

Technique modern

Autor: Cristopher McIntyre Garcia

Mini-Cours

- Jour 1: Introduction à l'image digital
- Jour 2: Techniques de base de la vision artificiel et au traitement d'image
- Jour 3: Filtre et technique traditionnel avancé
- Jour 4: Introduction a l'intelligence artificiel (ML et DL)
- Jour 5: Technique modern - Conversation de nos courrent

Inpainting

<https://arxiv.org/abs/1806.03589>

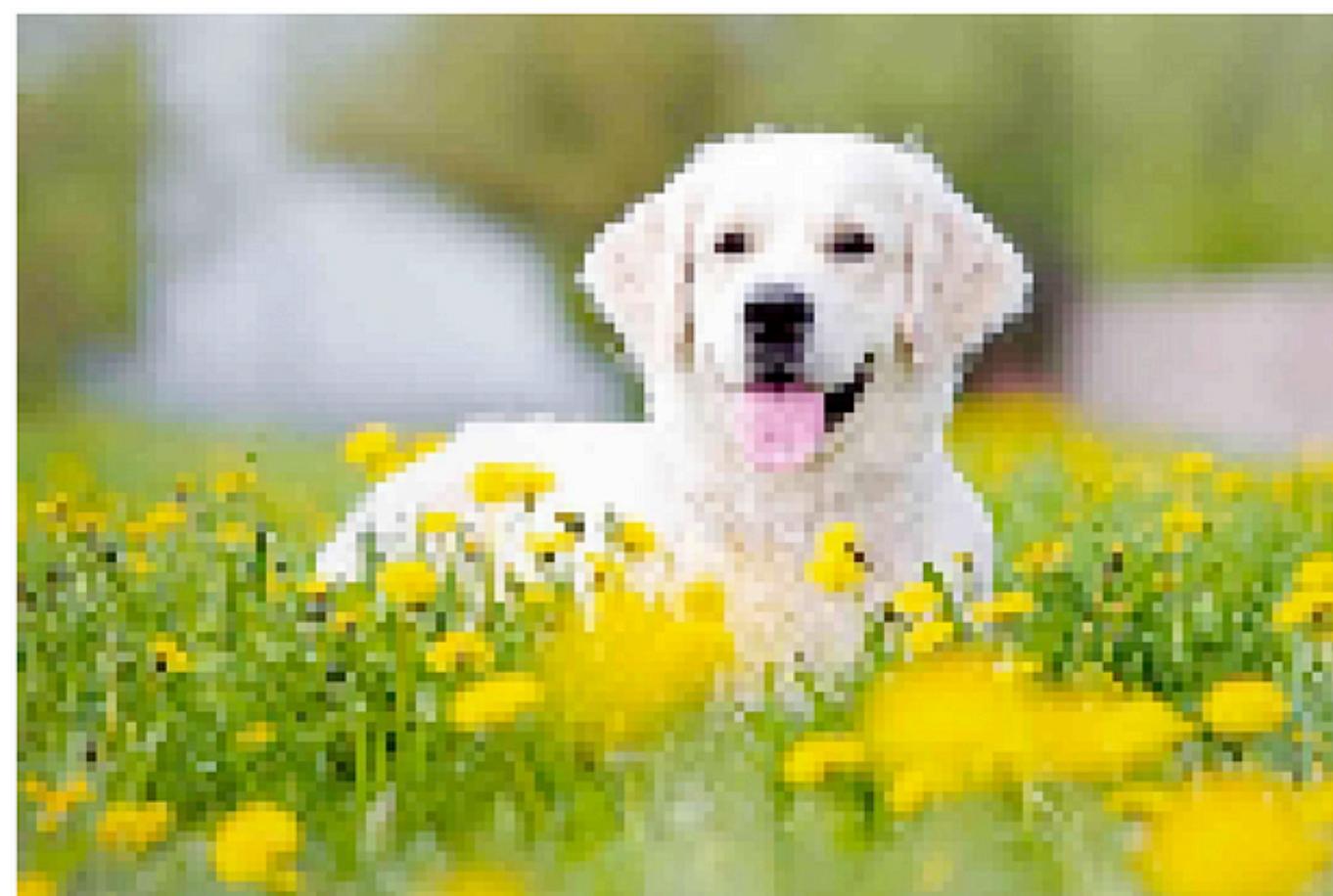
- Reconstruire les parties manquantes ou endommagées des images
- Remplir les trous d'une manière continue et indiscernable



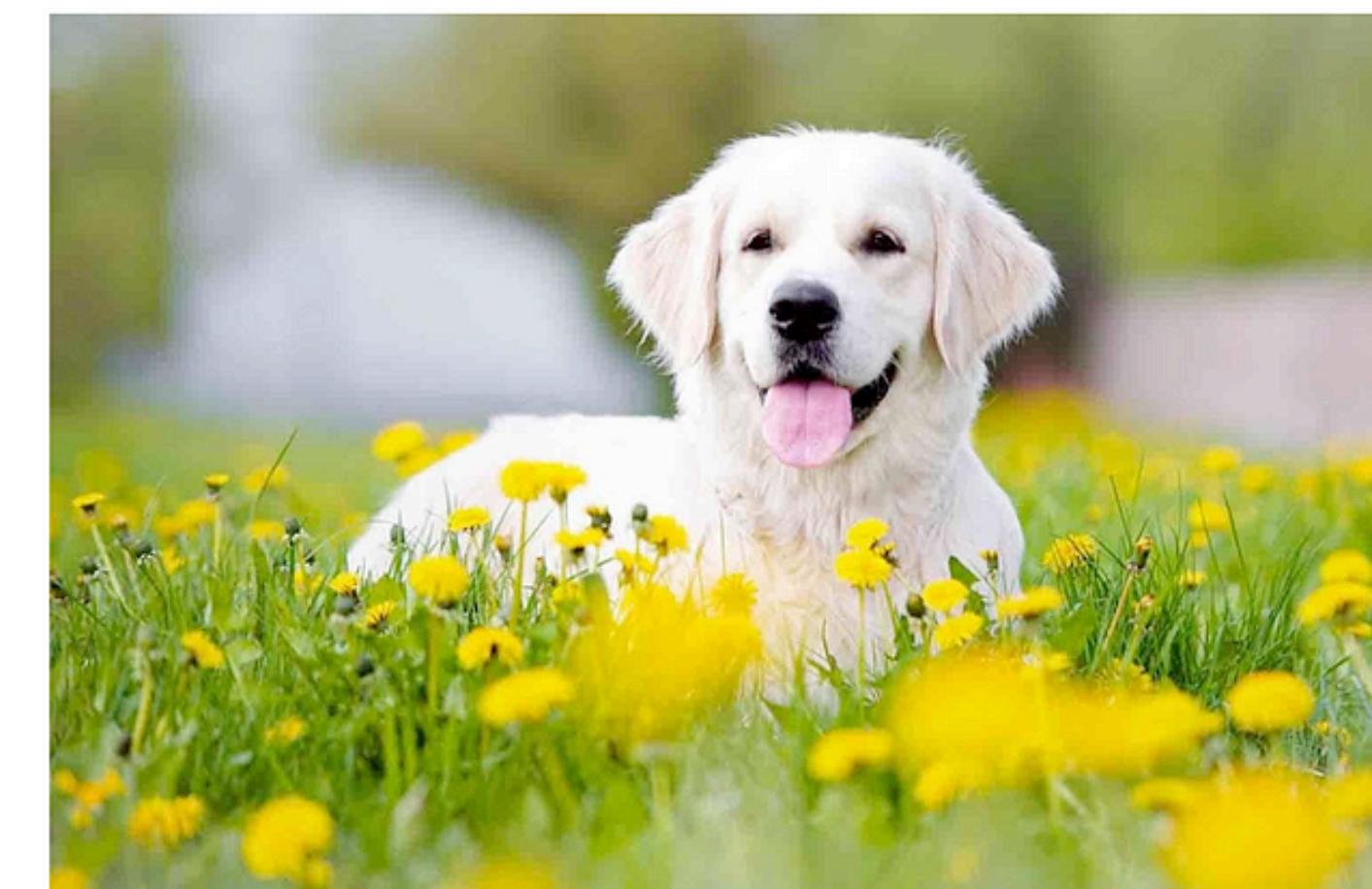
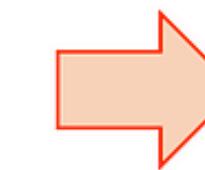
Super résolution

<https://arxiv.org/abs/2012.09841>

- Améliorer la résolution d'une image
- Transformer une image à basse résolution en une image à haute résolution
- Préserver et améliorer les détails fins et les textures



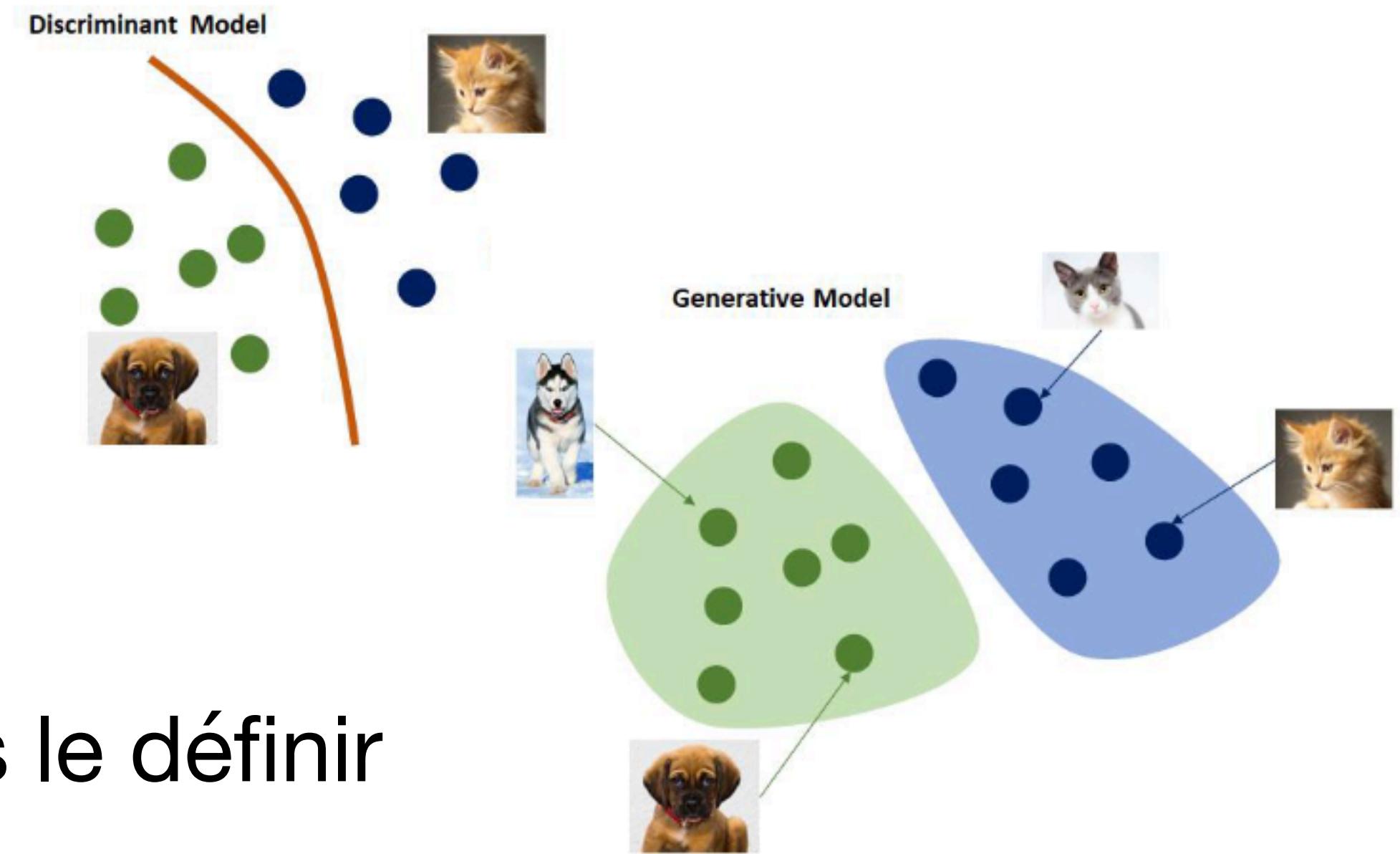
Low Resolution



High Resolution

Modèle génératif

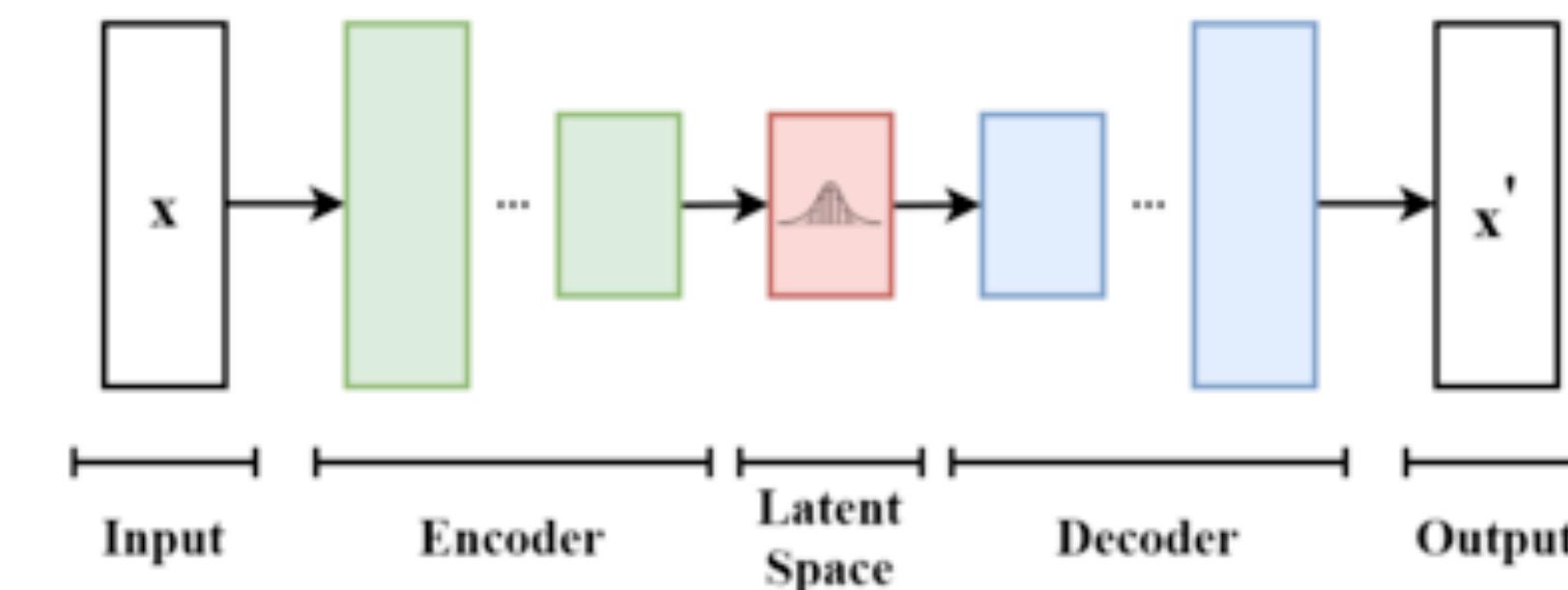
- Apprentissage non supervisé
- Estimation explicite de la densité
- Estimation implicite de la densité
 - Échantillonner à partir de $P_{model}(x)$ sans le définir
 - Échantillonner du vecteur z à partir d'une distribution fixe
 - Correspondre z à un x dans l'espace des données



Autoencodeur variationnel

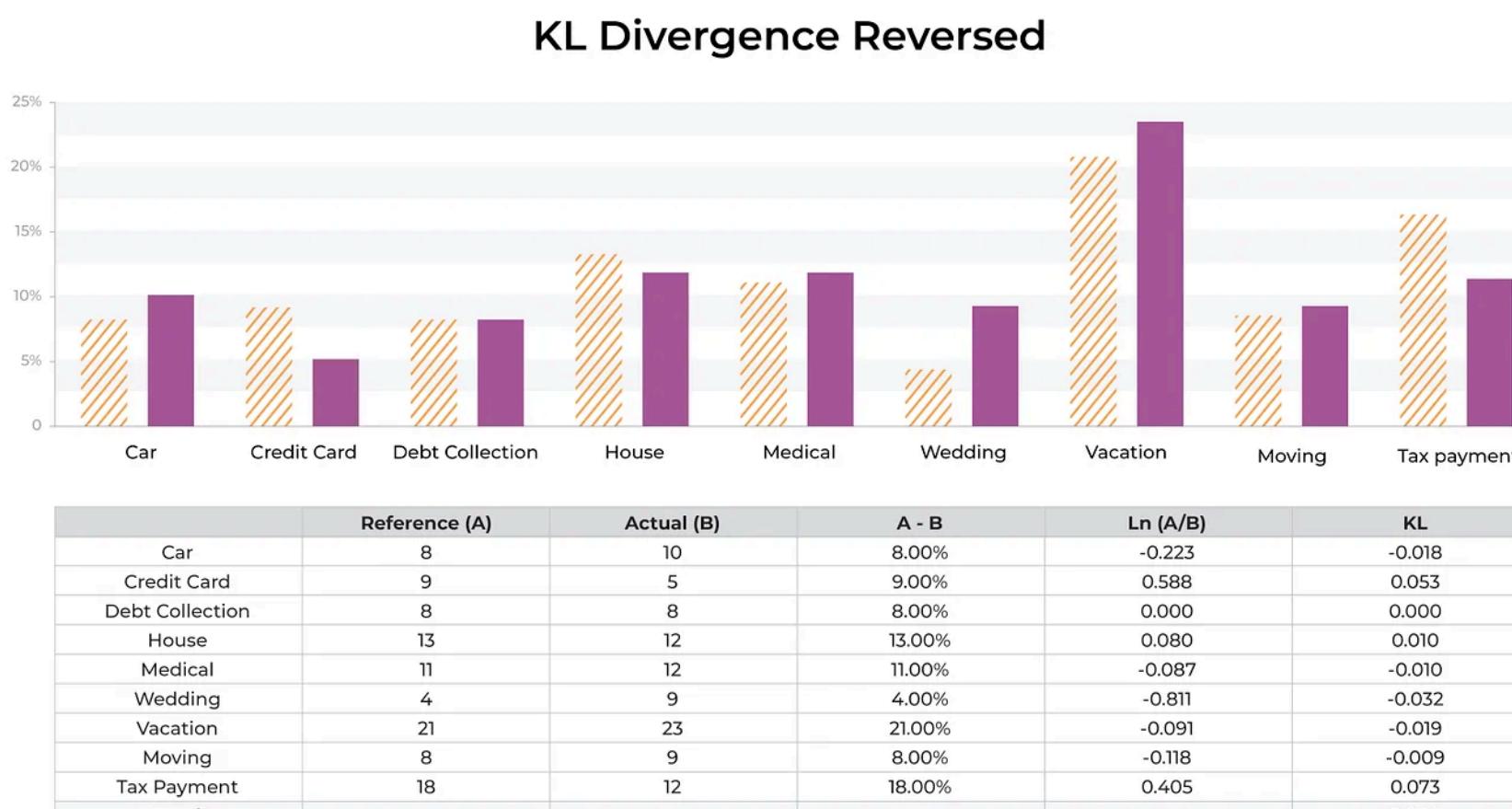
<https://arxiv.org/abs/1312.6114>

- Encodeur et d'un décodeur
 - Comprime les données dans l'espace latent
 - Reconstruit les données
 - Distributions probabilistes
 - Stochasticité

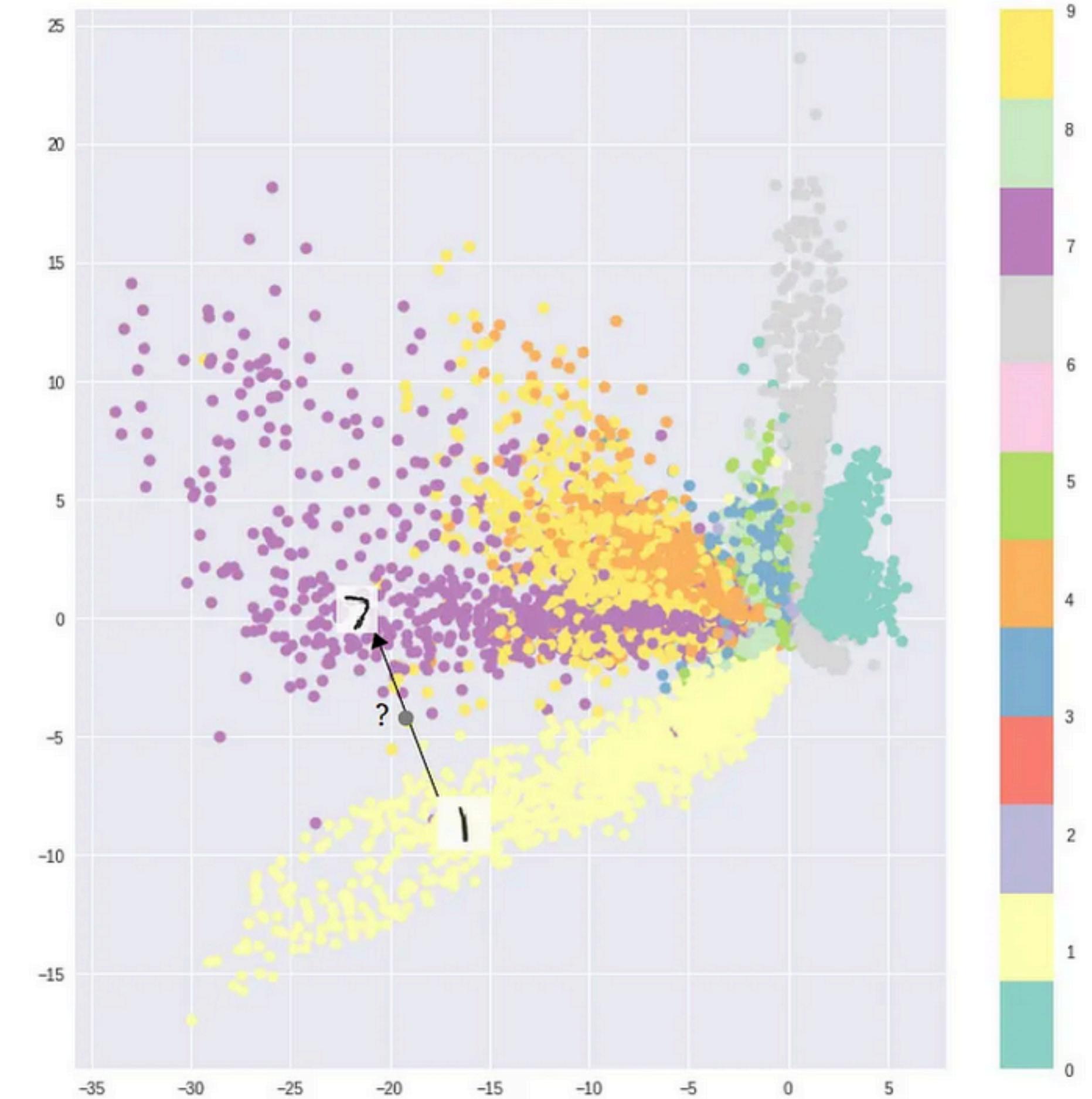


Autoencodeur variationnel

- Optimise deux objectifs principaux
 - Perte de reconstruction
 - Divergence de Kullback-Leibler (KL)

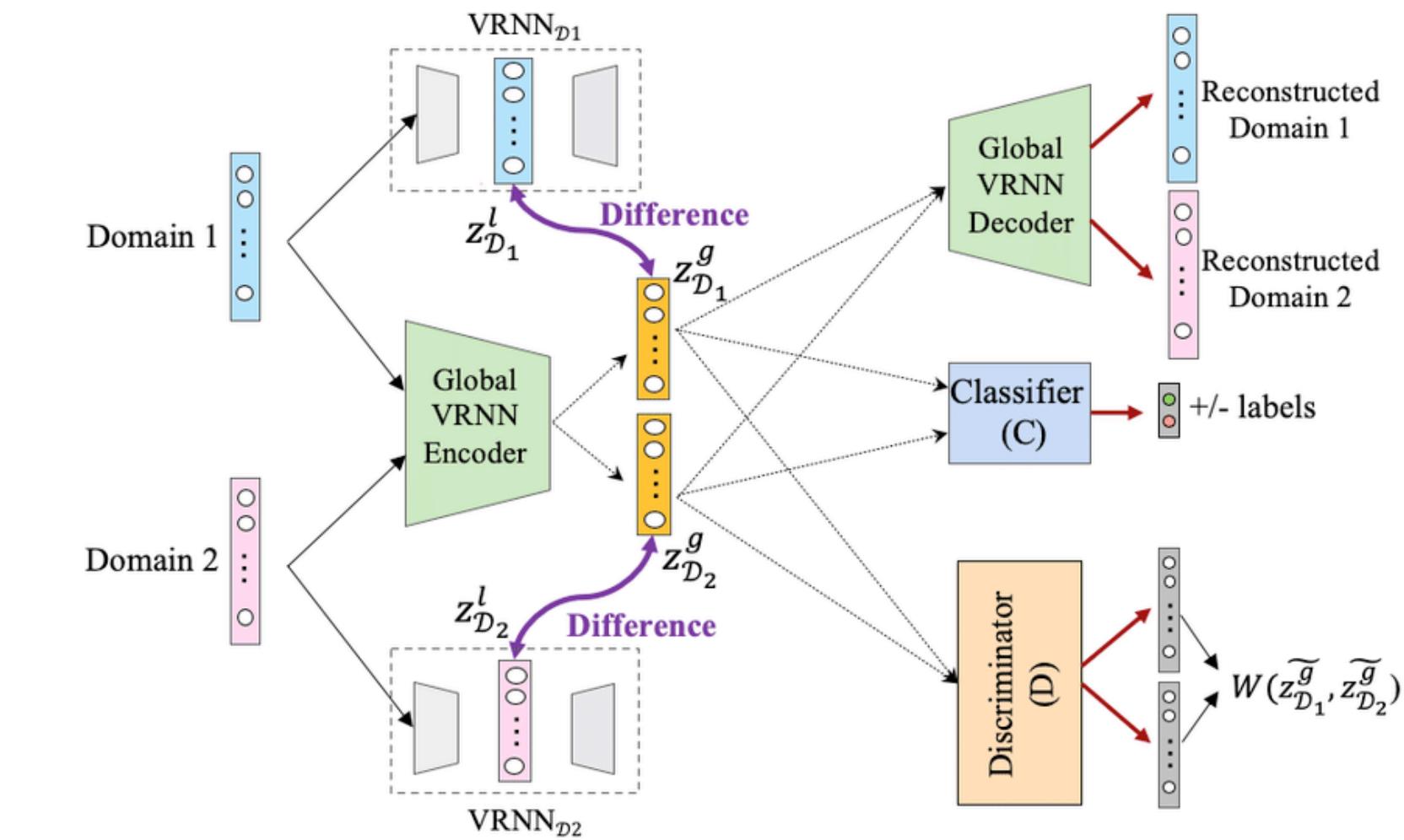
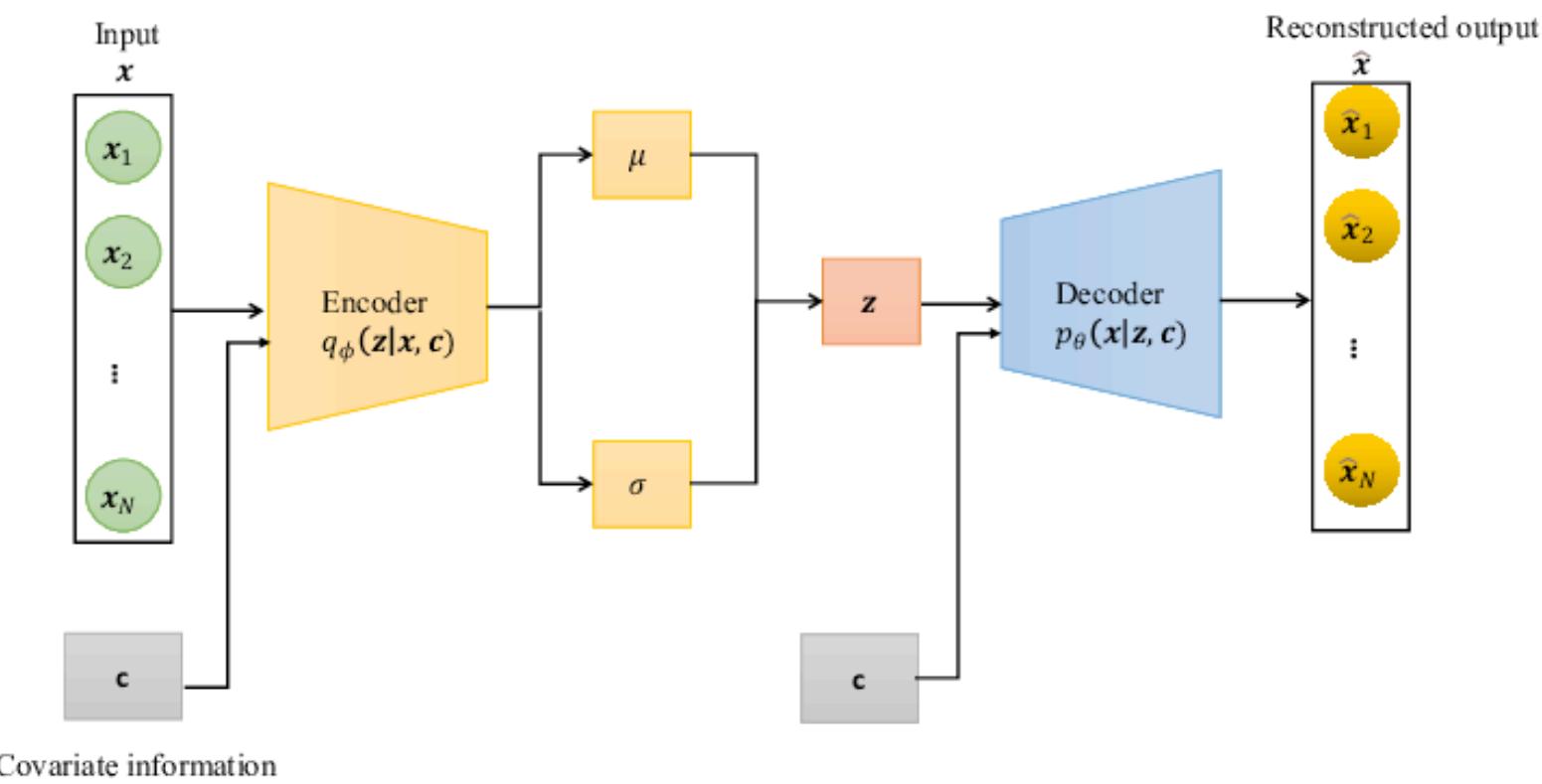


Production



Variations des VAE

- VAE conditionnelle (CVAE)
- Bêta-VAE
- VAE démêlée (DVAE)
- RNN variationnel (VRNN)
- Vectoriel quantifié-VAE (VQ-VAE)



Résultats du VAE



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

(a) 2-D latent spa

(b) 5-D latent sp

(c) 10-D latent sp

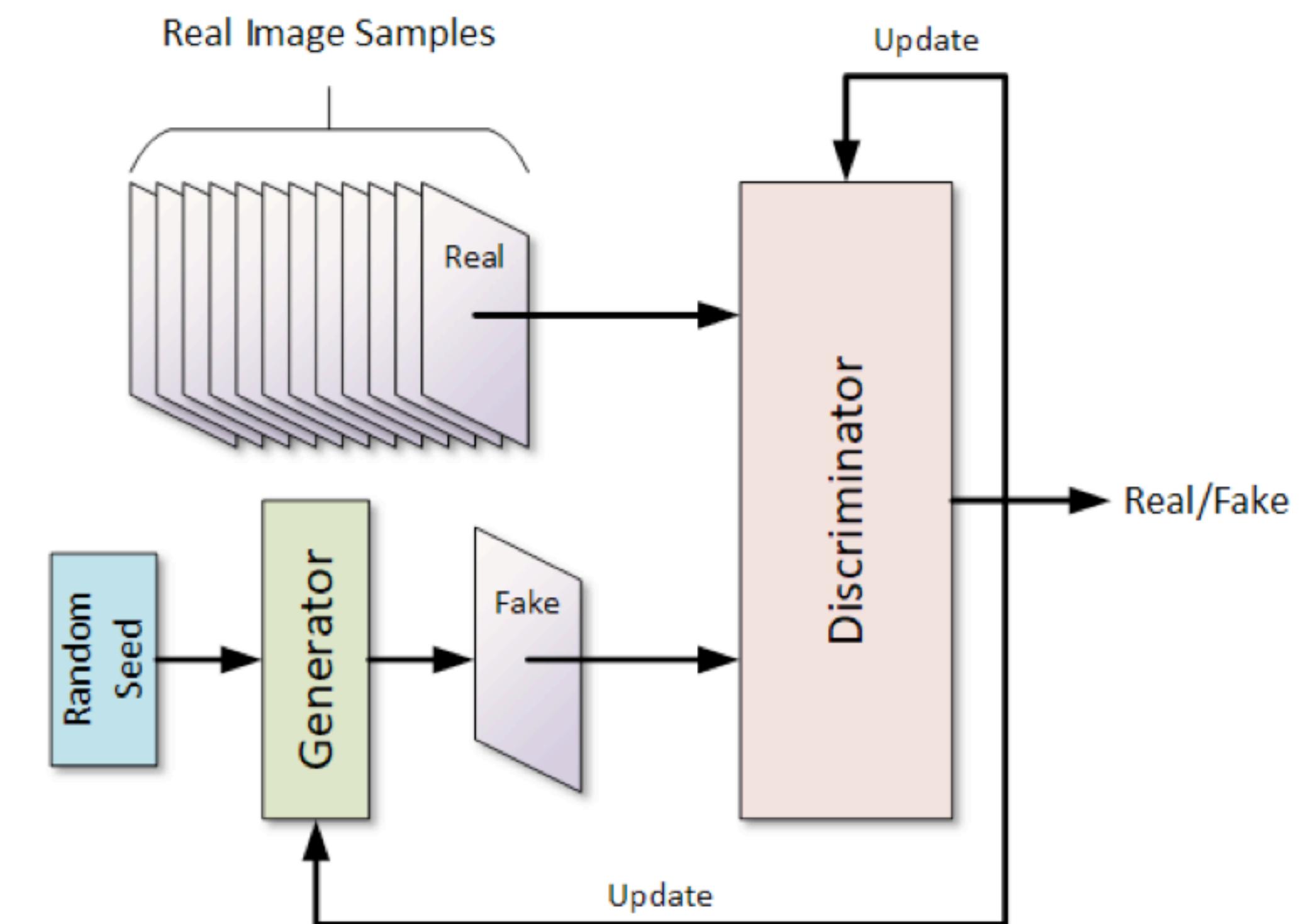
(d) 20-D latent space

Réseaux adversoriels génératifs

- Jeu à somme nulle
- Générateur
 - Crée de nouvelles instances de données à partir d'un vecteur
 - L'objectif est de générer des résultats qui ne se distinguent pas des données authentiques
- Discriminateur
 - Évalue l'authenticité des données générées et des données réelles

Mécanisme de base des GAN

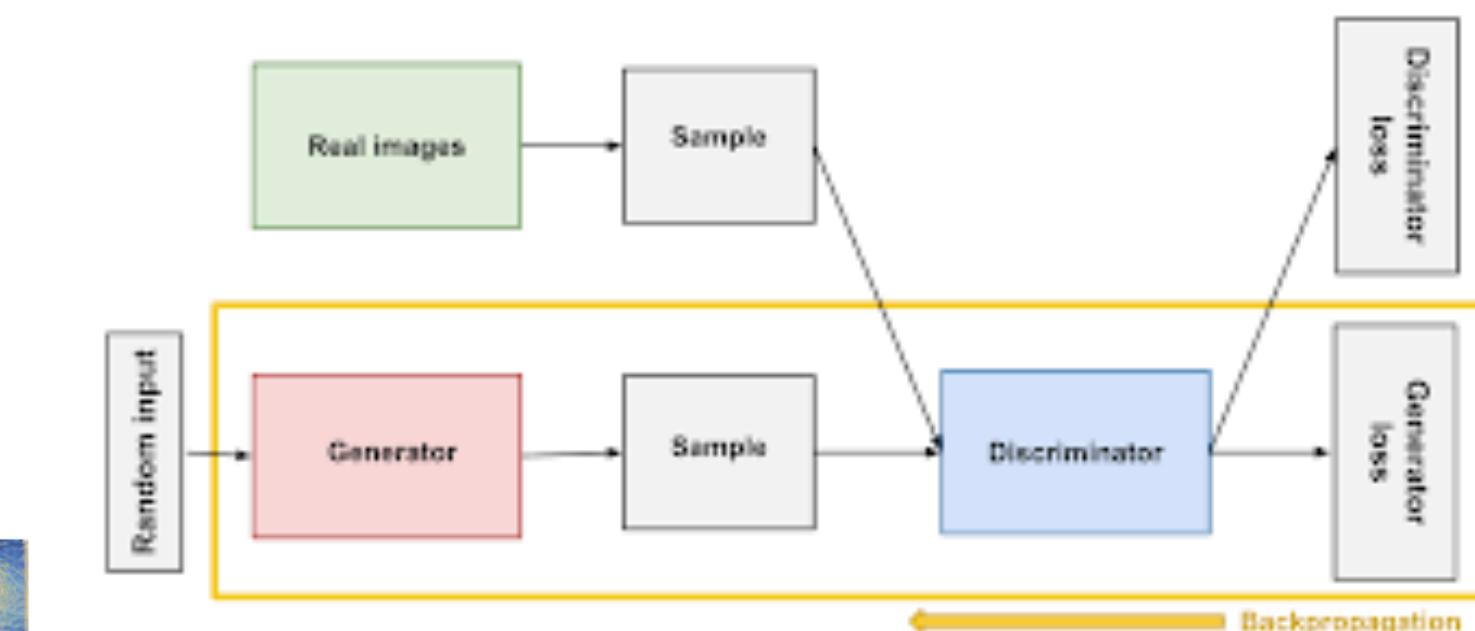
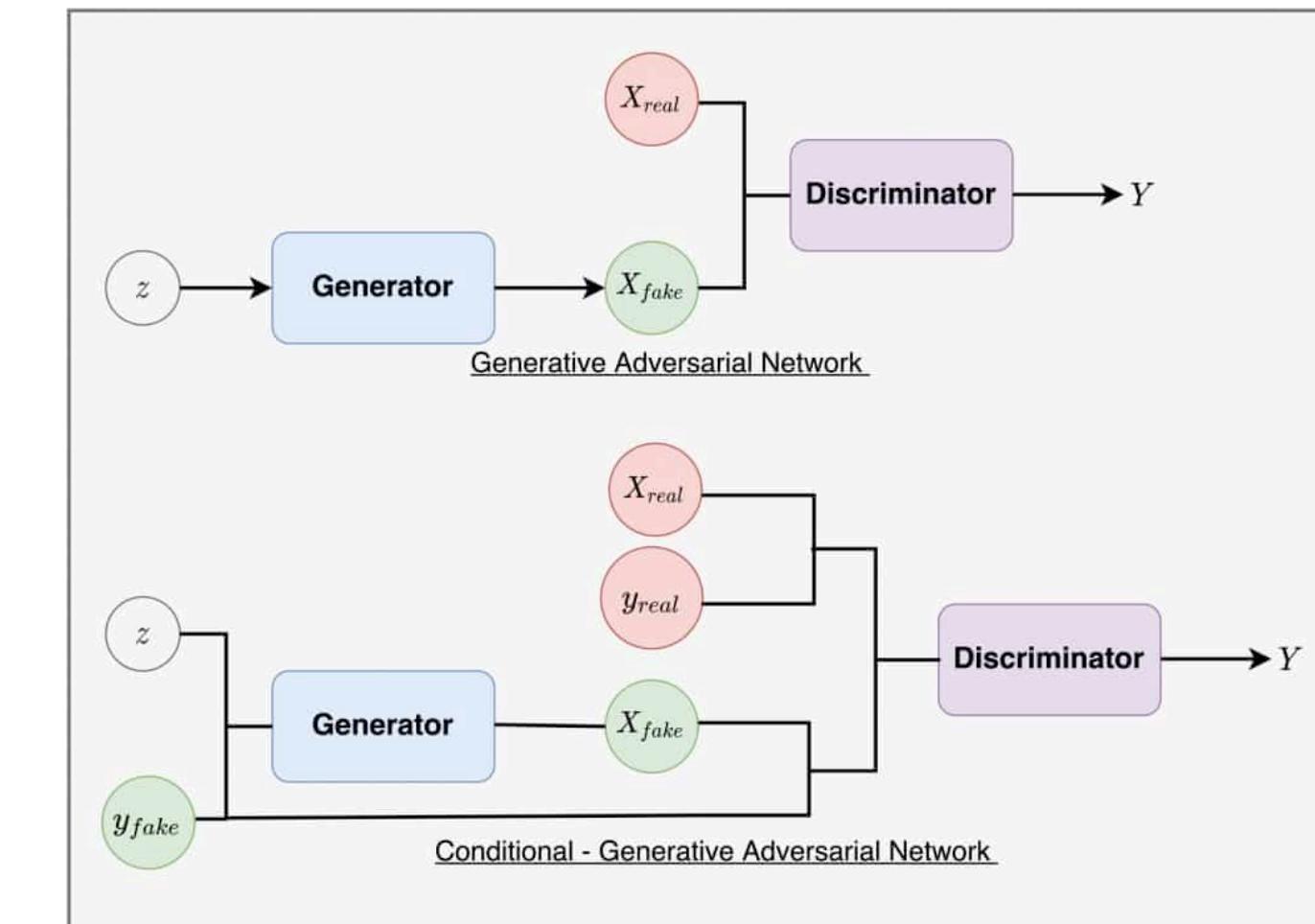
- Le générateur vise à tromper le discriminateur en améliorant sa capacité à produire des données réalistes
- Le discriminateur apprend à mieux distinguer les vraies données des fausses données créées par le générateur



Variations des GAN

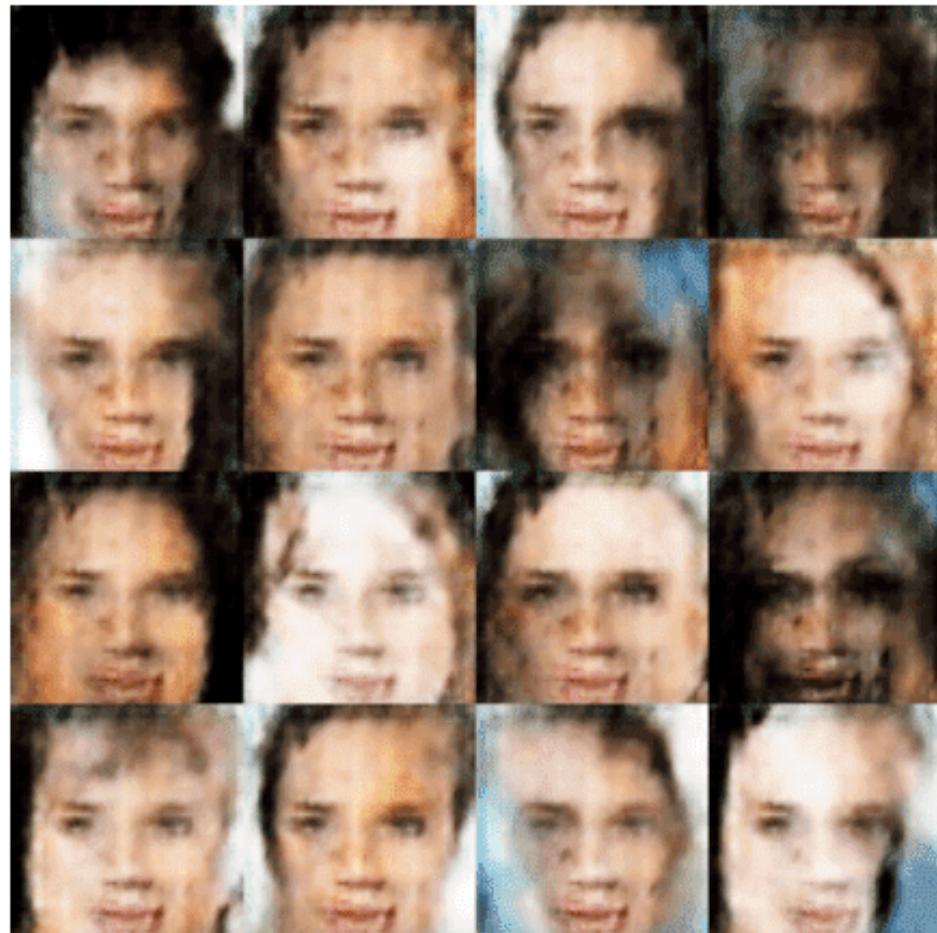
<https://arxiv.org/abs/1812.04948v3>

- GAN conditionnel (cGAN)
- GAN Convolutionnel Profond (DCGAN)
- Wasserstein GAN (WGAN)
- CycleGAN
- StyleGAN
- BigGAN

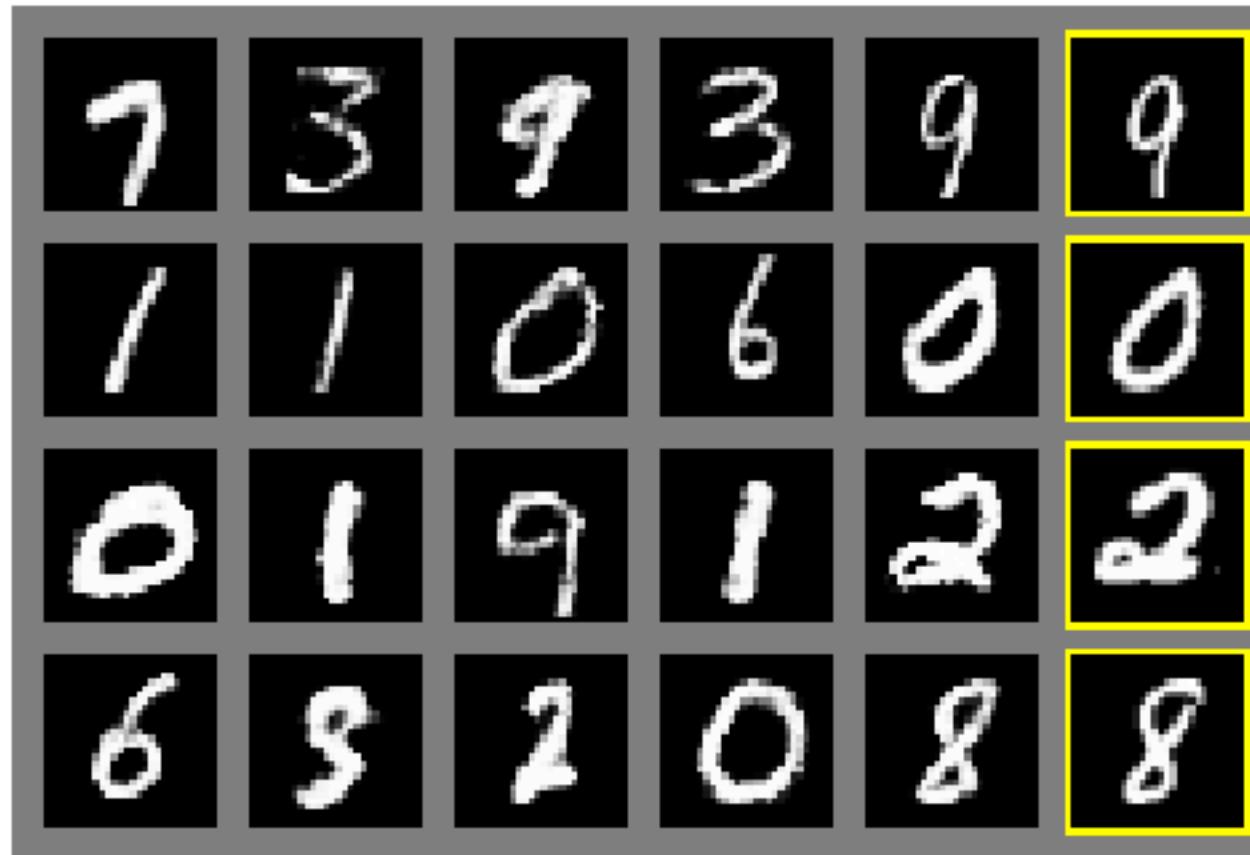
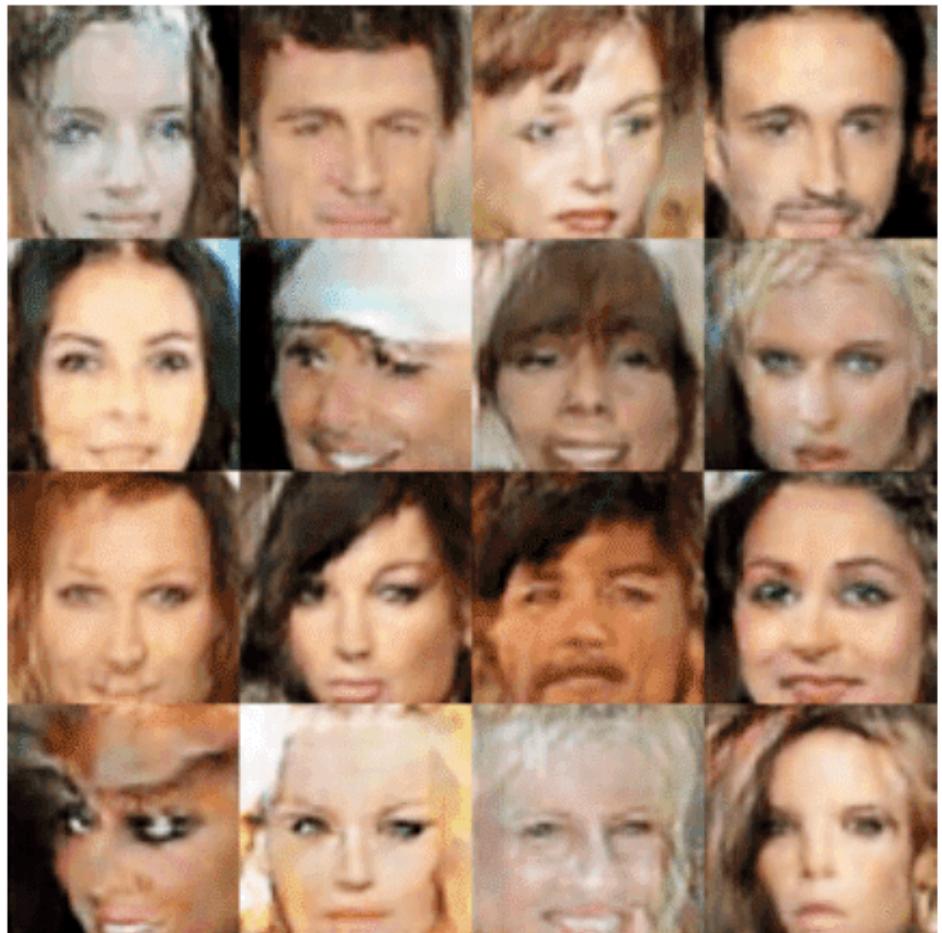


Résultats des GAN

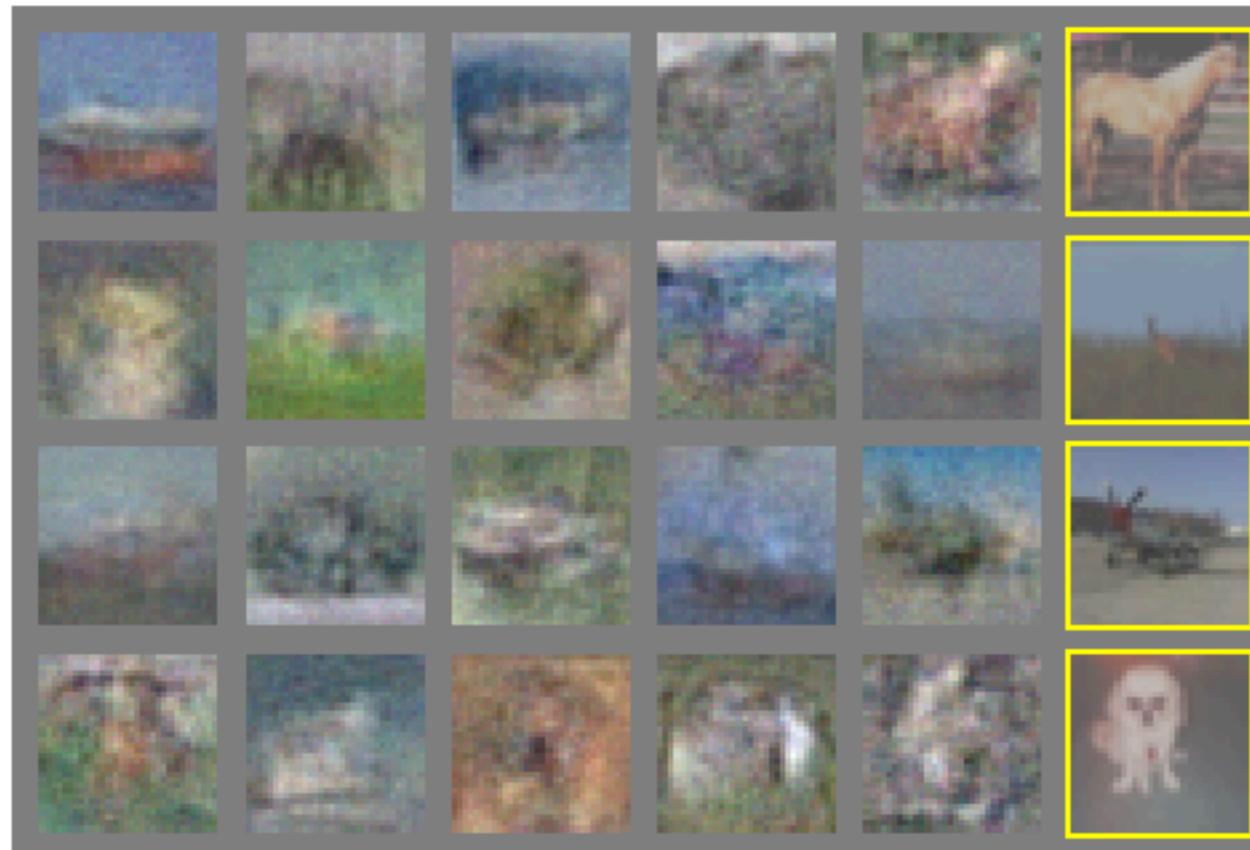
DCGAN



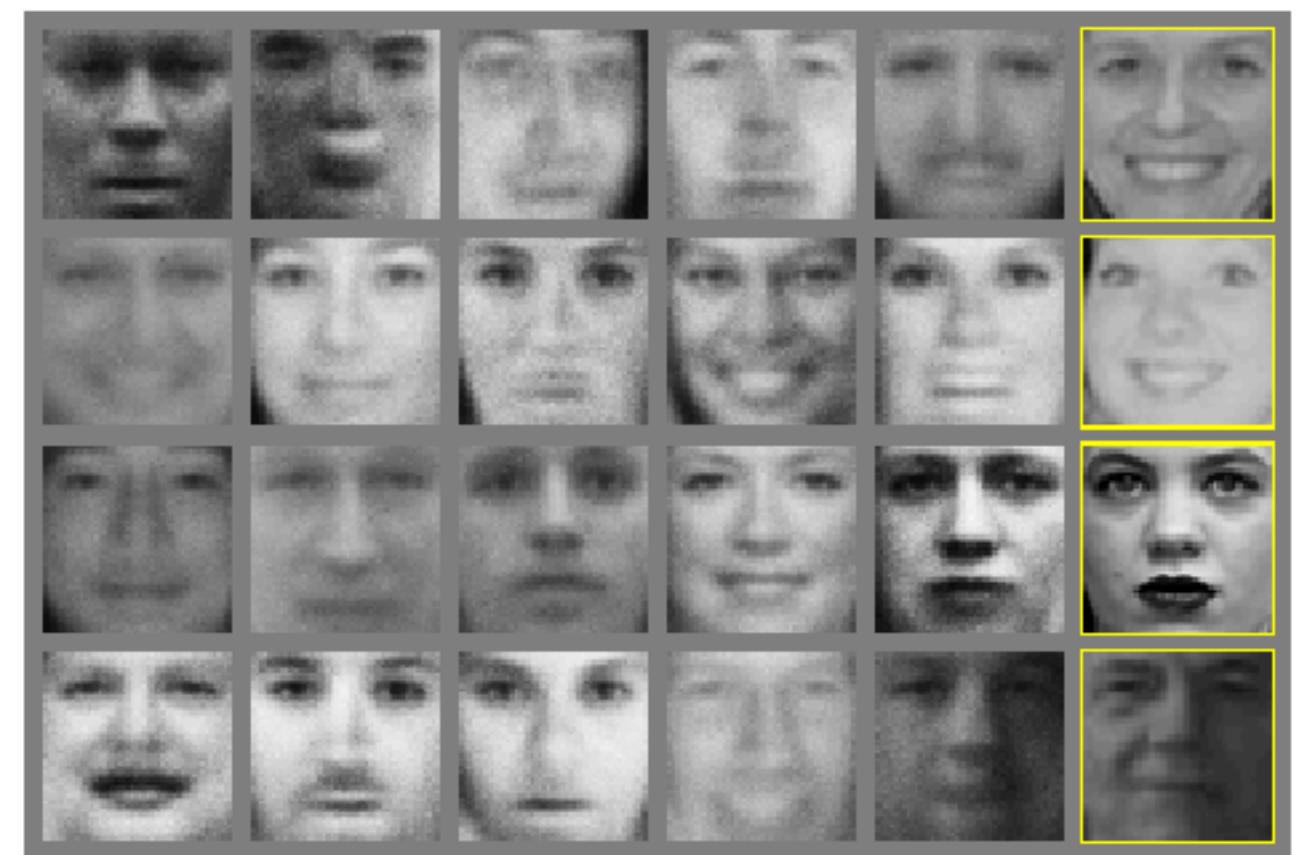
PacDCGAN2



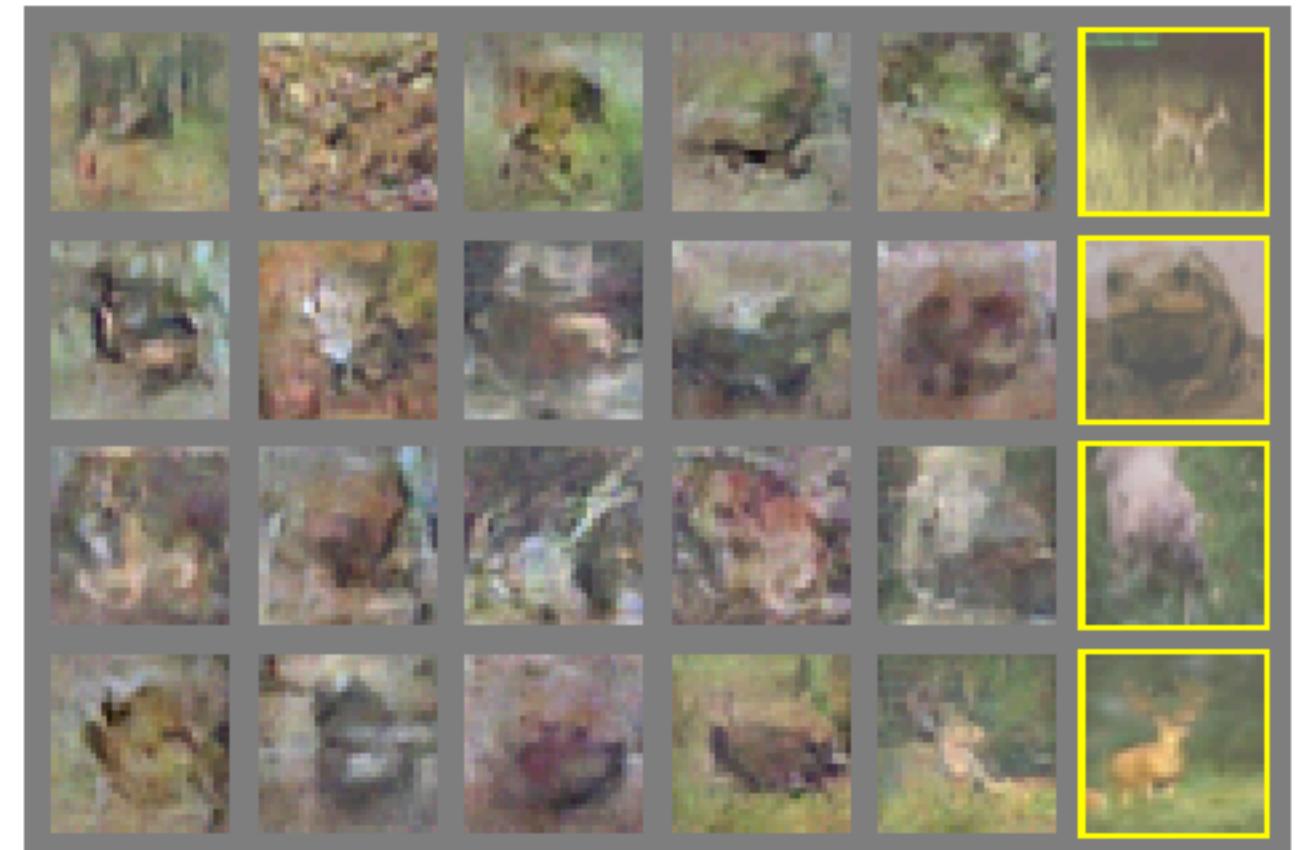
a)



c)



b)

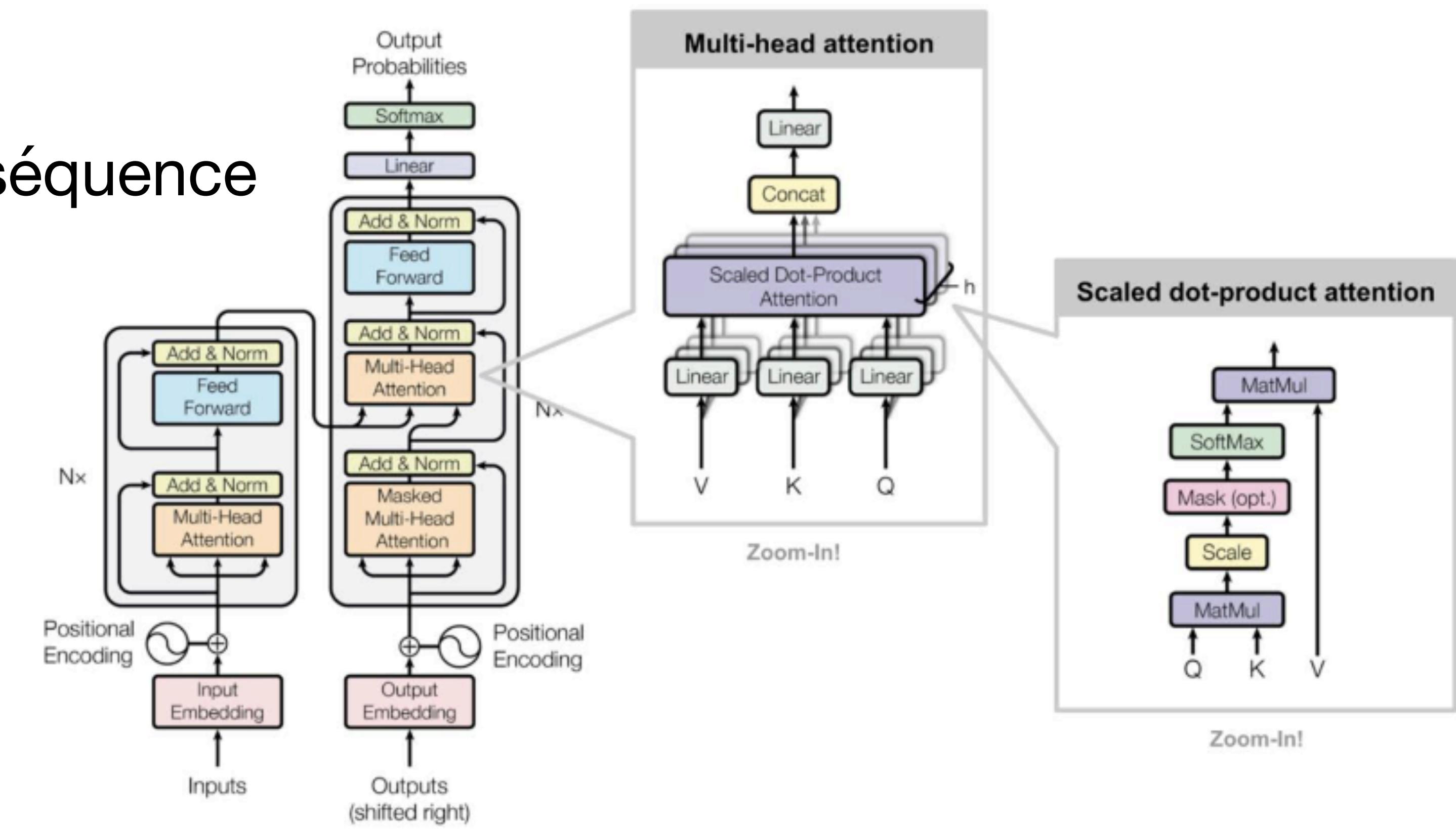
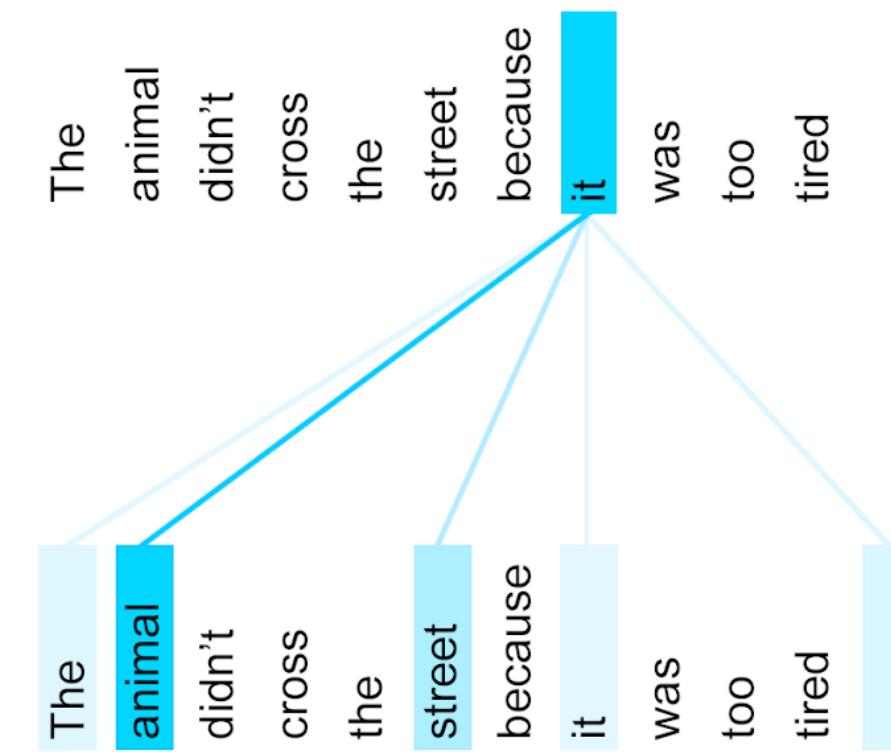


d)

Transformer

<https://arxiv.org/abs/1706.03762>

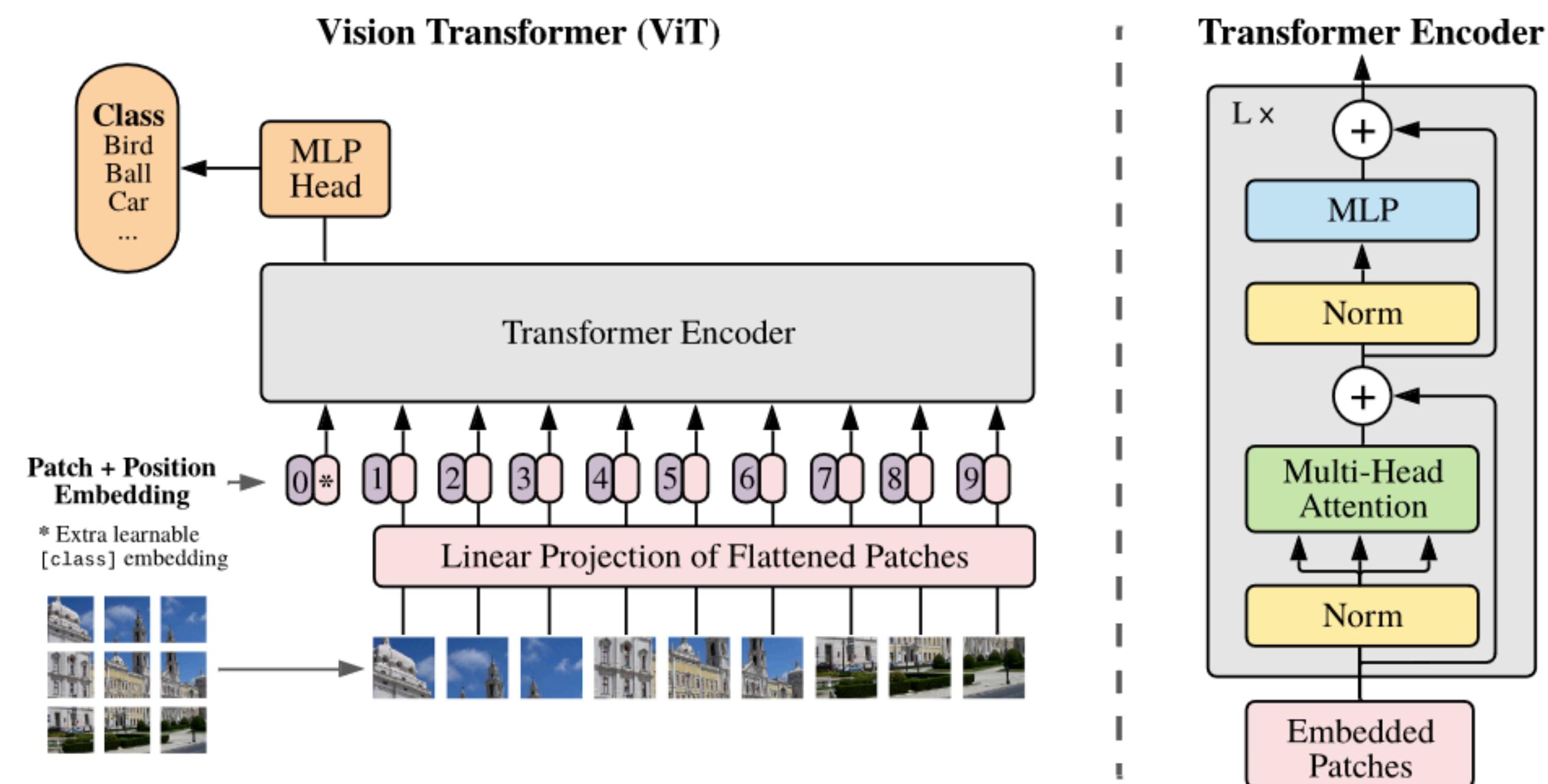
- Attention à soi
- Intégration contextuelle d'une séquence
- Traiter séquences en parallèle
- Dépendances à long terme



Transformer de vision

<https://arxiv.org/abs/2010.11929>

- Divisée en une série de patches de taille fixe
- Aplatie et transformée en une séquence de embeddings linéaires
- Codages positionnels
- Surpasser les CNN



Transformer vs CNN

- Traitement des données
 - Transformer: dépendances globales sont importantes
 - CNN: exploiter la connectivité locale
- Performance et scalabilité
- Efficacité du calcul
 - Les CNN sont généralement plus efficaces

Apprentissage auto-supervisé

- Générer des pseudo-étiquettes à partir des données elles-même
- Tâche prédictive
 - Prédire une partie d'une image à partir d'une autre partie
 - Prédire la rotation d'une image

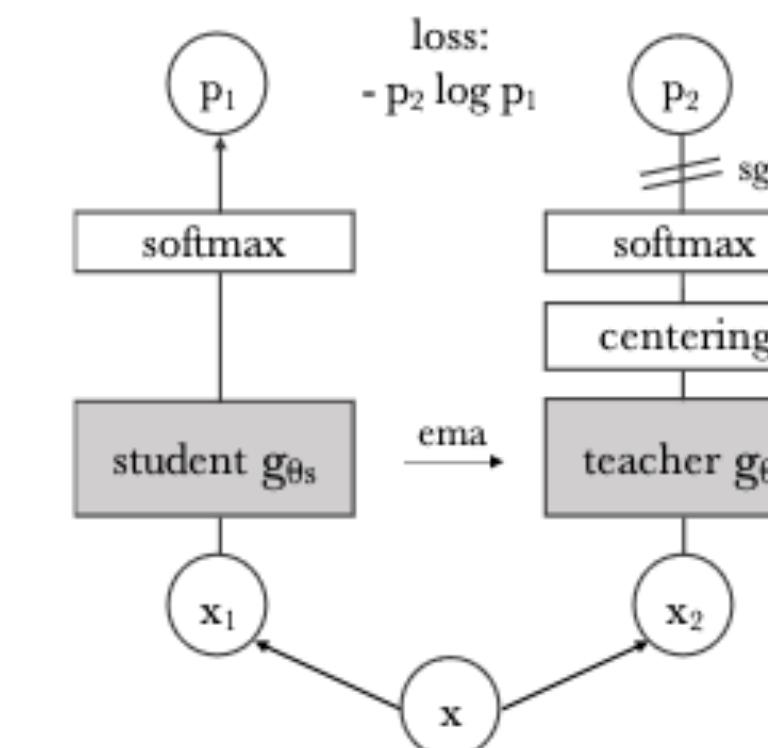
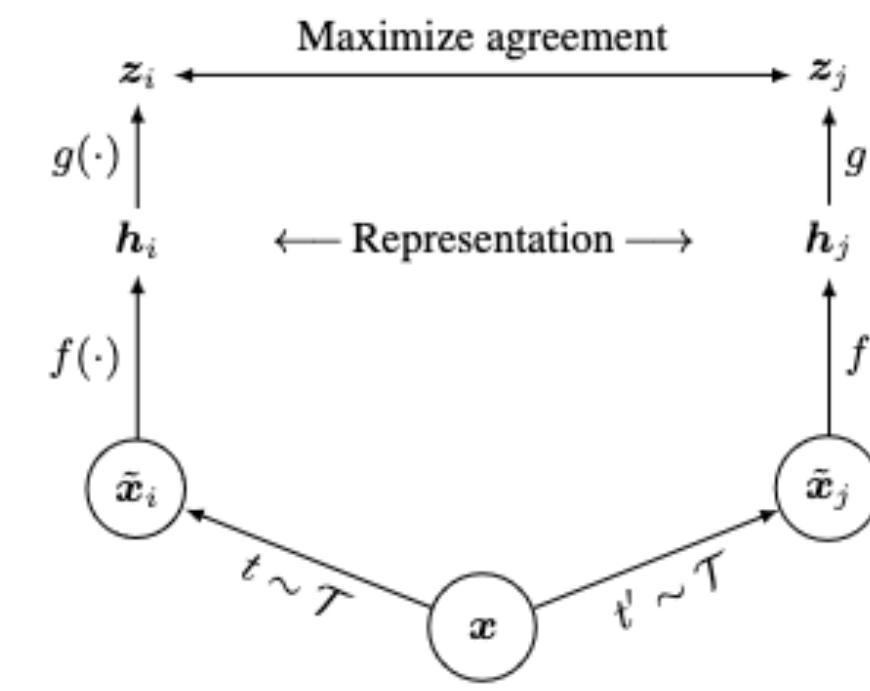
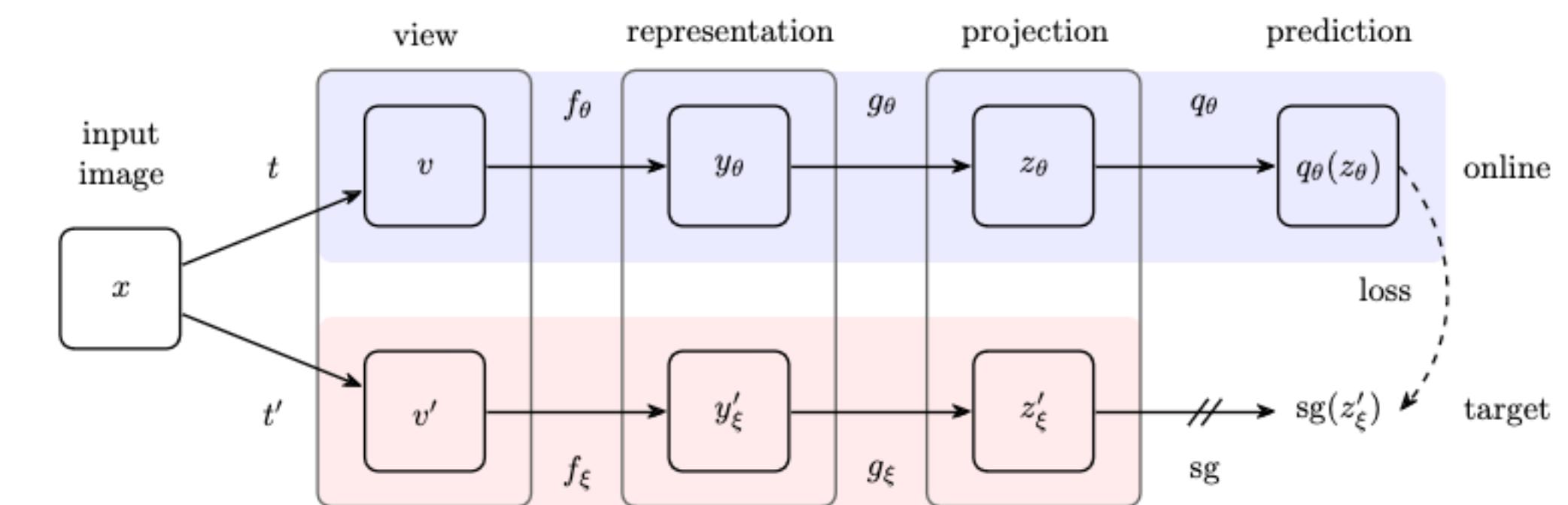
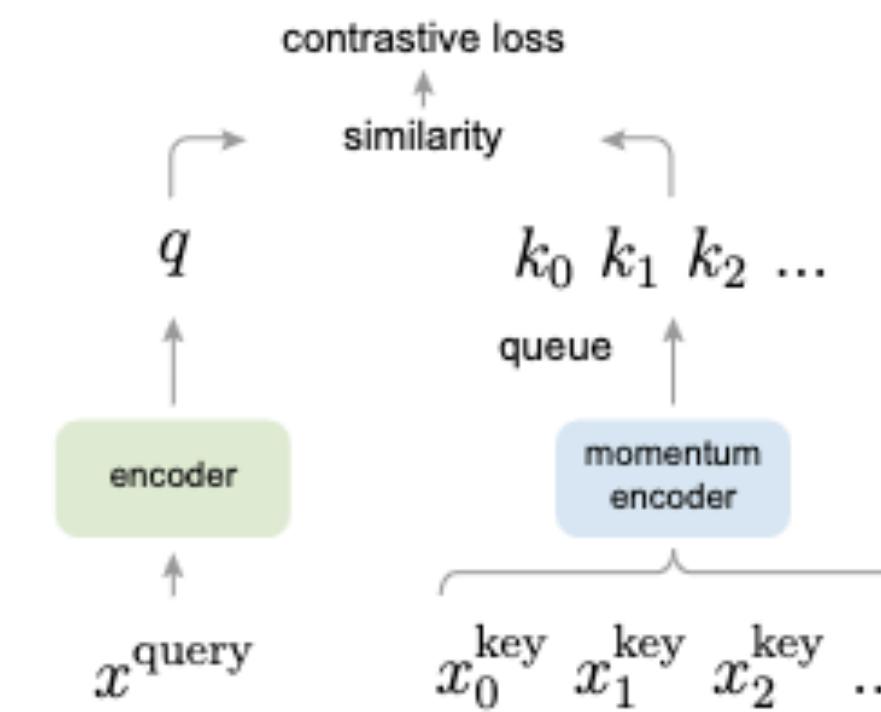
Apprentissage auto-supervisé

- Deux principales composantes
 - Tâche prétexte: Il s'agit d'une tâche créée artificiellement à partir des données elles-mêmes pour entraîner le modèle.
 - Tâche downstream: Une fois que le modèle a appris de la tâche prétexte, il est affiné sur un plus petit ensemble de données étiquetées pour une tâche spécifique

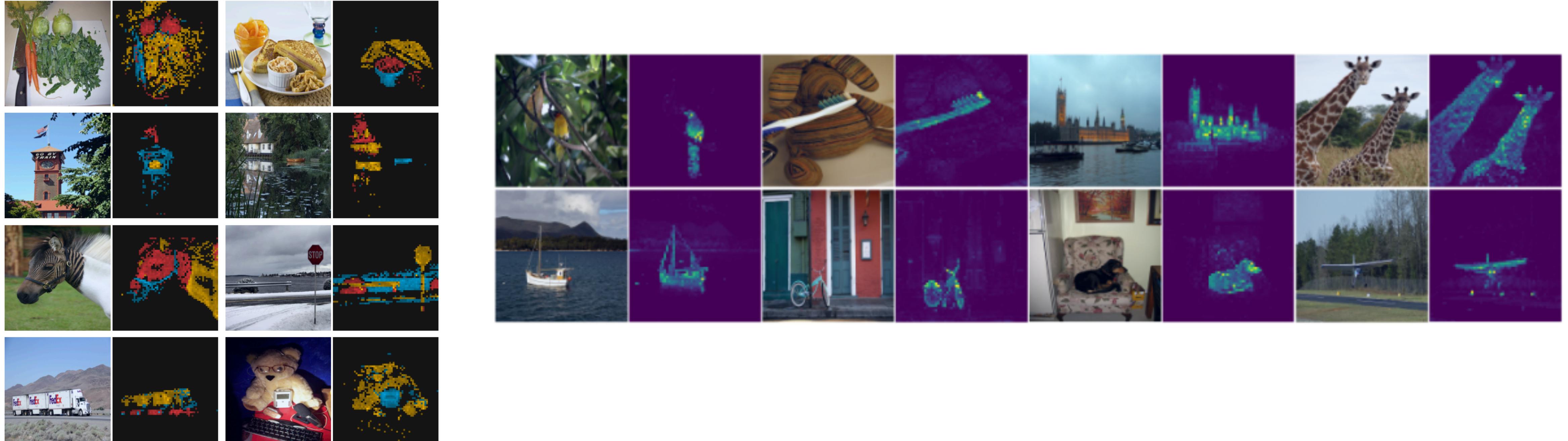
Apprentissage auto-supervisé

<https://arxiv.org/abs/2104.14294>

- MoCo
 - Contrastif
- BYOL
- SimCLR
- DINO
 - Transformer



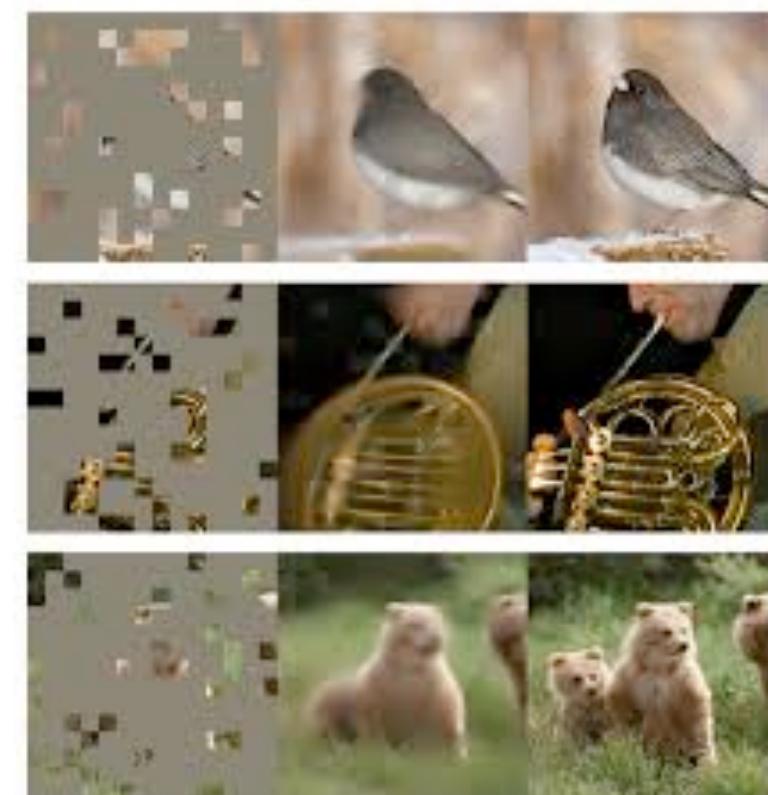
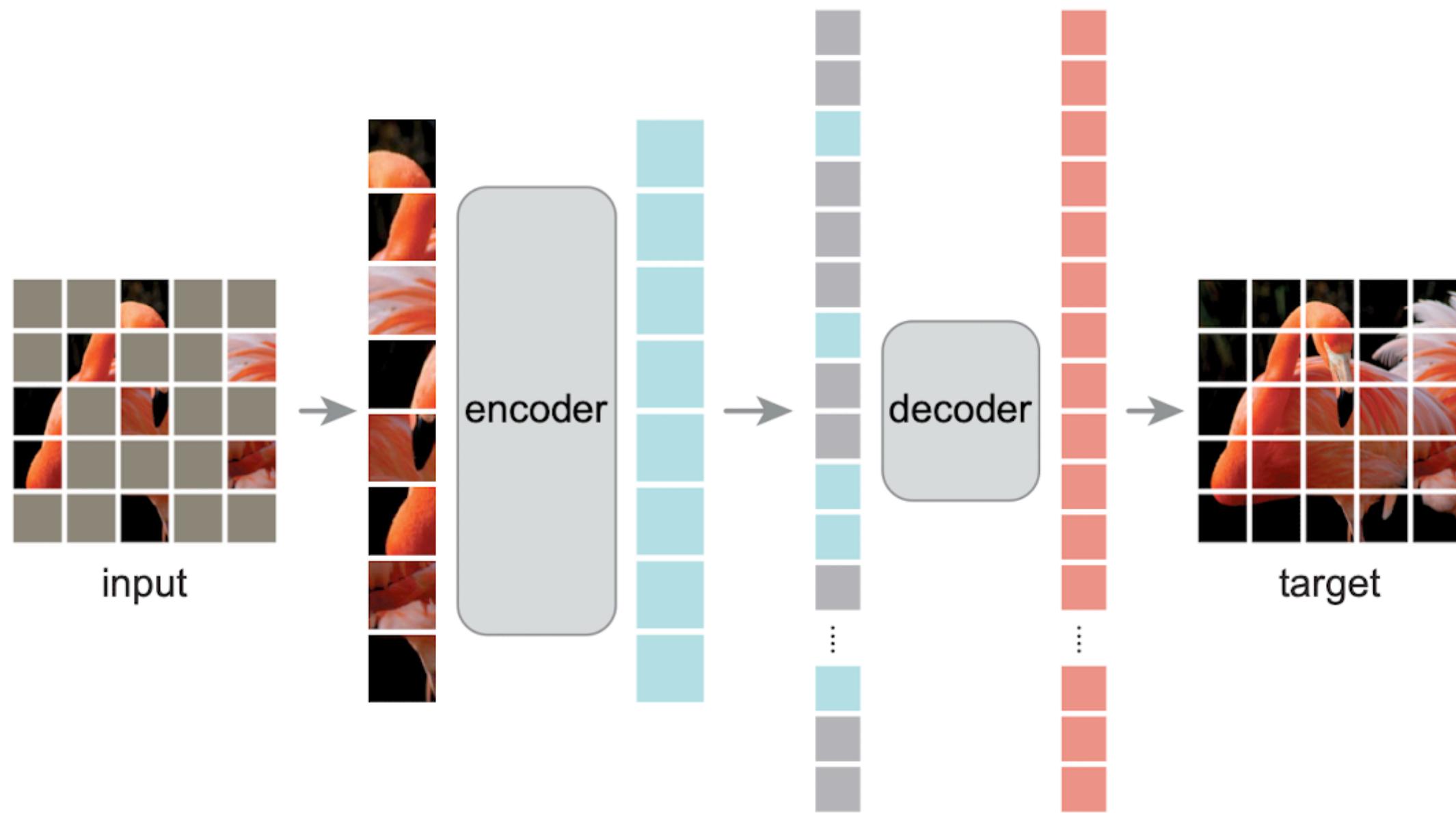
Apprentissage auto-supervisé et Transformer



Masked auto encoders

<https://arxiv.org/abs/2111.06377>

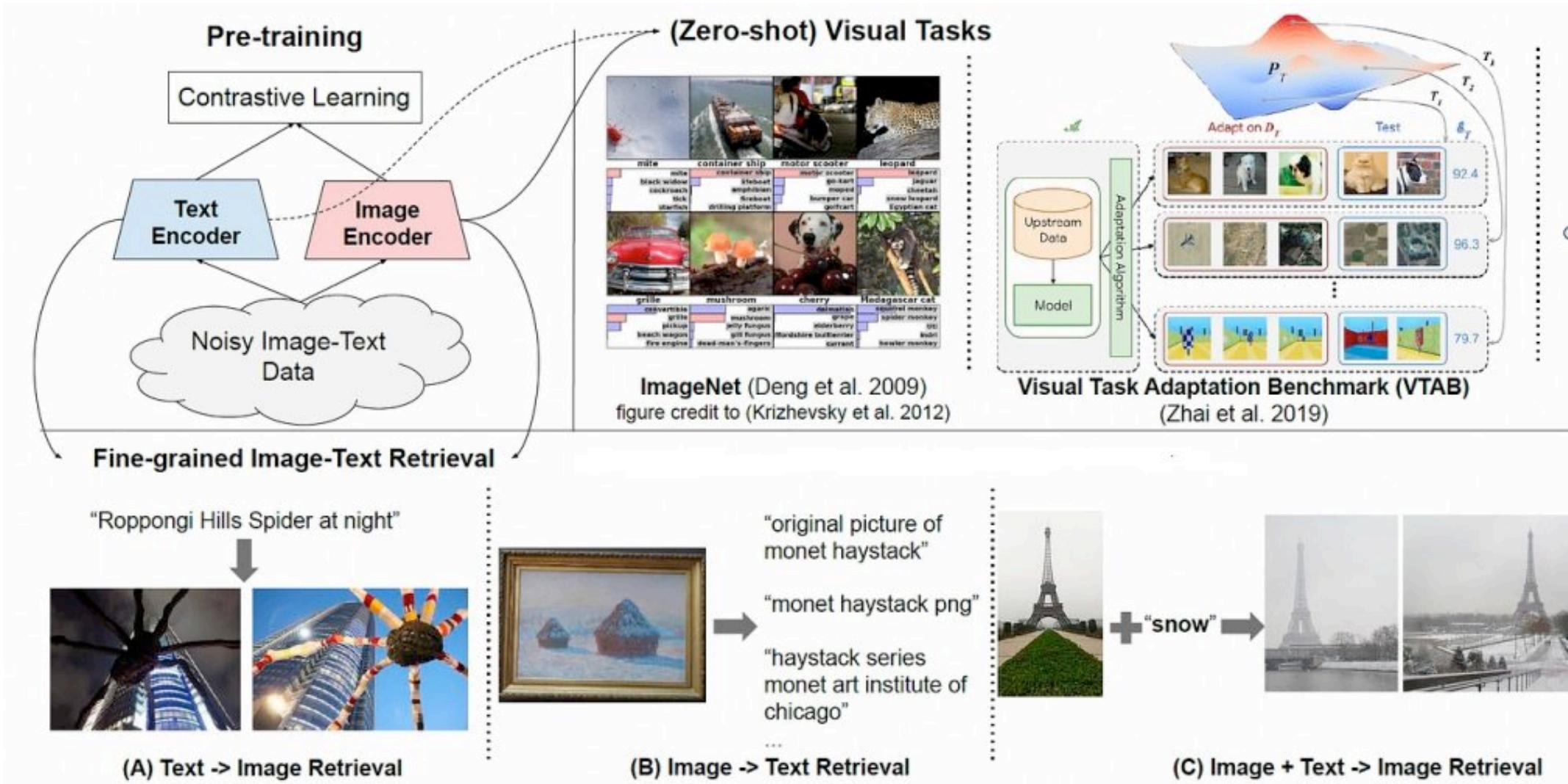
- Tâche de préformation
- Prévoir les zones manquantes



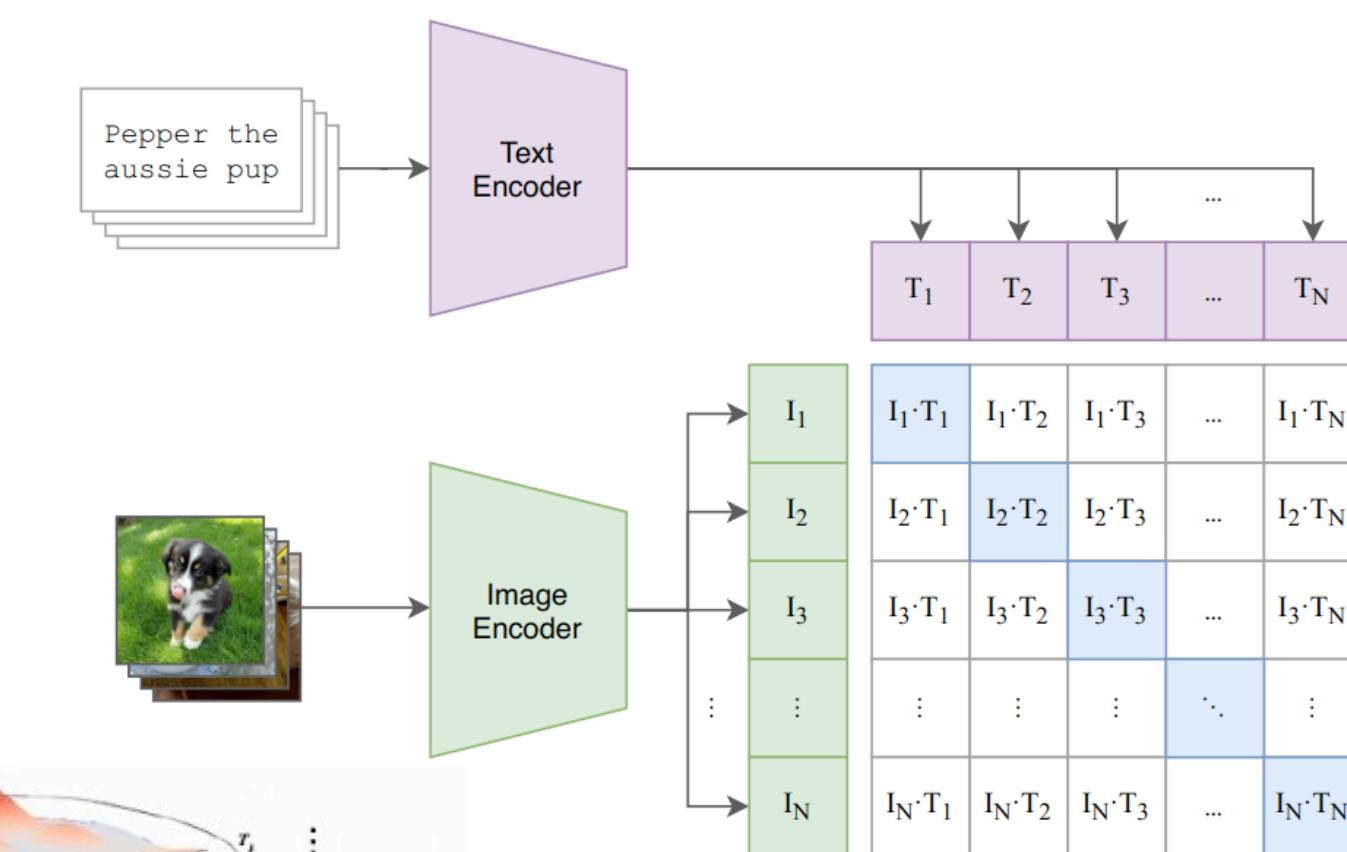
CLIP

<https://arxiv.org/abs/2103.00020>

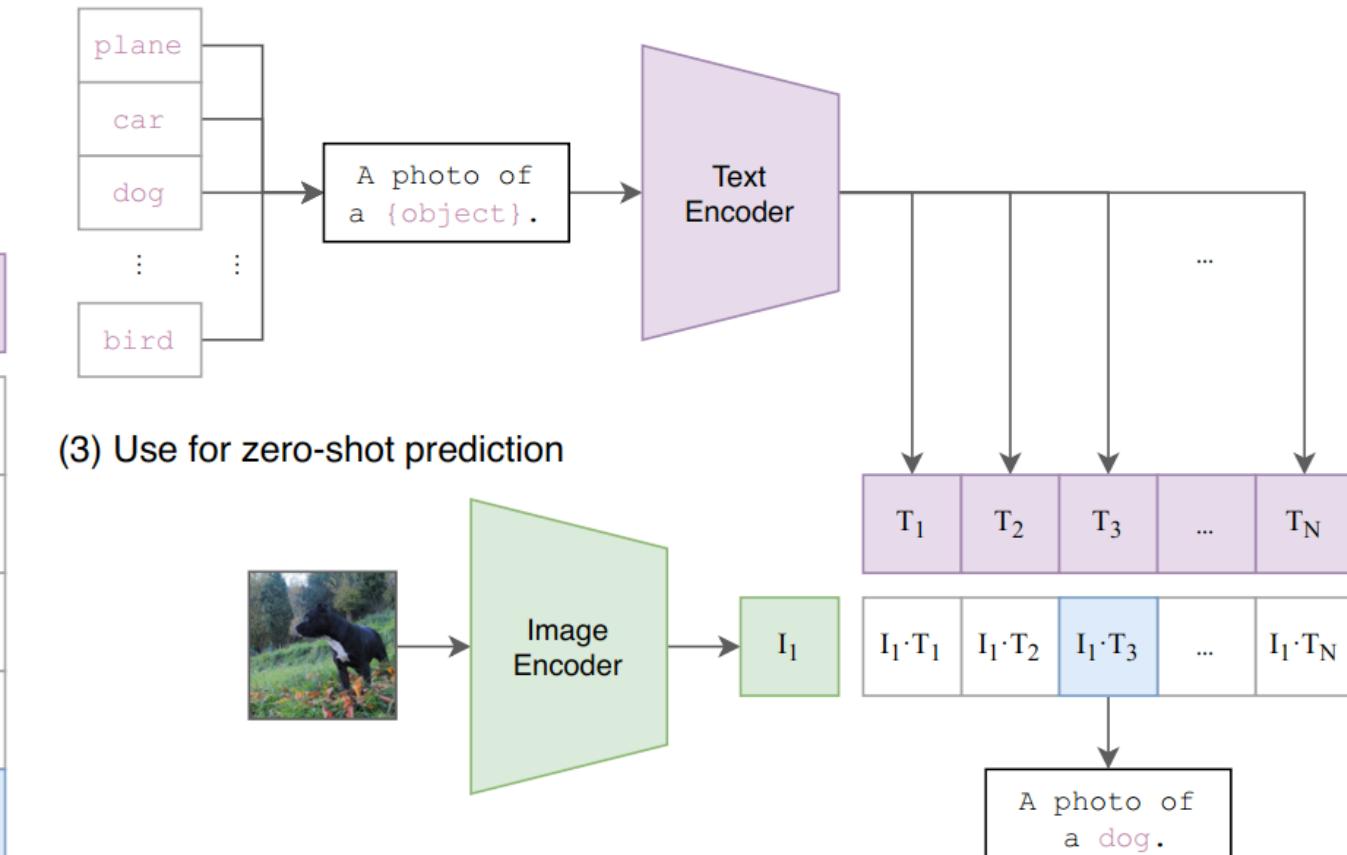
- Modèles multimodaux
- Text + Image



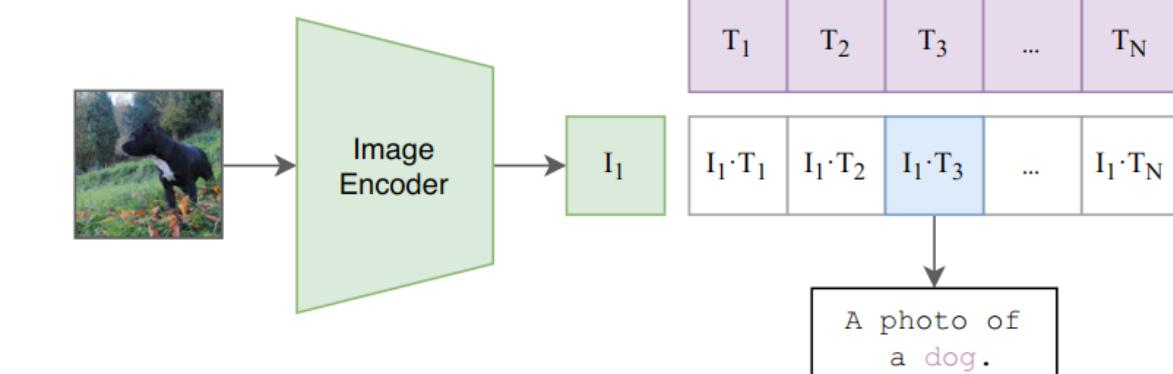
(1) Contrastive pre-training



(2) Create dataset classifier from label text

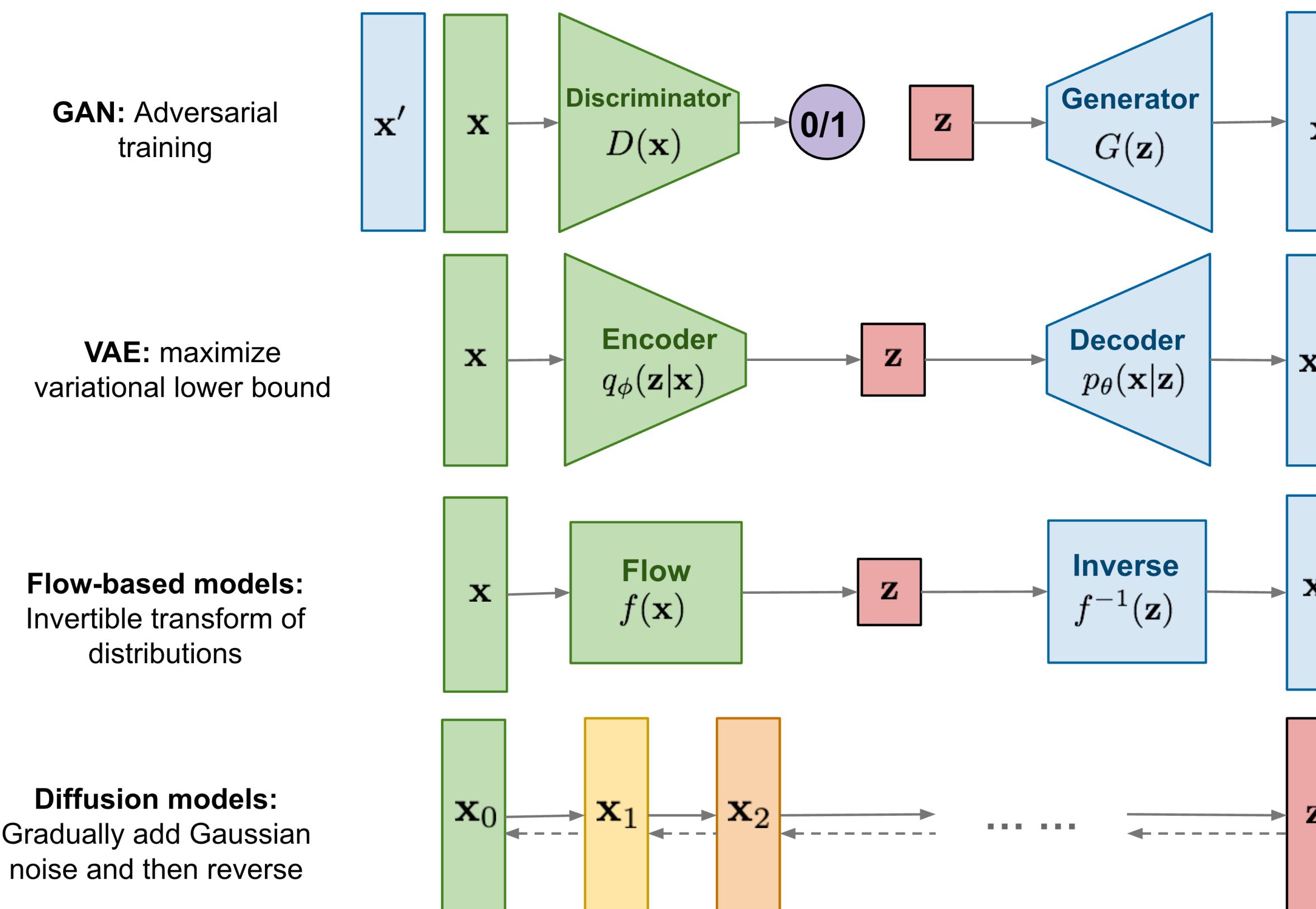


(3) Use for zero-shot prediction



Modèles de diffusion

<https://arxiv.org/abs/2006.11239>

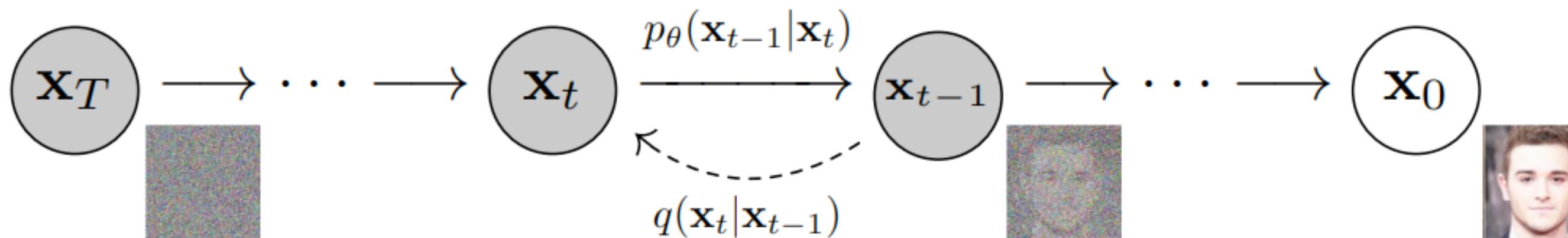
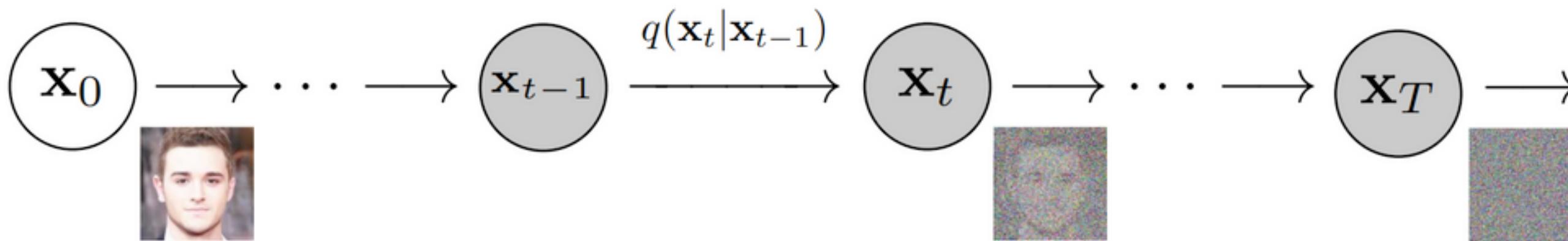


Modèles de diffusion

- Processus de progression : Ajoute progressivement du bruit gaussien à une image ou à un signal jusqu'à ce qu'il ne reste plus que du bruit.
- Processus inverse : Un réseau neuronal apprend à supprimer le bruit étape par étape, en reconstruisant les données originales à partir du bruit.
- Objectif de formation : Minimiser la différence entre les données d'origine et les données reconstruites à partir du bruit.

Modèles de diffusion

- Processus global



$$q(\mathbf{x}_{1:T} | \mathbf{x}_0) := \prod_{t=1}^T q(\mathbf{x}_t | \mathbf{x}_{t-1}) := \prod_{t=1}^T \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I})$$

Discussion sur les technologies actuelles

- Quelle est la technologie la plus récente?
- Comment rester informé et à jour?
- Comment lire les revues scientifiques?
- Quelles sont les meilleures ressources pour la recherche?

Fin

- Merci