

# Traitement et Analyse d'Images

---

## Modèles génératifs Auto-encodeur variationnel

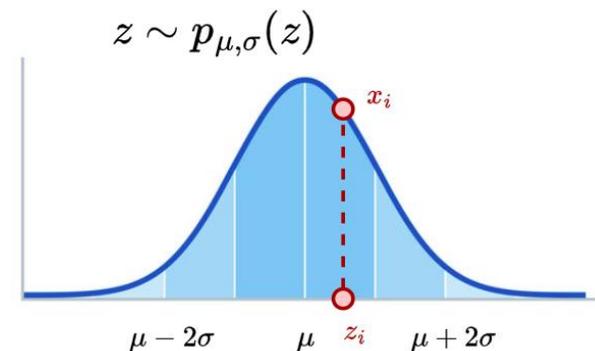
# Quel est l'intérêt des modèles génératifs ?

## ► Comment générer des visages synthétiques ?



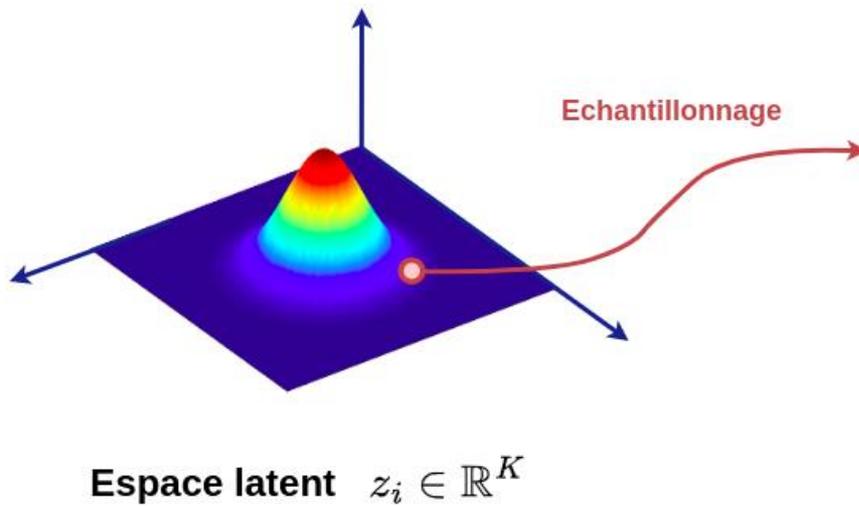
En modélisant la distribution correspondante  $p_{\theta}(\cdot)$  !

→ Est ce que des distributions classiques restent pertinentes



# Quel est l'intérêt des modèles génératifs ?

- ▶ Comment modéliser des distributions complexes ?

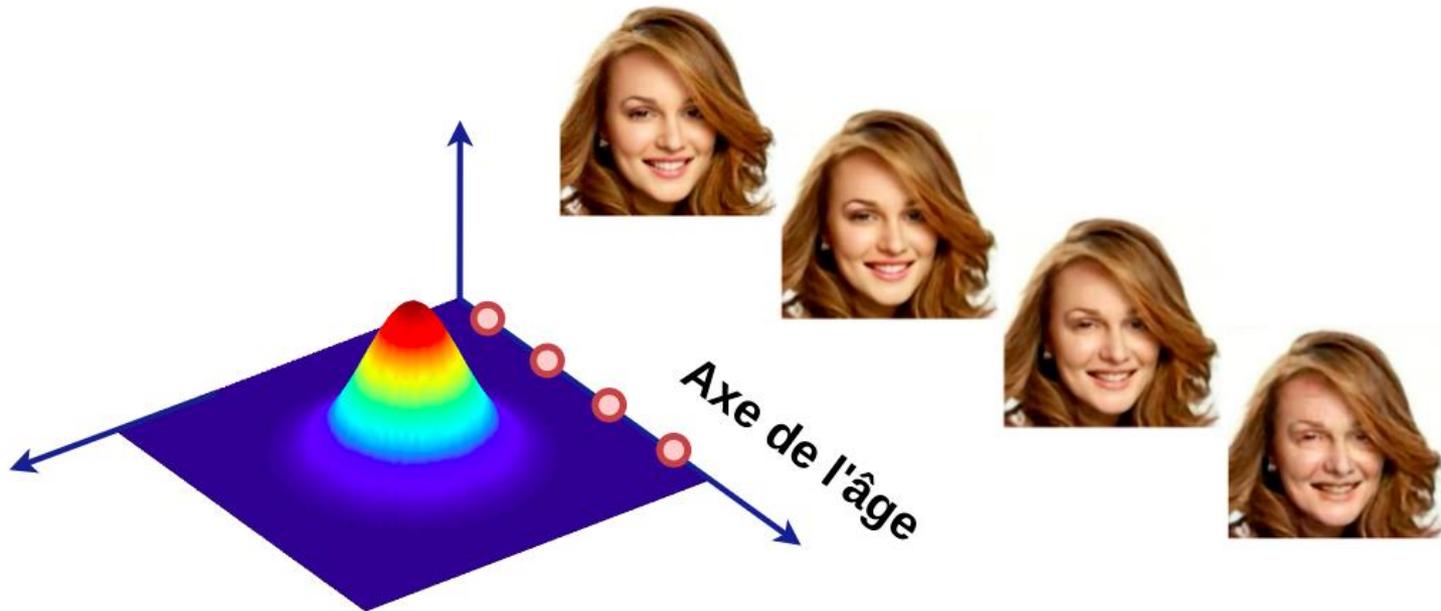


Distribution des visages

# Quel est l'intérêt des modèles génératifs ?

► Pour quel but ?

Une obsession : maîtriser l'espace latent !!

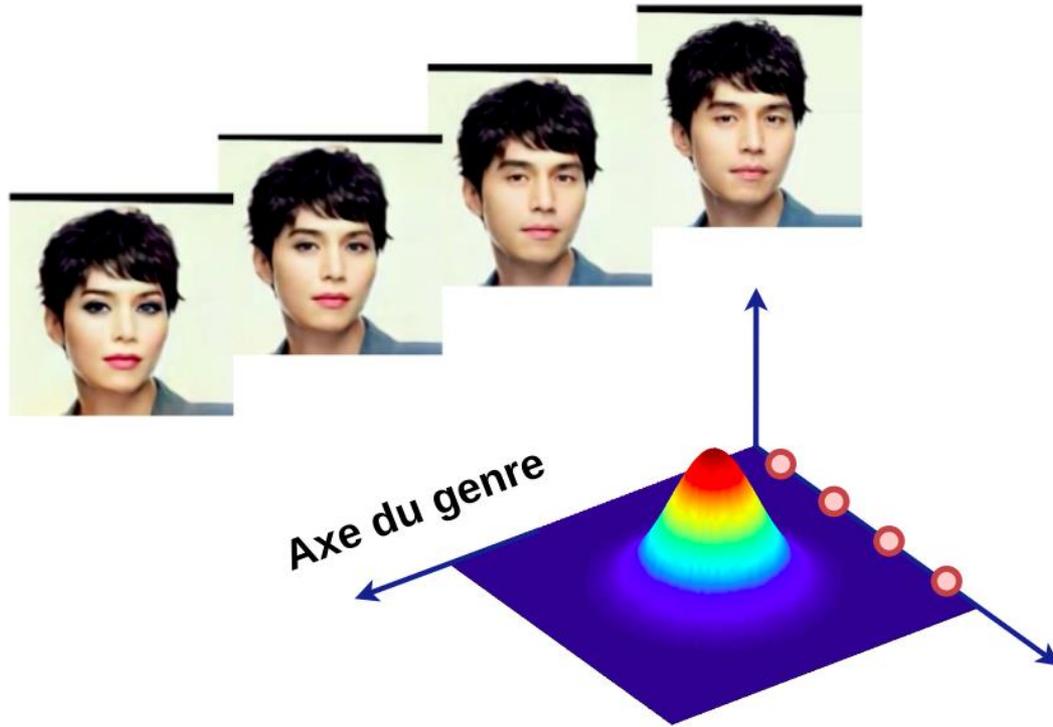


Espace latent  $z_i \in \mathbb{R}^K$

# Quel est l'intérêt des modèles génératifs ?

► Pour quel but ?

Une obsession : maîtriser l'espace latent !!



Espace latent  $z_i \in \mathbb{R}^K$

---

# Auto-encodeurs

---

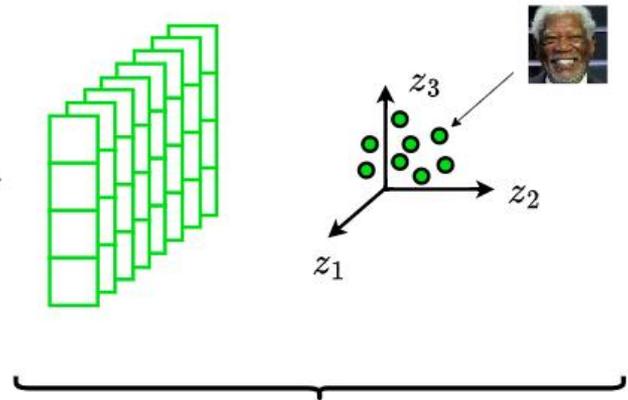
# Comment apprendre une distribution ?

- ▶ Projection dans un espace de représentation plus simple et de dimension inférieure



Espace d'entrée  $x_i \in \mathbb{R}^{N \times M}$

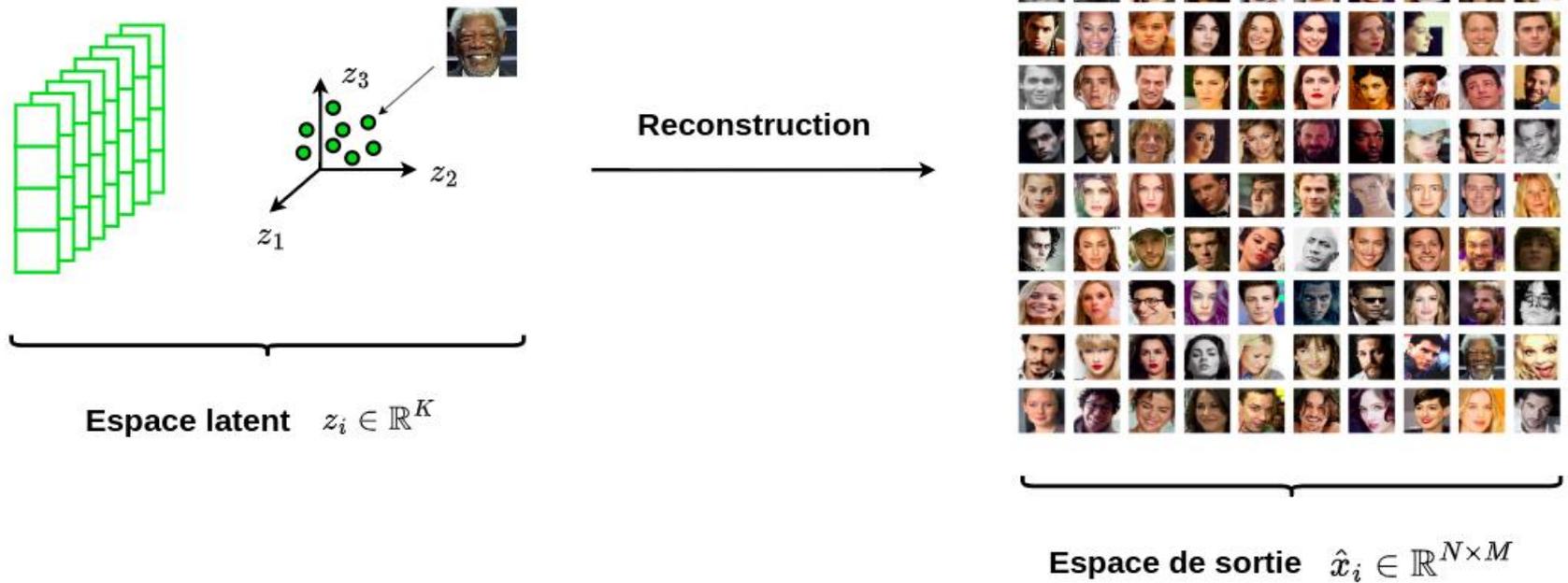
Projection



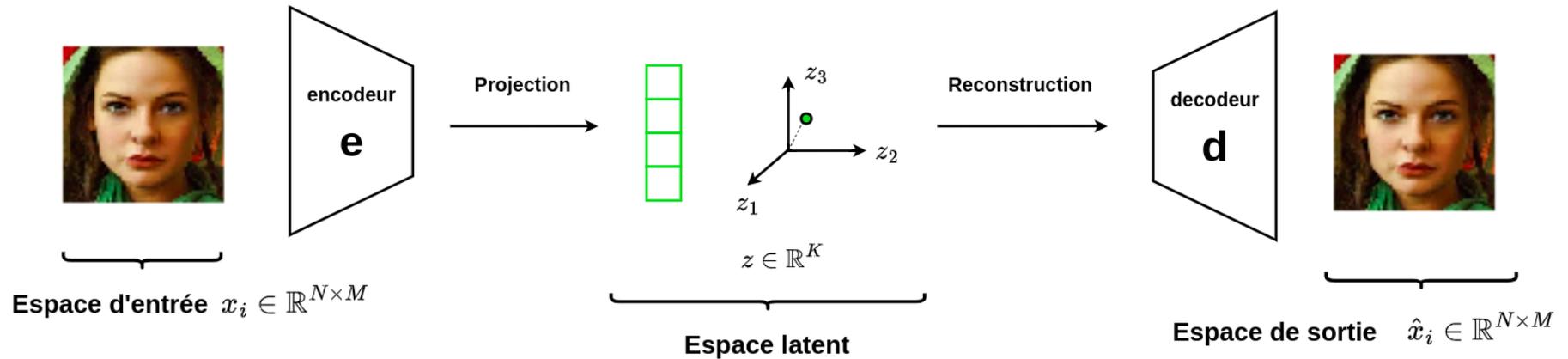
Espace latent  $z_i \in \mathbb{R}^K$

# Comment apprendre une distribution ?

- ▶ Comment disposer d'un espace de représentation pertinent ?



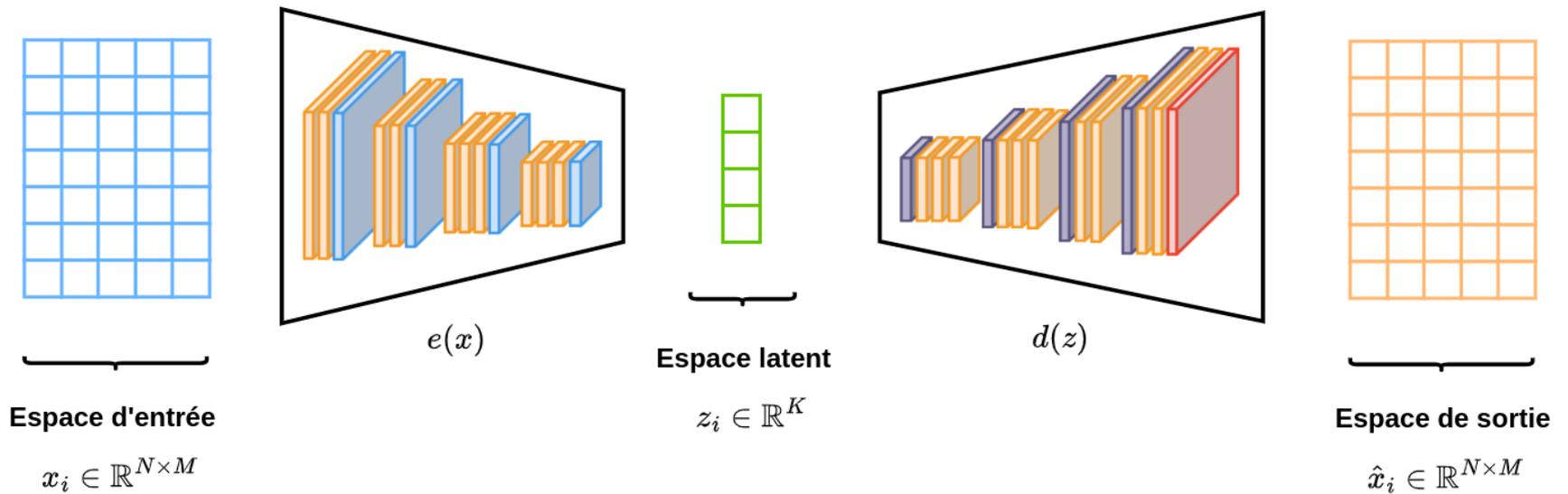
## ► Architectures standards



## ► Fonction de perte en apprentissage profond

$$\text{loss} = \|x - \hat{x}\|^2$$

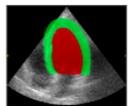
- ▶ Encodeur / Décodeur modélisé par des réseaux neuronaux (convolutionnels)



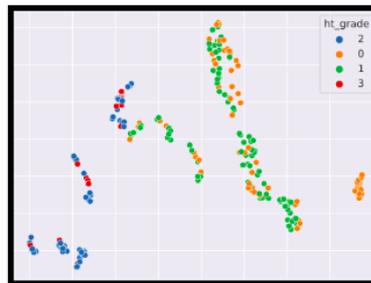
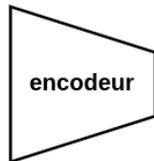
$$\text{loss} = \|x - d(e(x))\|^2$$

## ► Auto-encodeur ? Pour quoi faire ?

### ► Représentation de données



Patients



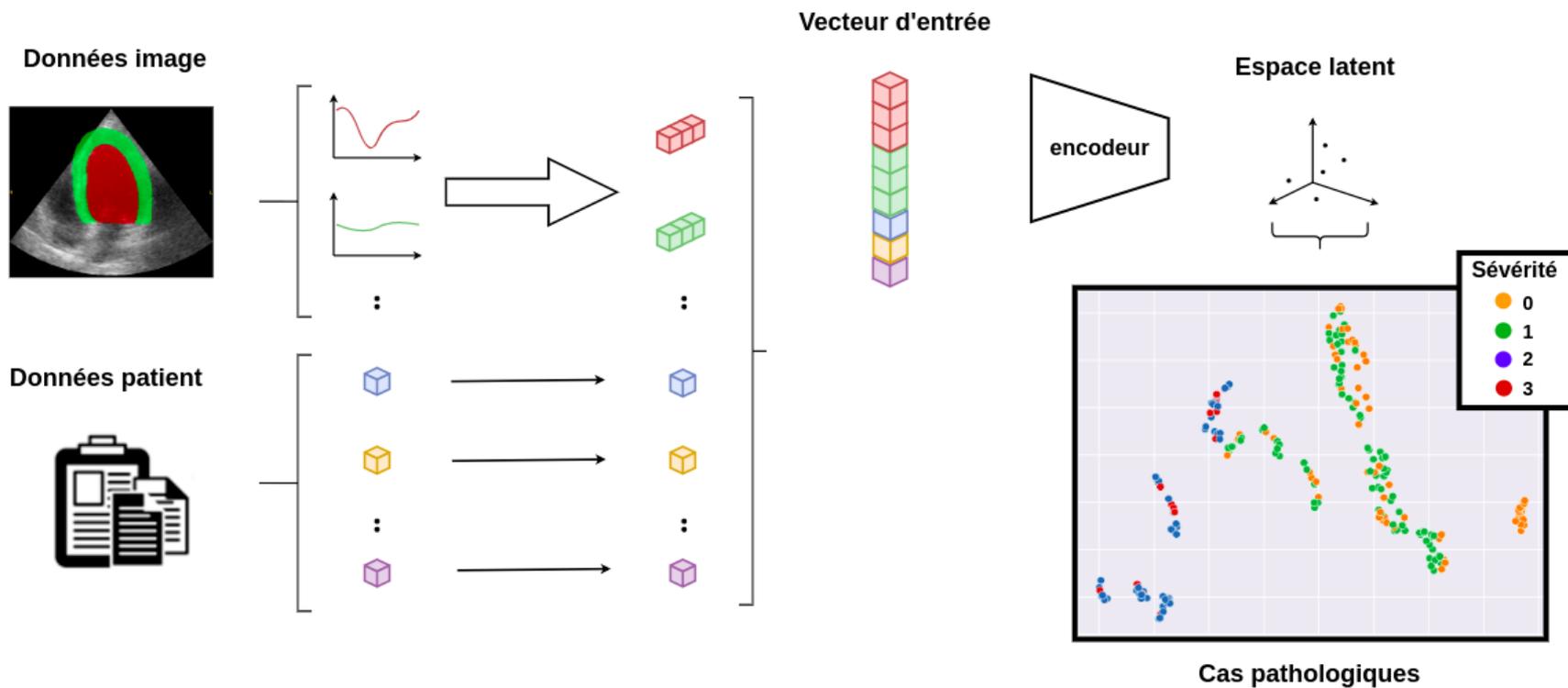
Representation de population

### ► Modèles génératifs

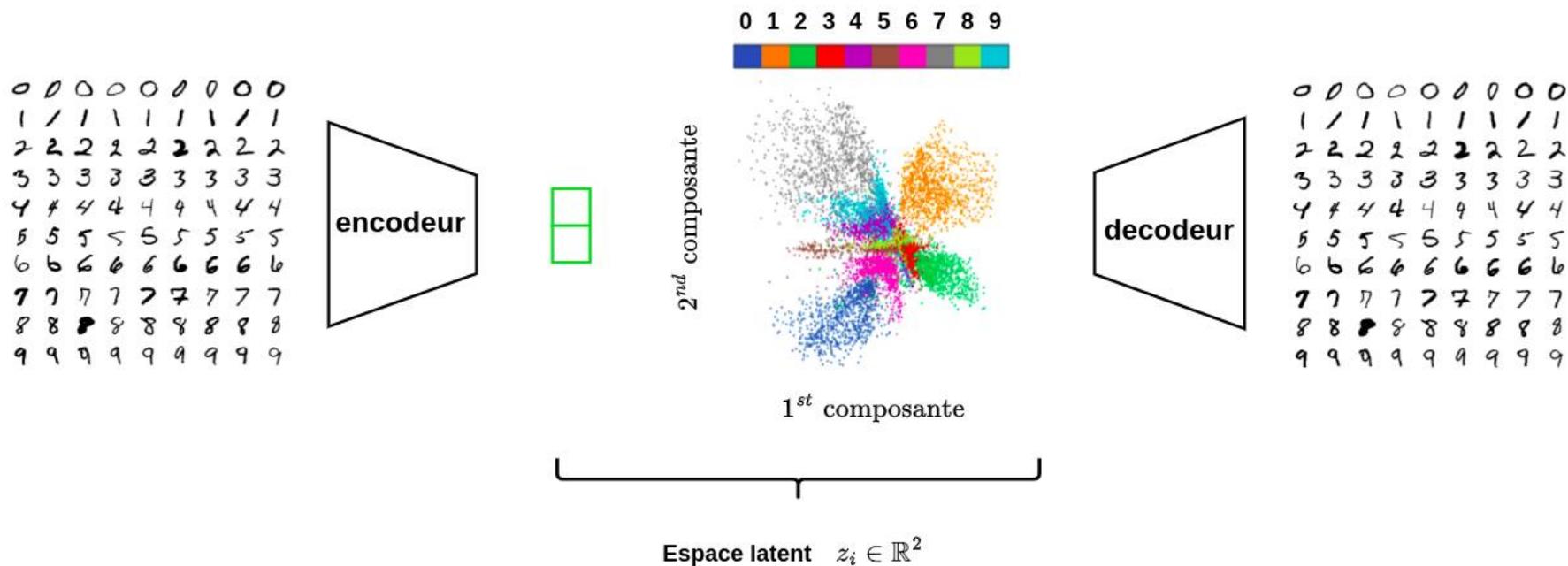


0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9

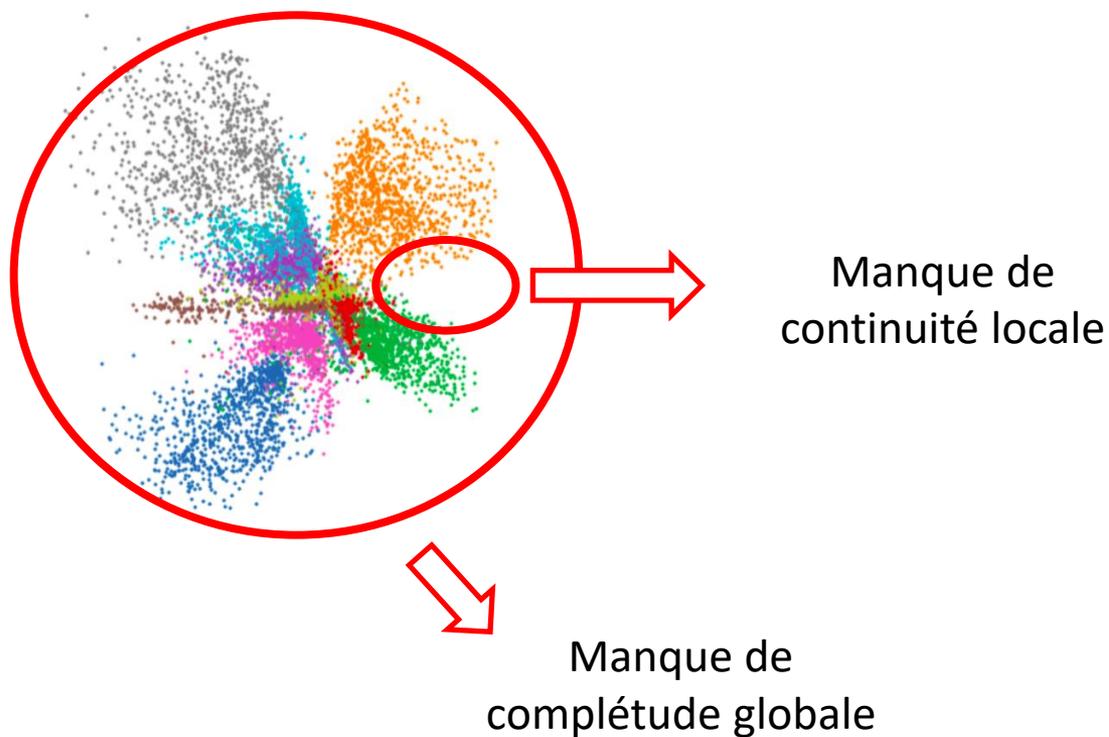
## ► Représentation de données



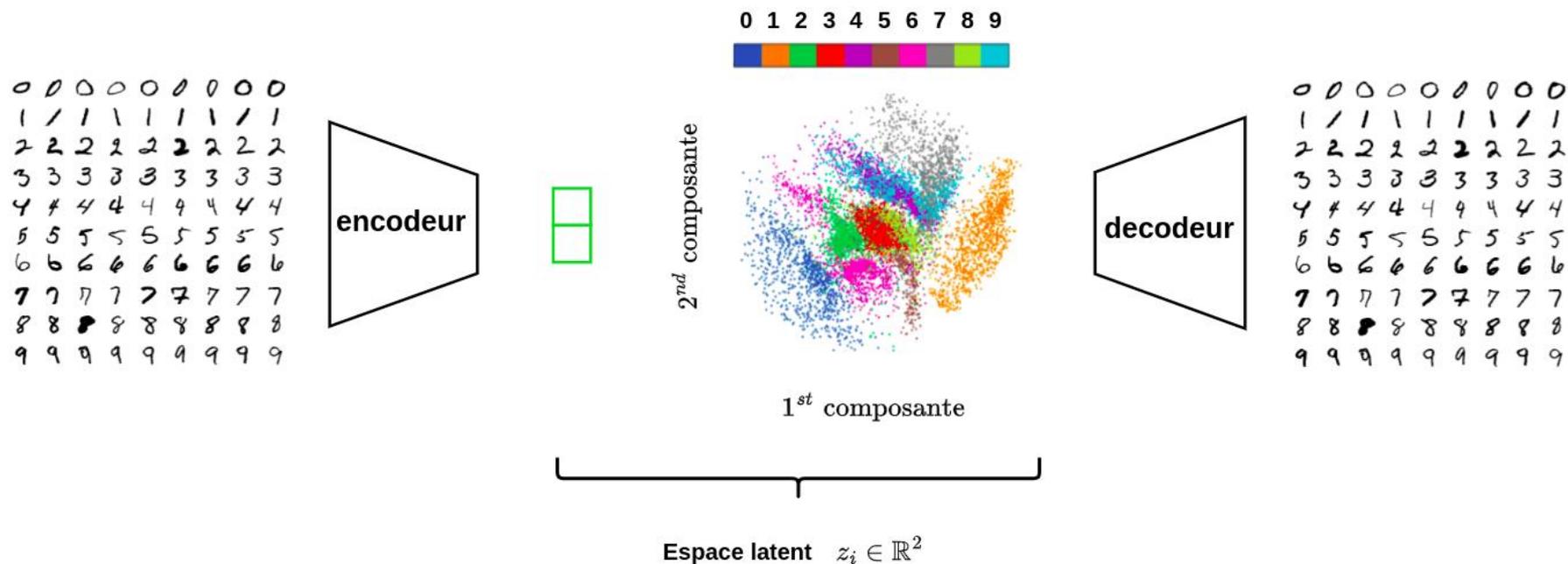
## ► Modèles génératifs



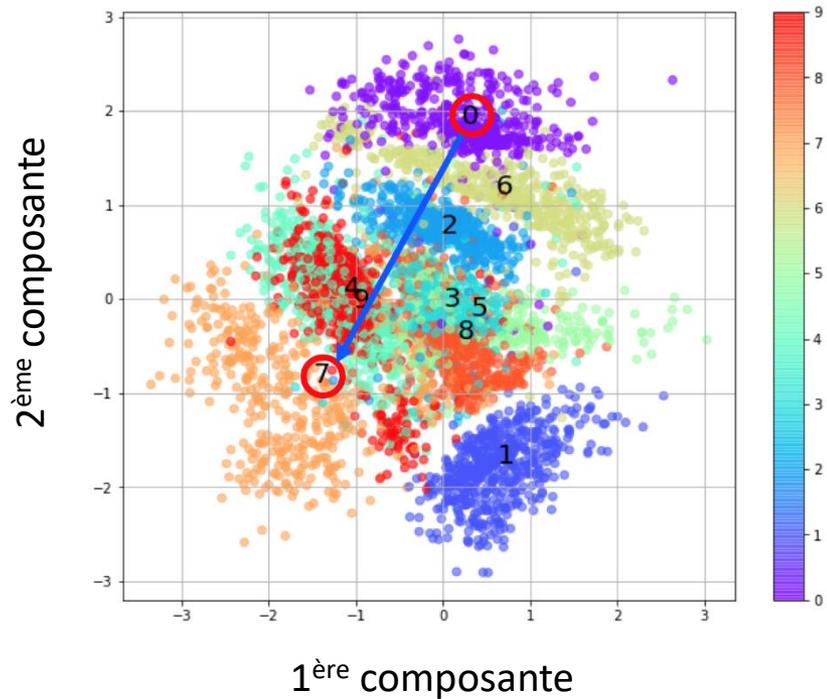
- Nécessité de mieux contrôler la structure de l'espace latent



- ▶ Modèle génératif avec de meilleures propriétés grâce au *cadre variationnel*

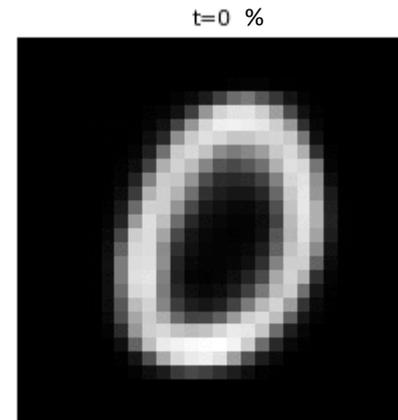


## ► Modèle génératif avec cadre variationnel



Interpolation linéaire dans l'espace latent

$$t \cdot z_0 + (1 - t) \cdot z_7, \quad 0 \leq t \leq 1$$



# Les auto-encodeurs variationnels

---

*L'ensemble des mathématiques sont décrites dans le blog suivant*

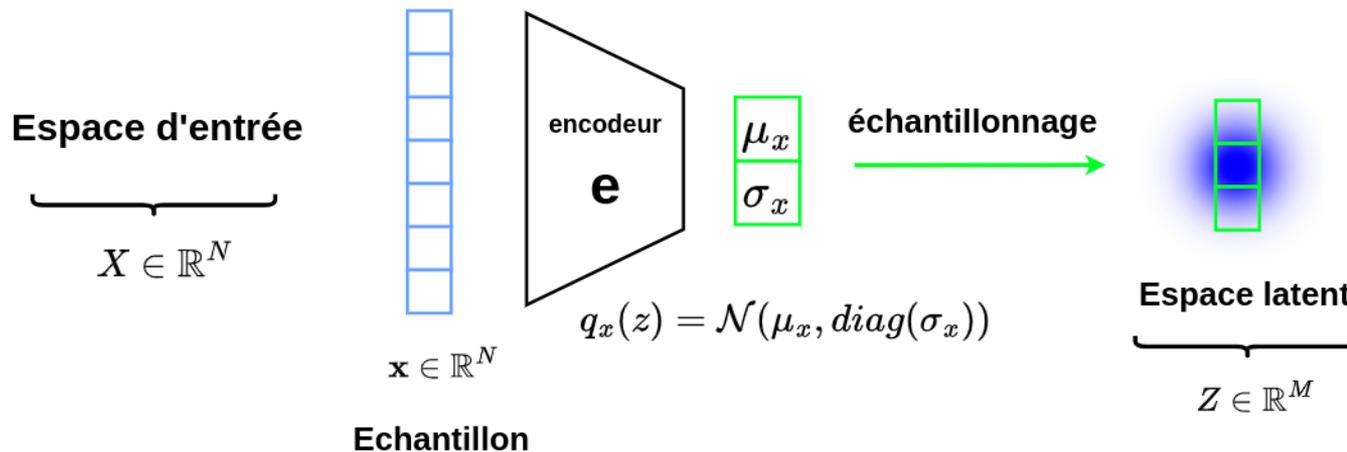
<https://creatis-myriad.github.io/tutorials/2022-09-12-tutorial-vae.html>

- ▶ Renforcement d'un espace latent structuré
  - Au travers d'un cadre probabiliste
  - En imposant des contraintes de continuité
  - En imposant des contraintes de complétude

► Cadre probabiliste: *continuité*

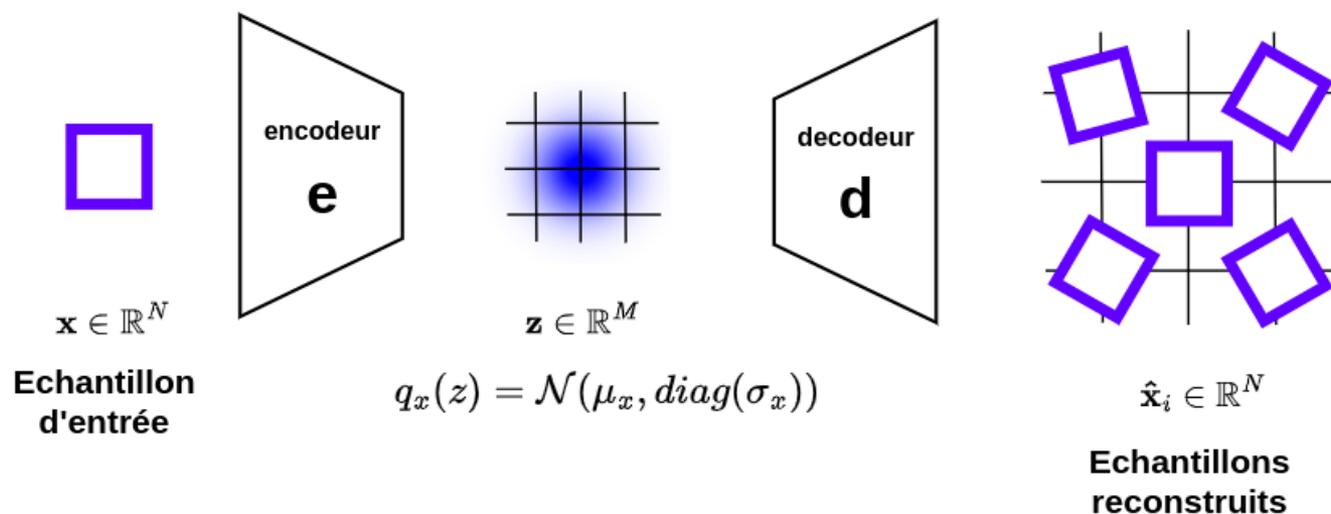
→ Introduction de régularisations locales de l'espace latent

→ Chaque donnée d'entrée  $x$  est encodée sous forme d'une distribution gaussienne  $q_x(z) = \mathcal{N}(\mu_x, \text{diag}(\sigma_x))$



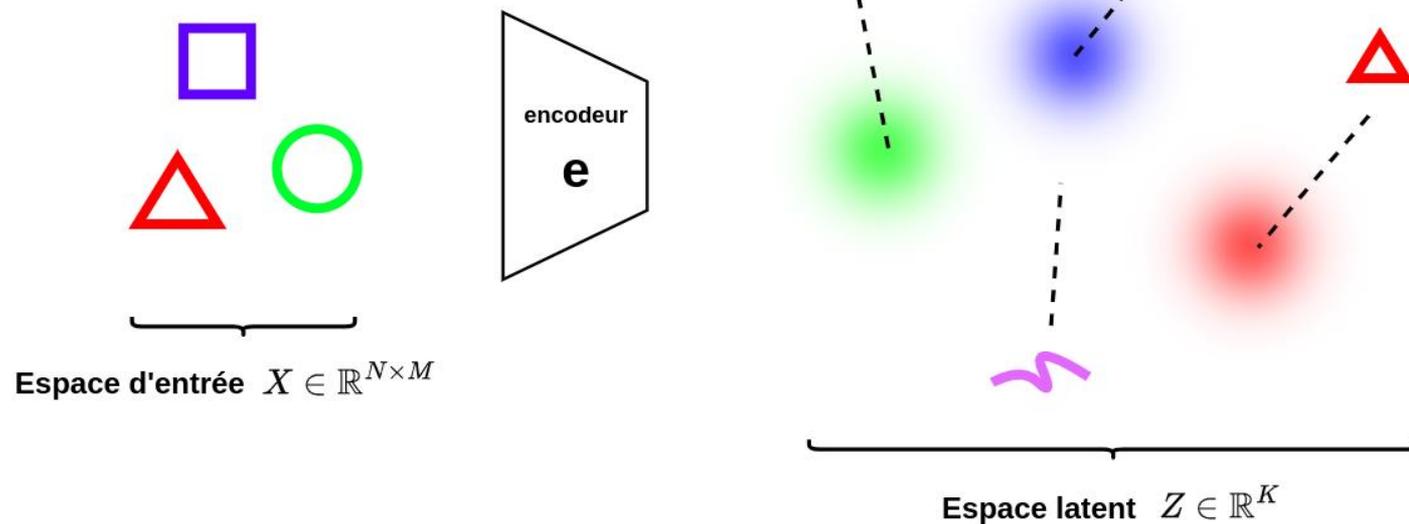
► Cadre probabiliste: *continuité*

→ L'échantillonnage d'une région locale de l'espace latent produit des résultats proches



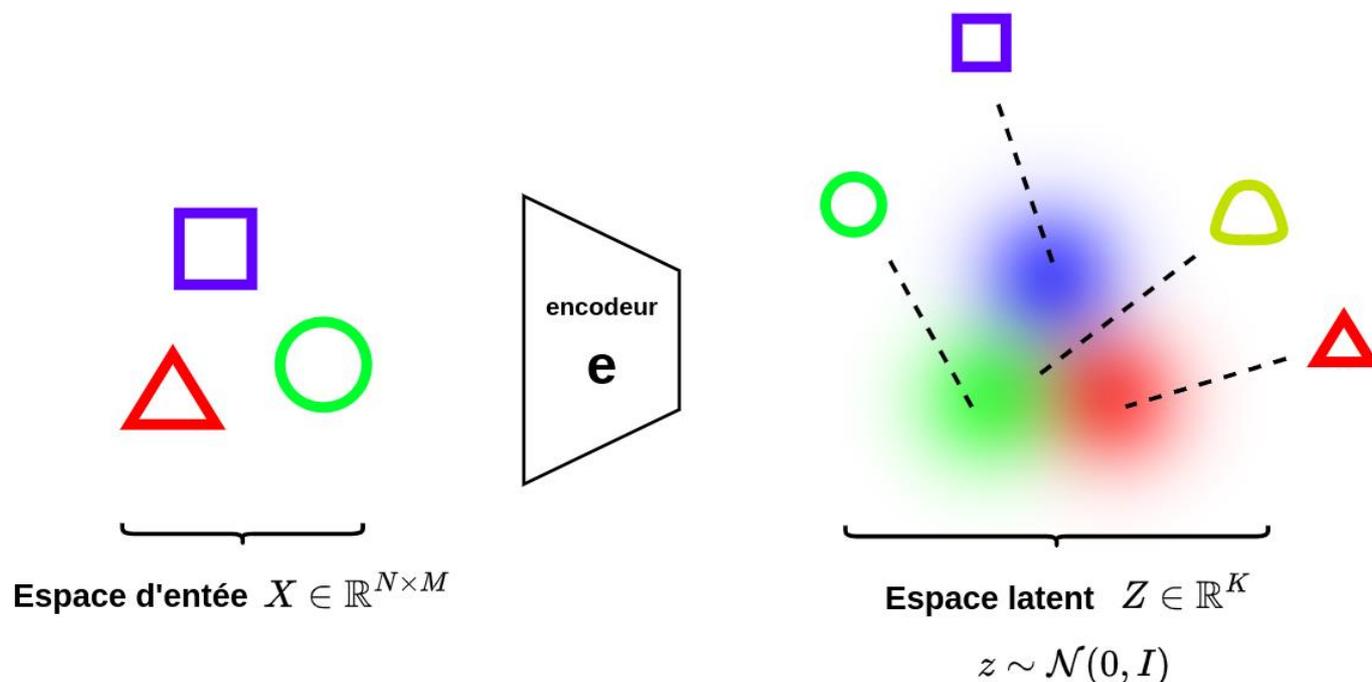
► Cadre probabiliste: *complétude*

→ Favoriser que tout point reconstruit dans l'espace latent donne des résultats cohérents



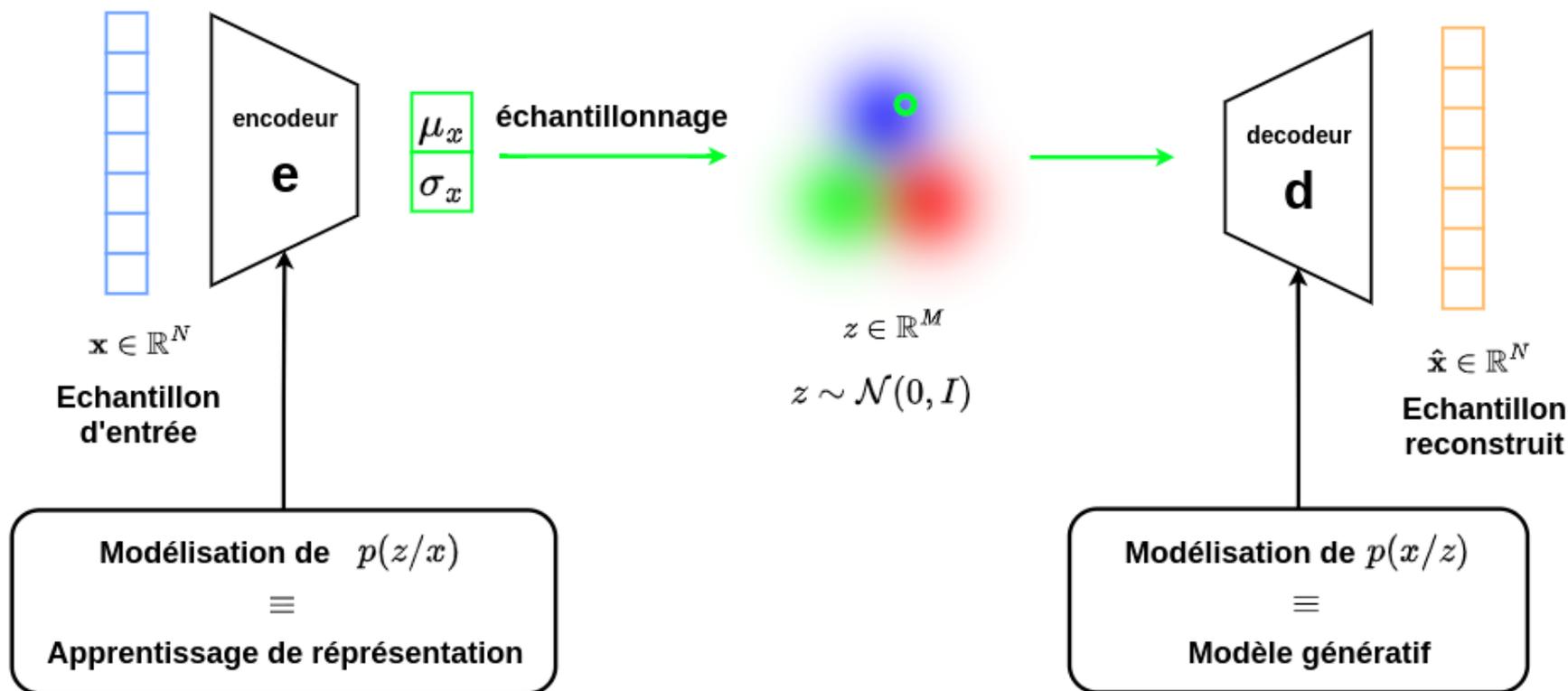
## ► Cadre probabiliste: *complétude*

- Imposer que toutes les distributions  $q_x(z)$  soient proches d'une distribution normal standard  $\mathcal{N}(0, I)$
- Variances proches de 1  $\Rightarrow$  limite la génération de distributions ponctuelles
- Moyennes proches de 0  $\Rightarrow$  favorise des distributions proches les unes des autres

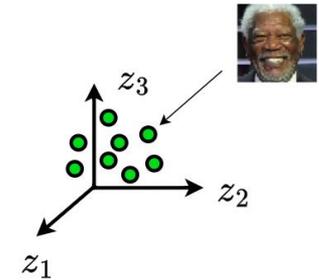
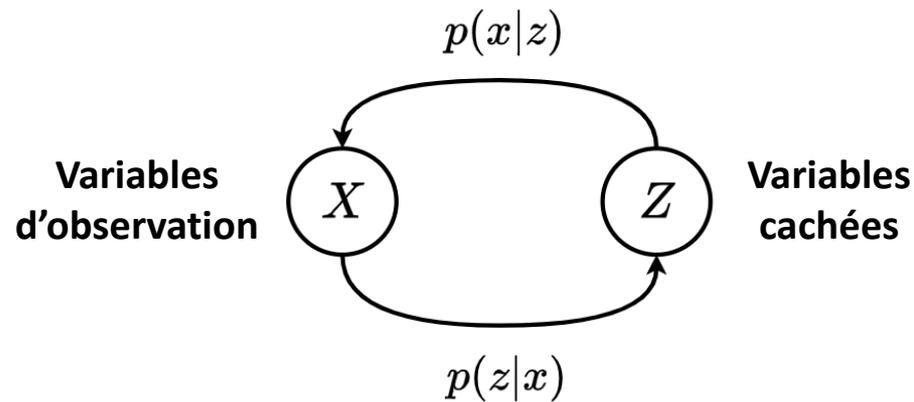


## ► Cadre probabiliste: *continuité & complétude*

### → Architecture des VAE



## ► Formulation mathématique



Approximation de  $p(z|x)$  par une technique d'inférence variationnelle

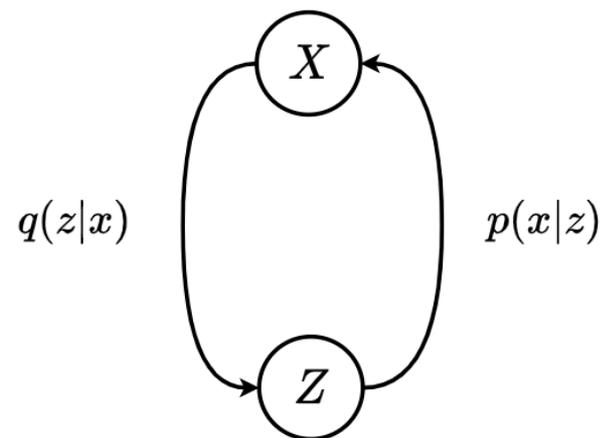
## ► Hypothèses

→  $q(z|x)$  est modélisée par une distribution gaussienne alignée sur les axes

→  $q(z|x) = \mathcal{N}(\mu_x, \sigma_x) = \mathcal{N}(g(x), \text{diag}(h(x)))$

$$(g^*, h^*) = \arg \min_{(g, h)} D_{KL}(q(z|x) \parallel p(z|x))$$

$D_{KL}(\cdot \parallel \cdot)$  fonction de divergence de Kullback-Liebler



## ► Processus d'optimisation

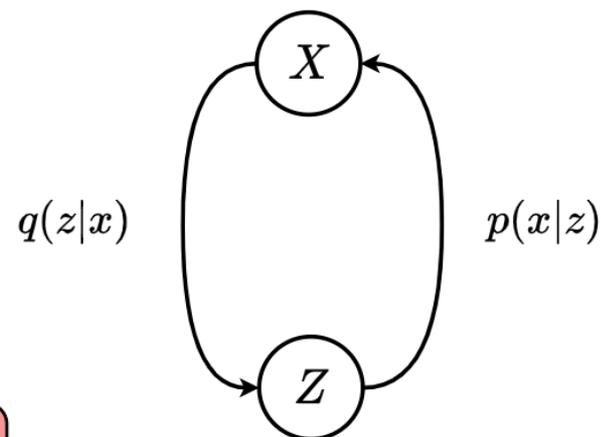
→ Maximisation de la limite inférieure de l'évidence (ELBO)

$$\mathcal{L} = \mathbb{E}_{z \sim q_x} [\log(p(x|z))] - D_{KL}(q(z|x) \parallel p(z))$$

→ En exploitant l'hypothèse gaussienne suivante

$$p(x|z) = \mathcal{N}(f(z), cI)$$

$$\mathcal{L} \propto \mathbb{E}_{z \sim q_x} [-\alpha \|x - f(z)\|^2] - D_{KL}(q(z|x) \parallel p(z))$$



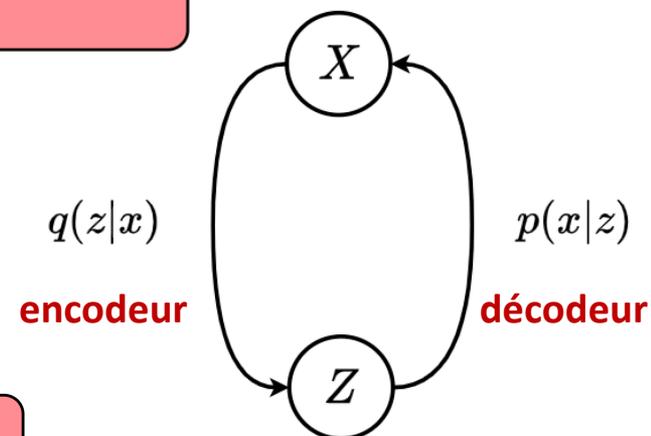
## ► Processus d'optimisation

$$(f^*, g^*, h^*) = \arg \min_{(f, g, h)} (\mathbb{E}_{z \sim q_x} [\alpha \|x - f(z)\|^2] + D_{KL}(q(z|x) \parallel p(z)))$$

## ► Fonction de perte en apprentissage profond

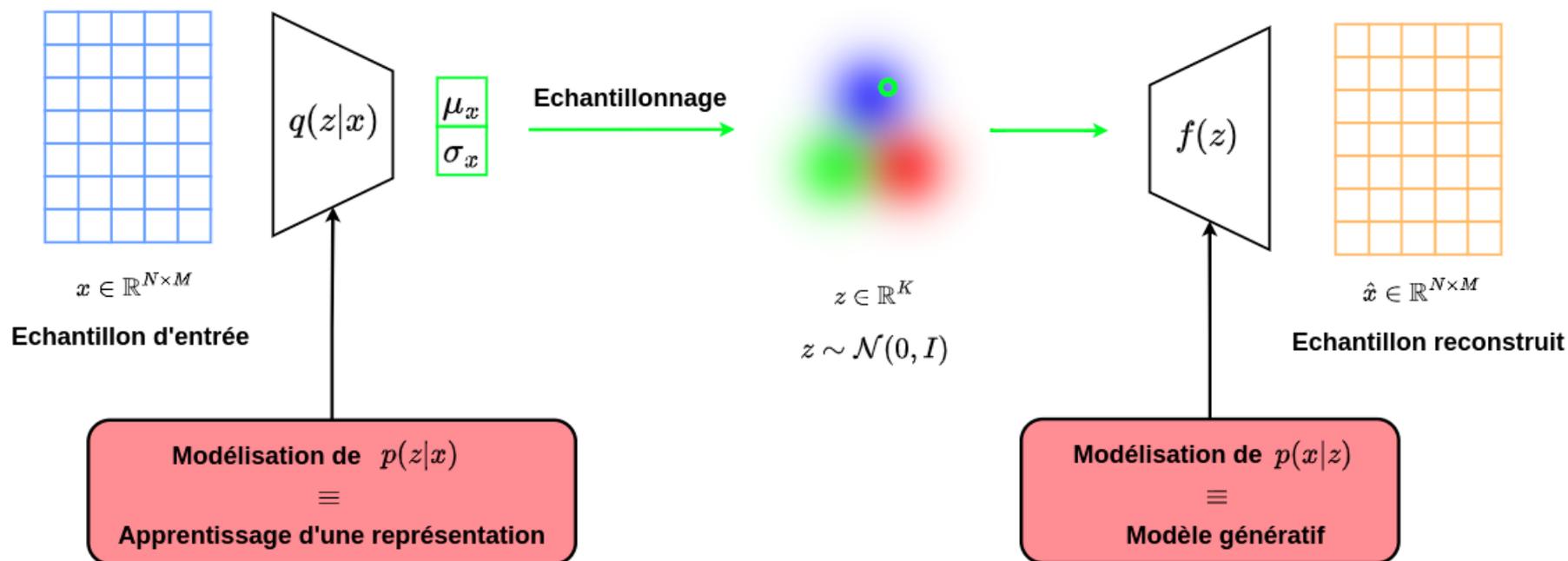
$$\text{loss} = \alpha \|x - f(z)\|^2 + D_{KL}(\mathcal{N}(g(x), \text{diag}(h(x))), \mathcal{N}(0, I))$$

- $g(\cdot)$  et  $h(\cdot)$  sont modélisés par un encodeur
- $f(\cdot)$  est modélisée par un décodeur



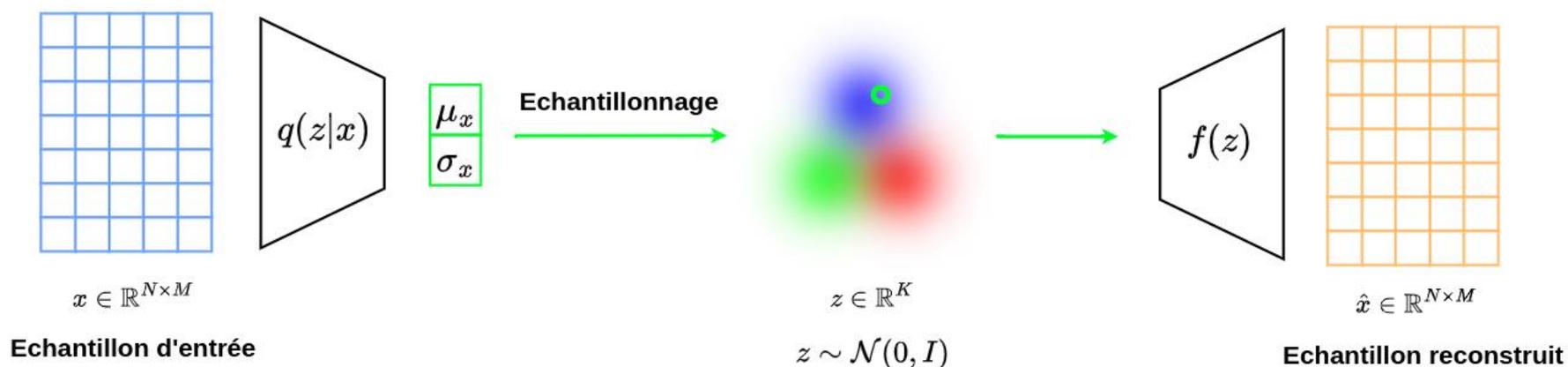
## ► Interprétation de la fonction de perte

$$\text{loss} = D_{KL}(\mathcal{N}(g(x), \text{diag}(h(x))), \mathcal{N}(0, I)) + \alpha \|x - f(z)\|^2$$



## ► Interprétation de la fonction de perte

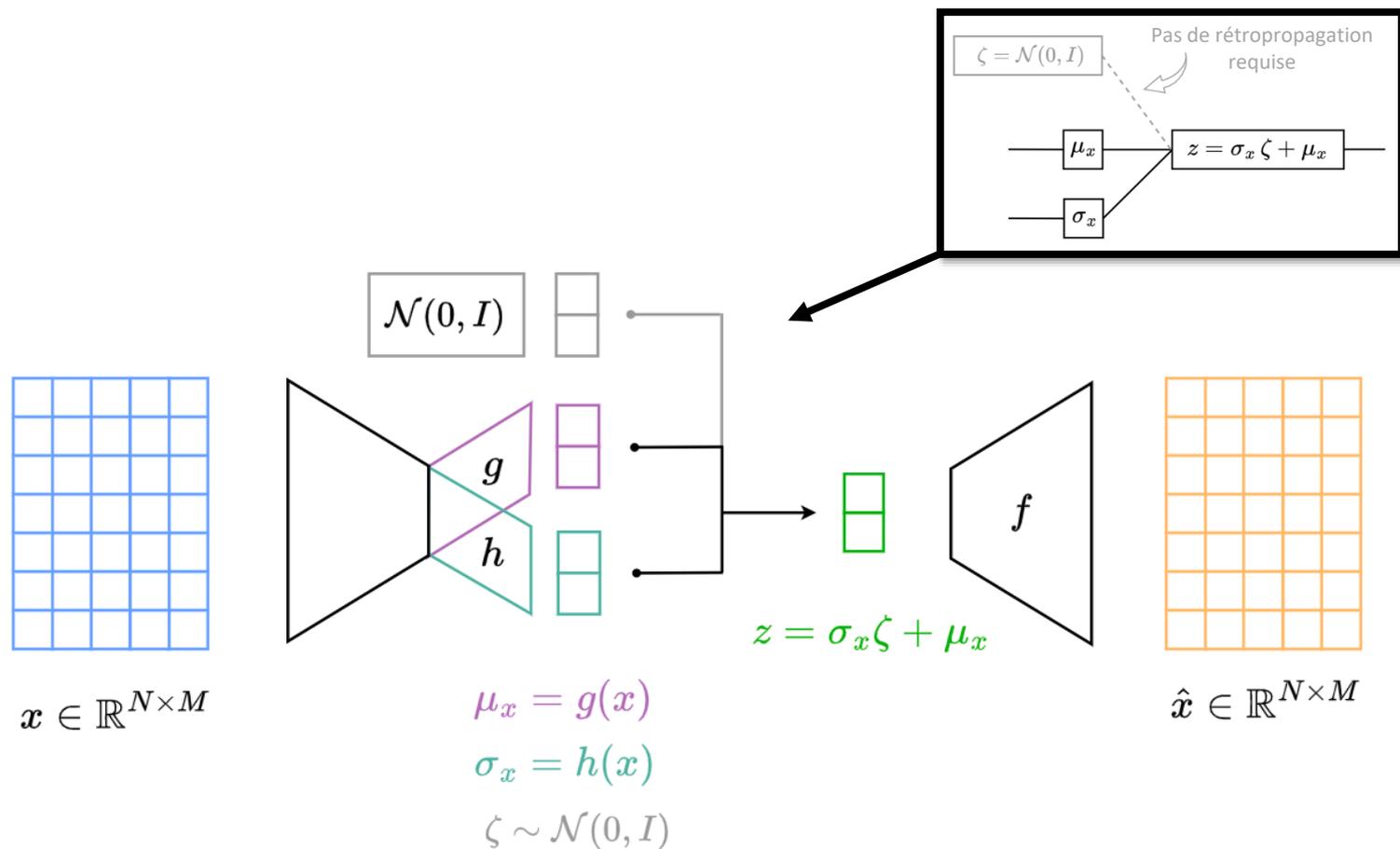
$$\text{loss} = D_{KL}(\mathcal{N}(g(x), \text{diag}(h(x))), \mathcal{N}(0, I)) + \alpha \|x - f(z)\|^2$$



→  $\mathcal{N}(g(x), h(x))$  impose une contrainte de **continuité** locale

→  $\mathcal{N}(\cdot, \mathcal{N}(0, I))$  impose une contrainte de **complétude** globale

## ► Astuce de reparamétrage



---

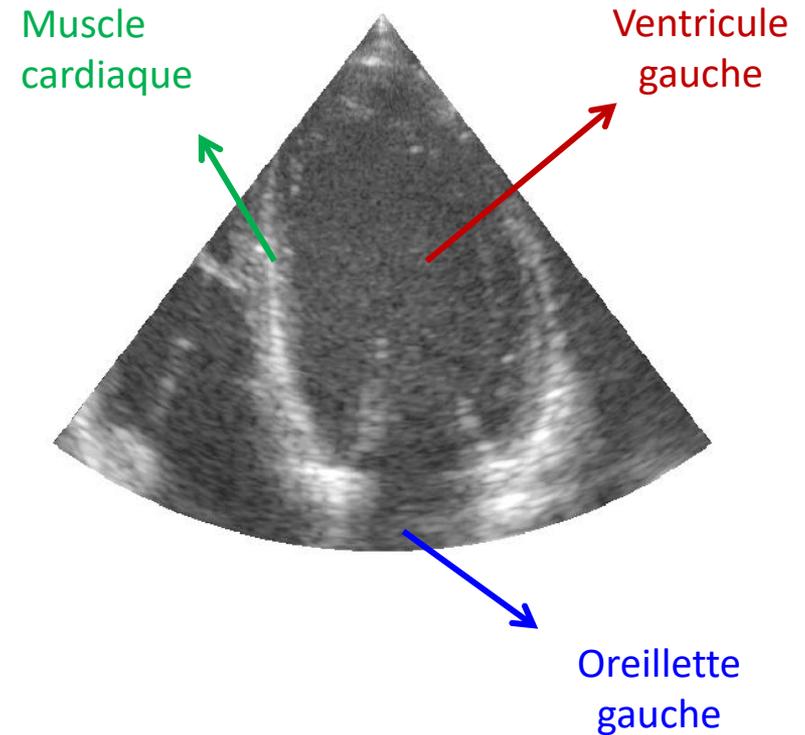
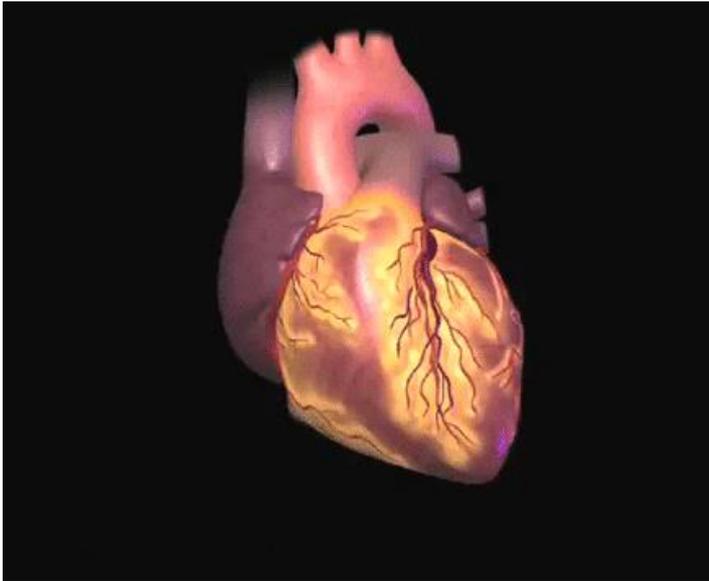
# Application pratique

---

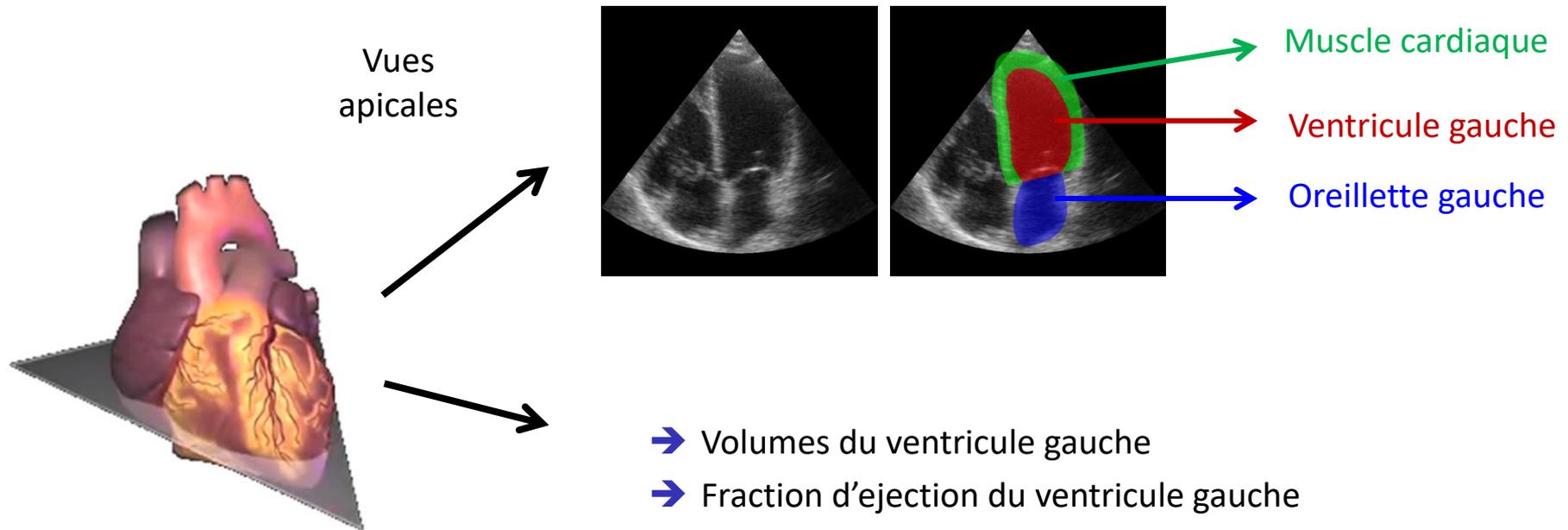
L'obsession est de maîtriser l'espace latent !!

## Besoin d'une segmentation précise et robuste des structures cardiaques

- Quantification d'indices cliniques à partir d'images échocardiographiques



## ► Indices cliniques anatomiques



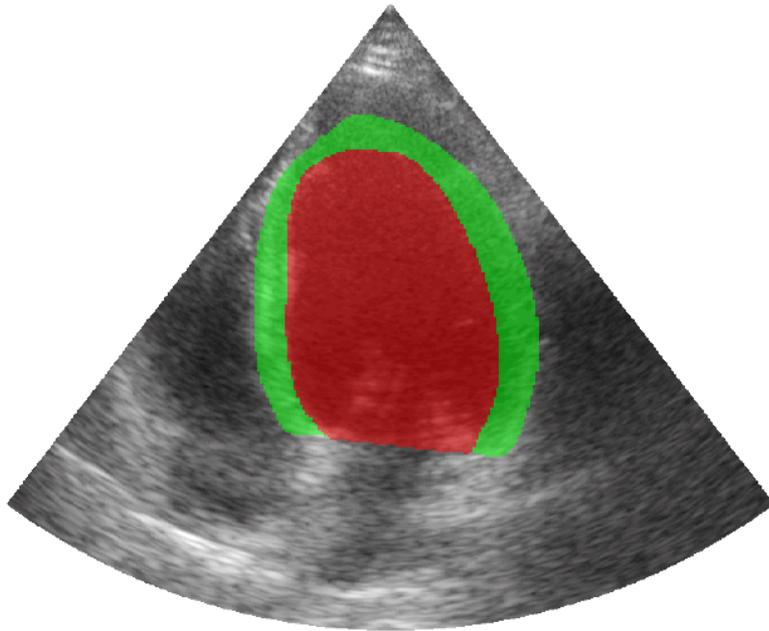
---

Comment garantir une cohérence temporelle des scores cliniques ?

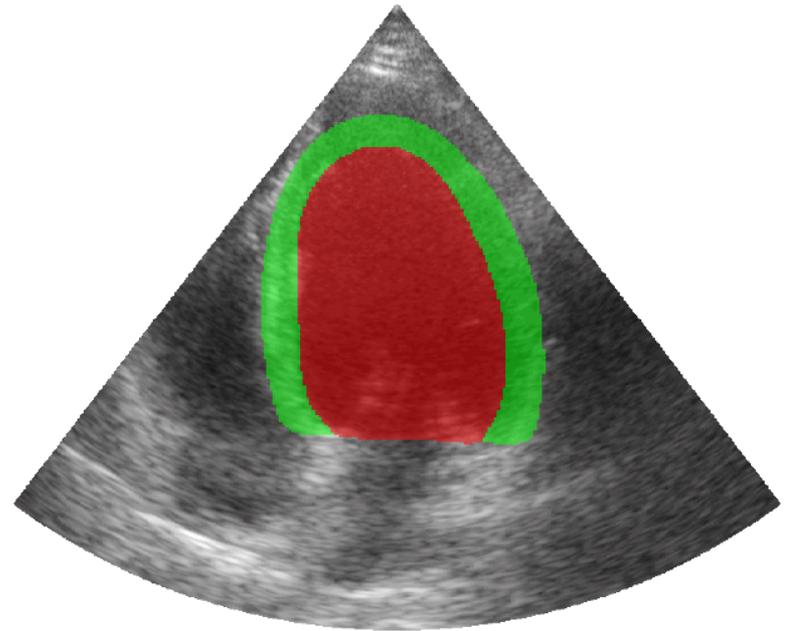
---

- ▶ Quantification d'indices cliniques à partir d'images échocardiographiques

Résultats issus d'un U-Net 2D



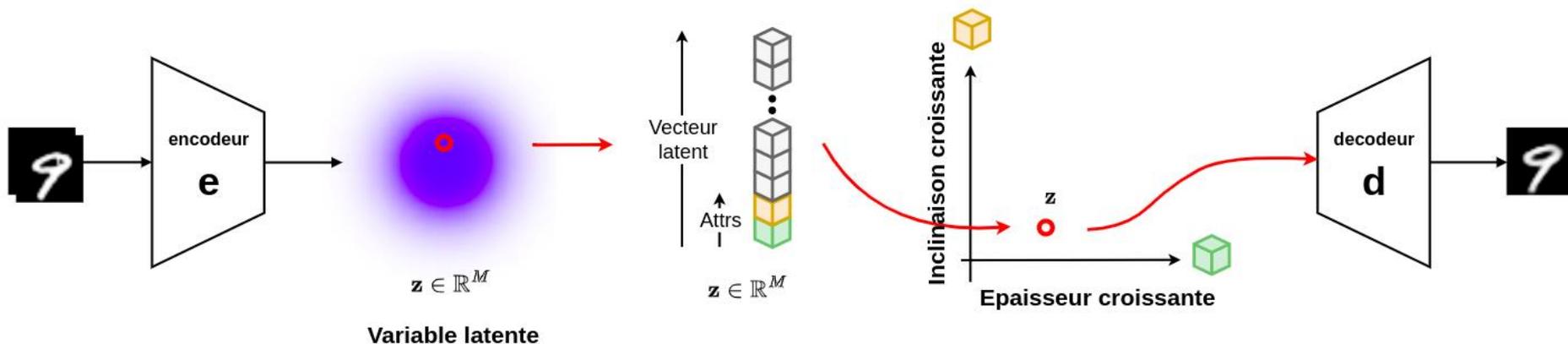
Ce que nous voulons



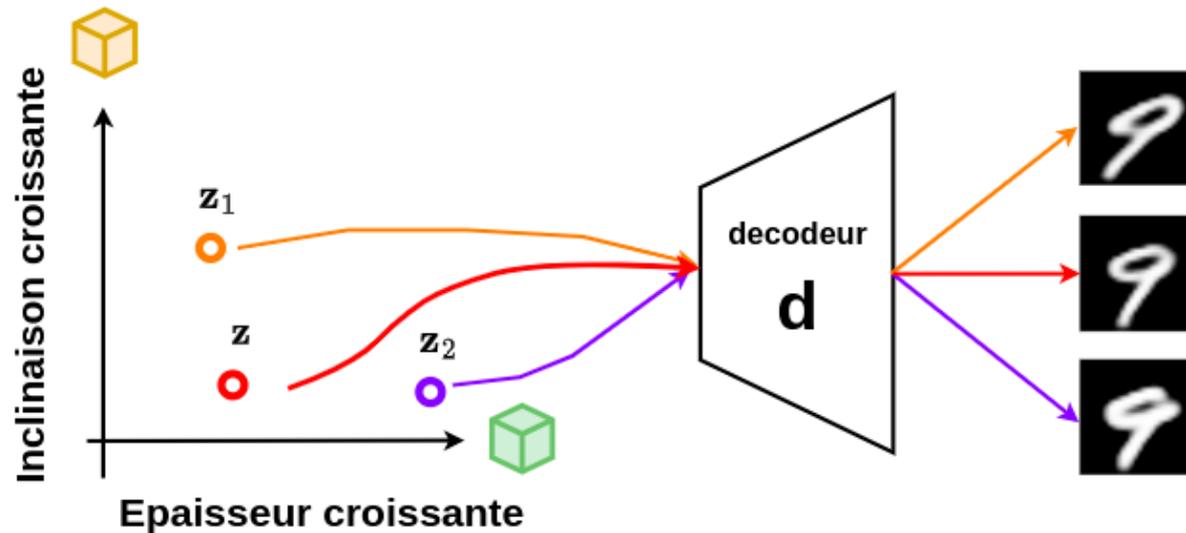
## ► AR-VAE : régularisation de l'espace latent des VAE basée sur les attributs

- Génération d'un espace latent structuré

- Les attributs spécifiques à valeur continue doivent être codés selon des dimensions spécifiques
- $Loss = VAE\ loss + Attribute\ Regularisation\ Loss$



- ▶ AR-VAE : régularisation de l'espace latent des VAE basée sur les attributs
  - Echantillonnage de l'espace latent structuré

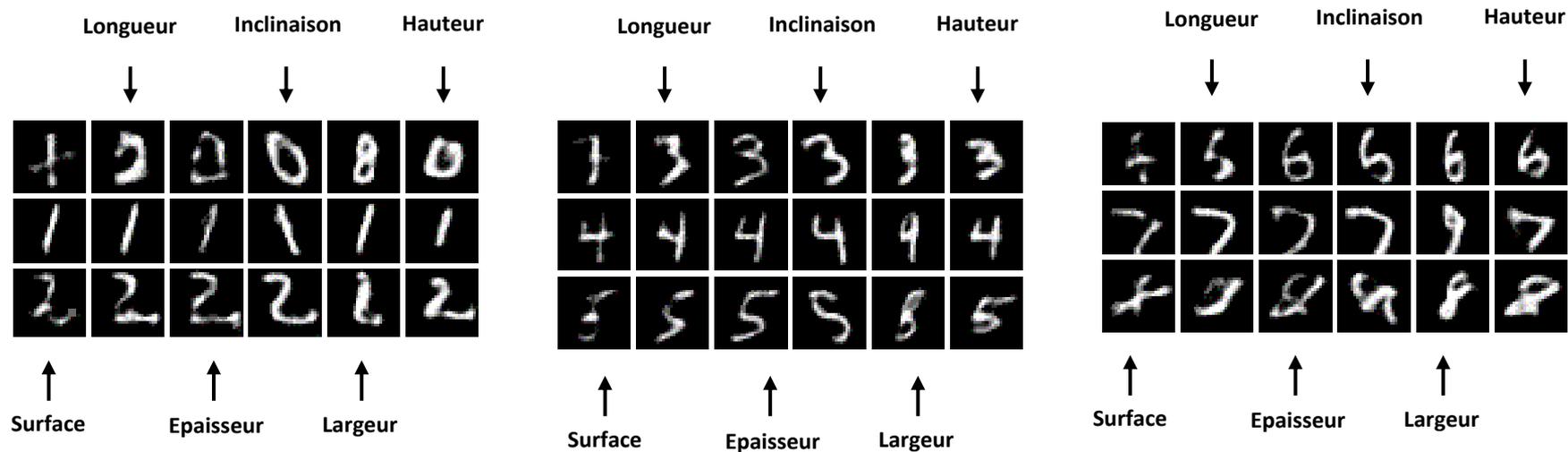


## ► AR-VAE : régularisation de l'espace latent des VAE basée sur les attributs

- Echantillonnage de l'espace latent structuré

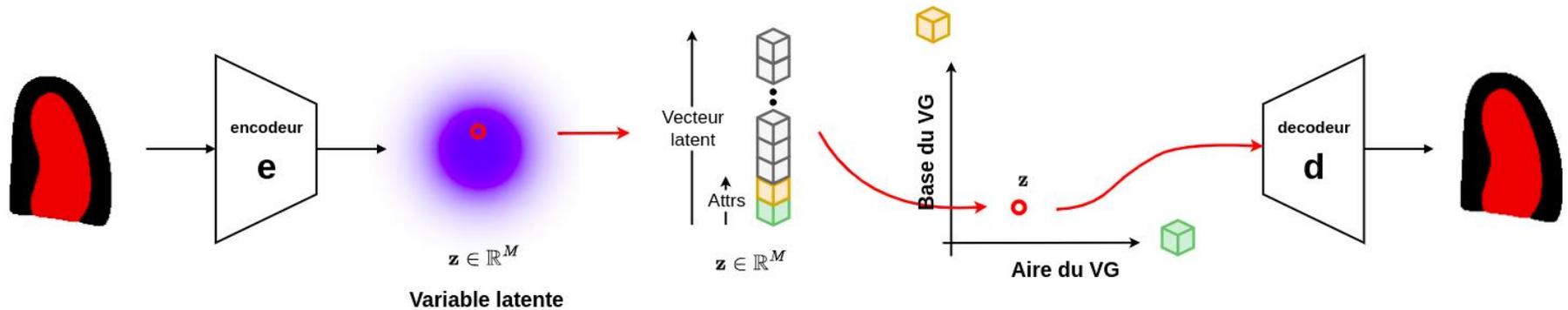
- ➔ Attribut spécifique: surface, longueur, épaisseur, inclinaison, largeur, hauteur

- ➔ Chaque colonne correspond à une traversée le long d'une dimension régularisée

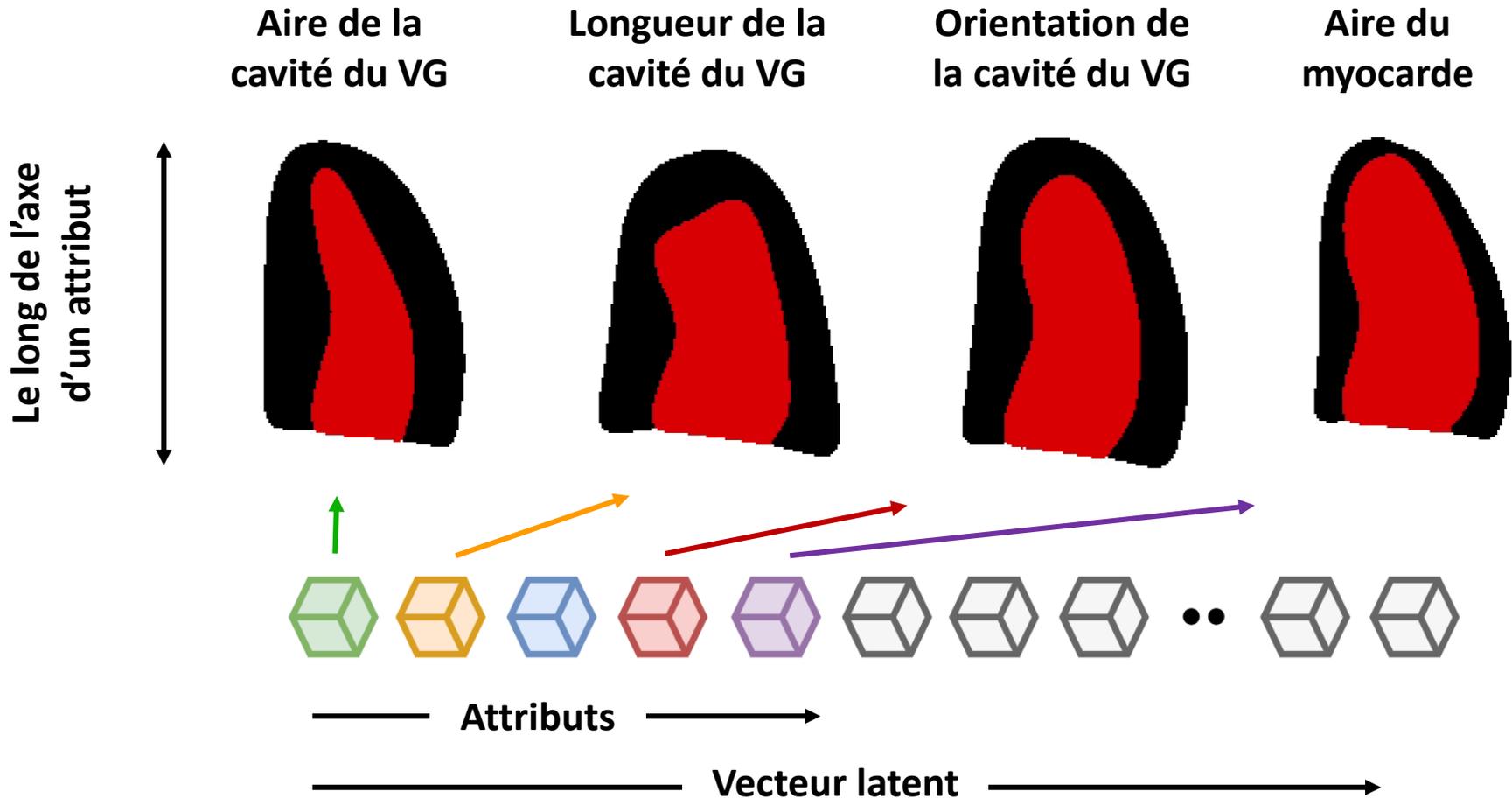


## ► Application à la description des formes cardiaques

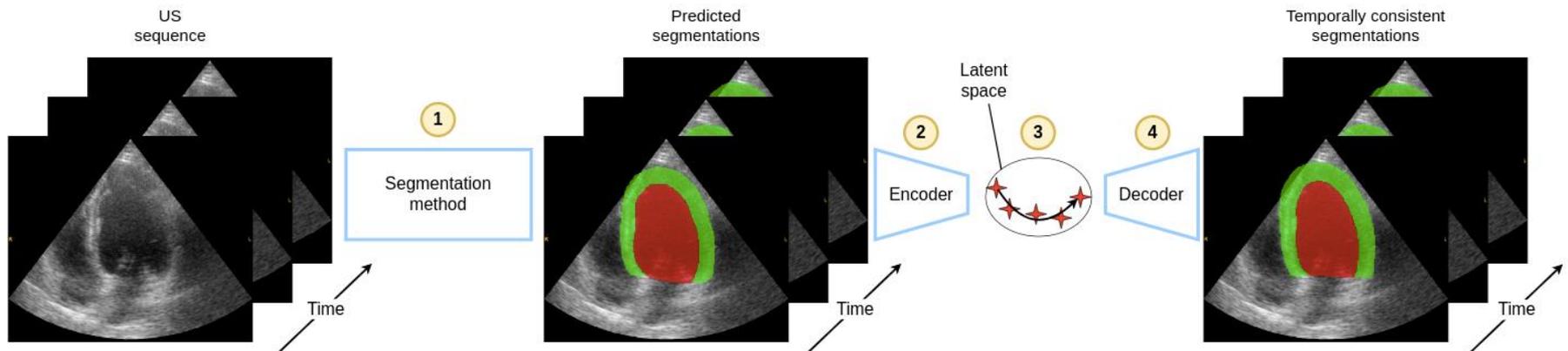
- Génération d'un espace latent structuré en fonction des attributs suivants
  - ➔ Cavité du ventricule gauche (VG) : surface, longueur, largeur basale, orientation
  - ➔ Surface du myocarde
  - ➔ Centre de la paroi épicardique



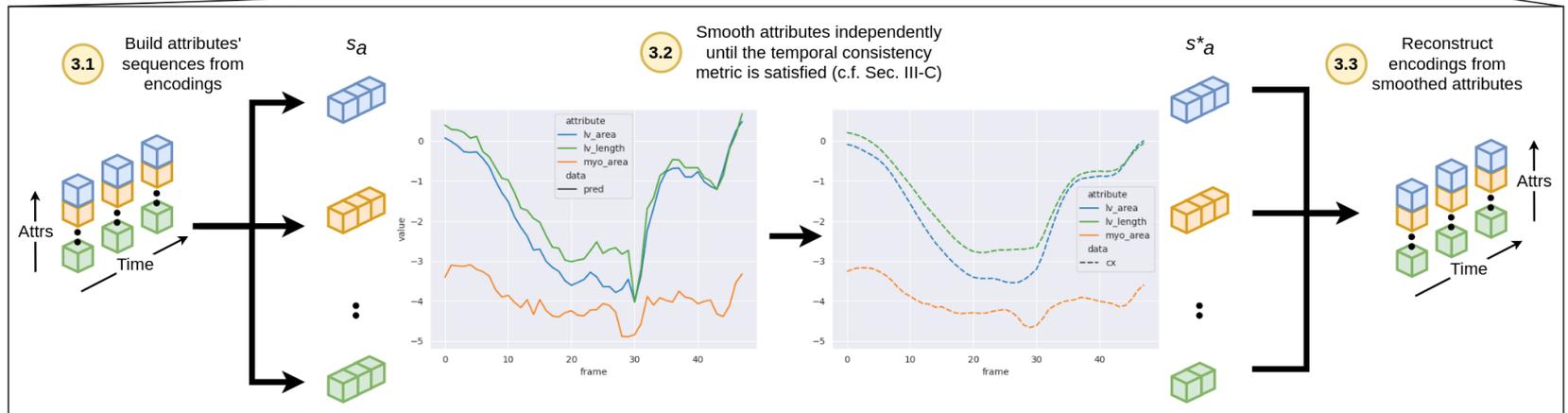
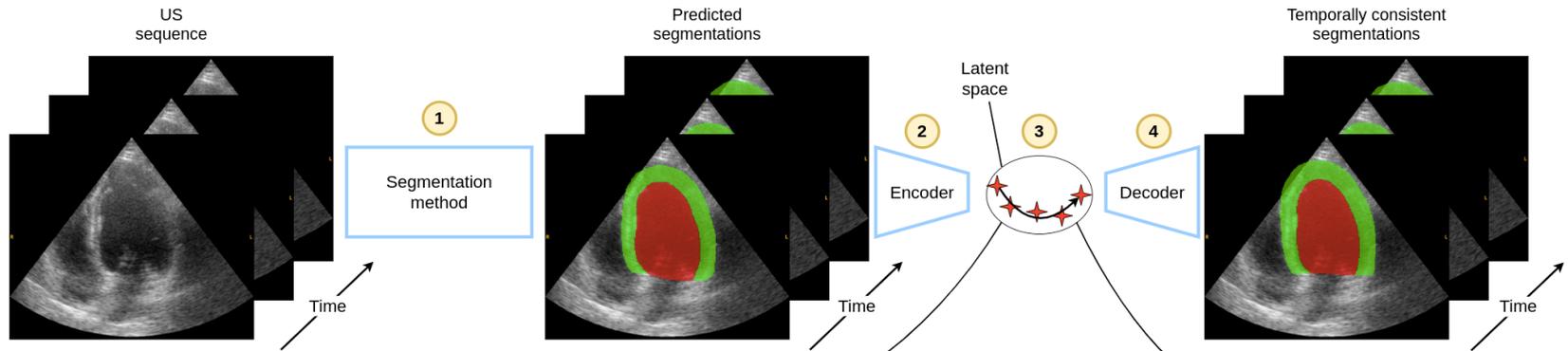
# Segmentation cardiaque avec une consistance temporelle



## ► Chaîne de traitement développée

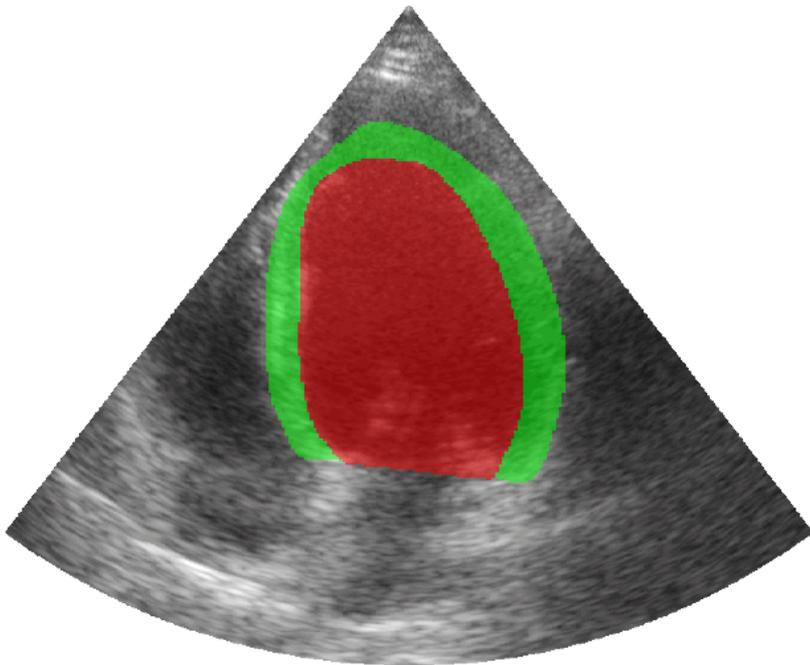


# Segmentation cardiaque avec une consistance temporelle

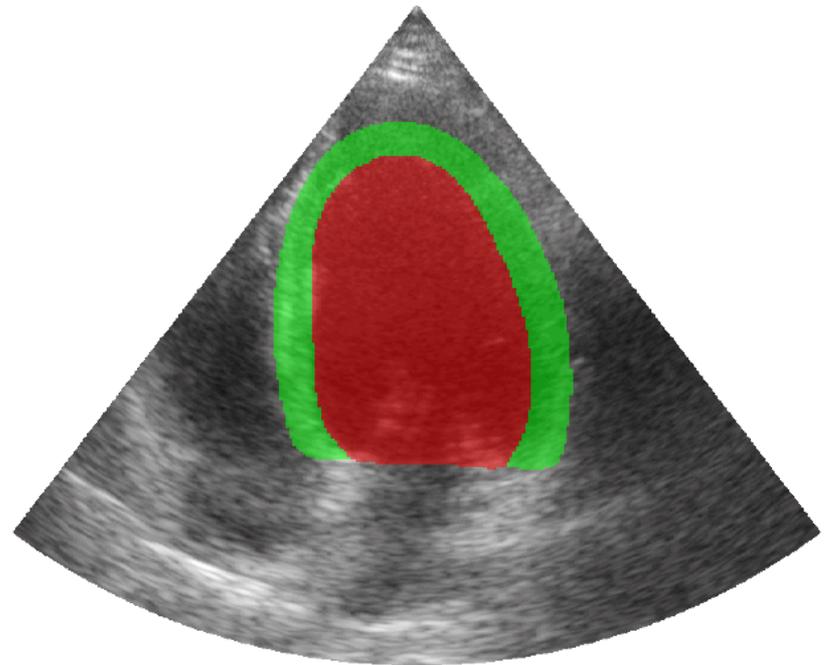


- ▶ Quelques exemples de résultats

Résultat issu d'un U-Net 2D

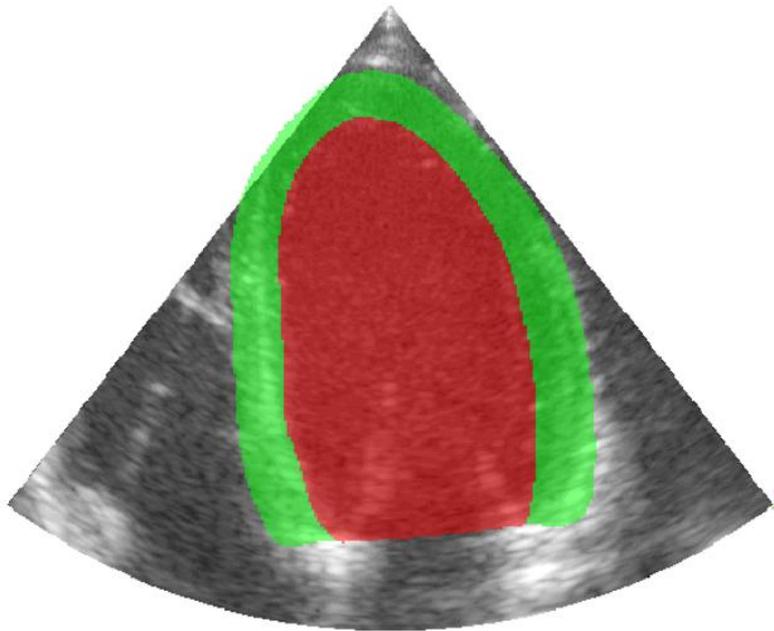


Résultat corrigé

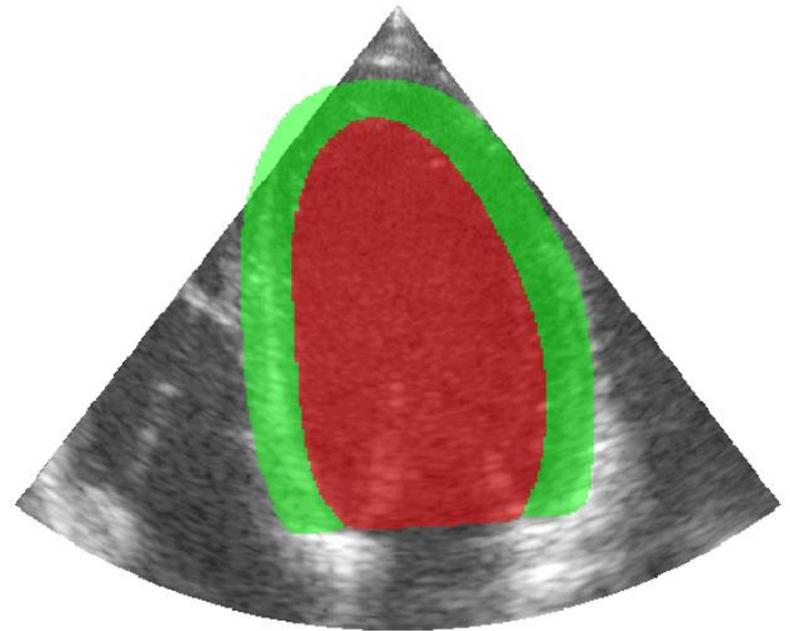


► Quelques exemples de résultats

Résultat issu d'un U-Net 2D



Résultat corrigé



---

That's all folks

---