

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 !



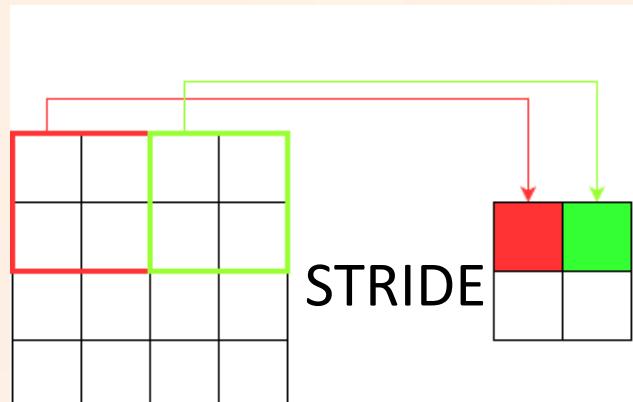
I) Rappels

Convolution

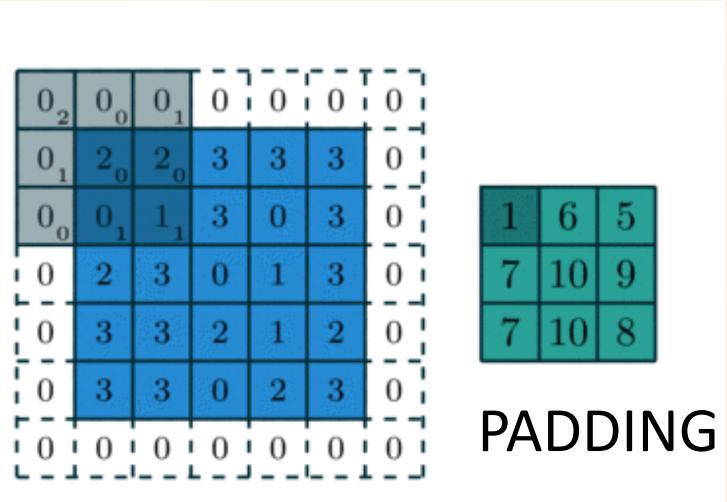


I(0,0)	I(1,0)	I(2,0)	I(3,0)	I(4,0)	I(5,0)	I(6,0)
I(0,1)	I(1,1)	I(2,1)	I(3,1)	I(4,1)	I(5,1)	I(6,1)
I(0,2)	I(1,2)	I(2,2)	I(3,2)	I(4,2)	I(5,2)	I(6,2)
I(0,3)	I(1,3)	I(2,3)	I(3,3)	I(4,3)	I(5,3)	I(6,3)
I(0,4)	I(1,4)	I(2,4)	I(3,4)	I(4,4)	I(5,4)	I(6,4)
I(0,5)	I(1,5)	I(2,5)	I(3,5)	I(4,5)	I(5,5)	I(6,5)
I(0,6)	I(1,6)	I(2,6)	I(3,6)	I(4,6)	I(5,6)	I(6,6)

Input image



Output image



PADDING

Max Pooling			
29	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

Average Pooling			
31	15	28	184
0	100	70	38
12	12	7	2
12	12	45	6

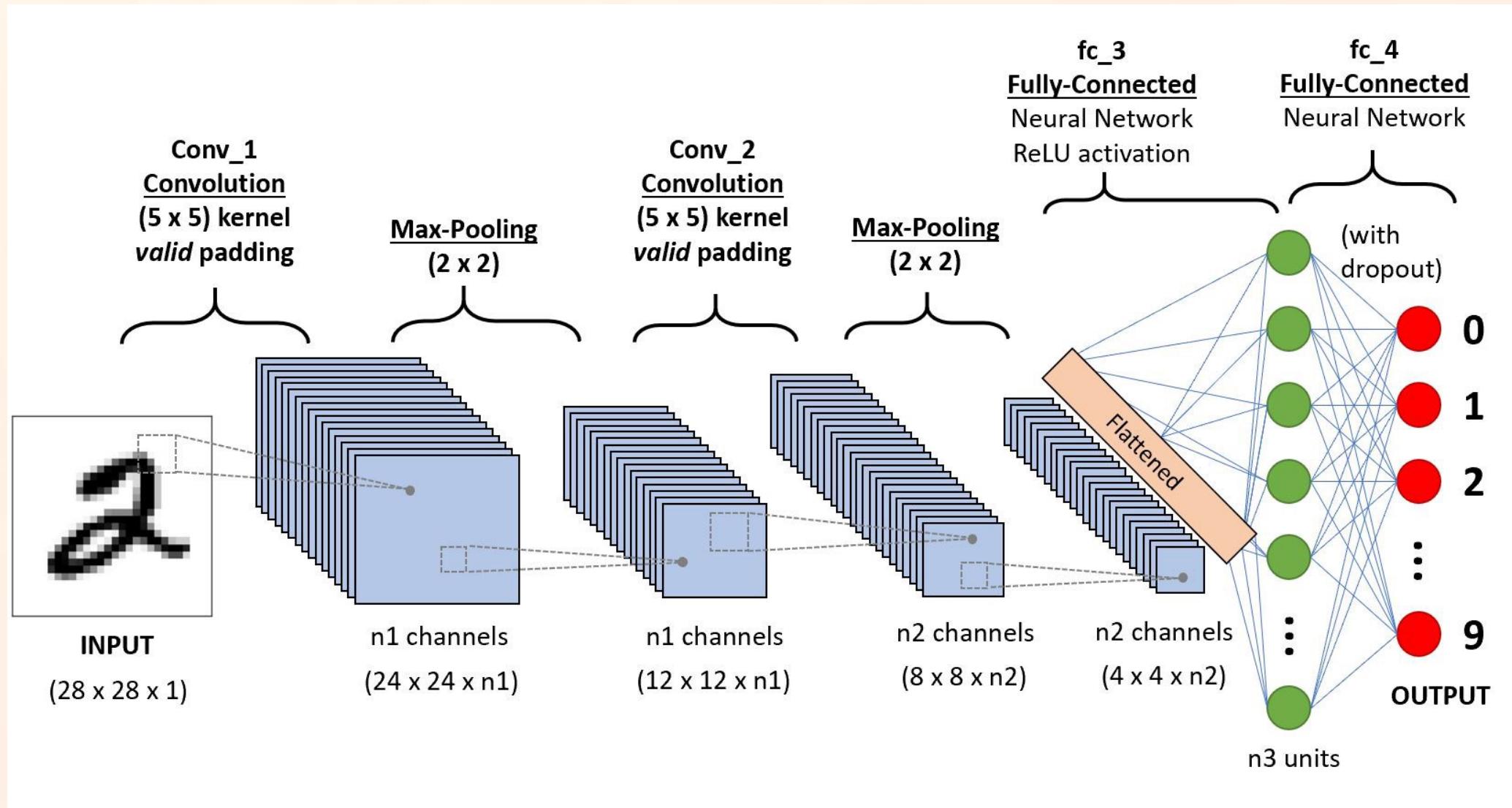
2 x 2
pool size

POOLING

2 x 2
pool size

G

CNN



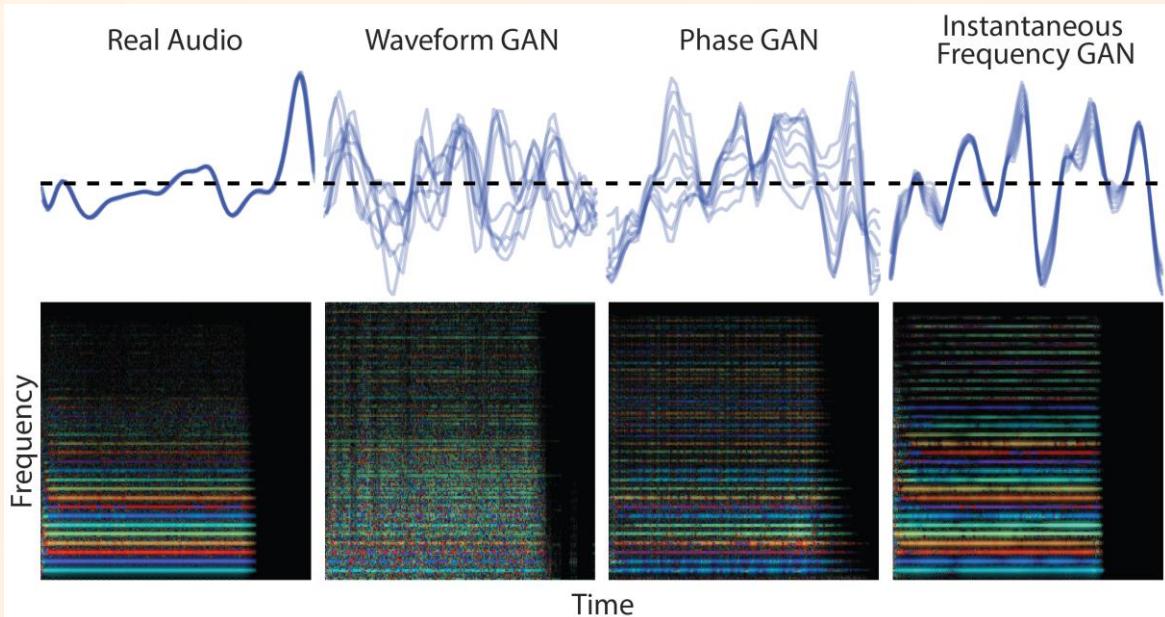


Questions?



II) Les GAN

A quoi ça sert un GAN?



IA ou humain?





Questions?



1) Principes de base

Objectif des GAN



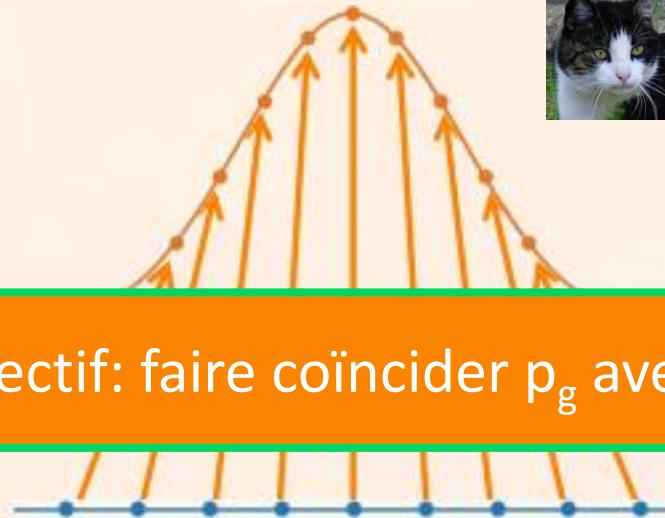
Bruit
Loi normale
 $z \in \mathbb{R}^{100}$
 p_z

GAN

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



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



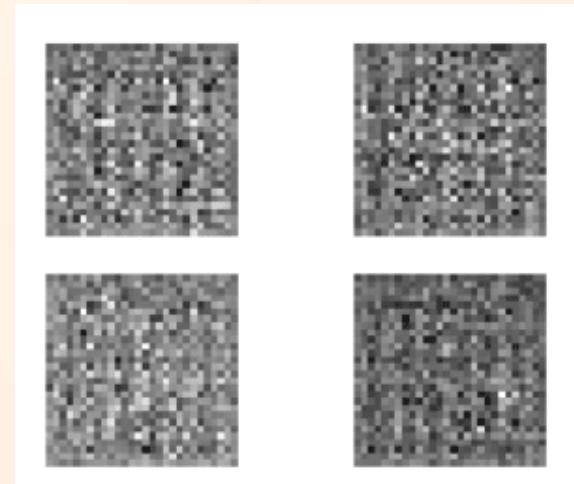
Objectif: faire coïncider p_g avec p_x



C'est l'heure du-du-du-duel



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



Comment comparer
cette image???

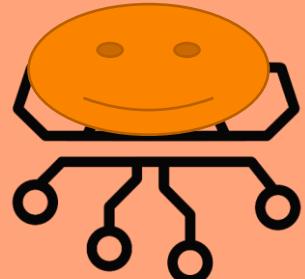
0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9

Données à
approcher

C'est l'heure du-du-du-duel



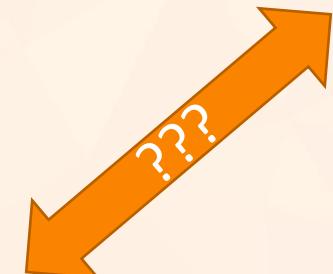
Le peintre paint



Le faux:



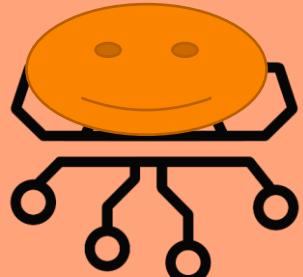
Un vrai dessin:



C'est l'heure du-du-du-duel



Le peintre paint



Un vrai dessin:



Le faux:



Le critique d'art:
WTF??

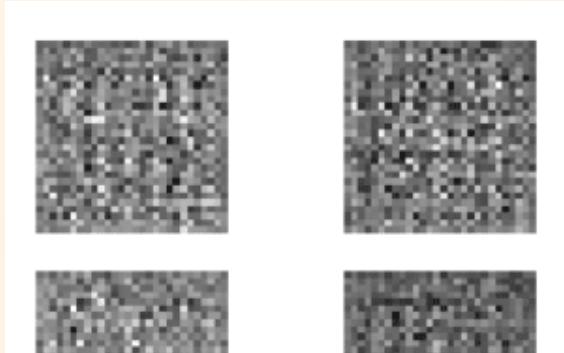


APPREND À DESSINER !!

C'est l'heure du-du-du-duel



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



0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9

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

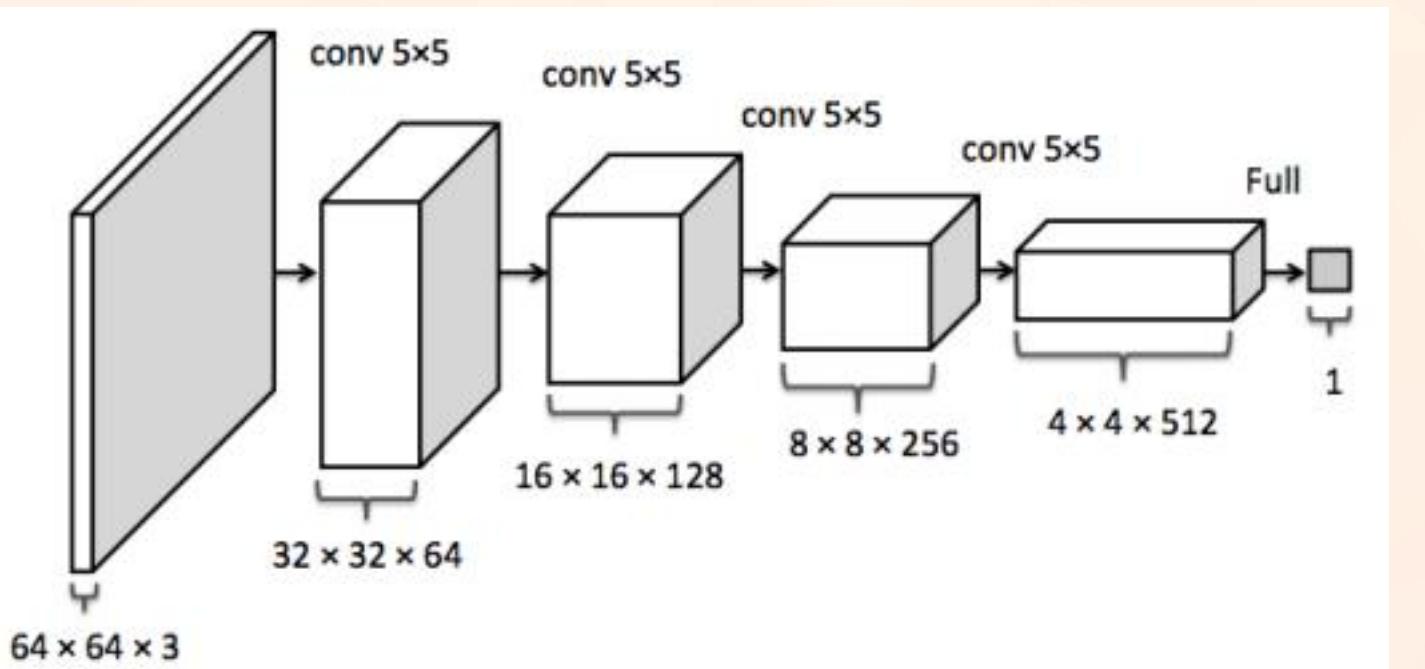


Questions?



2) En pratique

Le discriminateur



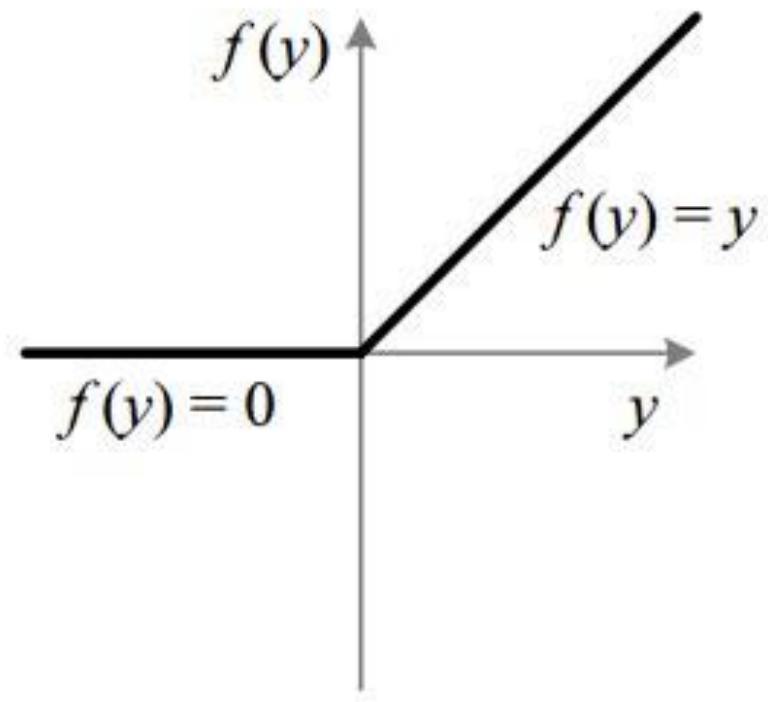
1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Image

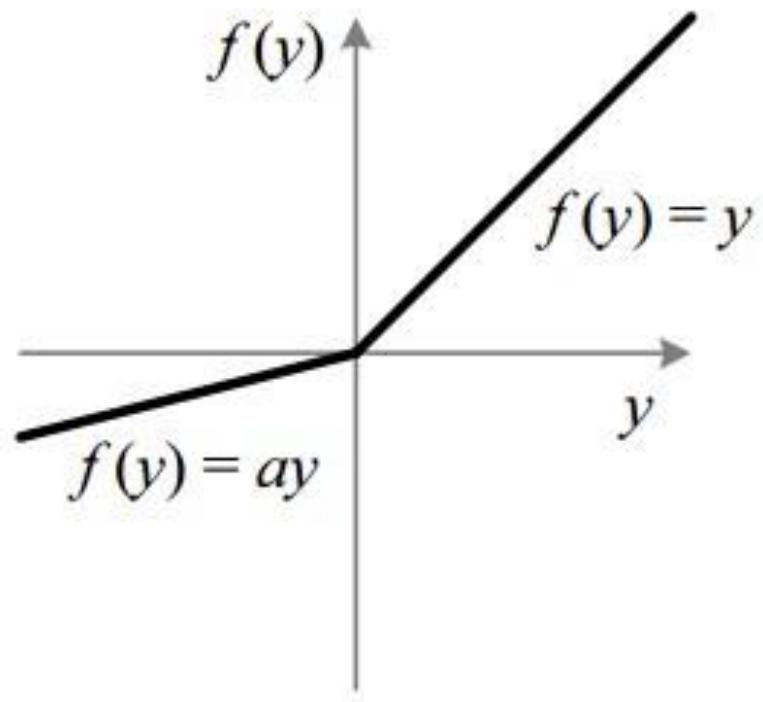
4		

Convolved Feature

Les activations

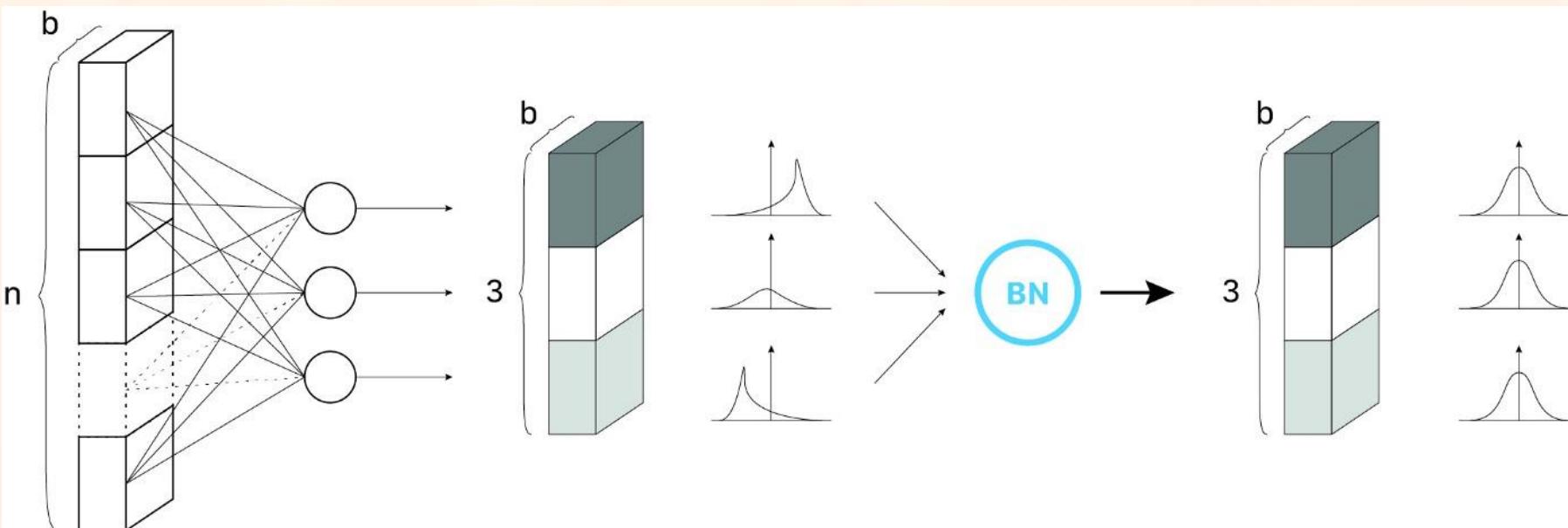


ReLU



LeakyReLU

La batchnorm



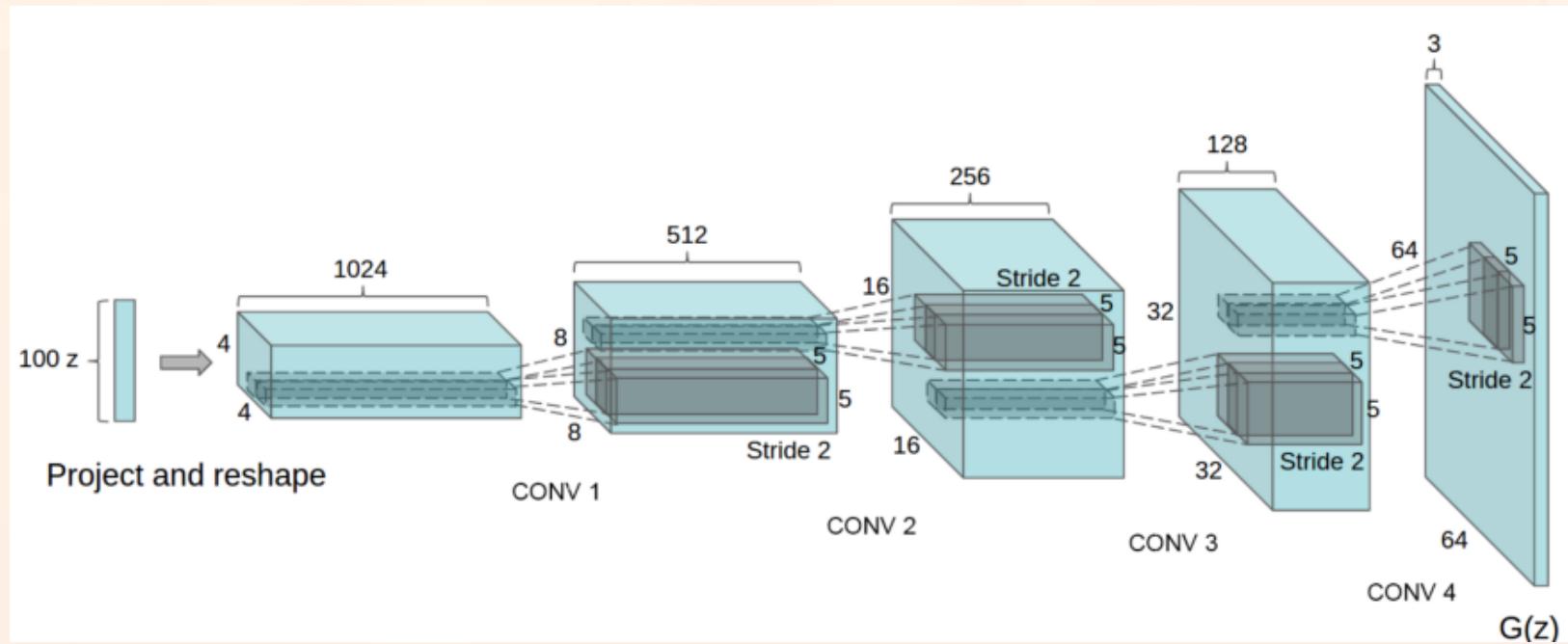
$$\mu = \frac{1}{batch_size} \sum X_i$$

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

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

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

Le générateur



Upsampling



Nearest Neighbor

1	2
1	2
3	4



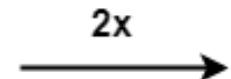
1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Input: 2 x 2

Output: 4 x 4

10	20
30	40

2x2



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:

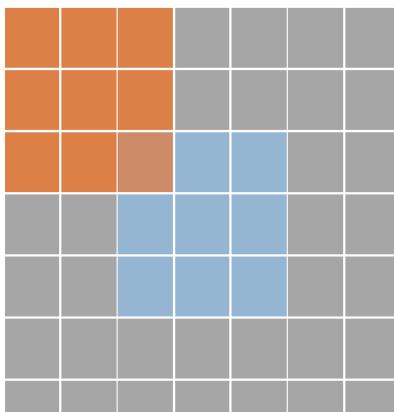
- $z = \text{stride} - 1$
- $p' = \text{kernel_size} - \text{padding} - 1$

Ajouter:

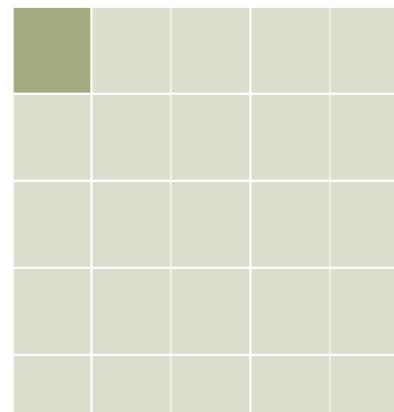
- z zeros entre les pixels
- p' zeros autour de l'image

Convolution normale

Type: transposed'conv - Stride: 1 Padding: 0



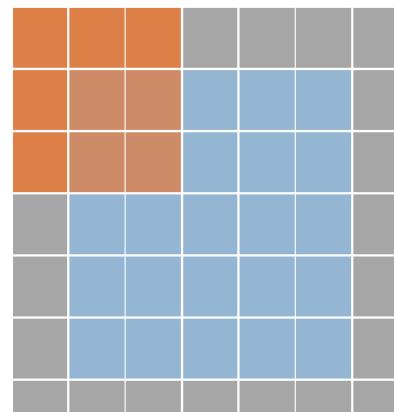
Input



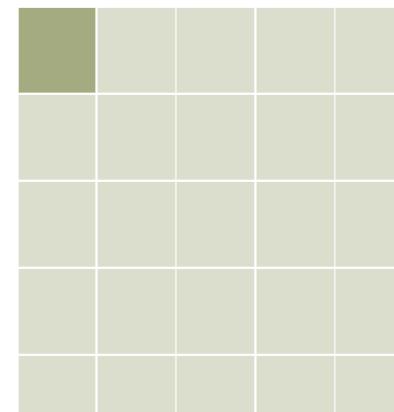
Output

$$z = 0, p' = 2$$

Type: transposed'conv - Stride: 1 Padding: 1



Input



Output

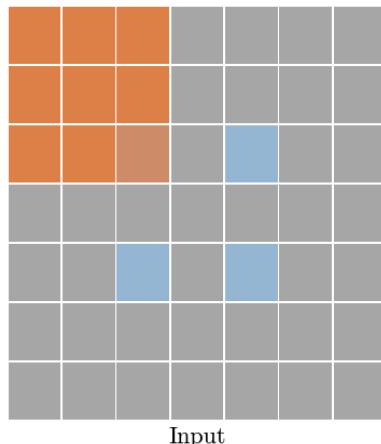
$$z = 0, p' = 1$$

Les convolutions transposées

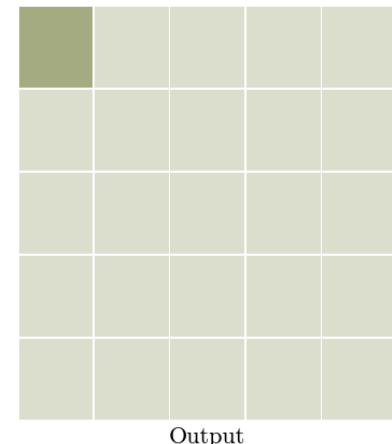


Avec du stride

Type: transposed'conv - Stride: 2 Padding: 0

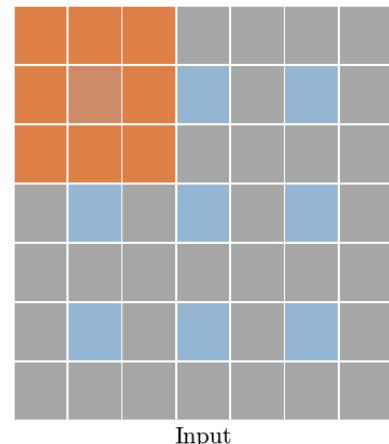


Input

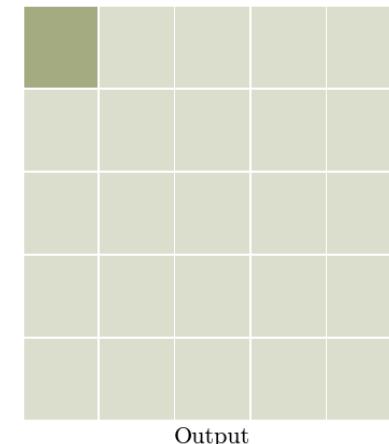


Output

Type: transposed'conv - Stride: 2 Padding: 1



Input

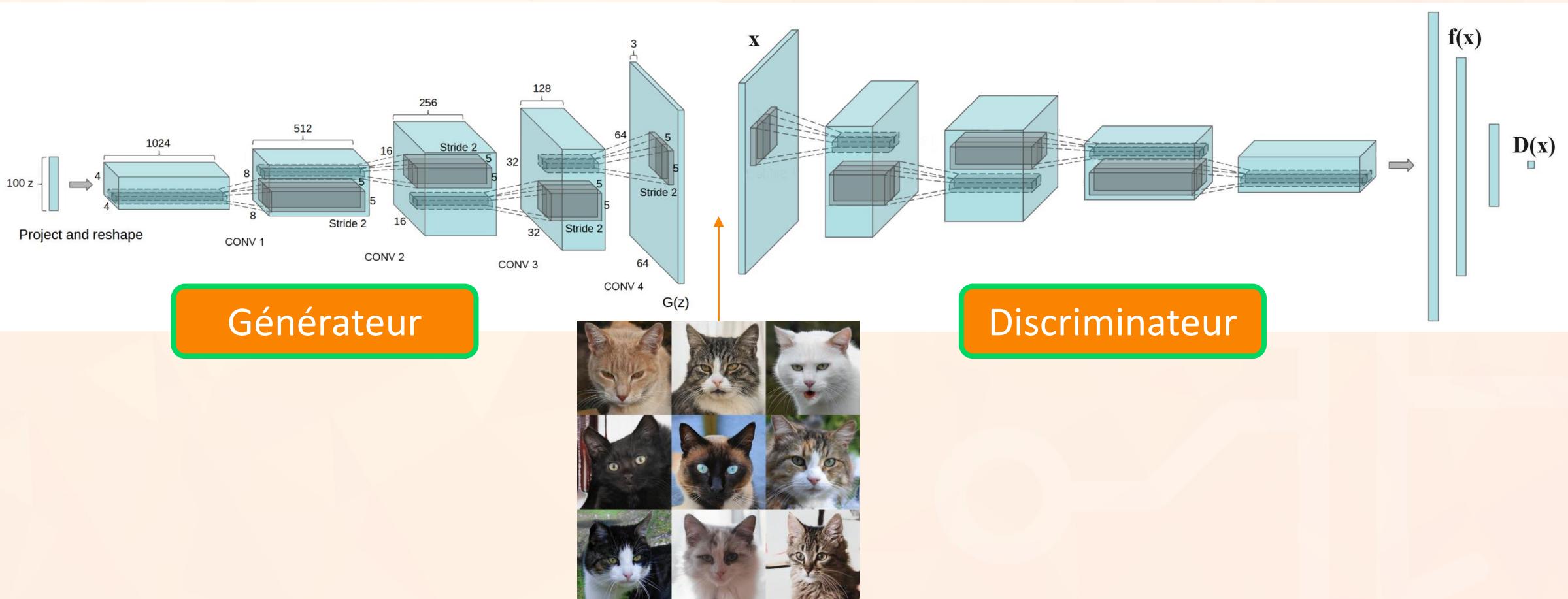


Output

$z = 1, p' = 2$

$z = 1, p' = 1$

Vue d'ensemble



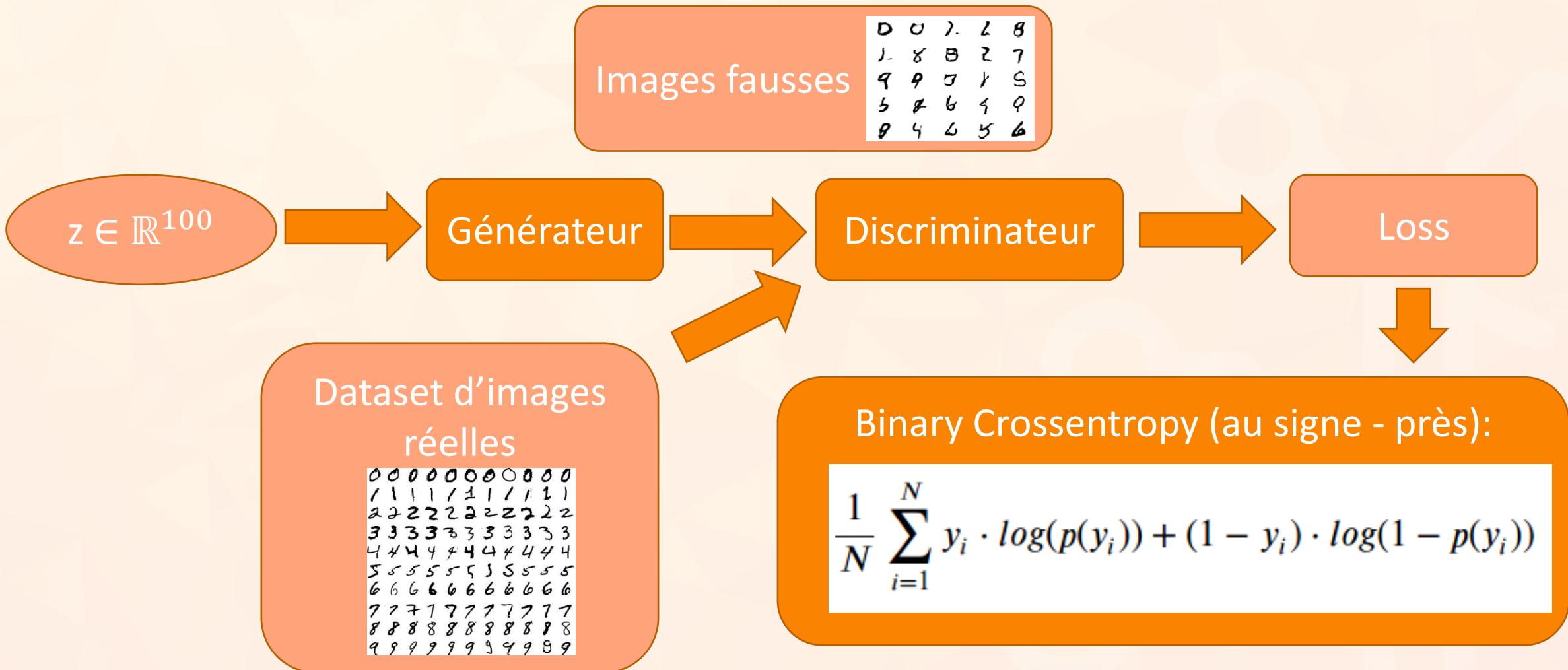


Questions?

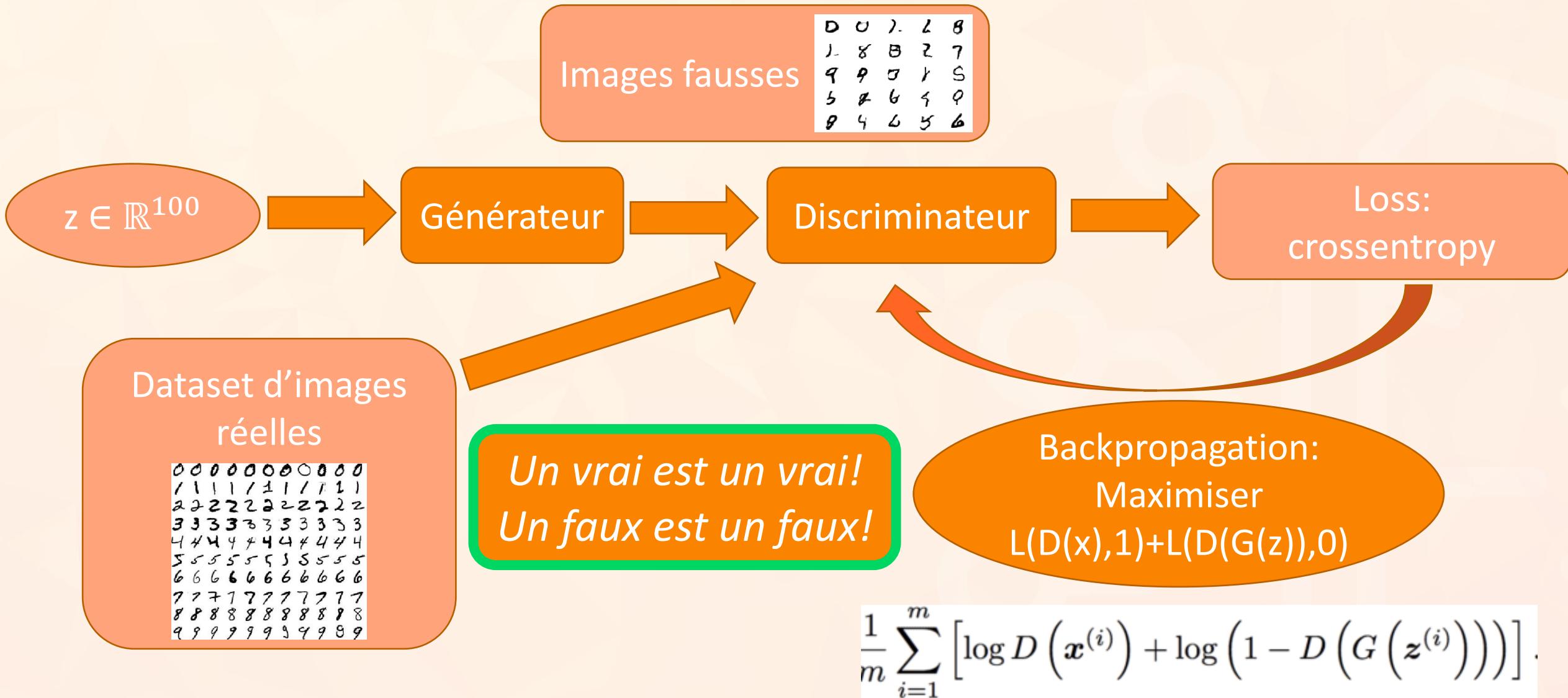


3) Entraînement

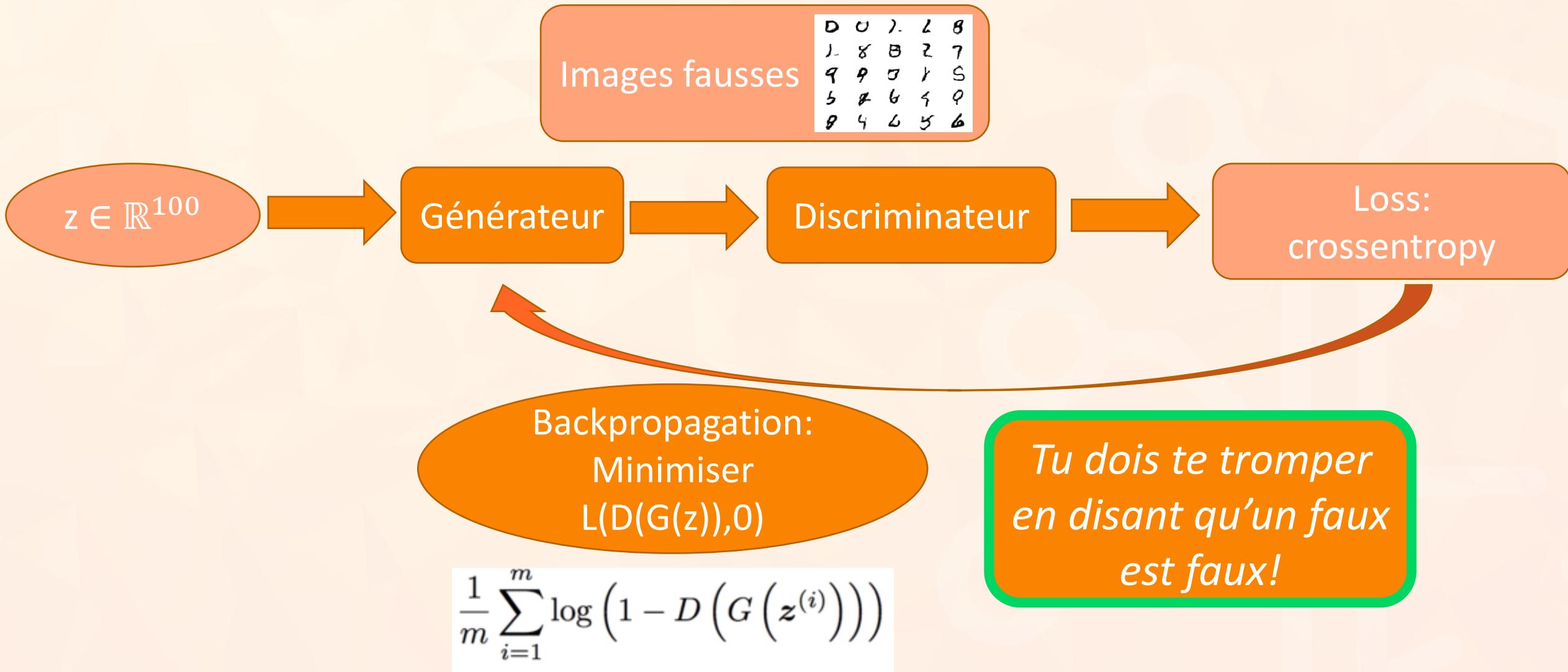
La loss



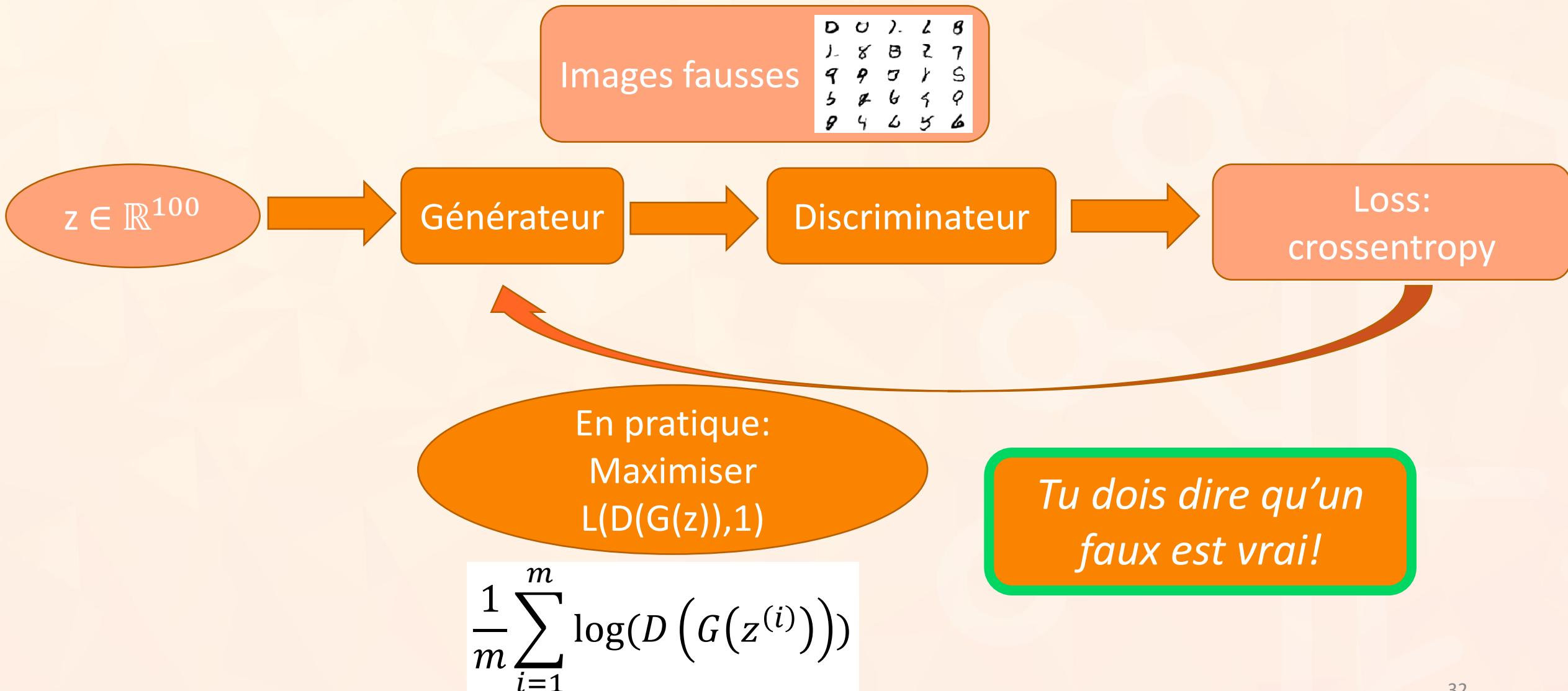
Le discriminateur apprend



Le générateur apprend



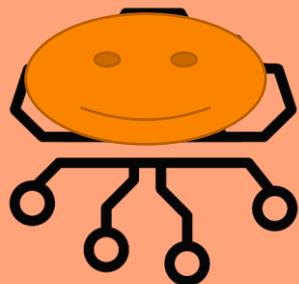
Le générateur apprend



Le combat final



Générateur



Discriminateur



Veut minimiser



Veut maximiser



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



Questions?

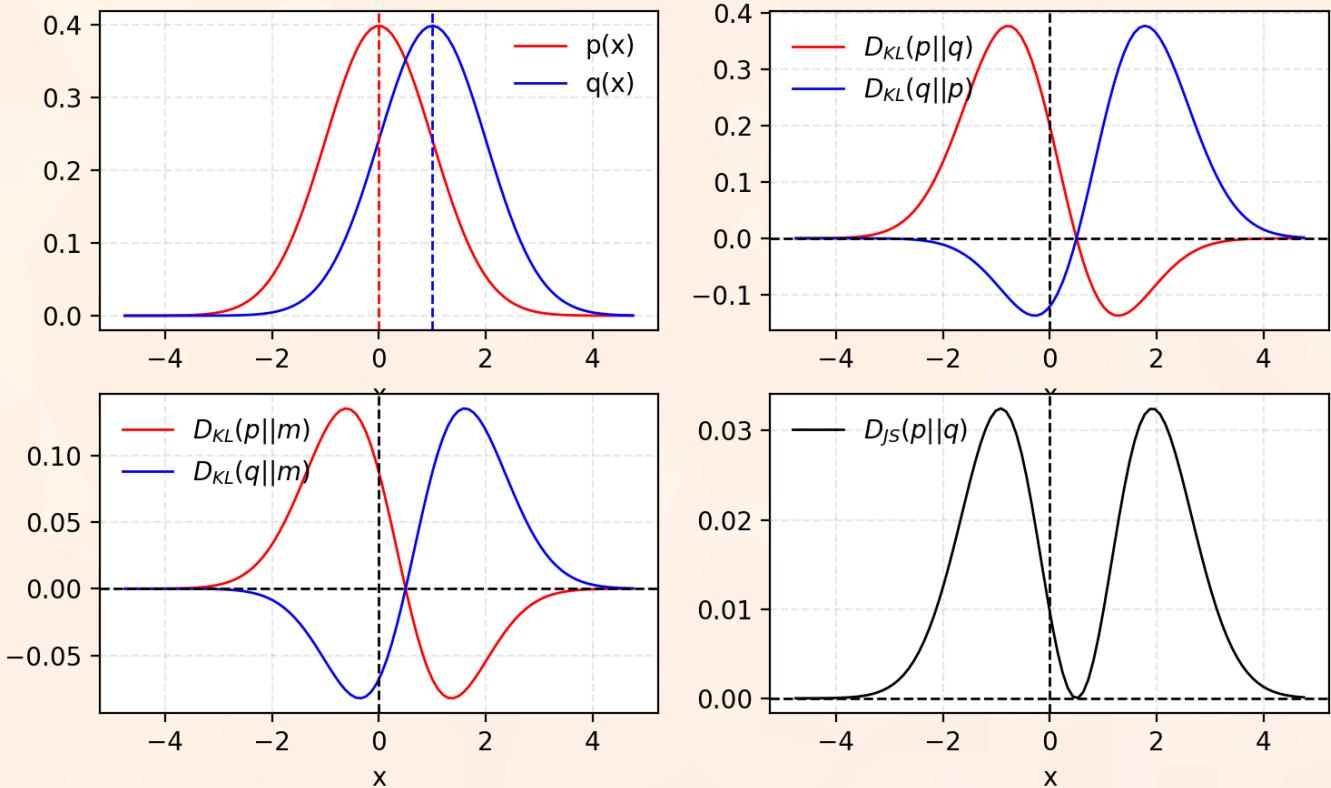


4) Interlude : maths

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



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



x faux = G(z) avec z vecteur latent

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



Espérance en intégrale

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

Valeur optimale



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



On dérive ce truc

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

Si $p_g = p_{data}$, $D^*(x) = \frac{1}{2}$
Et $L = -2\log 2 !!$

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

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



Questions?

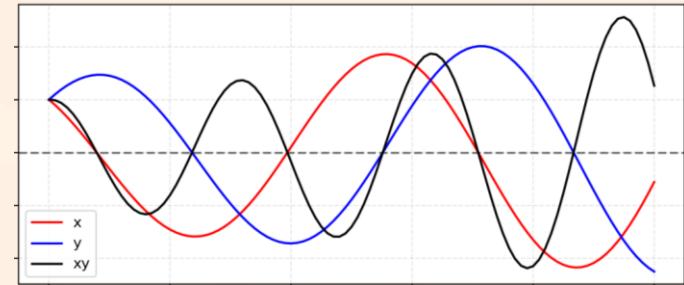


5) Des GAN et des problèmes

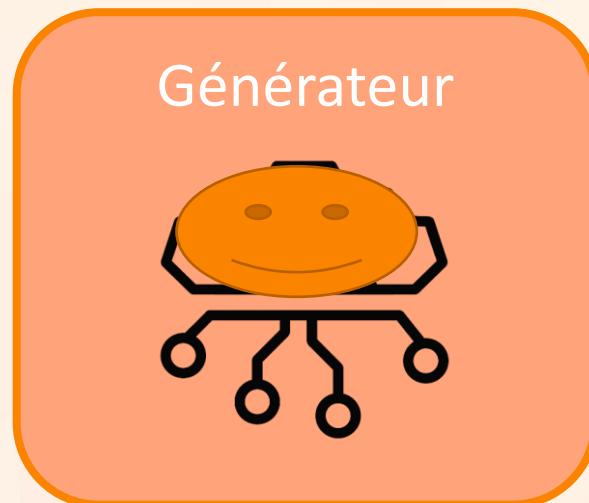
Non convergence



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

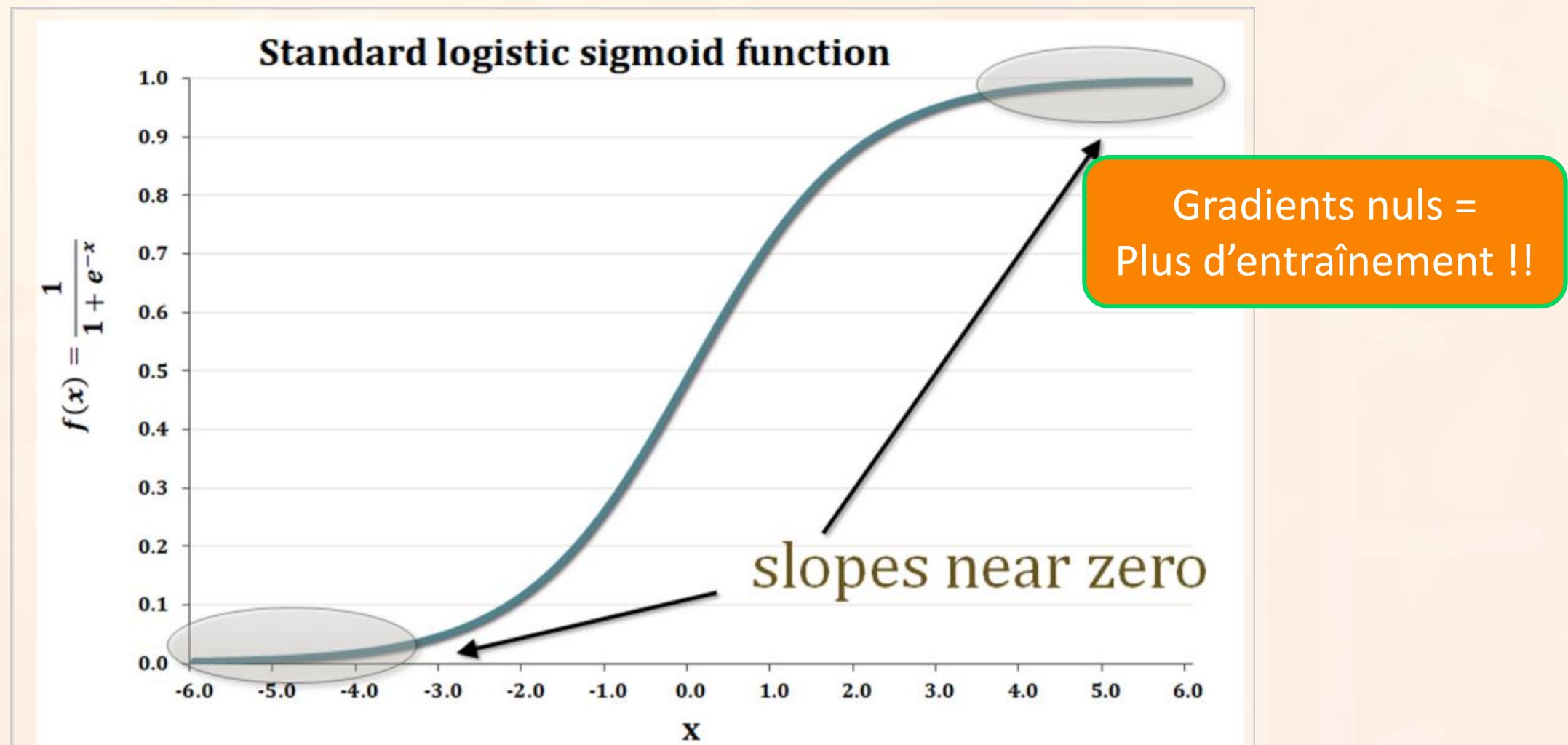


Problème dû à la descente de gradient

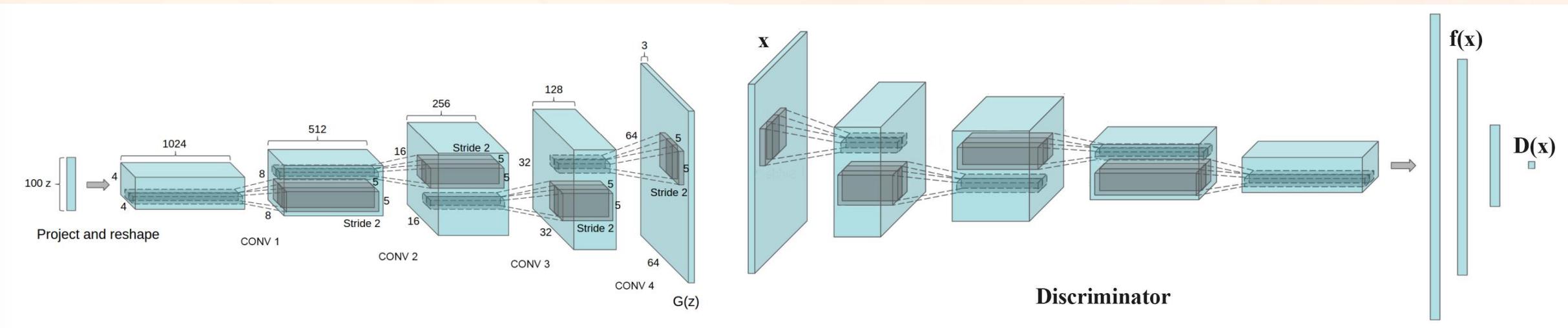


Equilibre fragile!

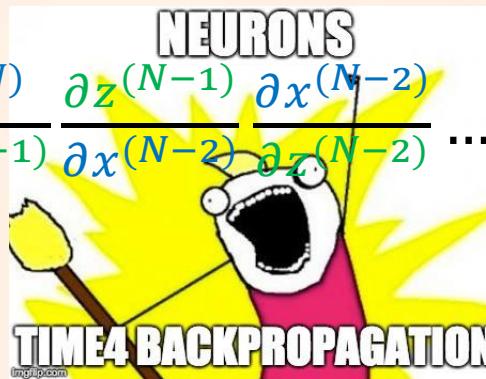
Vanishing gradient



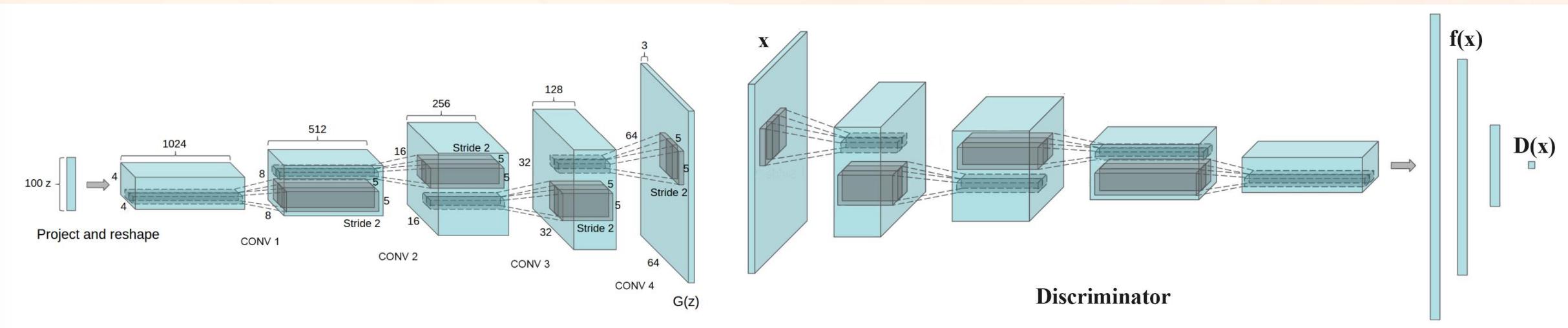
Vanishing gradient



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



Vanishing gradient



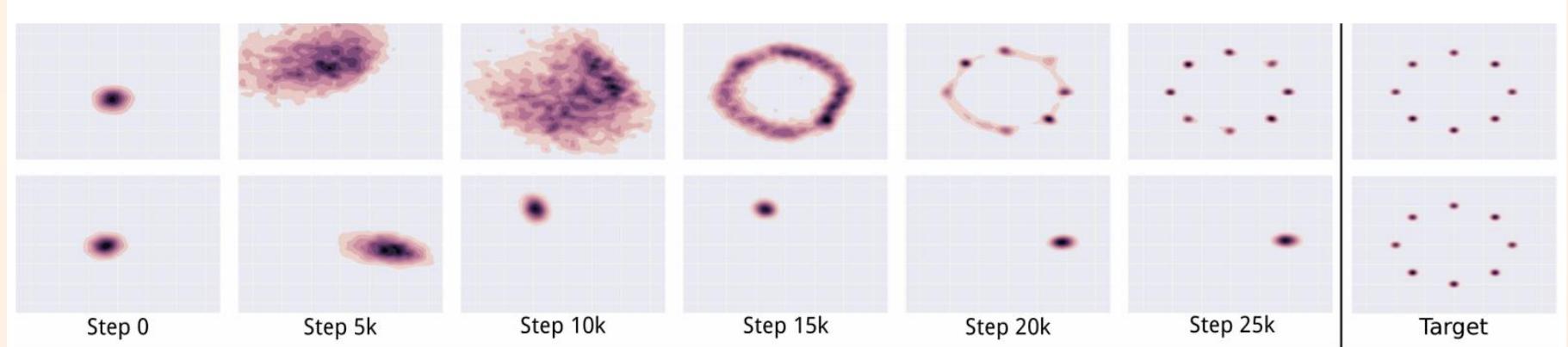
$$\frac{\partial L}{\partial w^{(i)}} = \underbrace{\frac{\partial L}{\partial x^{(N)}}}_{\sim 0} \frac{\partial x^{(N)}}{\partial z^{(N)}} \frac{\partial z^{(N)}}{\partial x^{(N-1)}} \frac{\partial x^{(N)}}{\partial z^{(N-1)}} \frac{\partial z^{(N-1)}}{\partial x^{(N-2)}} \frac{\partial x^{(N-2)}}{\partial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

Tout tend vers 0!!

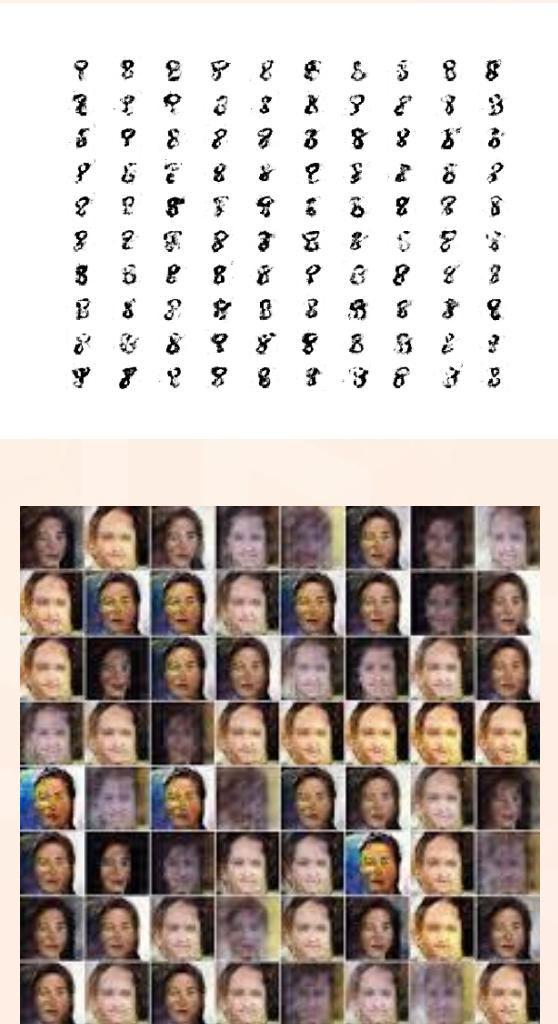
Mode collapse



«Bon » entraînement



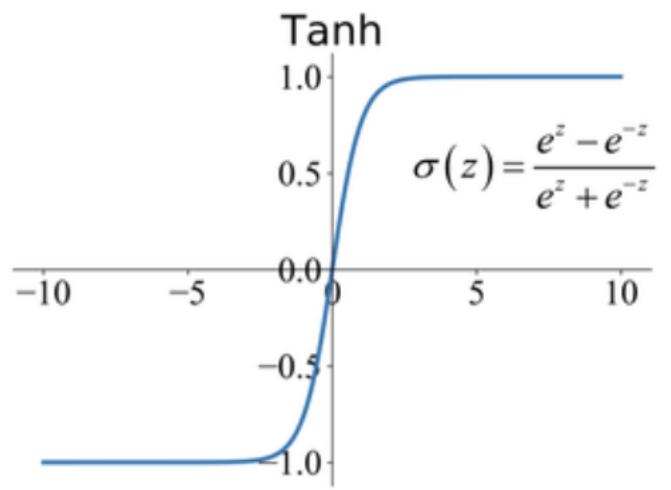
Mode collapse



Les petites astuces



Normaliser entre -1 et 1



Discriminator/Critic

GAN

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

WGAN

$$\nabla_w \frac{1}{m} \sum_{i=1}^m [f(x^{(i)}) - f(G(z^{(i)}))]$$

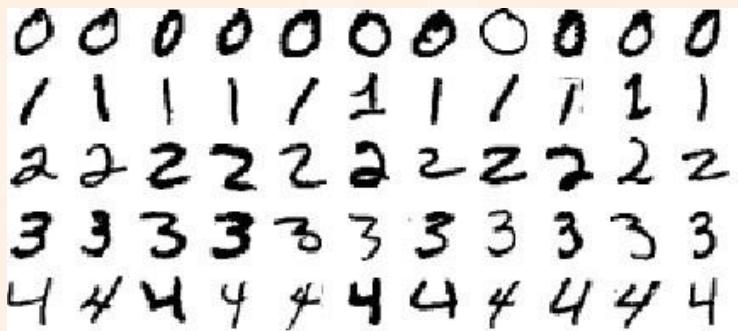
$$E_{x \sim p_{data}(x)}[0.9 \log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$

One-sided label smoothing

Feature matching

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

Ajouter labels



Generator

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log (D(G(\mathbf{z}^{(i)})))$$

$$\nabla_{\theta} \frac{1}{m} \sum_{i=1}^m f(G(z^{(i)}))$$

WGAN

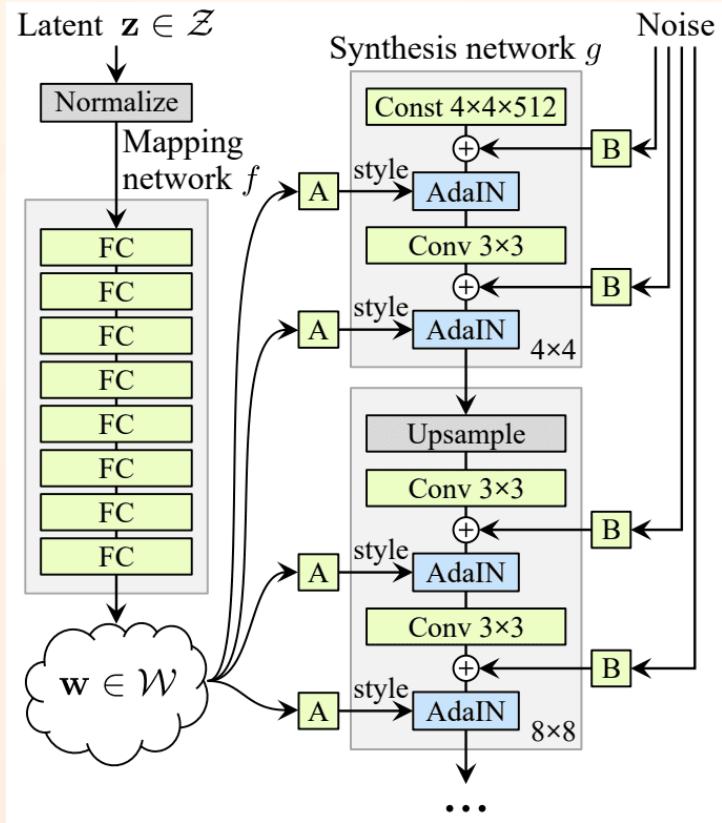


Questions?

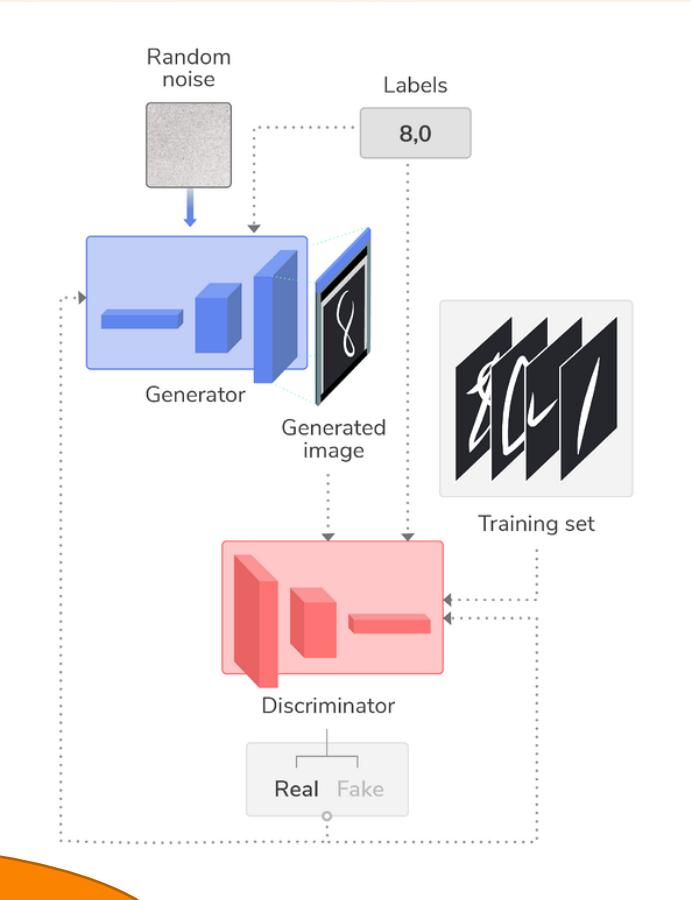


6) Y en a encore !

Génération +

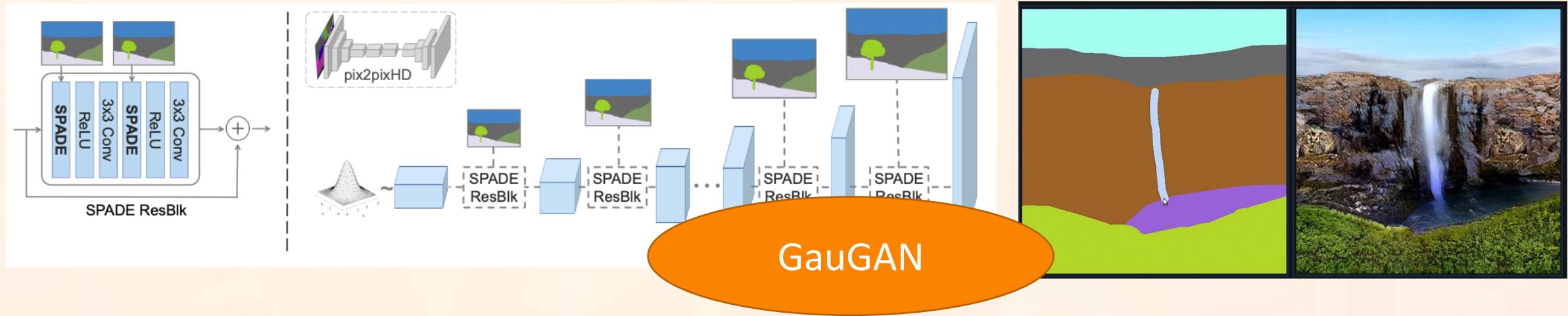


Style GAN

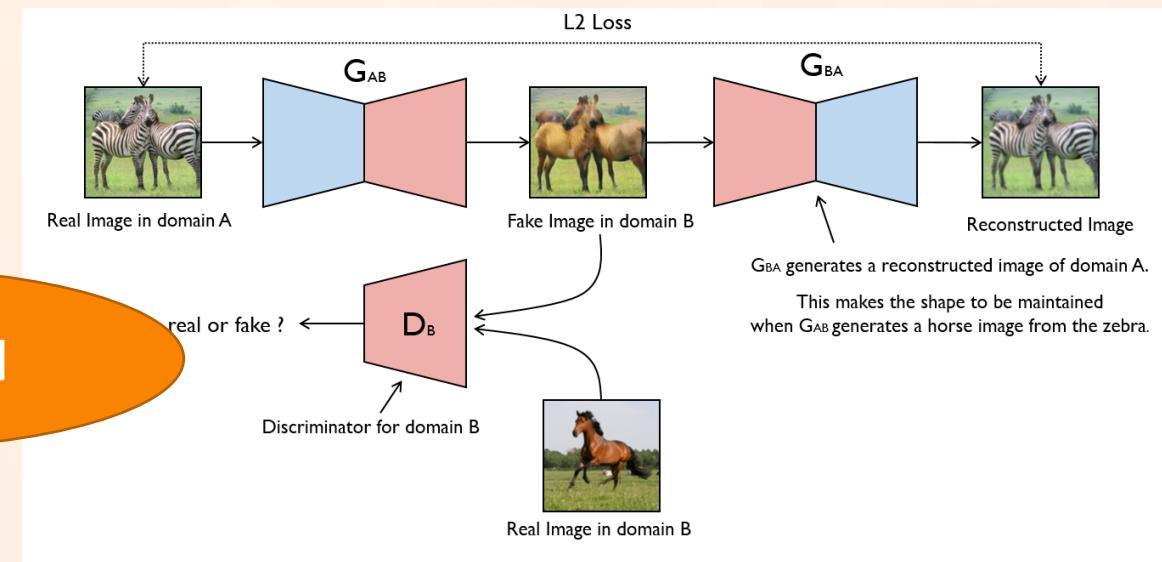


Conditional GAN

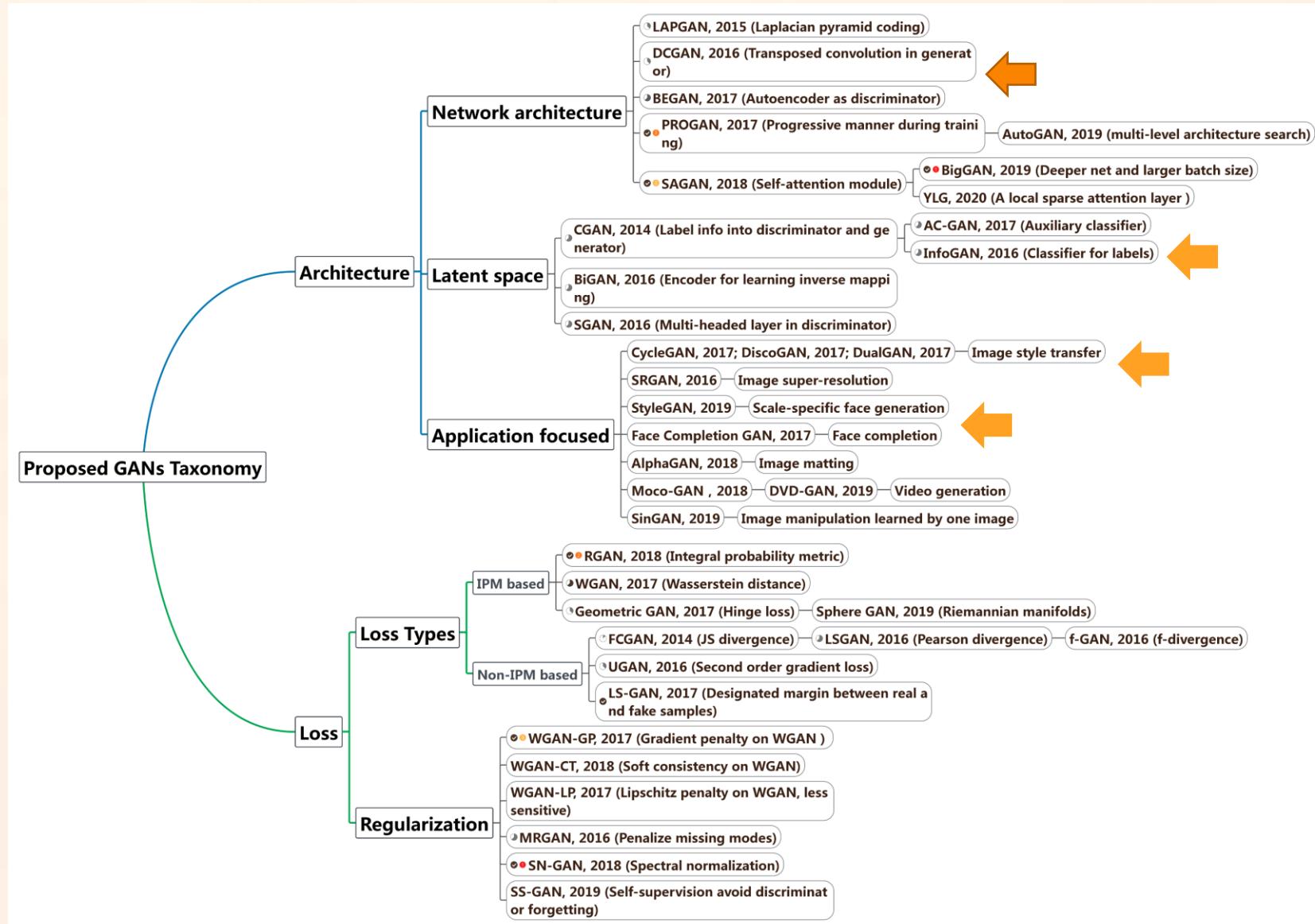
Autres trucs marrants



Cycle GAN



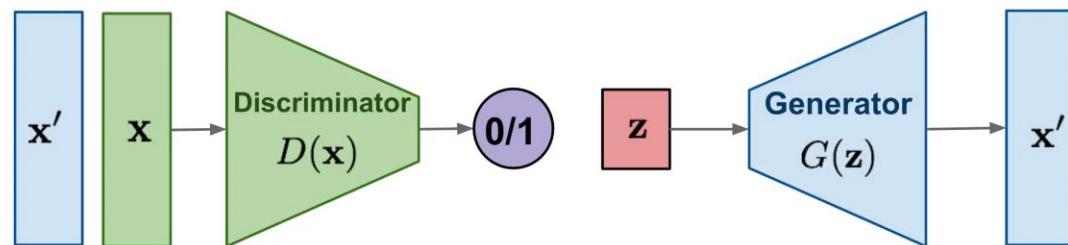
Et pleins d'autres encore



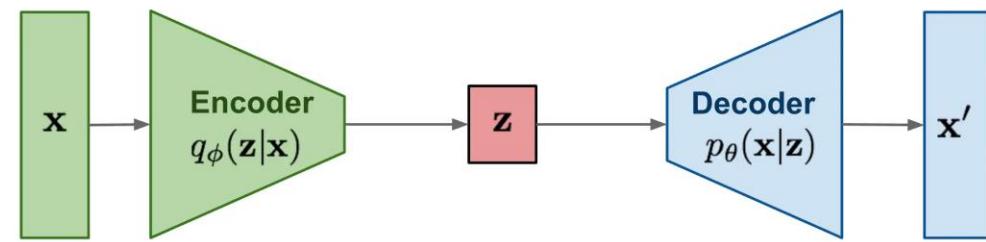
Pas que des GANs



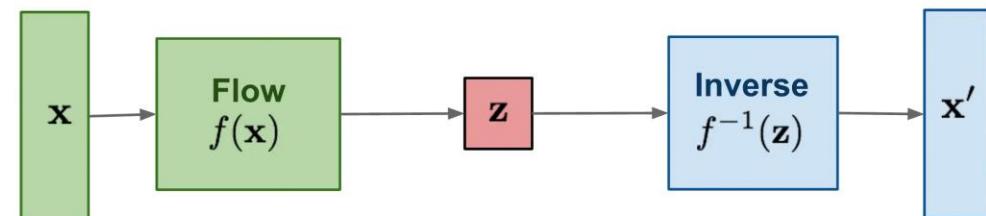
GAN: Adversarial training



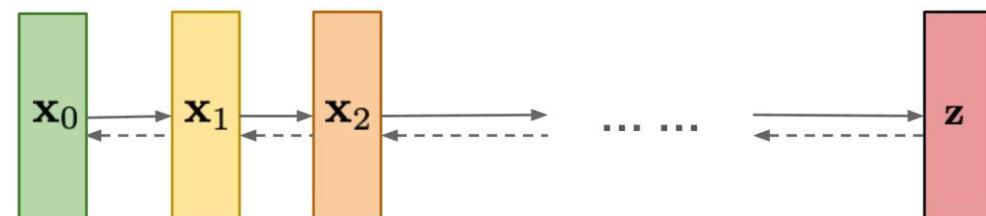
VAE: maximize variational lower bound



Flow-based models:
Invertible transform of distributions



Diffusion models:
Gradually add Gaussian noise and then reverse



VAE



vibrant portrait painting of Salvador Dalí with a robotic half face
a shiba inu wearing a beret and black turtleneck
a close up of a hand with leaves growing from it



an espresso machine that makes coffee from human souls, artstation
panda mad scientist mixing sparkling chemicals, artstation
a corgi's head depicted as an explosion of a nebula



a dolphin in an astronaut suit on saturn, artstation
a propaganda poster depicting a cat dressed as french emperor napoleon holding a piece of cheese
a teddy bear on a skateboard in times square

DALL-E 2



Questions?



Merci de votre
attention !

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