen-Shannon Divergence

Réécriture

$$\min_G \max_D \mathrm{L(D,G)} = E_{x\sim p_{data}(x)}[logD(x)] + E_{z\sim p_z(z)}[log(1 - D(G(z)))]$$

x faux = G(z) avec z vecteur latent

$$\min_G \max_D \mathrm{L(D,G)} = E_{x\sim p_{data}(x)}[logD(x)] + E_{x\sim p_g(x)}[log(1 - D(x))]$$

Espérance en intégrale

$$L(G,D) = \int_x (p_{data}(x)log(D(x)) + p_g(x)log(1 - D(x)))dx$$

Valeur optimale

$$L(G,D) = \int_x (p_{data}(x) log(D(x)) + p_{\ }(x) log(1 - D(x)))dx$$

ruc

$$D^*(x) = \frac{p_{d}}{p_{data}}$$

p_{data} , D*(x) = ½

-2log2 !!

En théorie, ça marche !!

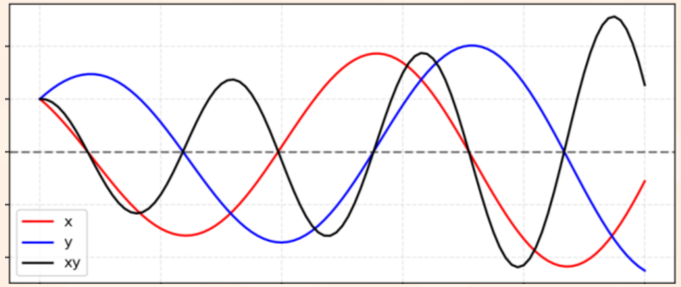
Or aussi (trust me): $L(G,D) = 2D_{JS}(p_{data}||p_g) - 2log2$

GAN loss = mesure similarité par JS Divergence entre p_{data} et p_g quand D est optimal !!!

Questions?

5) Des GAN et des problèmes

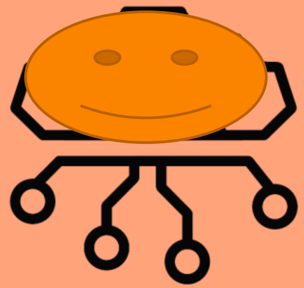
Non convergence



Problème dû à la descente de gradient

G mauvais : D apprend pas
D mauvais : G apprend pas

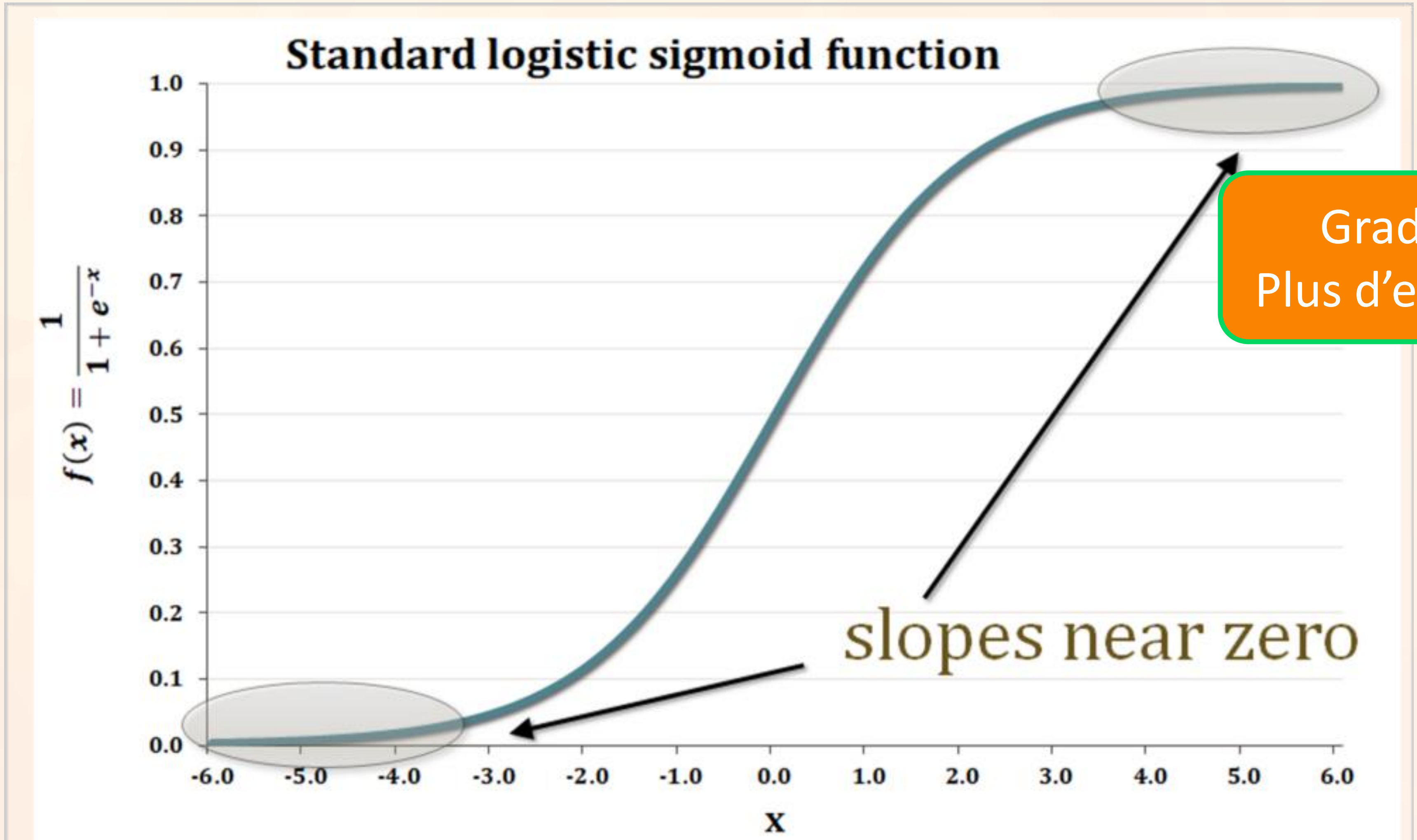
Générateur



Discriminateur

Equilibre fragile!

Vanishing gradient



Gradients nuls =
Plus d’entraînement !!

(Flashback) Backpropagation

$$y = g(z) = g(W.x + b) \approx t \qquad L : x, \theta \longmapsto \frac{1}{2}(y-t)^2$$

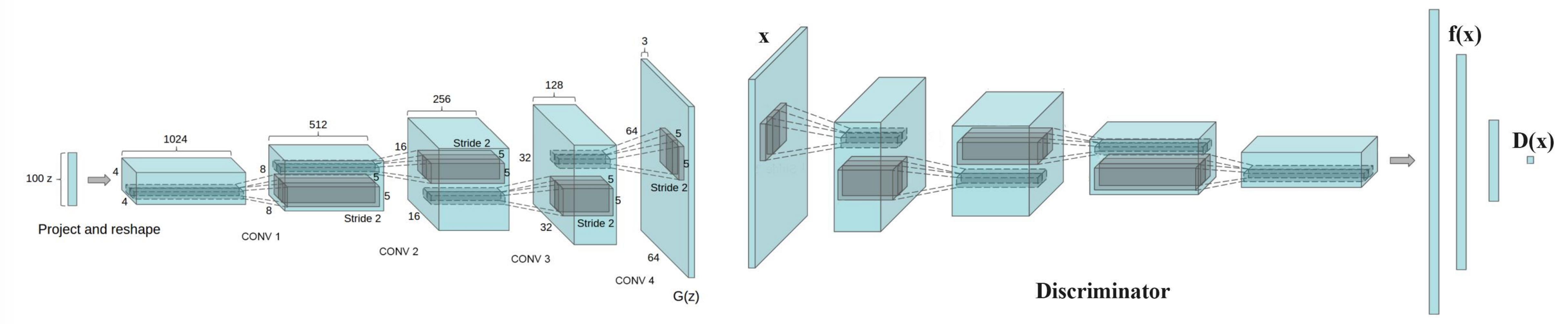
$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial w} = (y-t)g'(z)x \longrightarrow \nabla_W L = (y-t)g'(z).x$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial b} = (y-t)g'(z) \longrightarrow \nabla_b L = (y-t)g'(z)$$

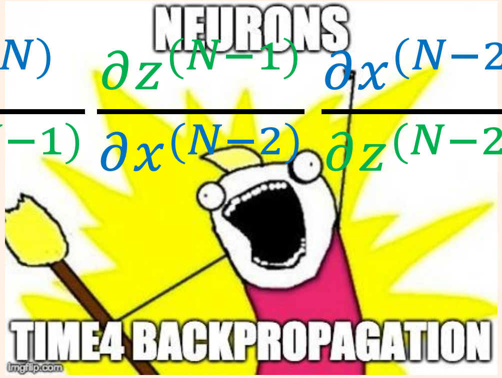
Qu'en est-il pour un réseau de taille N ?

$$\frac{\partial L}{\partial w^{(i)}} = \frac{\partial L}{\partial x^{(N)}}\frac{\partial x^{(N)}}{\partial z^{(N)}}\frac{\partial z^{(N)}}{\partial x^{(N-1)}}\frac{\partial x^{(N)}}{\partial z^{(N-1)}}\frac{\partial z^{(N-1)}}{\partial x^{(N-2)}}\frac{\partial x^{(N-2)}}{\partial z^{(N-2)}}\cdots\frac{\partial x^{(i+2)}}{\partial z^{(i+1)}}\frac{\partial z^{(i+1)}}{\partial x^{(i+1)}}\frac{\partial x^{(i+1)}}{\partial z^{(i)}}\frac{\partial z^{(i)}}{\partial w^{(i)}}$$

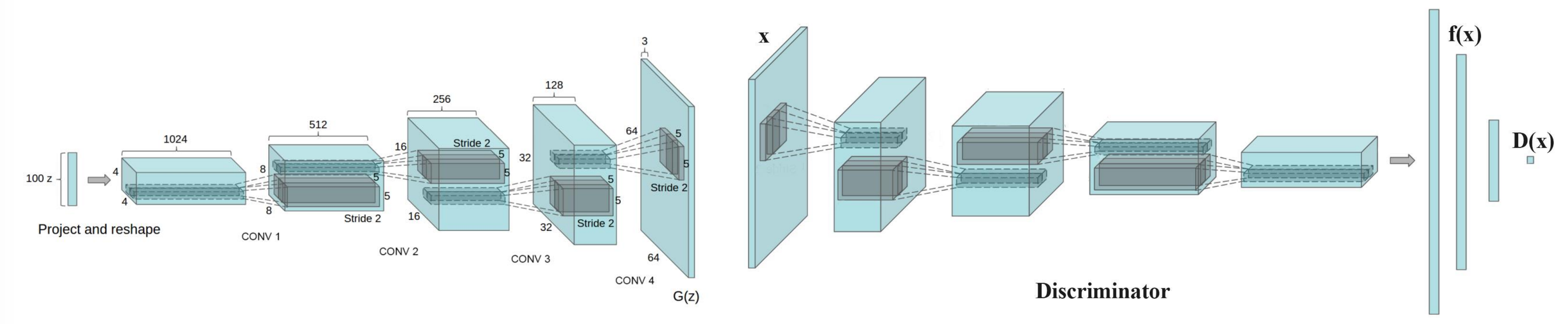
Vanishing gradient



$$\frac{\partial L}{\partial w^{(i)}} = \frac{\partial L}{\partial x^{(N)}} \frac{\partial x^{(N)}}{\partial z^{(N)}} \frac{\partial z^{(N)}}{\partial x^{(N-1)}} \frac{\partial x^{(N)}}{\partial z^{(N-1)}} \frac{\partial z^{(N-1)}}{\partial x^{(N-2)}} \frac{\partial x^{(N-2)}}{\partial z^{(N-2)}} \cdots \frac{\partial x^{(i+2)}}{\partial z^{(i+1)}} \frac{\partial z^{(i+1)}}{\partial x^{(i+1)}} \frac{\partial x^{(i+1)}}{\partial z^{(i)}} \frac{\partial z^{(i)}}{\partial w^{(i)}}$$

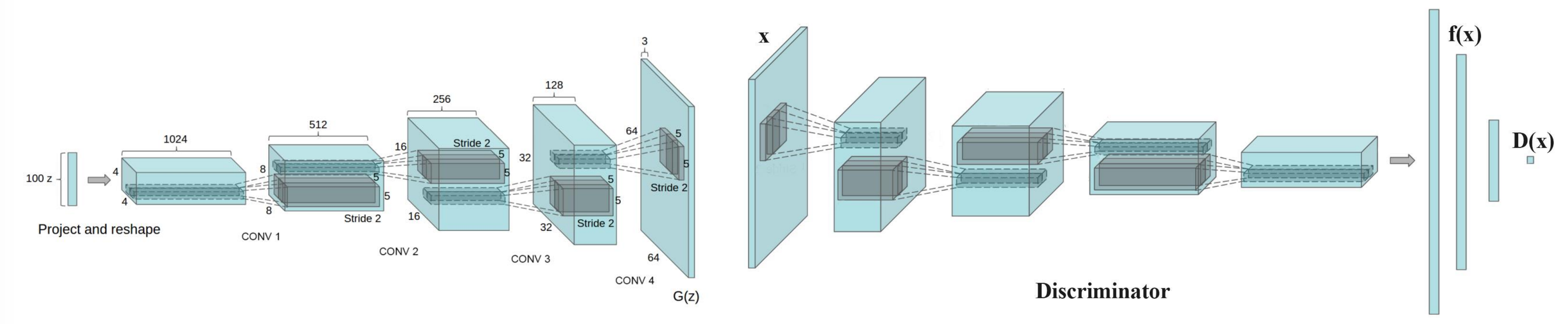


Vanishing gradient



$$\frac{\partial L}{\partial w^{(i)}} = \underbrace{\frac{\partial L}{\partial x^{(N)}}}_{\sim 0} \frac{\partial x^{(N)}}{\partial z^{(N)}} \frac{\partial z^{(N)}}{\partial x^{(N-1)}} \frac{\partial x^{(N)}}{\partial z^{(N-1)}} \frac{\partial z^{(N-1)}}{\partial x^{(N-2)}} \frac{\partial x^{(N-2)}}{\partial z^{(N-2)}} \cdots \frac{\partial x^{(i+2)}}{\partial z^{(i+1)}} \frac{\partial z^{(i+1)}}{\partial x^{(i+1)}} \frac{\partial x^{(i+1)}}{\partial z^{(i)}} \frac{\partial z^{(i)}}{\partial w^{(i)}}$$

Vanishing gradient

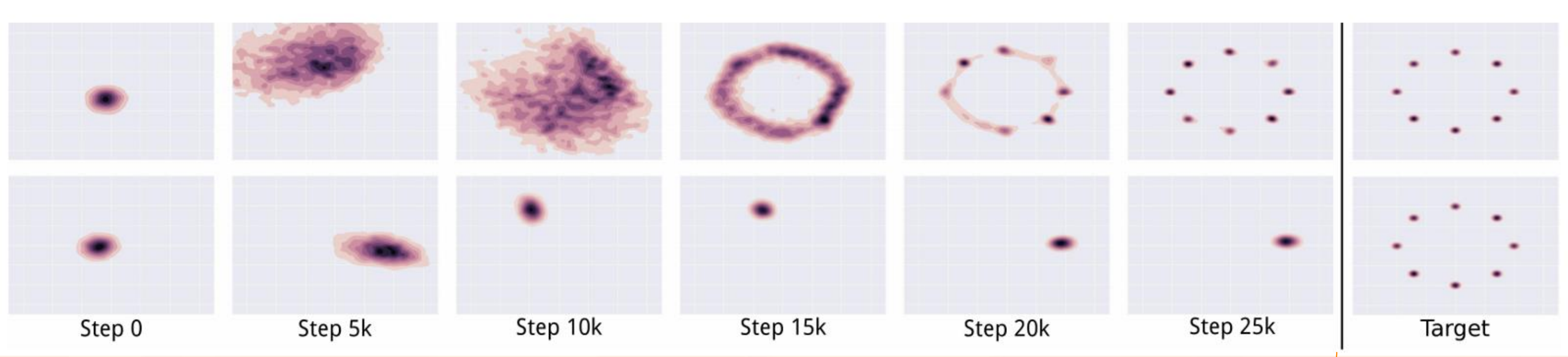


$$\frac{\partial L}{\partial w^{(i)}} \approx 0$$

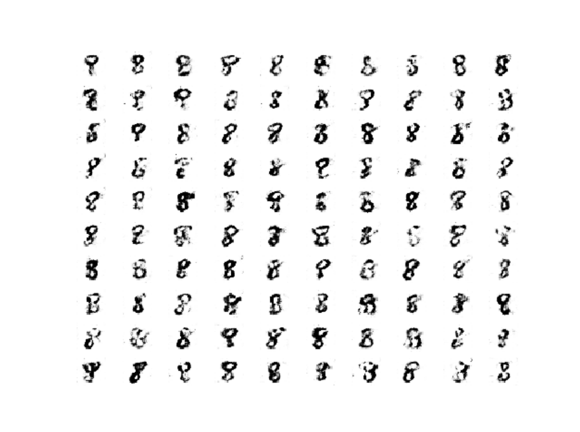
Gradients très faibles = entraînement inefficace

Mode collapse

«Bon » entraînement

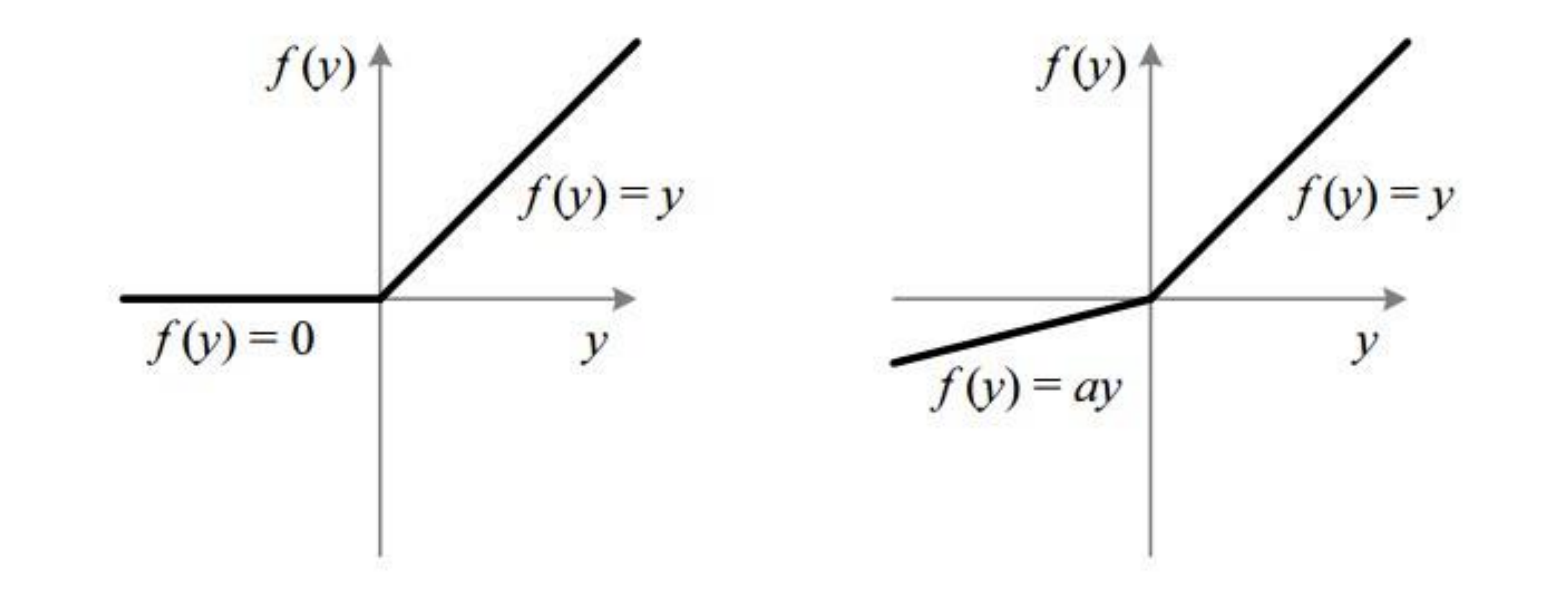


Mode collapse



6) Des solutions?

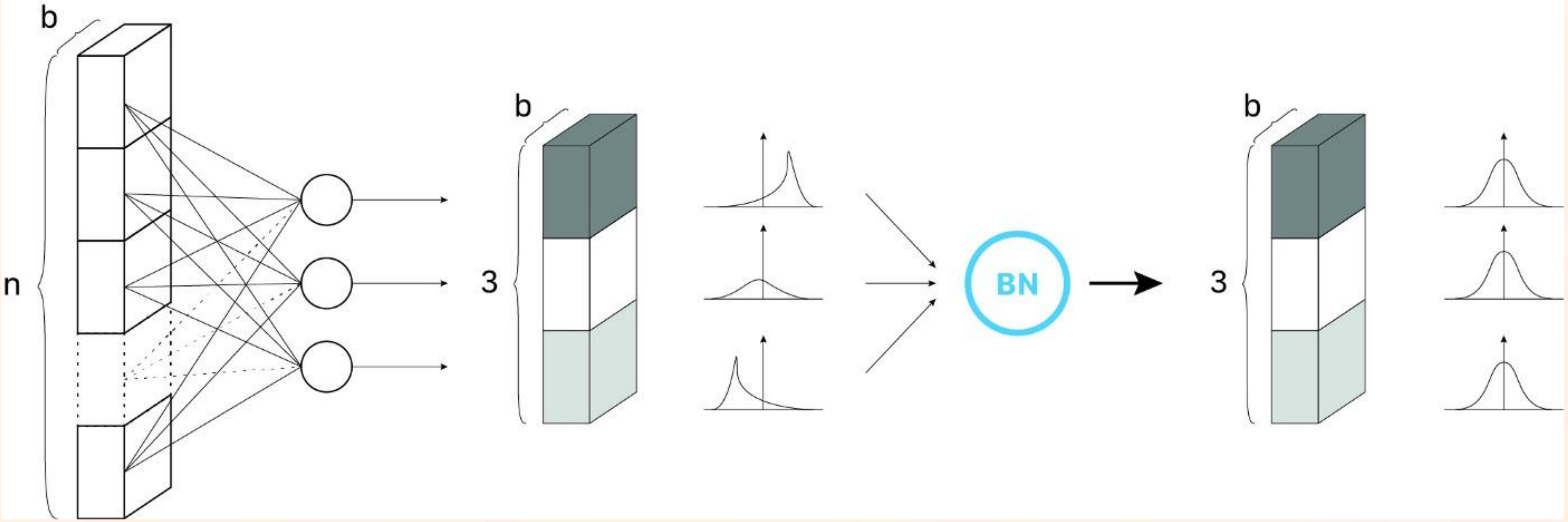
Les activations



ReLU

LeakyReLU

La batchnorm



$$\mu = \frac{1}{batch\_size} \sum X_i$$

$$\sigma^2 = \frac{1}{batch\_size} \sum (X_i - \mu)^2$$

$$\hat{X}_i = \frac{X_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}$$

$$Y_i = \alpha \hat{X}_i + \beta$$

Récap

El Discriminator

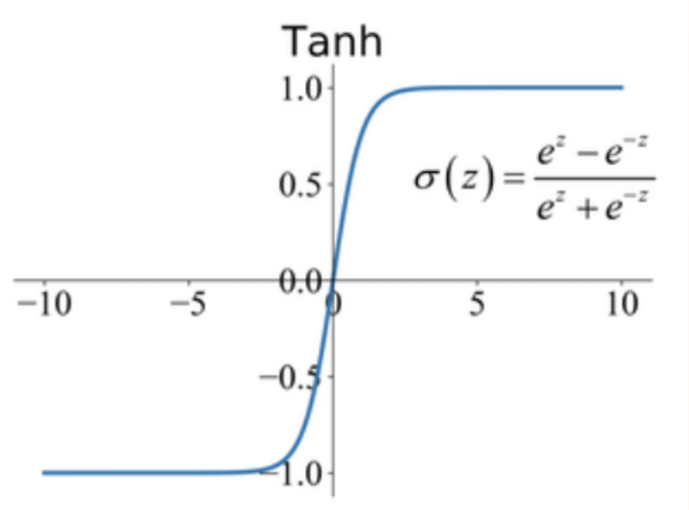
→ Conv2D → Batch Normalization → LeakyReLU →

El Generator

→ Conv2D Transpose → Batch Normalization → ReLU →

Les petites astuces

Normaliser entre -1 et 1



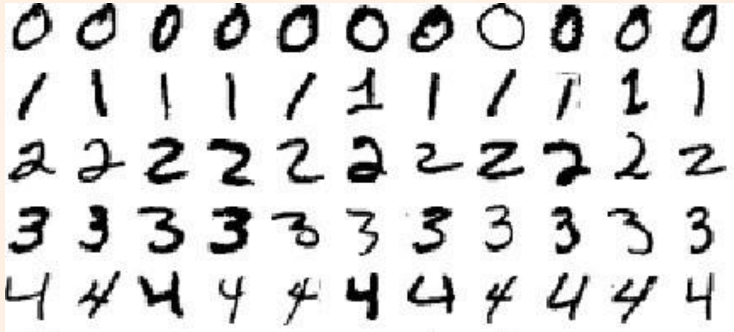
$$E_{x \sim p_{data}(x)}[0.9 log D(x)] + E_{z \sim p_z(z)}[log(1 - D(G(z)))]$$

One-sided label smoothing

Feature matching

$$||\mathbb{E}_{x \sim p_{data}} \mathbf{f}(\mathbf{x}) - \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} \mathbf{f}(G(\mathbf{z}))||_2^2.$$

Ajouter labels



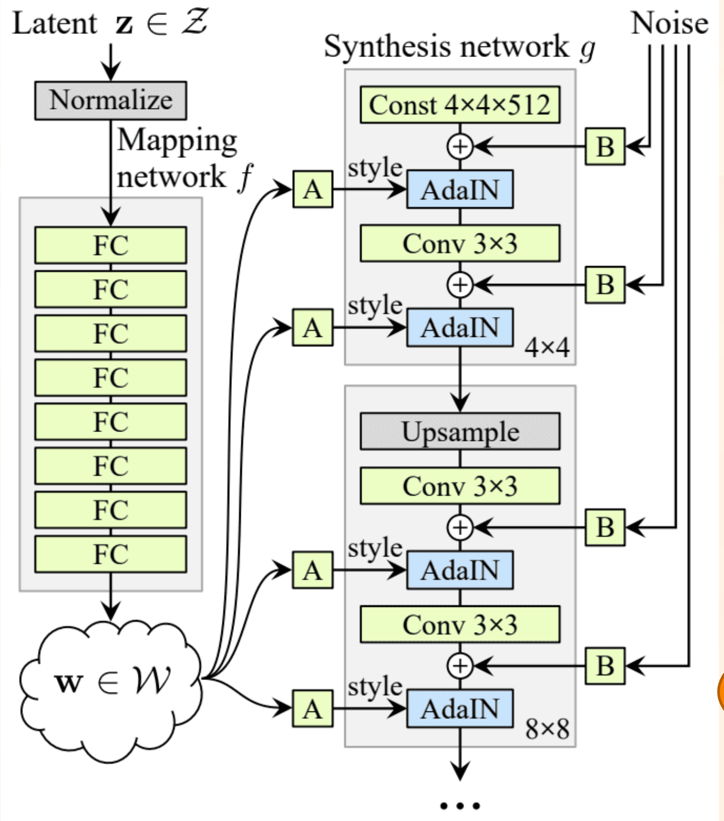
	Discriminator/Critic	Generator
GAN	$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right) \right]$	$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)$
WGAN	$\nabla_w \frac{1}{m} \sum_{i=1}^{m} \left[f\left(x^{(i)}\right) - f\left(G\left(z^{(i)}\right)\right) \right]$	$\nabla_\theta \frac{1}{m} \sum_{i=1}^{m} f\left(G\left(z^{(i)}\right)\right)$

WGAN

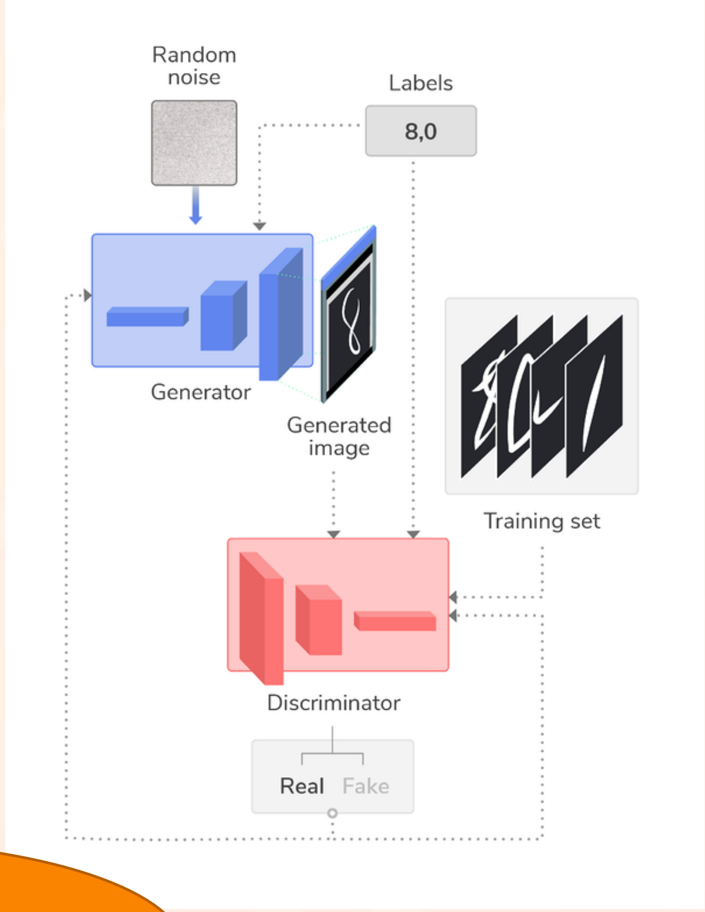
Questions?

7) Y en a encore !

Génération +

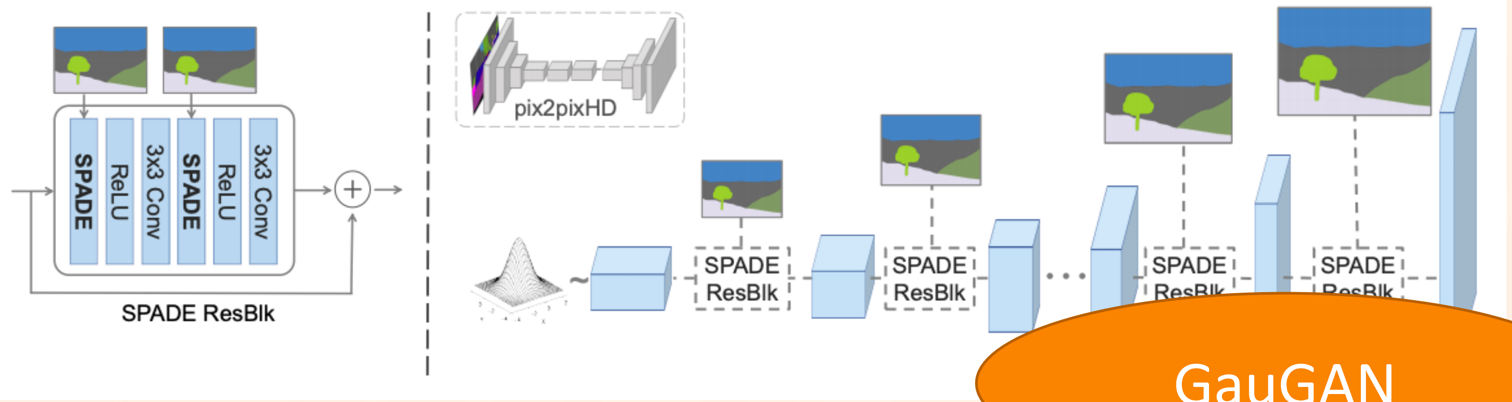


Style GAN



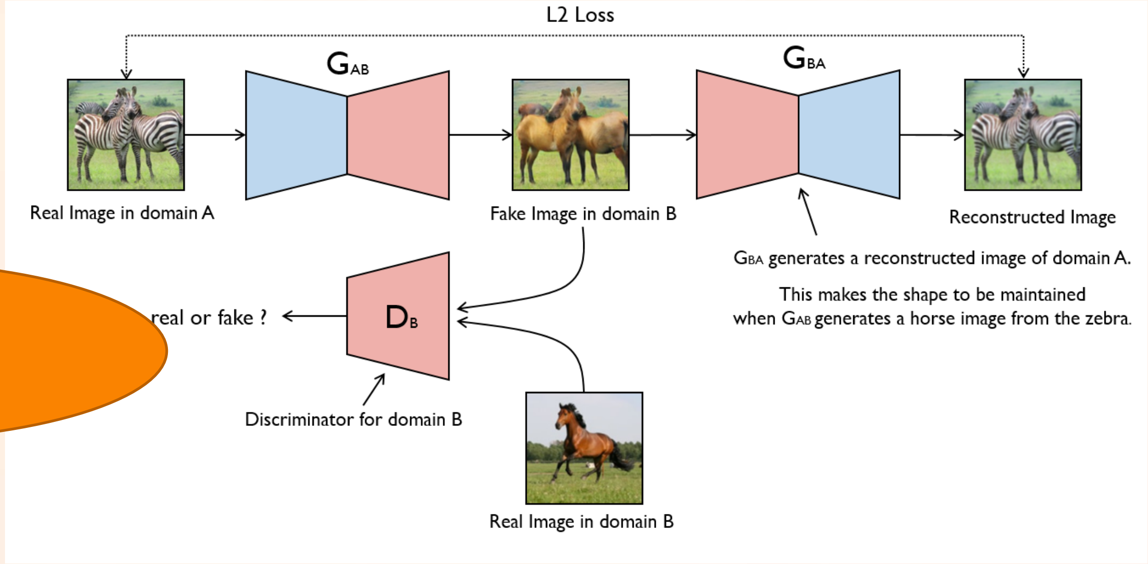
Conditional GAN

Autres trucs marrants

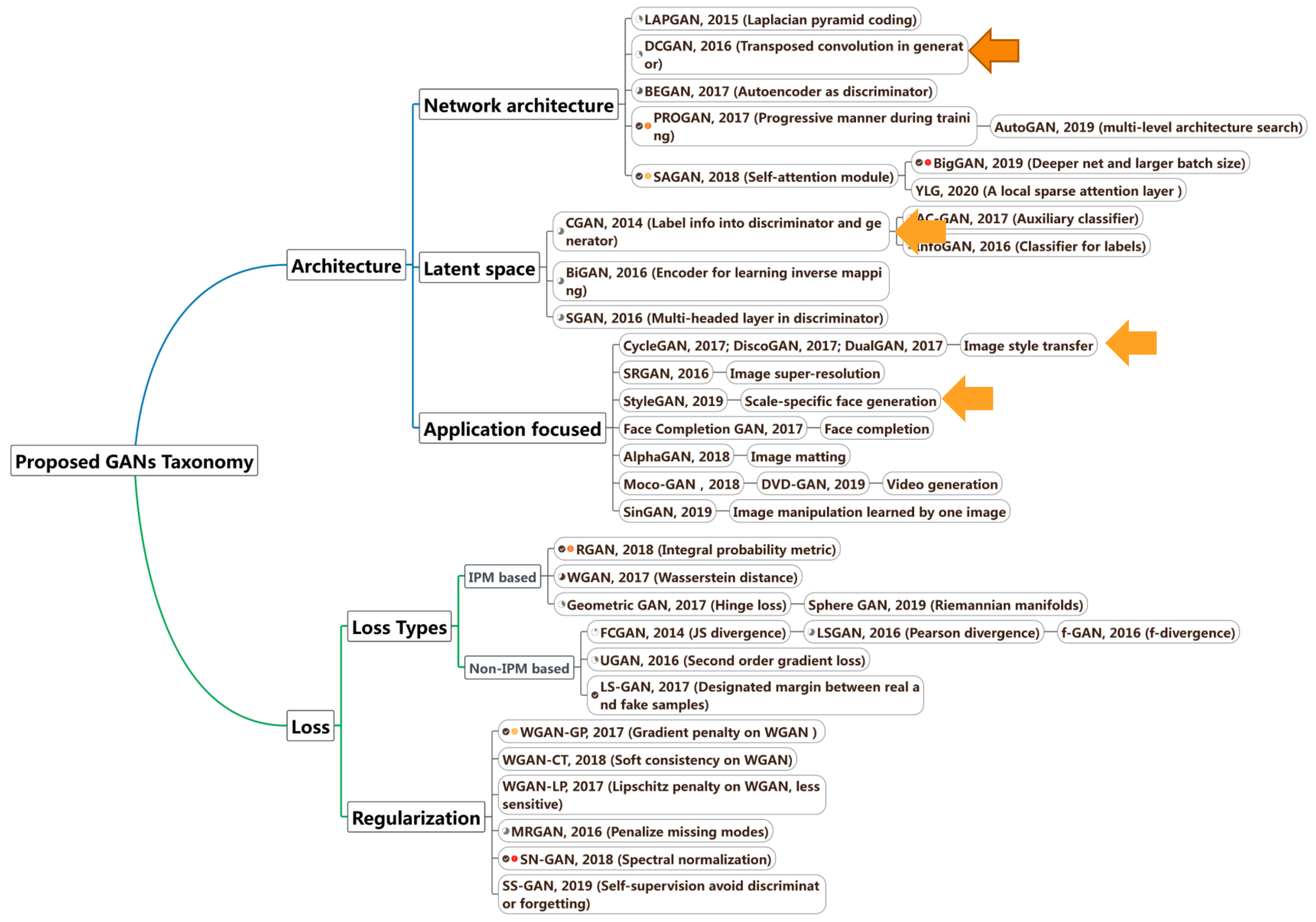


GauGAN

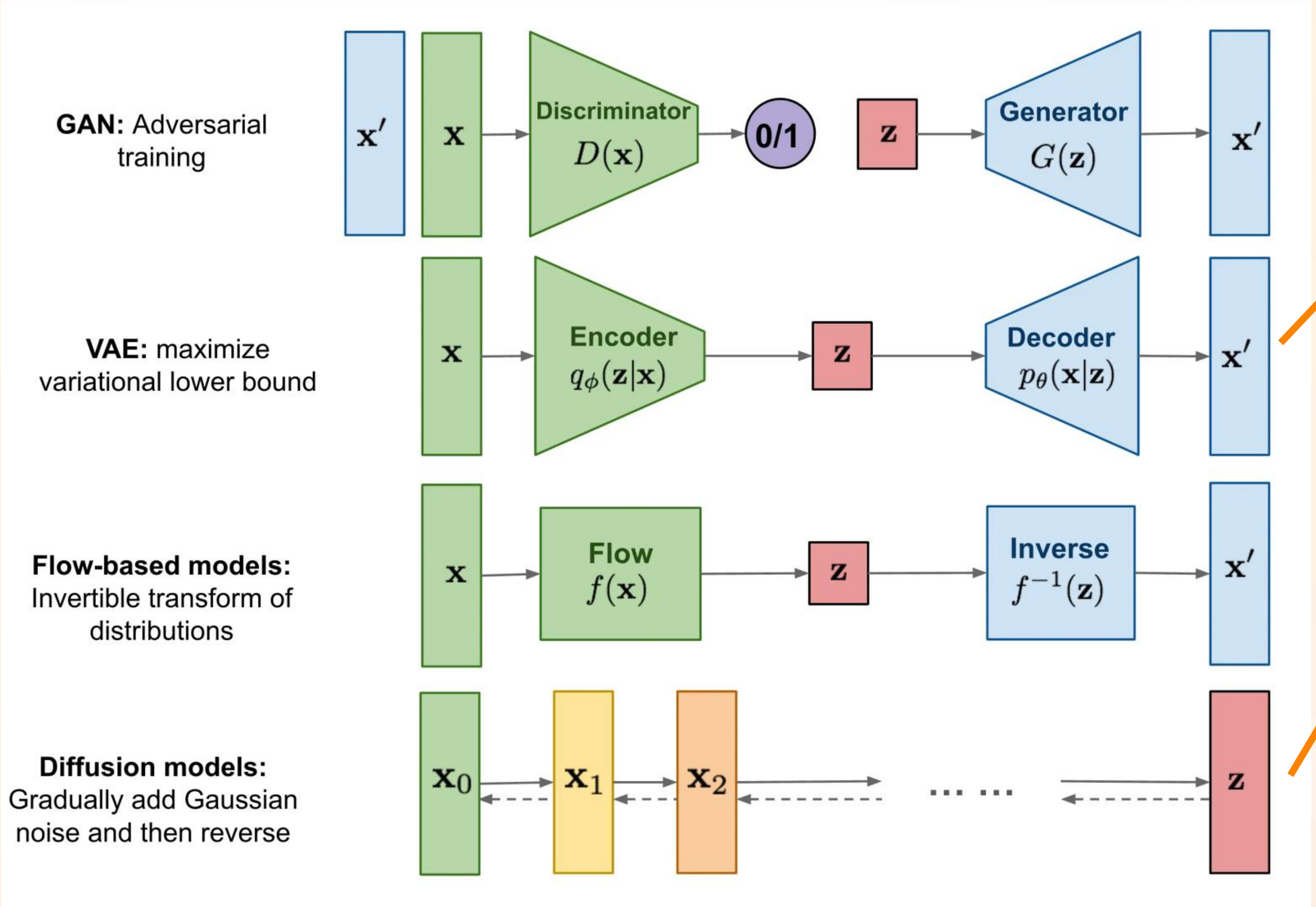
Cycle GAN



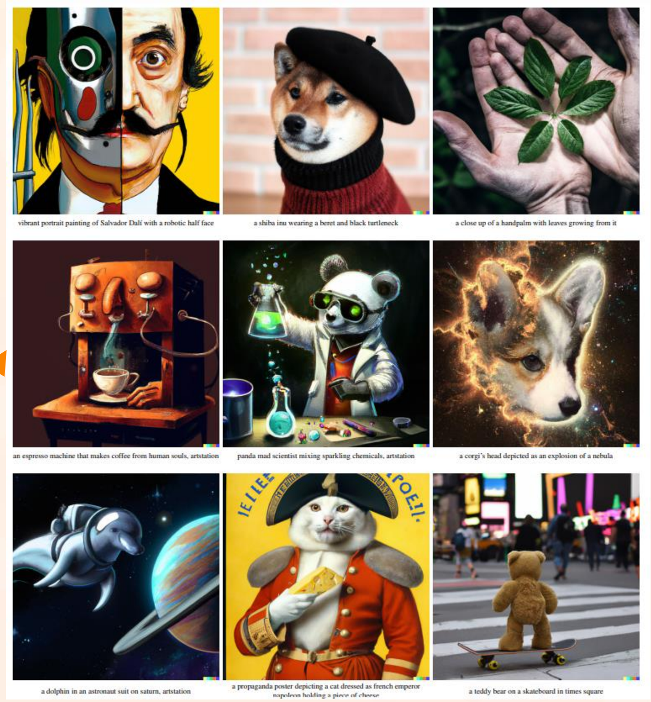
Et pleins d'autres encore



Pas que des GANs



VAE



DALL-E 2

Questions?

Merci de votre attention !

TP GAN : Devenir un artiste sans prendre de cours (de dessin) !

On va générer des chats trop mignons !!!

(=^-ω-^=)

<u>RDV Jeudi 13/10 de 14h à 16h en EA-007</u>