Generative models

Pr. Olivier Bernard
Lab. CREATIS – Univ. of Lyon, France

Pr. Christian Desrosier
ETS, Montreal

CREATIS
UNIVERSITÉ DE LYON

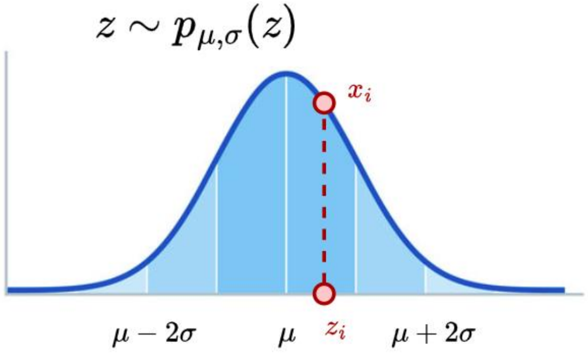
ÉTS Le génie pour l'industrie
ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
Université du Québec

What is the interest of generative models ?

- **How to generate synthetic faces ?**

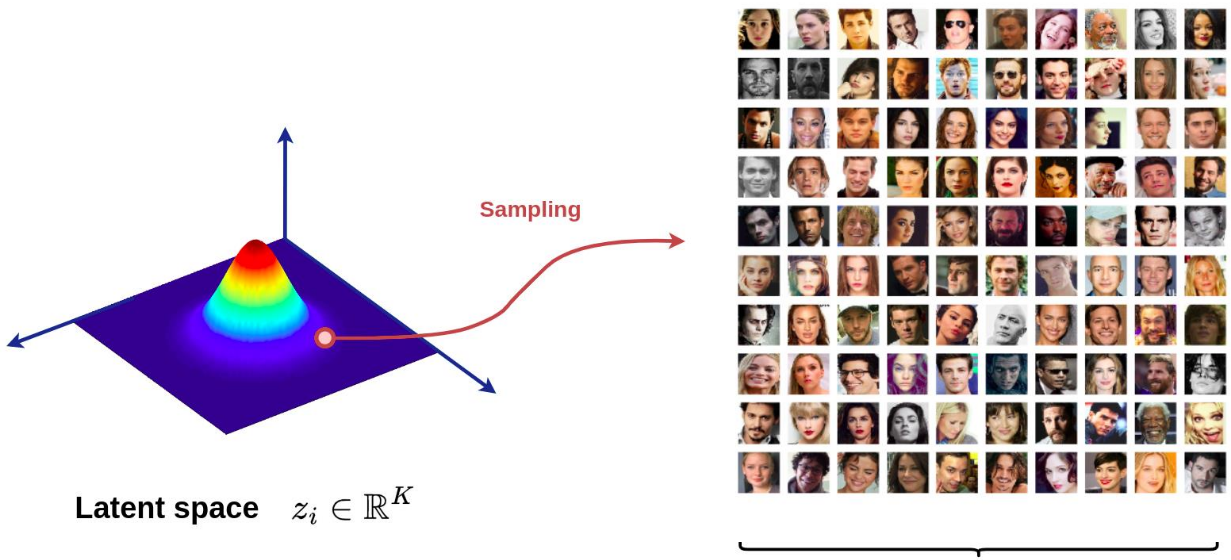
By modeling the corresponding distribution $p_\theta(\cdot)$!

→ **Reminder: normal distribution**



What are the interest of generative models ?

- **How to model complex distributions ?**



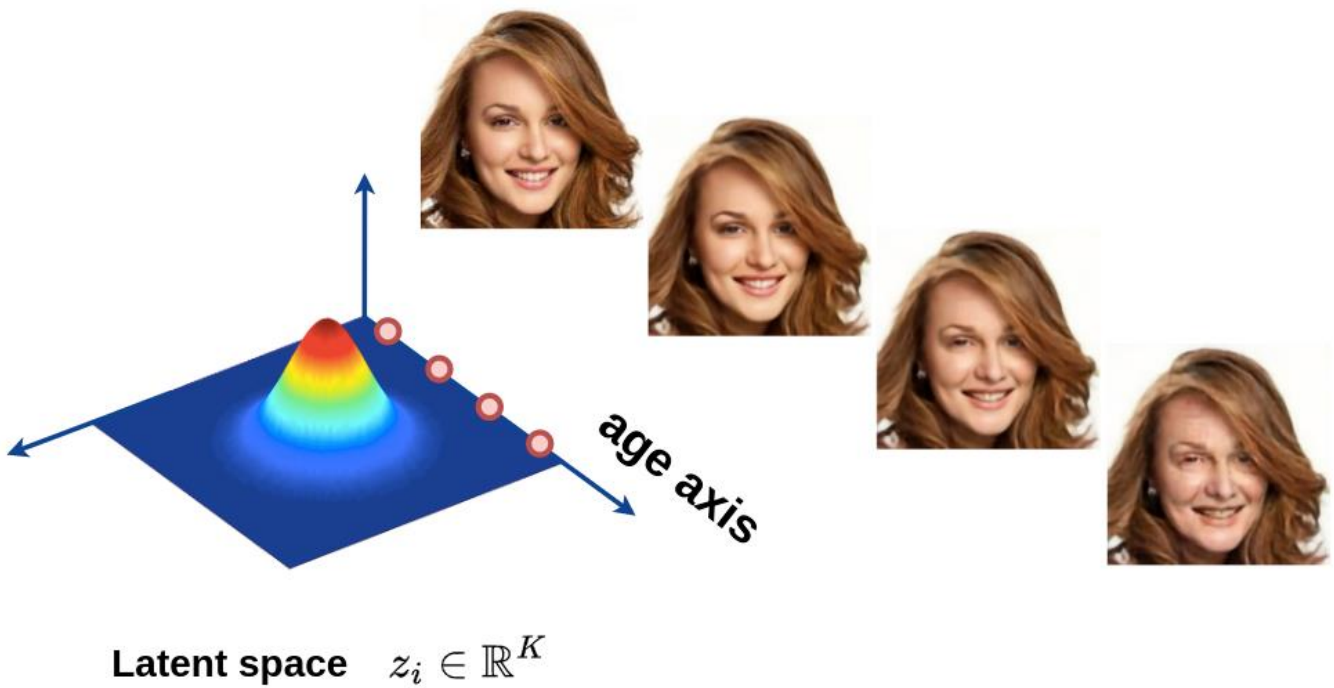
Latent space $z_i \in \mathbb{R}^K$

Face distribution

What are the interest of generative models ?

- **What for ?**

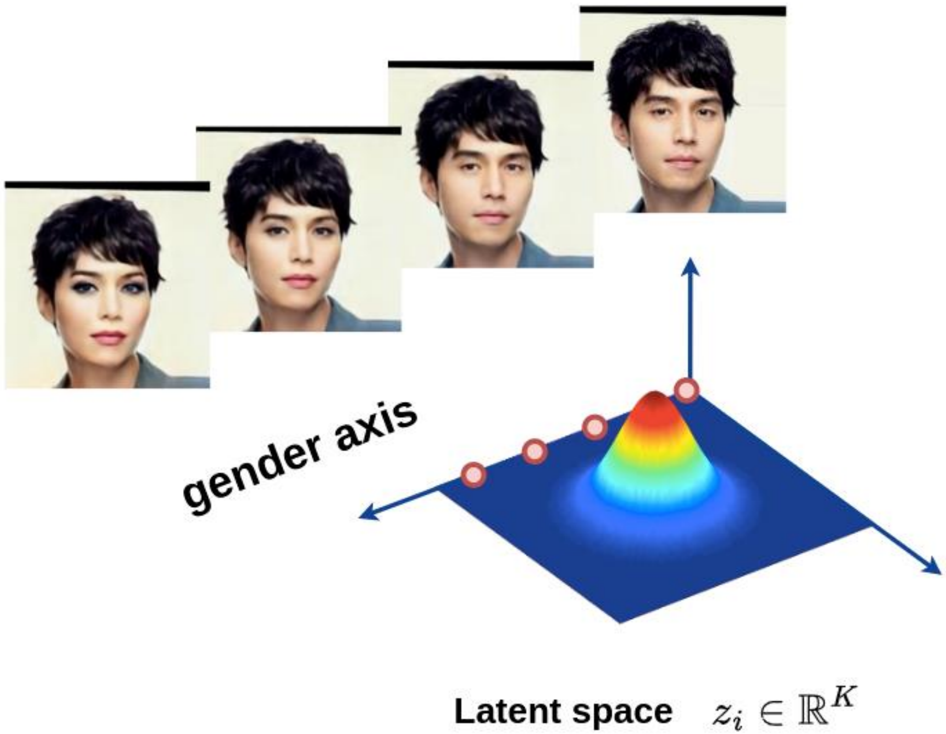
One obsession is to master the latent space !!!



What are the interest of generative models ?

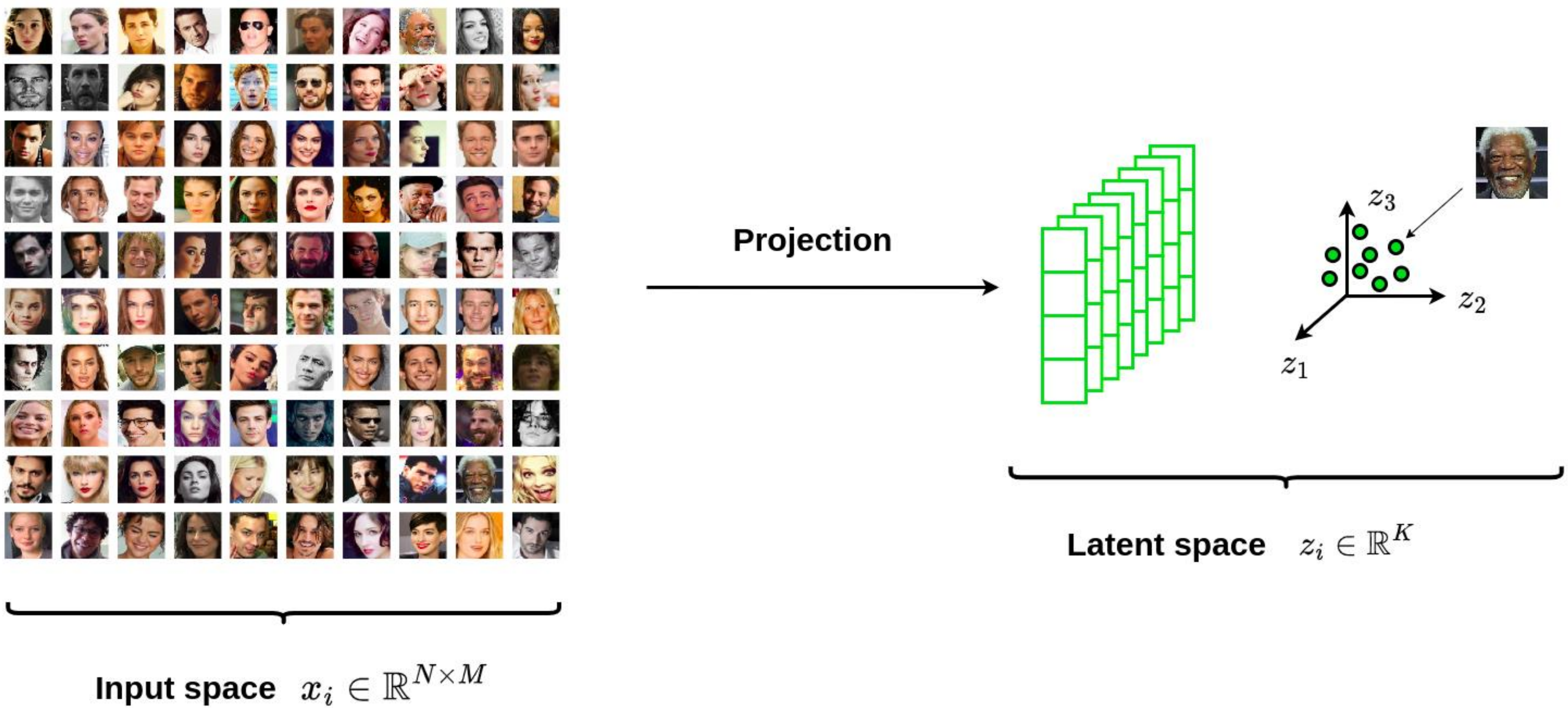
- **What for ?**

One obsession is to master the latent space !!!



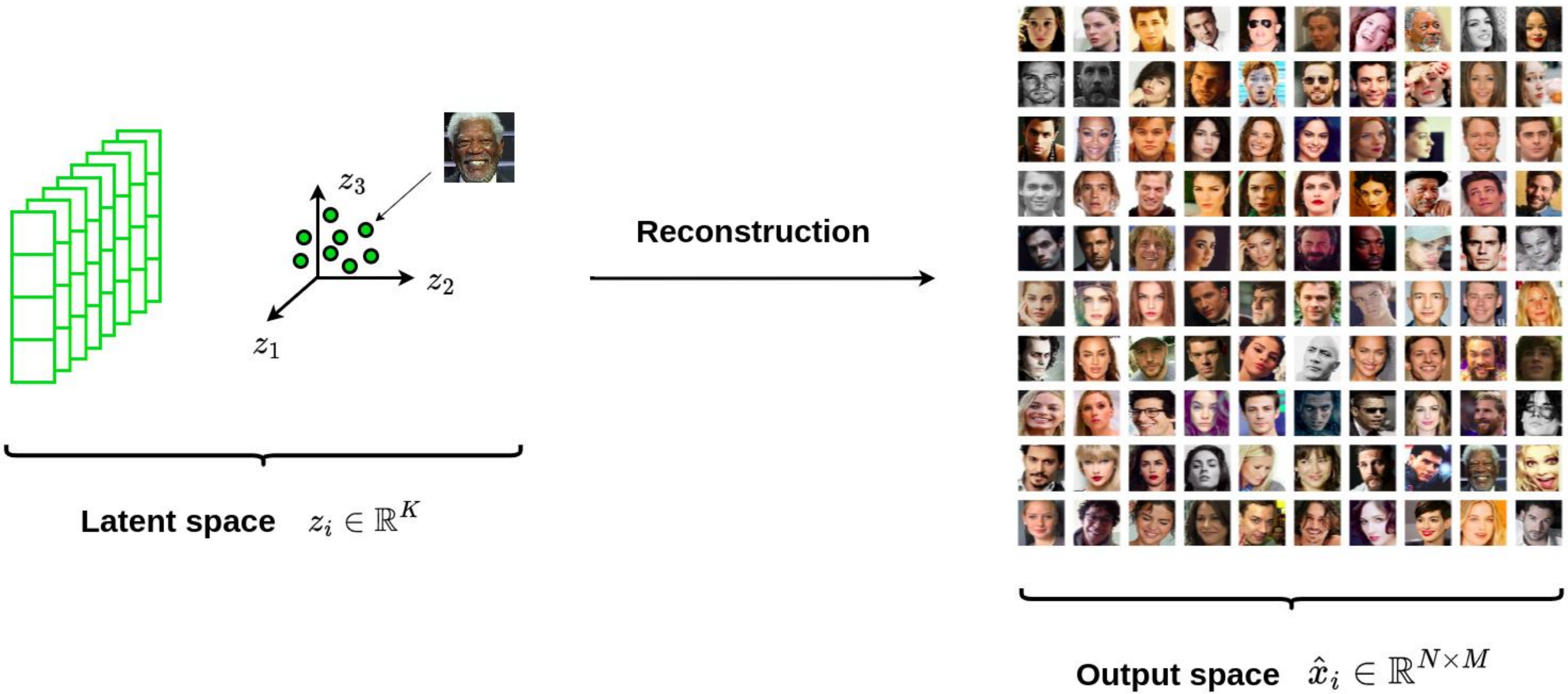
Auto-encoders

- **Projection into a simpler, lower-dimensional representation space**



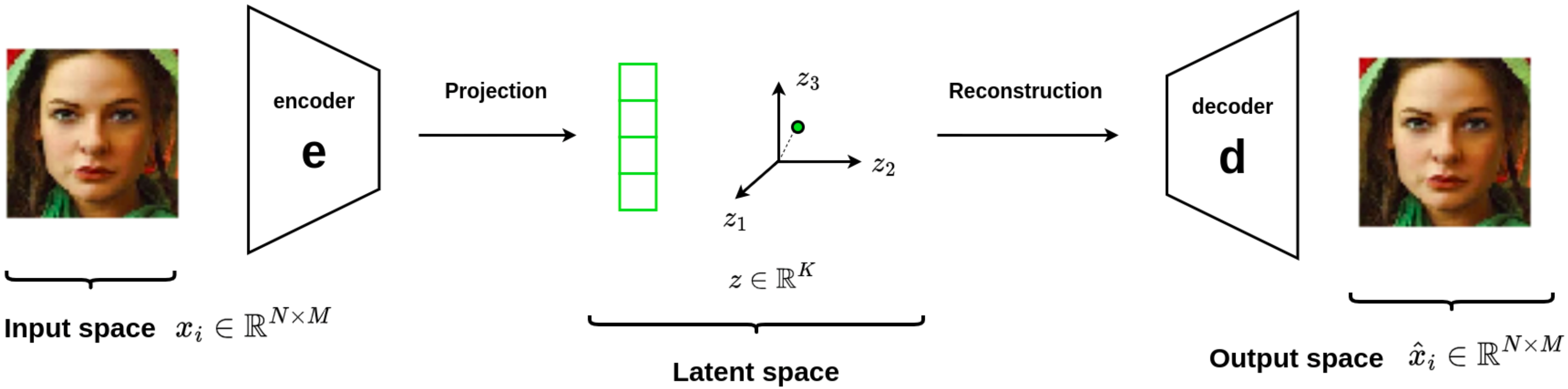
How to learn a complex distribution ?

- **How to have a relevant representation space ?**



Auto-encoder framework

▶ Standard architecture

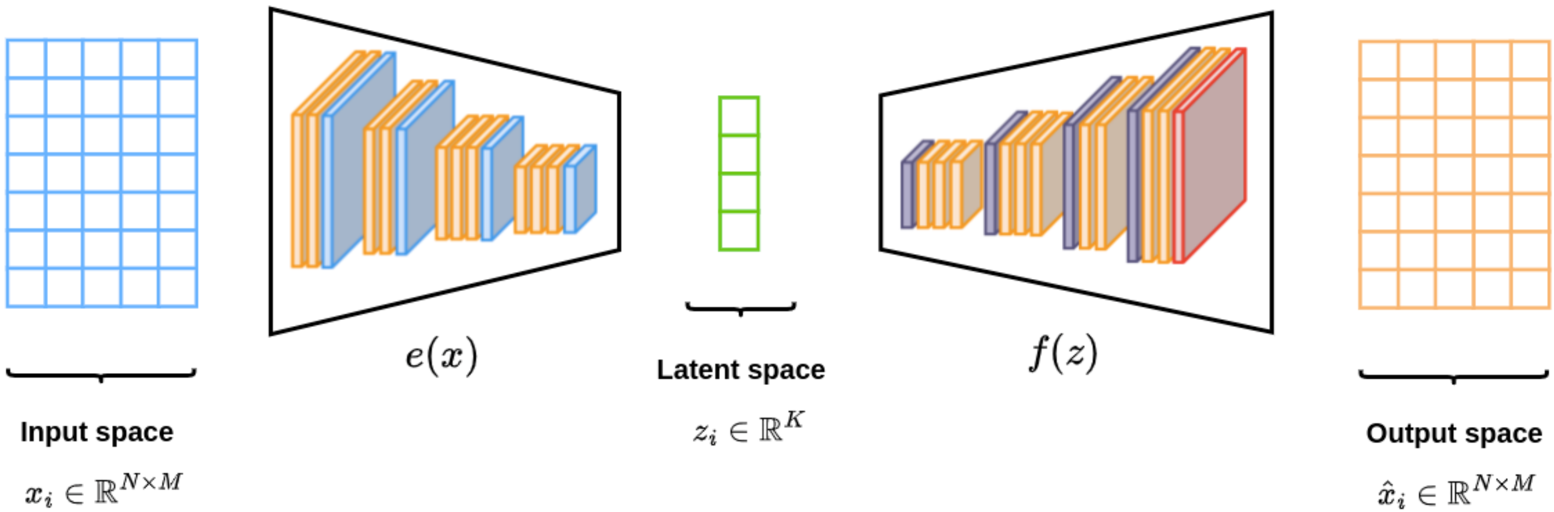


▶ Deep learning loss function

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

Deep learning implementation

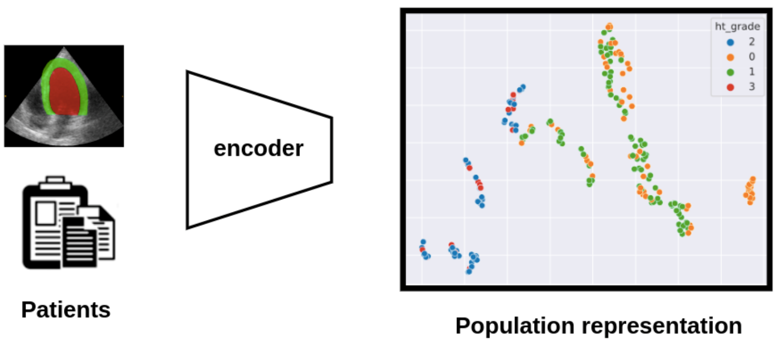
- **Encoder / Decoder modeled through (convolutional) neural networks**



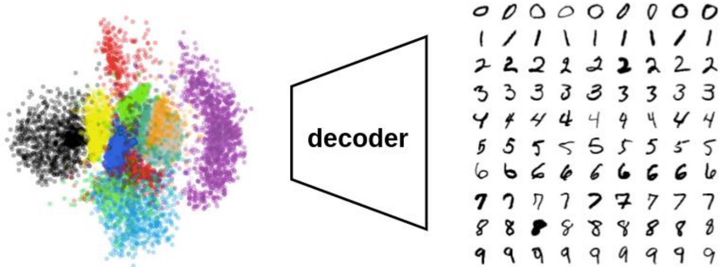
$$\text{loss} = \|x - f(e(x))\|^2$$

Auto-encoder ? For what purpose ?

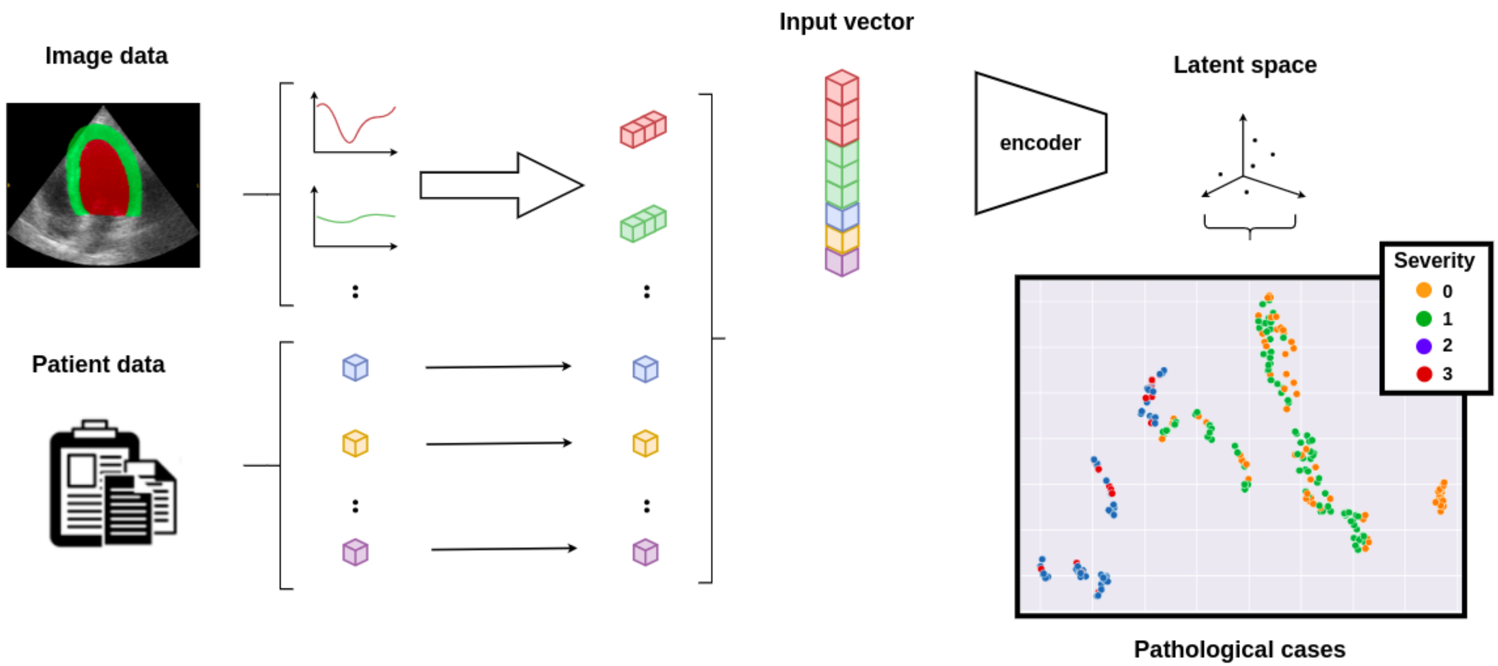
Data representation



Generative model

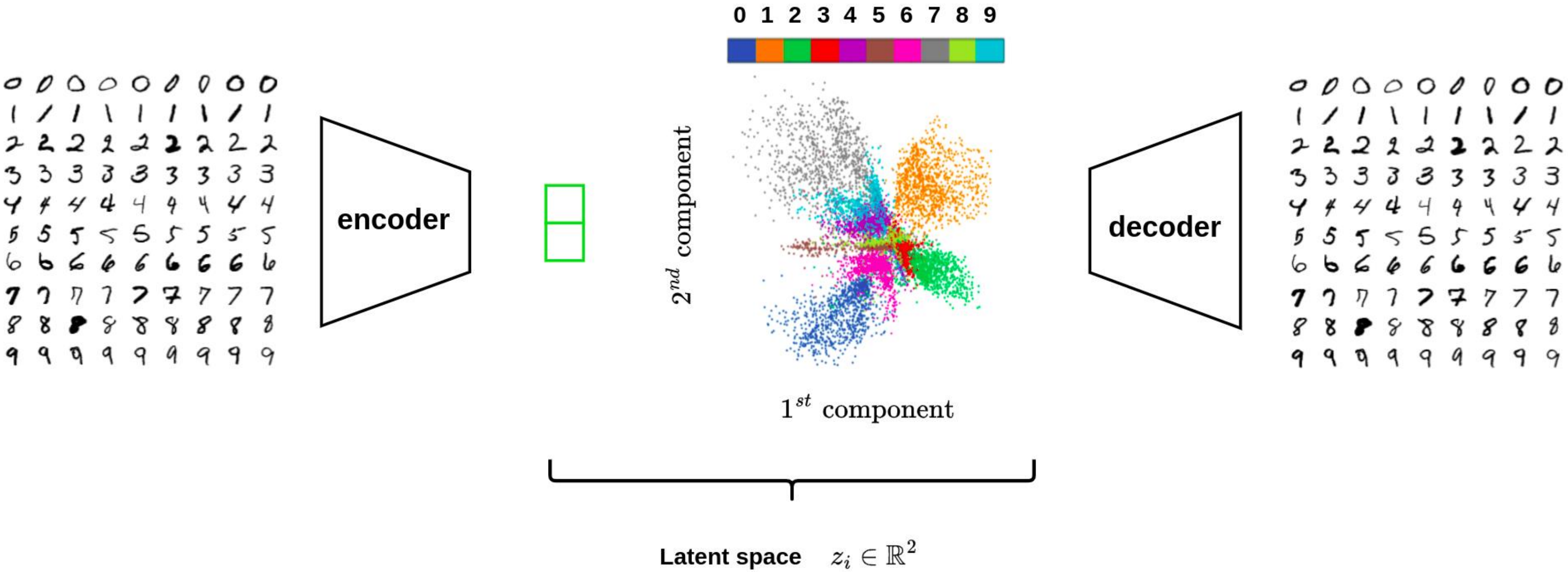


▶ Data representation

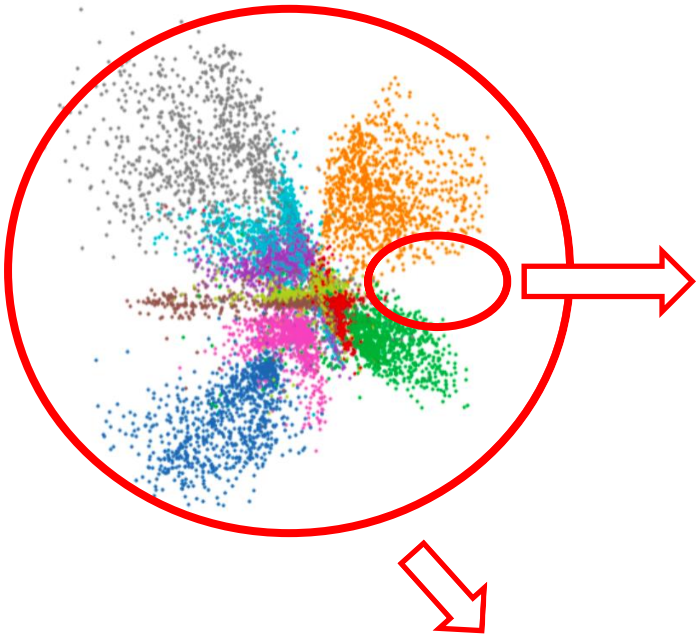


Interest of auto-encoders

- **Generative model**



- Needs to better control the structure of the latent space

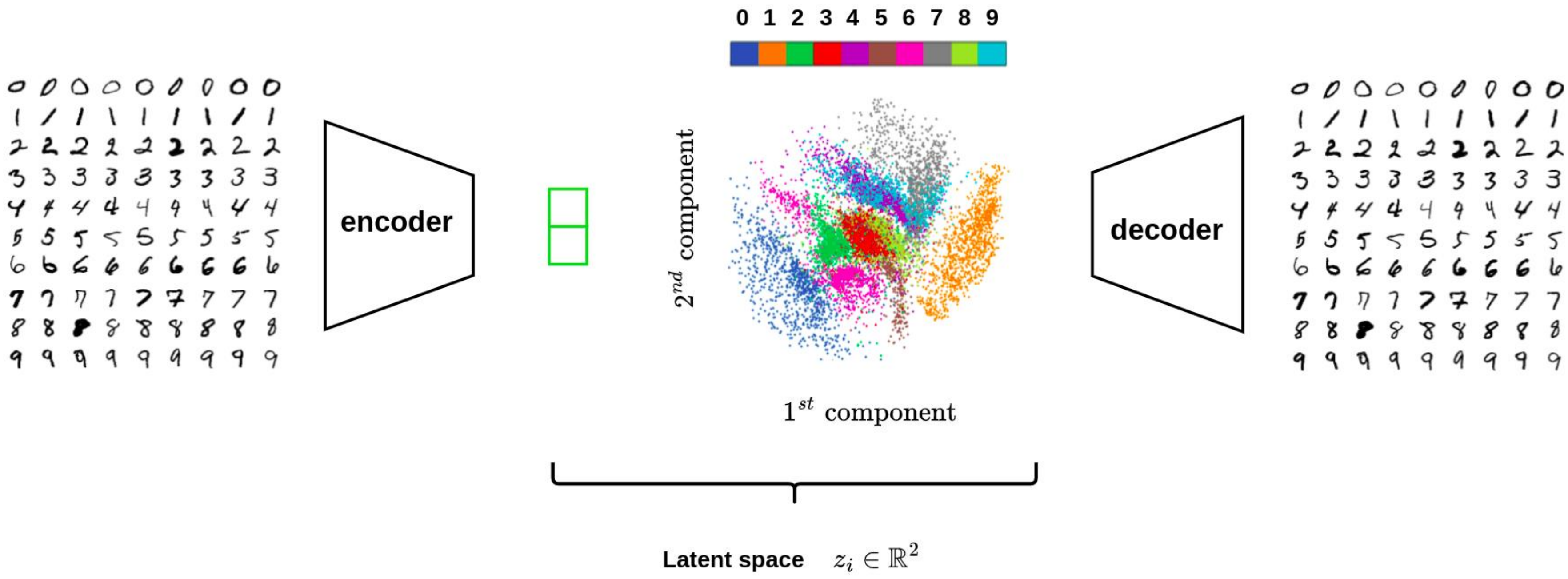


Lack of local continuity

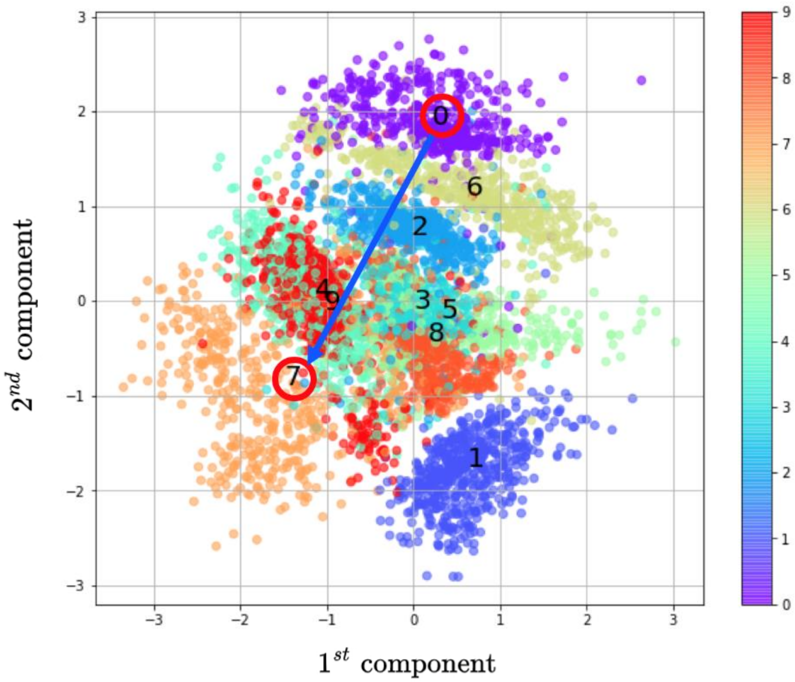
Lack of global completeness

Interest of auto-encoders

- **Generative model with better properties thanks to *variational framework***

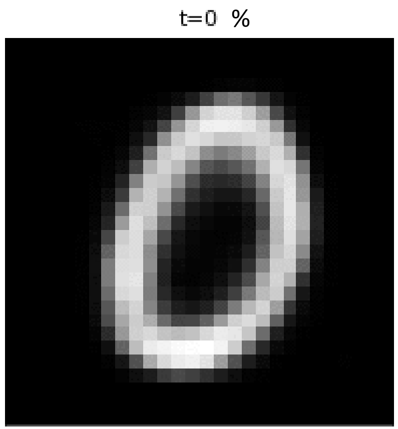


- **Generative model with variational framework**



Linear interpolation into the latent space

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



Variational autoencoders

All the mathematical details are given there !

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

Key concepts

- **Enforcing a structured latent space**

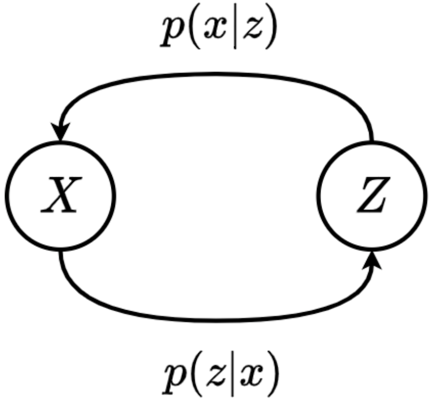
  → **Through a probabilistic framework**

  → **By imposing continuity**

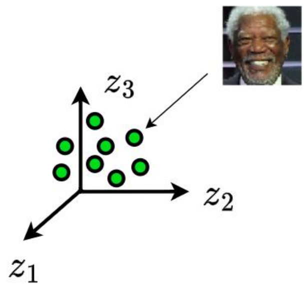
  → **By imposing completeness**

Mathematical formulation

Observation variables



Hidden variables



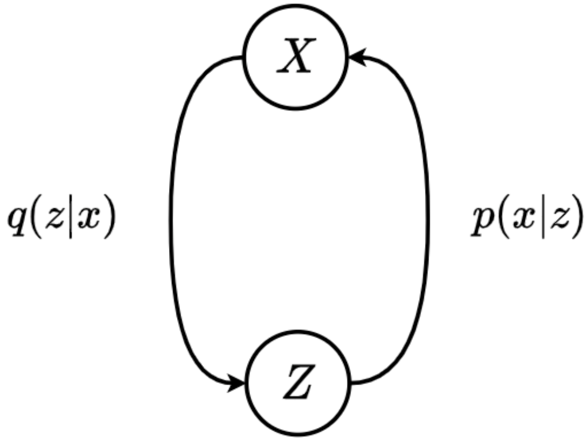
Approximation of $p(z|x)$ through a variational inference technique

Probabilistic framework

▶ Hypotheses

➔ $q(z|x)$ **is modeled by an axis-aligned Gaussian distribution**

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



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

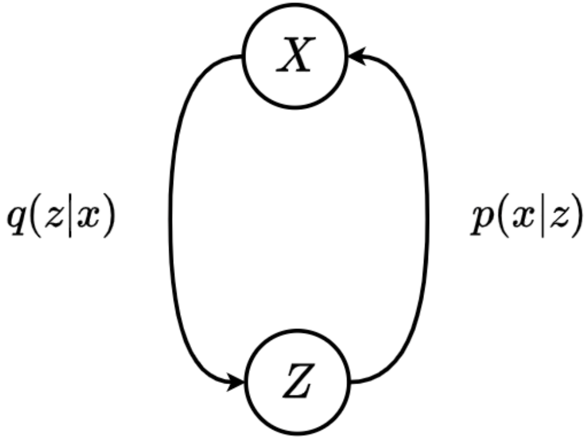
$D_{KL}(\cdot \parallel \cdot)$ *Kullback-Liebler divergence function*

Probabilistic framework

► Optimization process

→ Maximization of the Evidence Lower Bound (ELBO)

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



→ By exploiting gaussian assumption

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

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

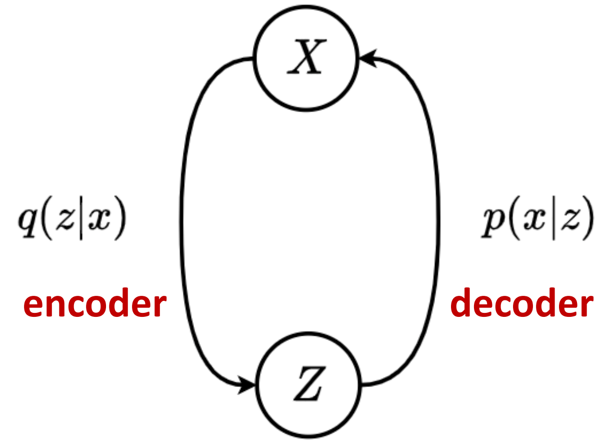
Probabilistic framework

Optimization process

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

Deep learning loss function

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

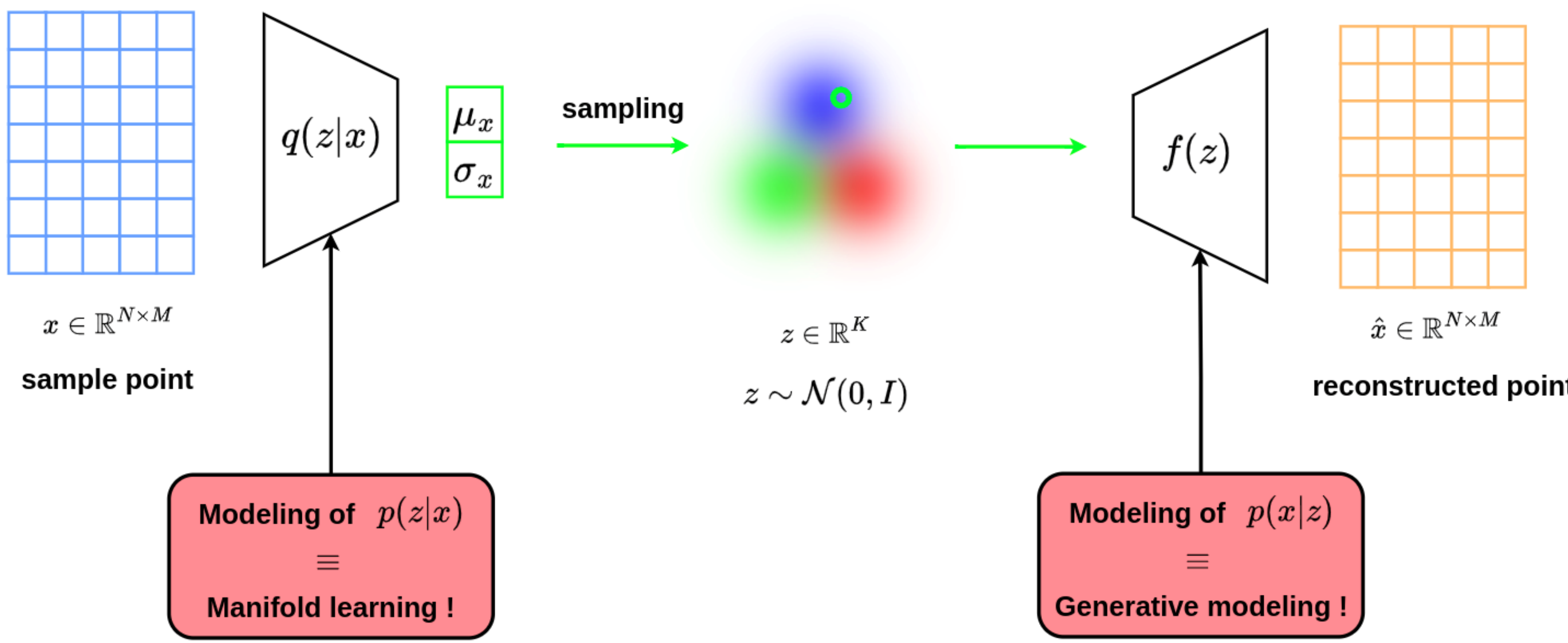


- ➔ $g(\cdot)$ **and** $h(\cdot)$ **are modeled through an encoder**
- ➔ $f(\cdot)$ **is modeled through a decoder**

Probabilistic framework

- **Loss interpretation**

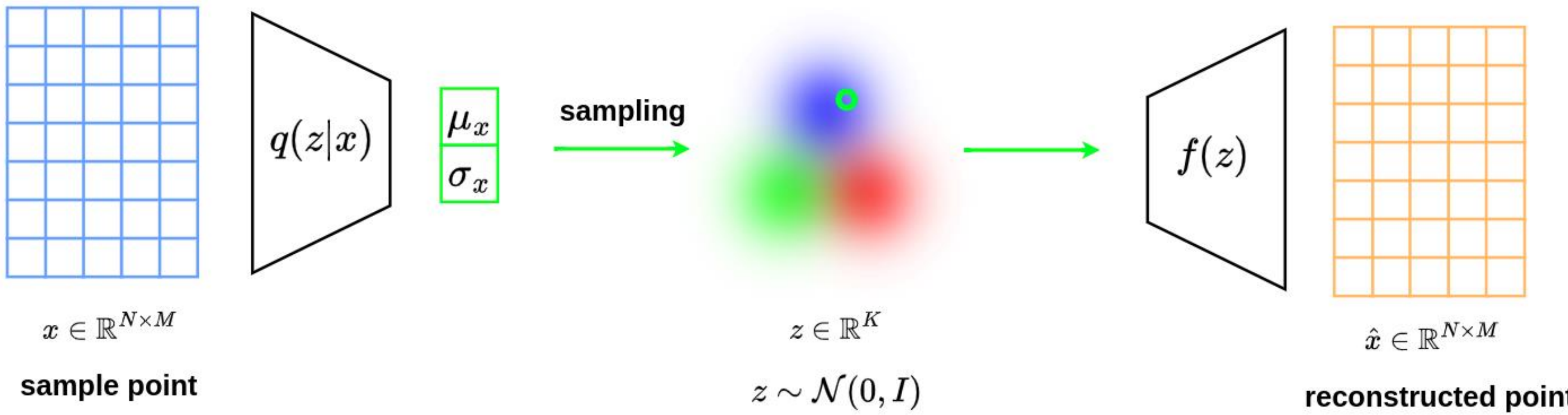
$$\text{loss} = D_{KL}\left(\mathcal{N}\left(g(x), diag\left(h(x)\right)\right), \mathcal{N}\left(0, I\right)\right) + \alpha\|x - f(z)\|^2$$



Probabilistic framework

- **Loss interpretation**

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

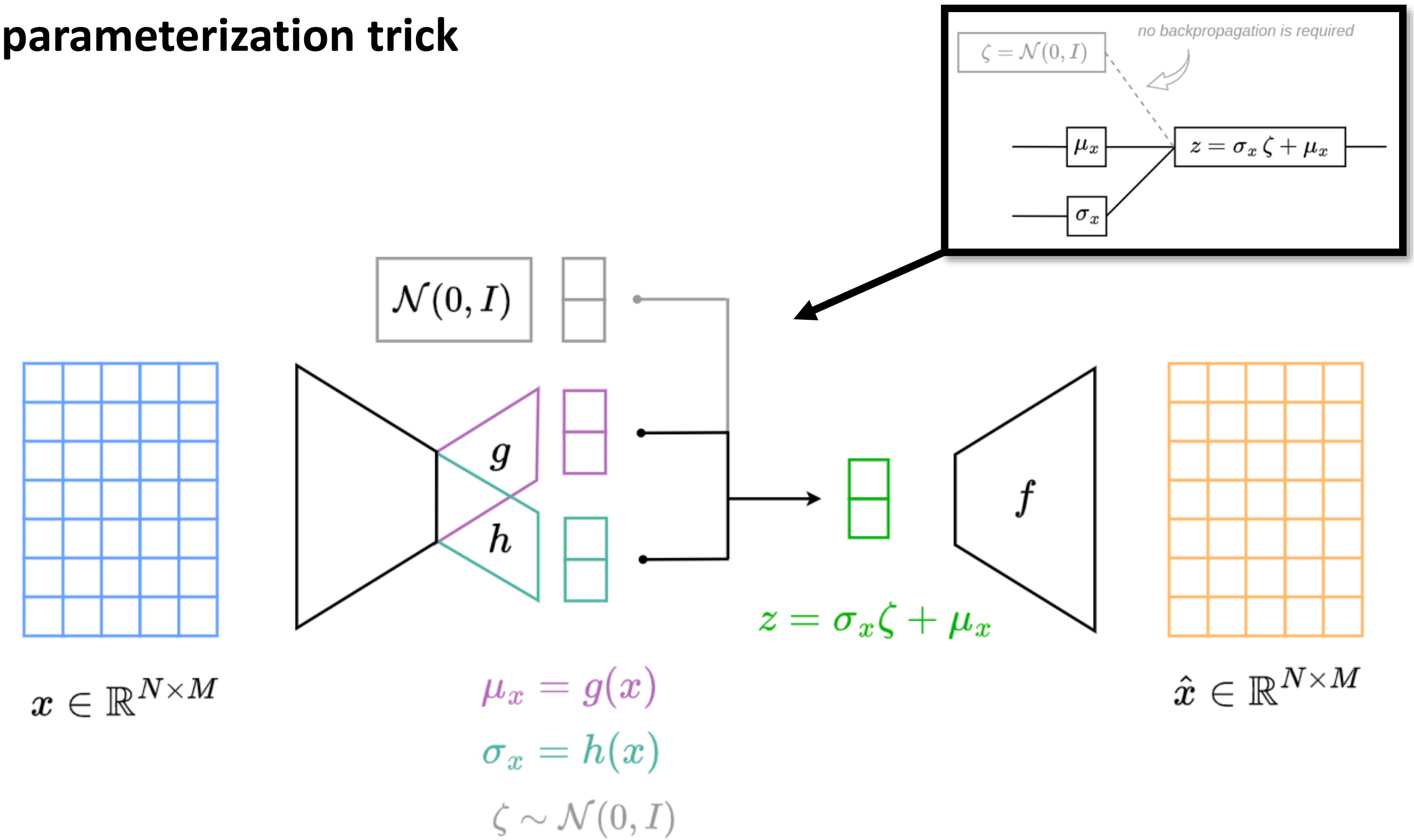


→ $\mathcal{N}\left(g(x), h(x)\right)$ **imposes local *continuity***

→ $\mathcal{N}\left(\cdot, \mathcal{N}(0, I)\right)$ **imposes global *completeness***

Deep learning implementation

- **Reparameterization trick**

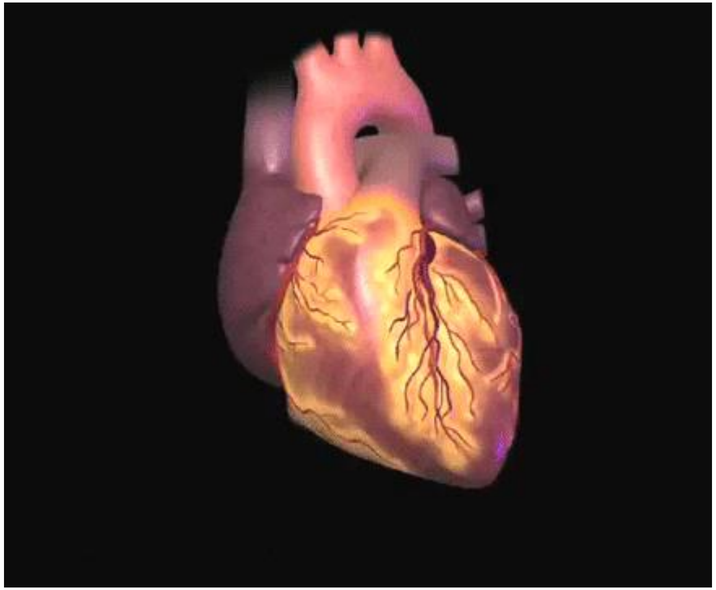


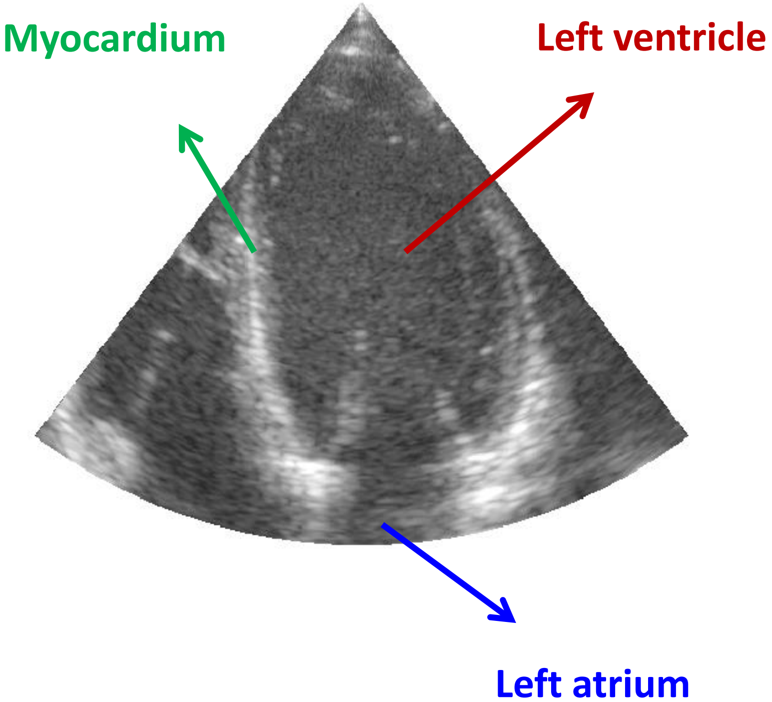
Practical applications

The obsession is to master the latent space !!!

Needs for accurate and robust segmentation of cardiac structures

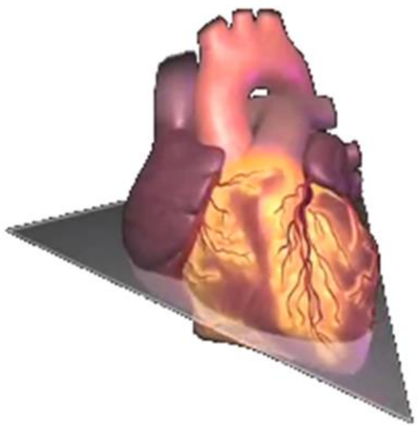
- **Quantification of clinical indices from echocardiographic images**



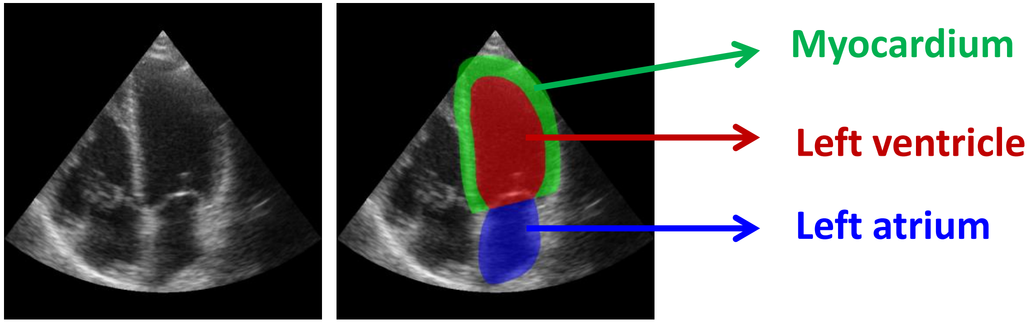


Needs for accurate and robust segmentation of cardiac structures

- **Anatomical clinical indices**



Apical views



- **Left ventricle volumes**
- **Left ventricle ejection fraction**

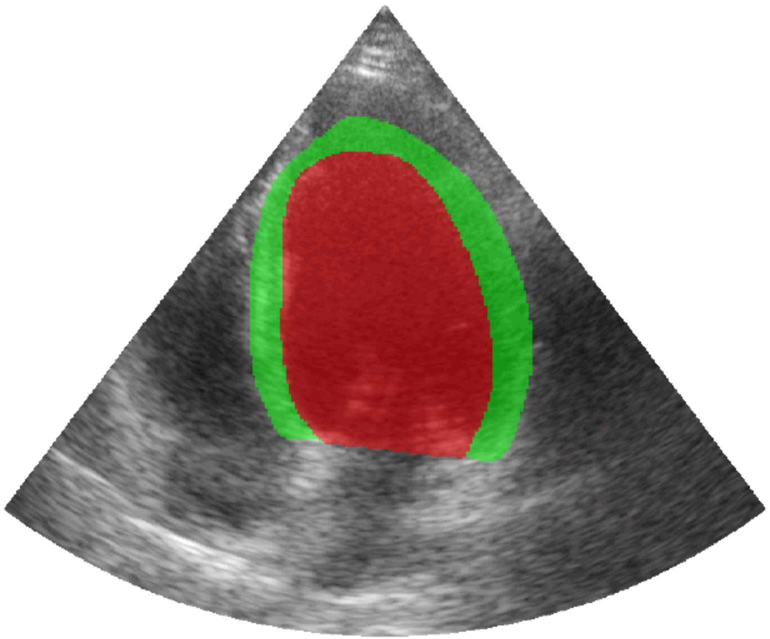
How to guarantee temporal consistency ?

Cardiac segmentation with temporal consistency

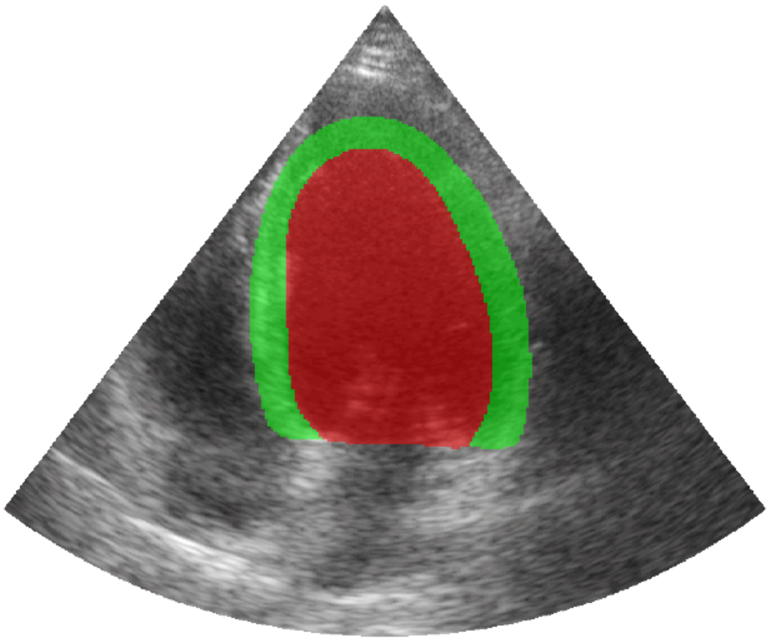
[Painchaud, IEEE TMI, 2022]

- **Quantification of clinical indices from echocardiographic images**

What we have with a 2D U-Net

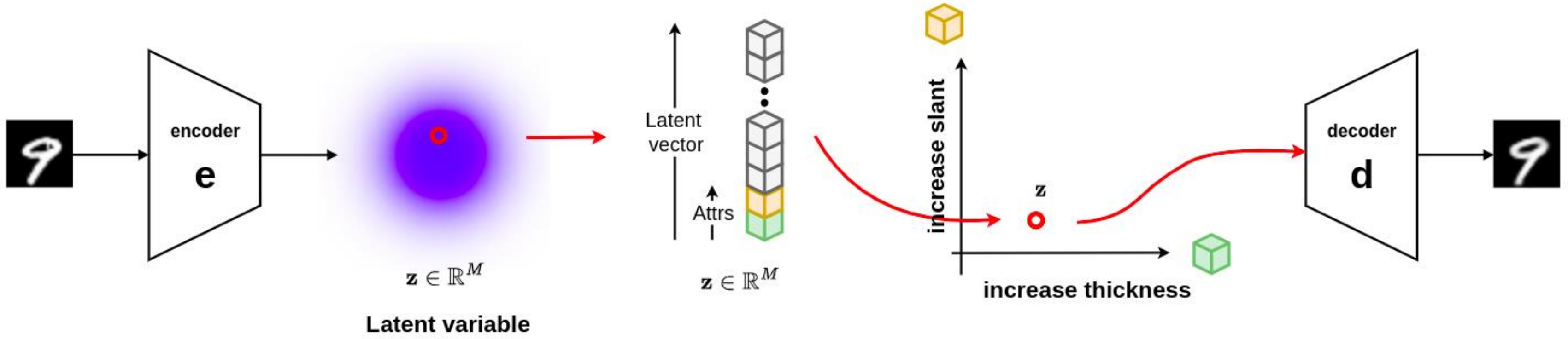


What we want



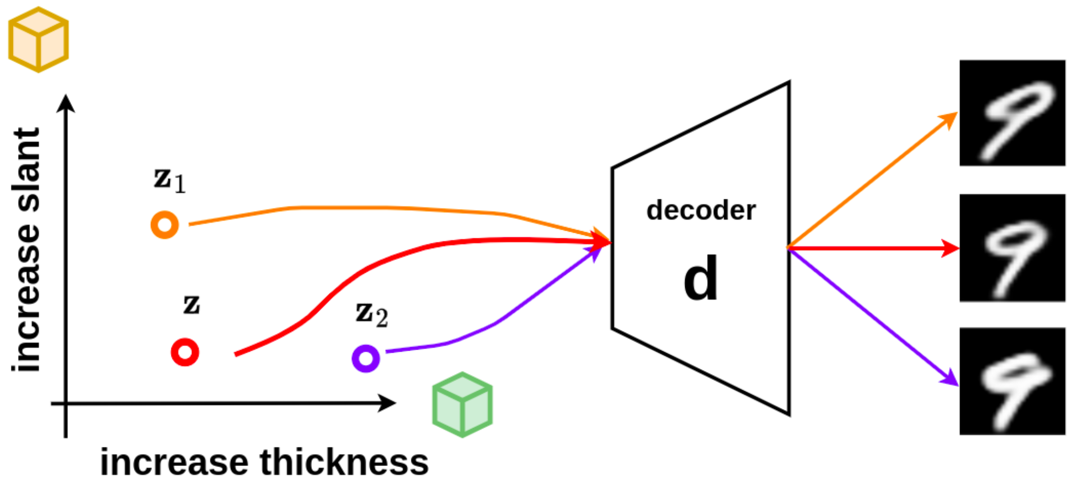
- **AR-VAE: attribute-based regularization of VAE latent space**
  [Pati, Neural Comp. Appli., 2021]

  - **Generation of structured latent space**
    - ➔ **Specific continuous-valued attributes forced to be encoded along specific dimensions**
    - ➔ $\boldsymbol{Loss = VAE\ loss + Attribute\ Regularisation\ Loss}$



- **AR-VAE: attribute-based regularization of VAE latent space**
  [Pati, Neural Comp. Appli., 2021]

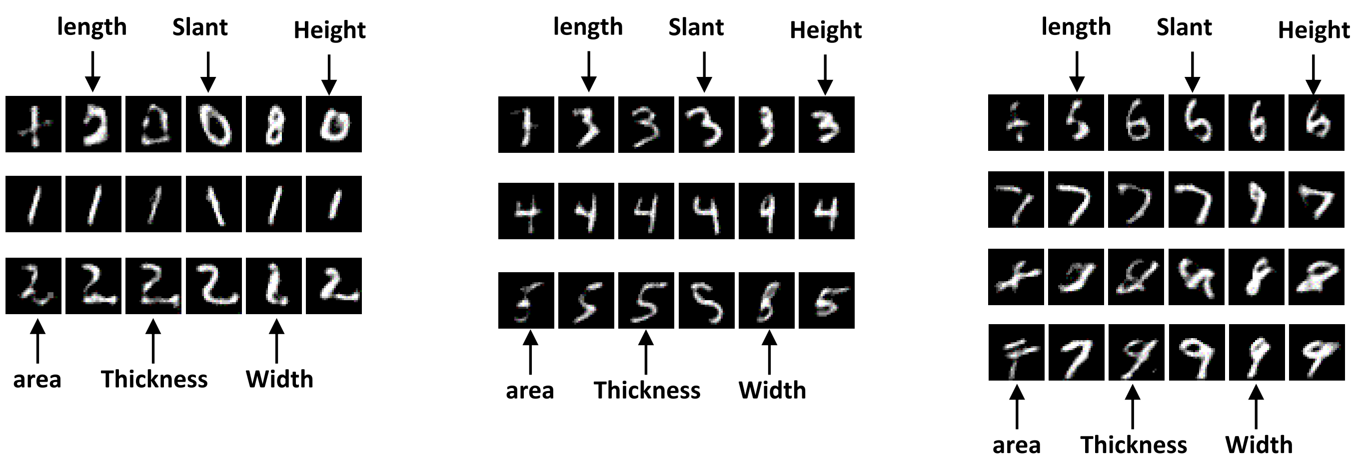
  - **Sampling of the structured latent space**



AR-VAE: attribute-based regularization of VAE latent space

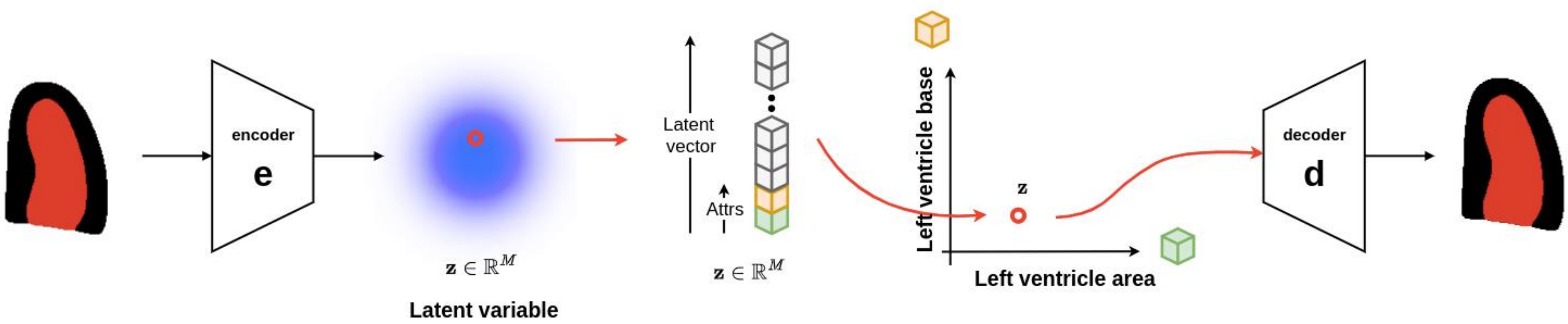
[Pati, Neural Comp. Appli., 2021]

- **Sampling of the structured latent space**
  - → Specific attribute (from left to right): area, length, thickness, slant, width, height
  - → Each column corresponds to traversal along a regularized dimension



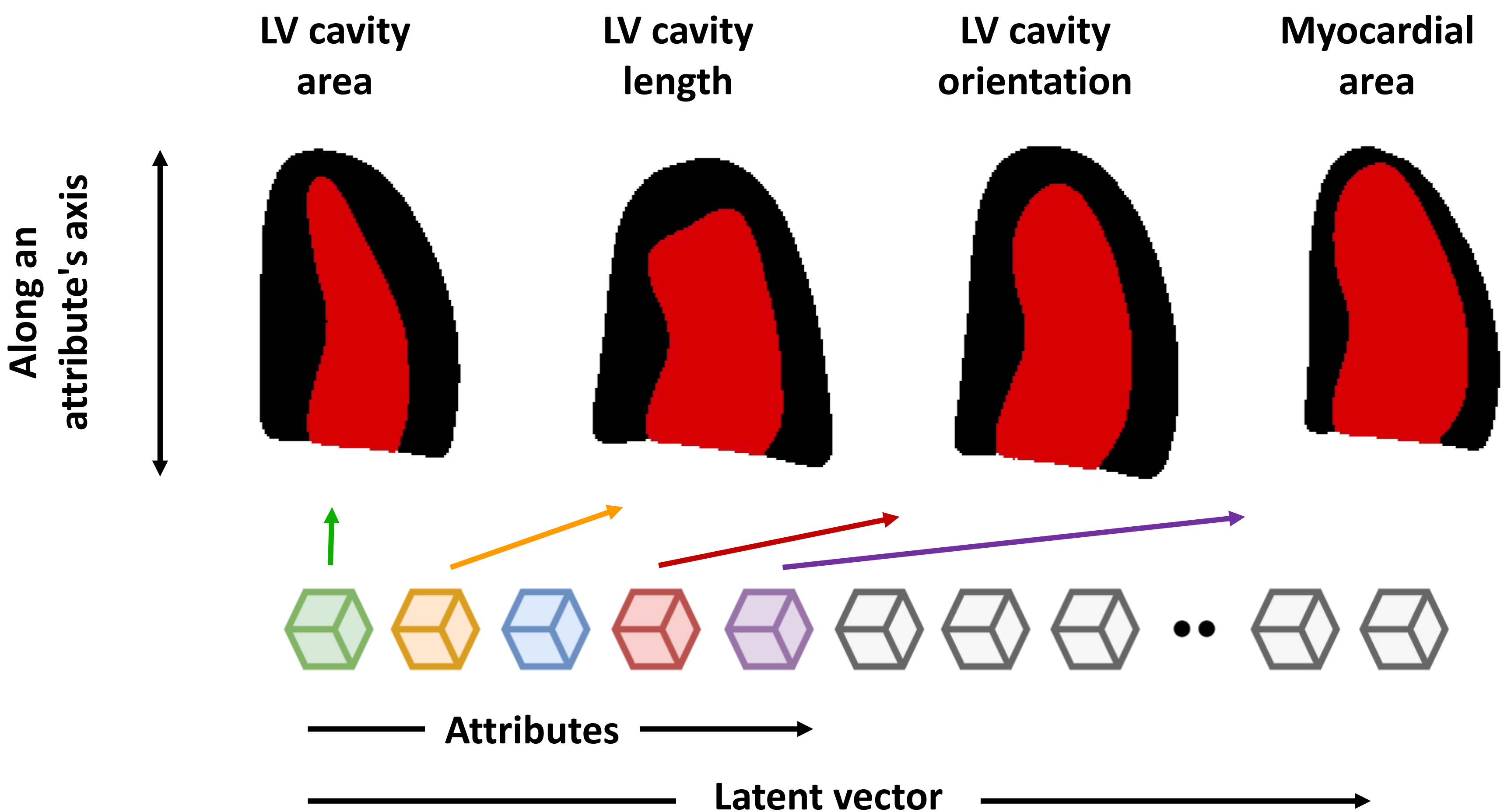
▶ Application to the description of the cardiac shapes

- **Generation of structured latent space according to the following attributes**
  - → **Left ventricle (LV) cavity: area, length, basal width, orientation**
  - → **Myocardial area**
  - → **Epicardial center**



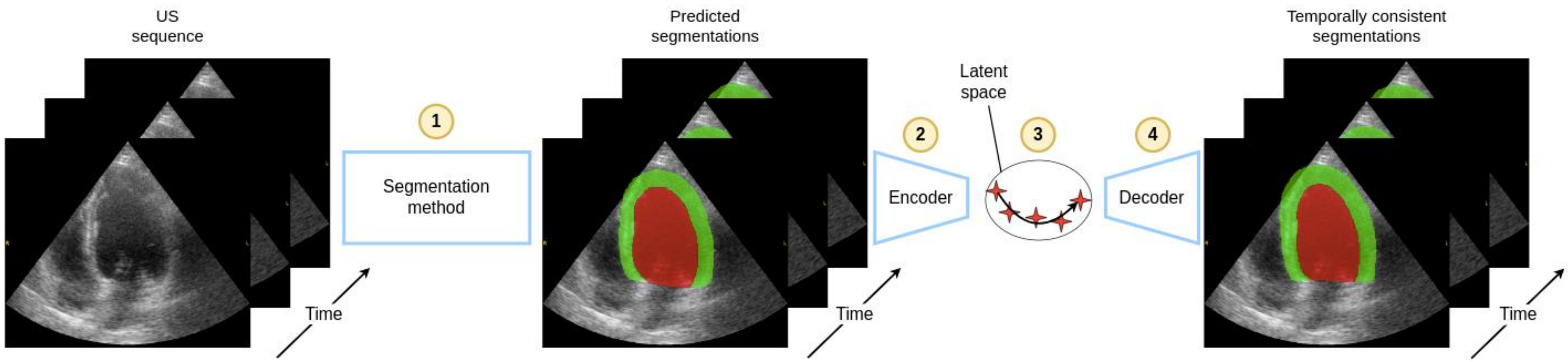
Cardiac segmentation with temporal consistency

[Painchaud, IEEE TMI, 2022]



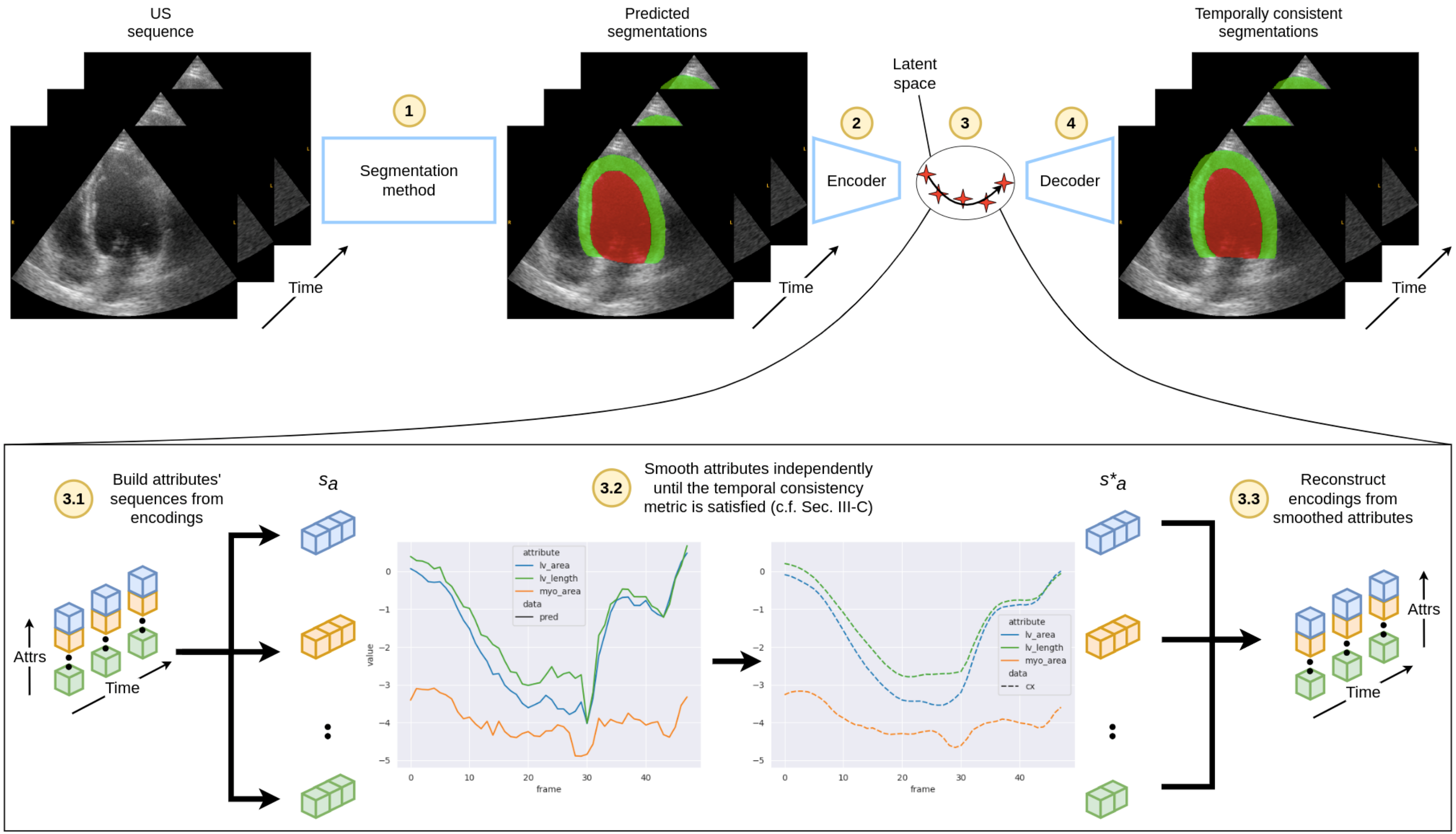
 [Painchaud, IEEE TMI, 2022]

▶ Proposed temporal pipeline



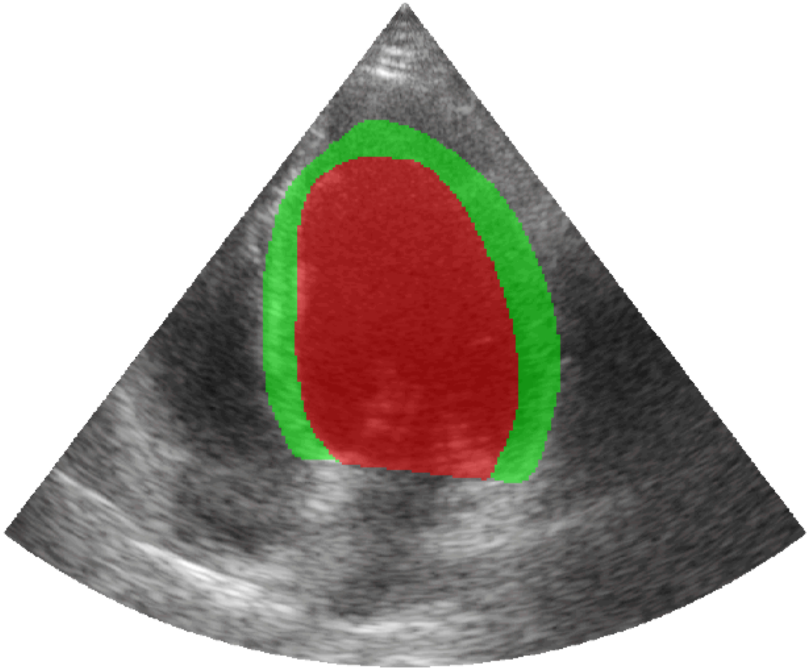
Cardiac segmentation with temporal consistency

[Painchaud, IEEE TMI, 2022]

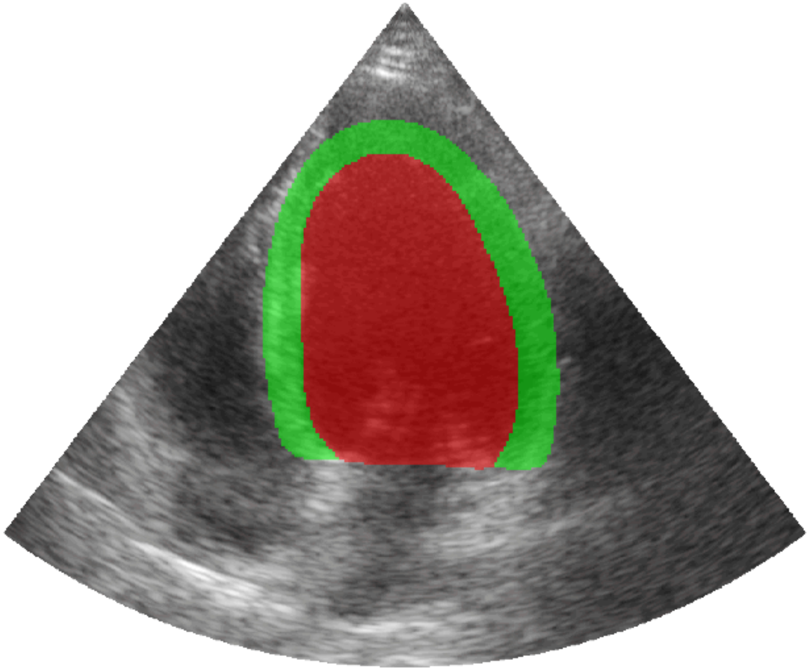


- Some post-processing examples

Original U-Net



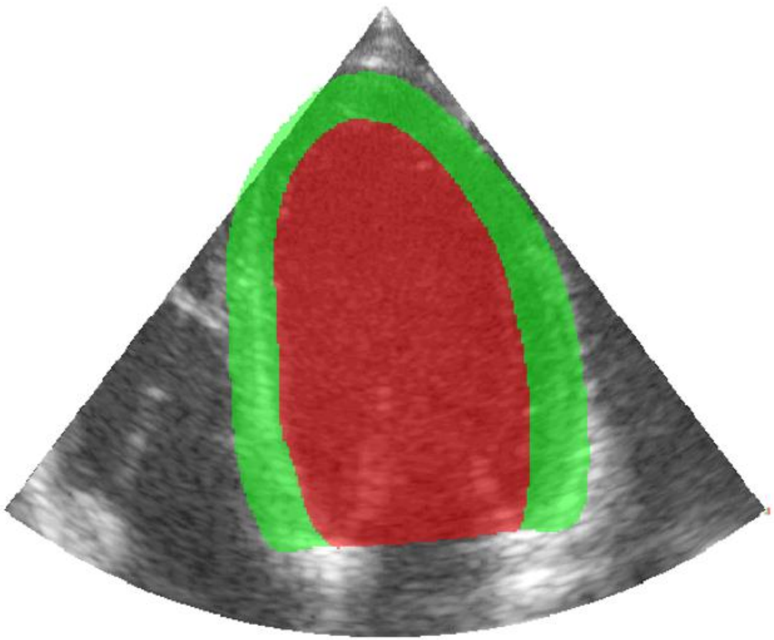
Post-processed U-Net



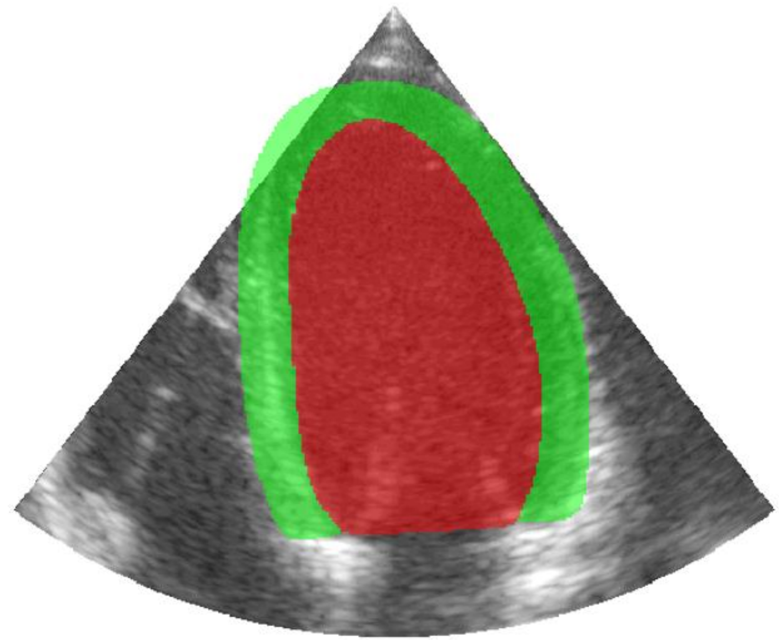
 [Painchaud, IEEE TMI, 2022]

- **Some post-processing examples**

Original U-Net



Post-processed U-Net



To conclude

To conclude

- **VAEs can be used effectively in medical imaging**
  - **Guarantee anatomical coherence** ✓
  - **Guarantee temporal consistency** ✓
  - **Estimation uncertainty for image segmentation** ✓
  - **Generative interest limited to simple distribution**

- **Useful tool for characterizing populations**
  - **Need to properly structure the learned latent space**
  - **Need to work on relatively large cohorts**

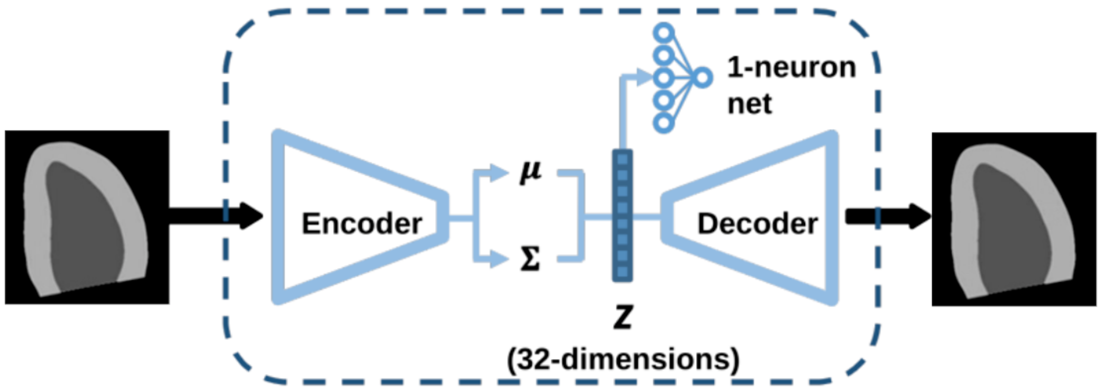
Appendix

How to guarantee the anatomical coherence ?

▶ Constrained Variational Auto Encoder

- **Approximation of a latent space with local linear properties**

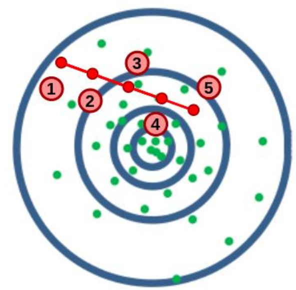
Use of a 1-neuron net to reinforce the linearity of the latent space

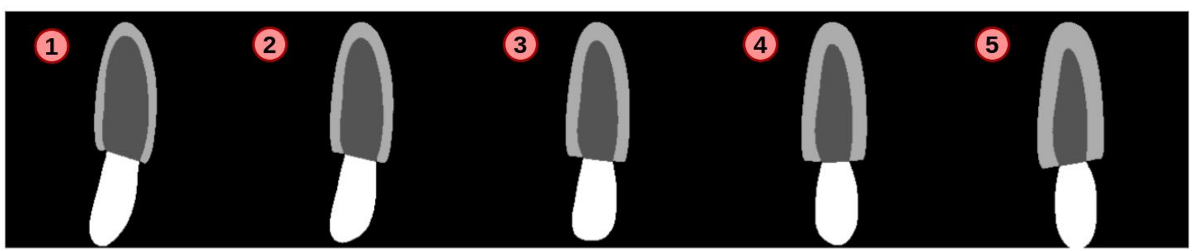


Constrained Variational Auto Encoder

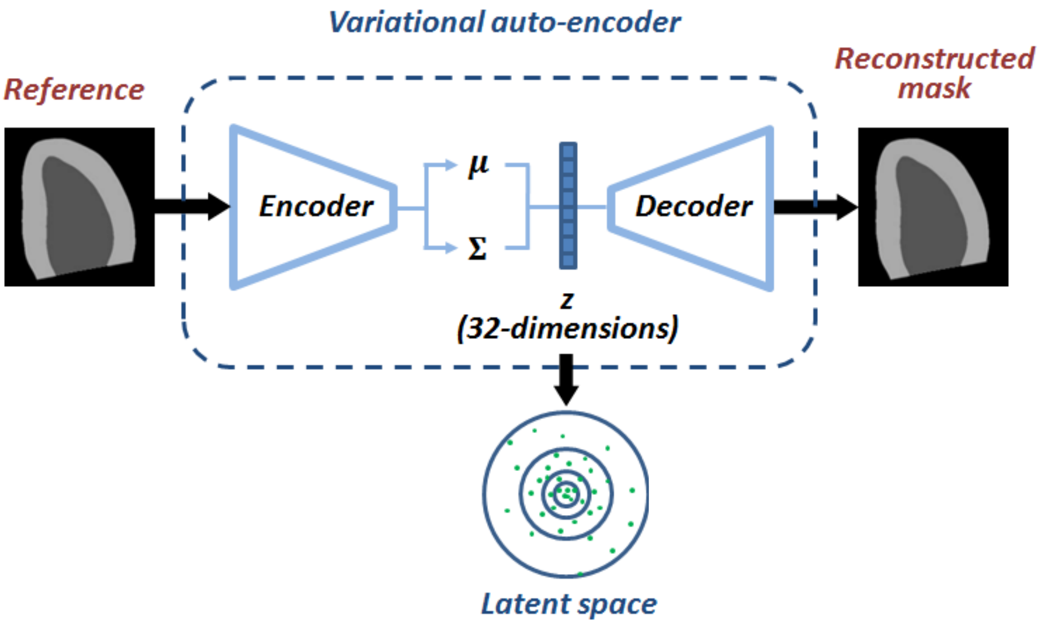
- **Approximation of a latent space with local linear properties**
  - → **Linear interpolation in the latent space makes sense**

Constrained latent space



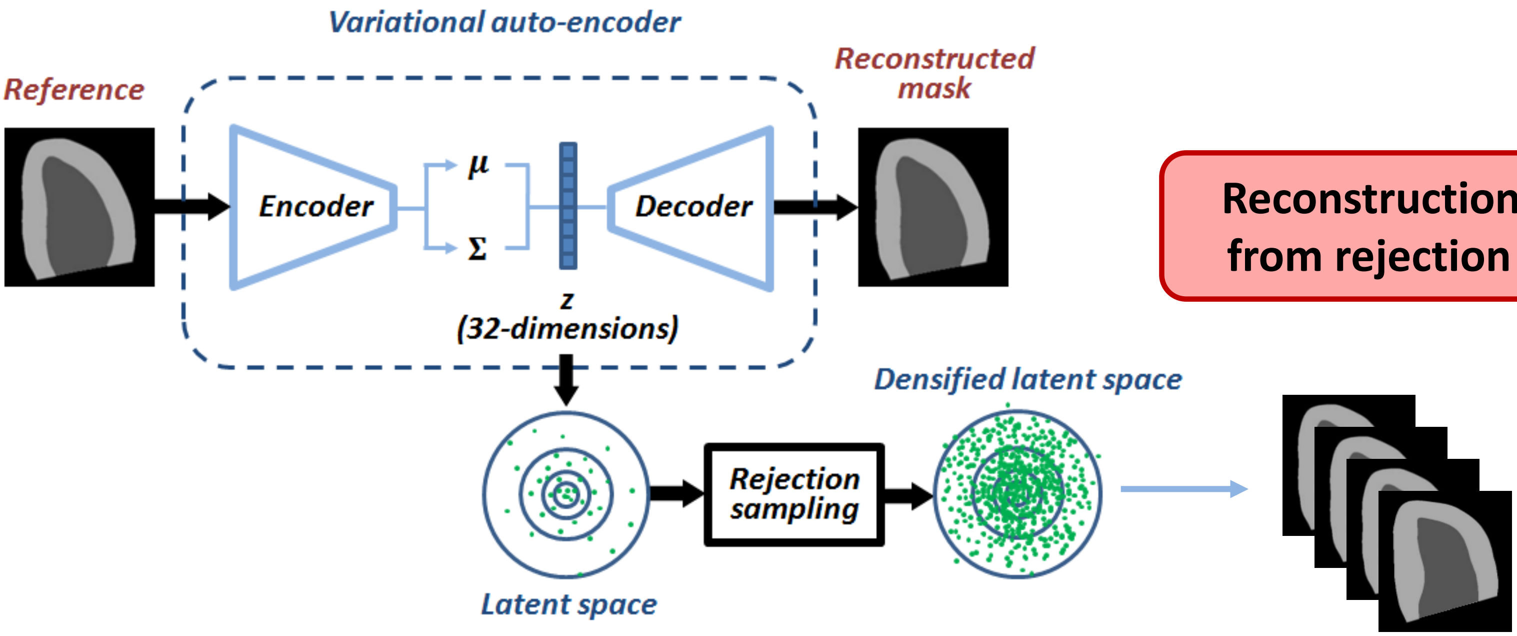


Cardiac segmentation with strong anatomical guarantees *[Painchaud, IEEE TMI, 2020]*



Efficient encoding of anatomical shapes in a latent space

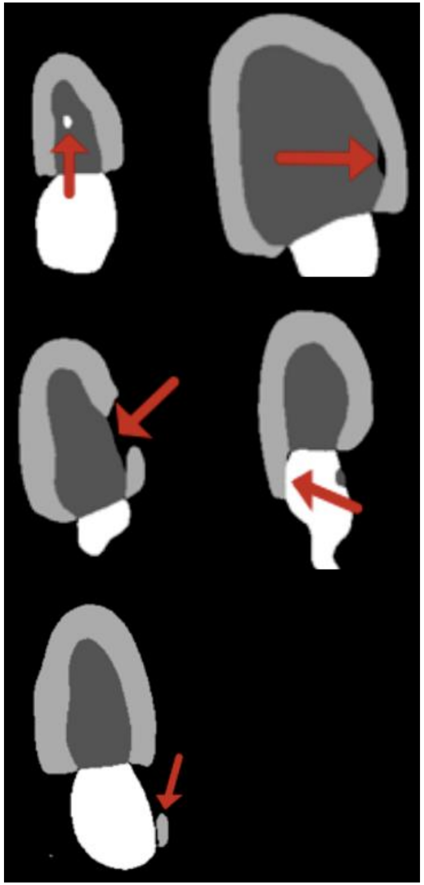
Cardiac segmentation with strong anatomical guarantees *[Painchaud, IEEE TMI, 2020]*



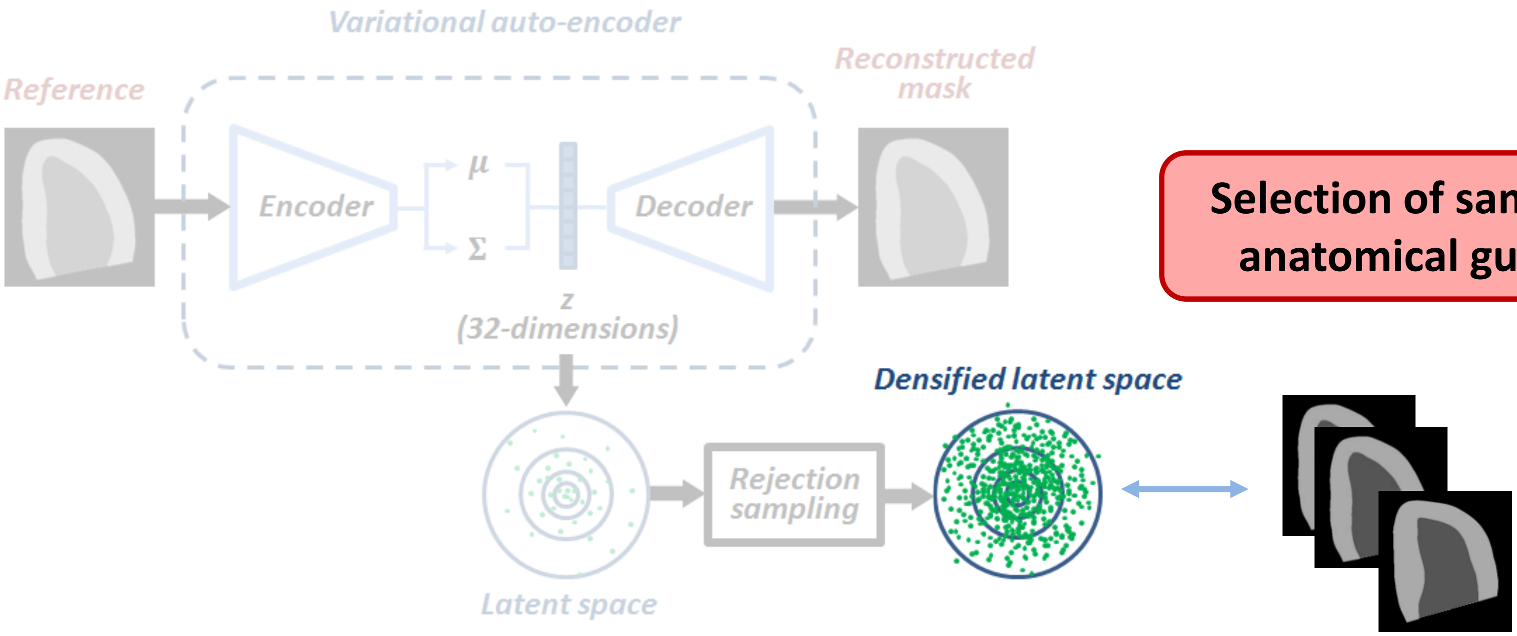
Reconstruction of masks from rejection sampling

Densified latent space with 5 million points

- **Definition of 12 anatomical metrics**
  - **(3 criteria) hole(s) in the LV, RV or LA**
  - **(2 criteria) hole(s) between LV and MYO or between LV and LA**
  - **(3 criteria) presence of more than one LV, MYO or LA**
  - **(2 criteria) size of the area by which the LV touches the background or the MYO touches the LA**
  - **(1 criterion) ratio between the minimal and maximal thickness of the MYO**
  - **(1 criterion) ratio between the width of the LV and the average thickness of the MYO**



Cardiac segmentation with strong anatomical guarantees *[Painchaud, IEEE TMI, 2020]*

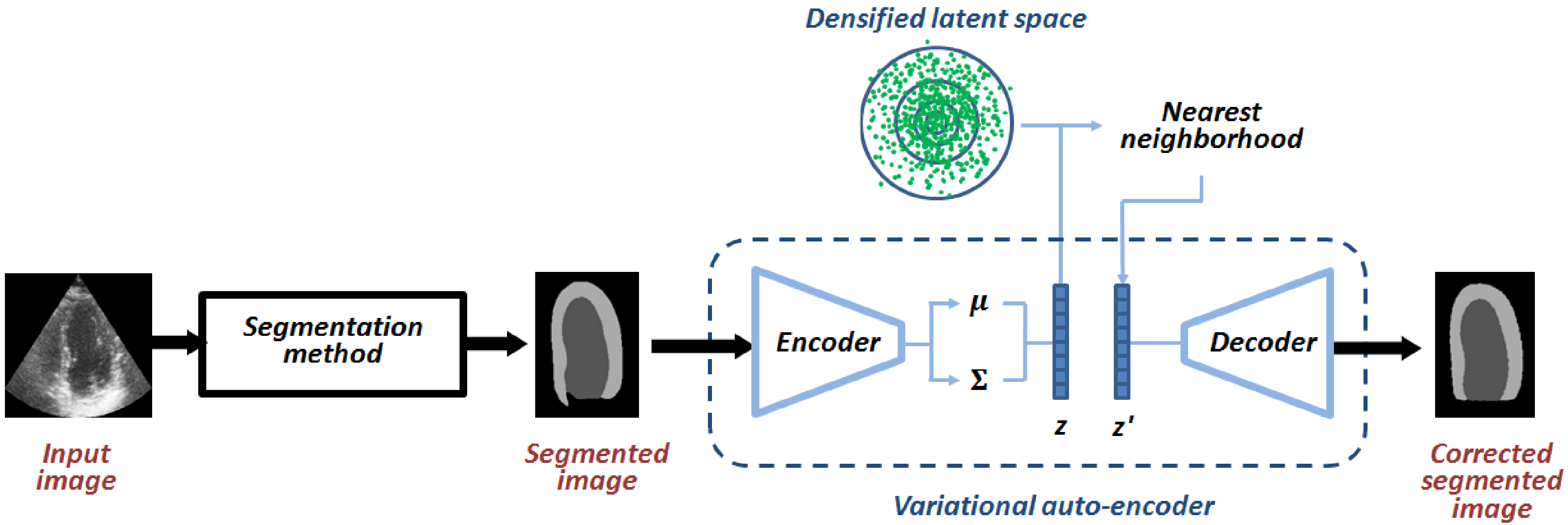


Selection of samples with anatomical guarantees

Densified latent space with 4 million points

Cardiac segmentation with strong anatomical guarantees [Painchaud, IEEE TMI, 2020]

Correction of segmentation to guarantee the plausibility of anatomical shapes → **Almost same accuracy as the original methods but with correct anatomical shapes**

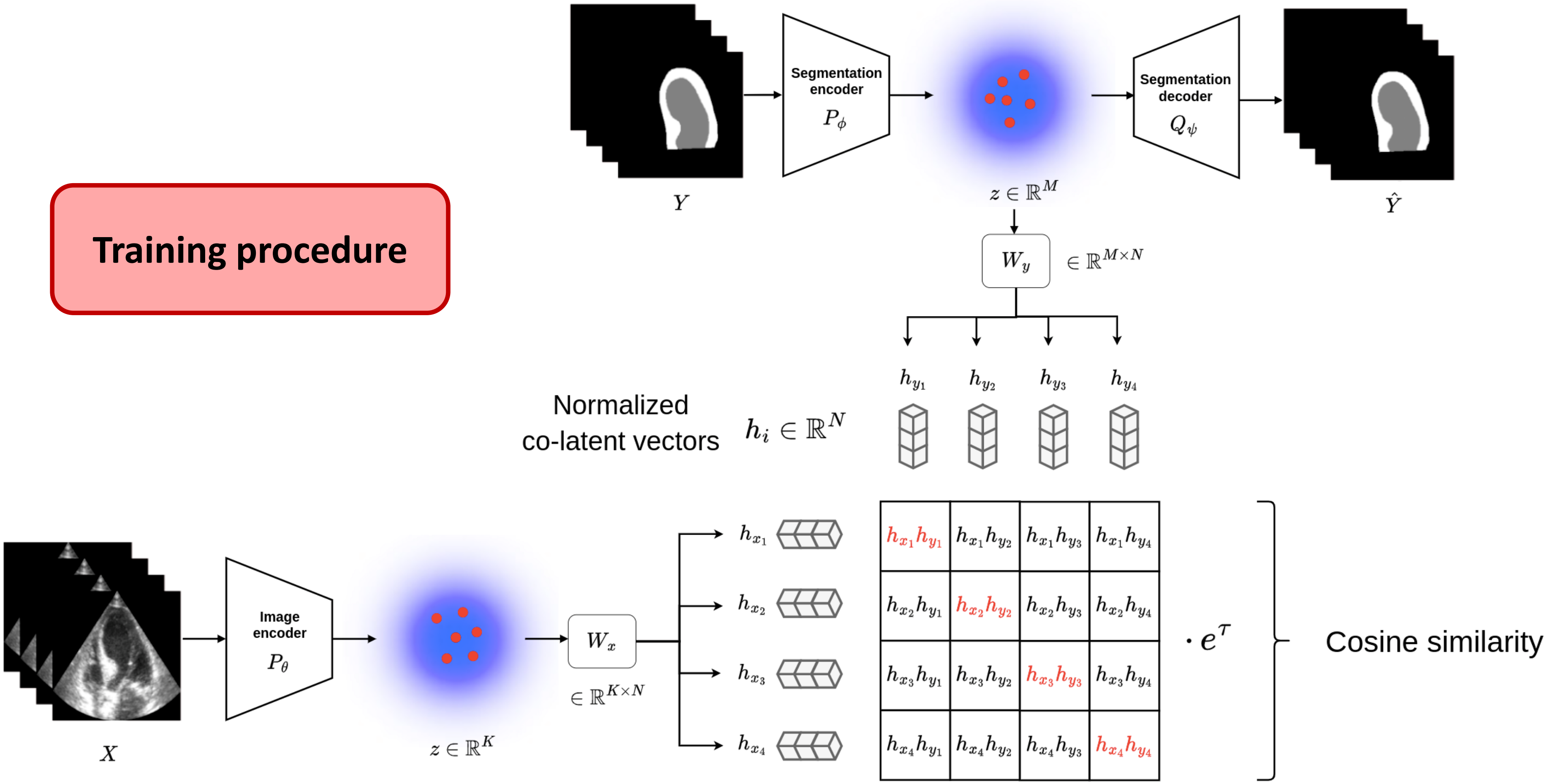


Uncertainty estimation for cardiac image segmentation

Uncertainty estimation for image segmentation

[MICCAI, 2022]

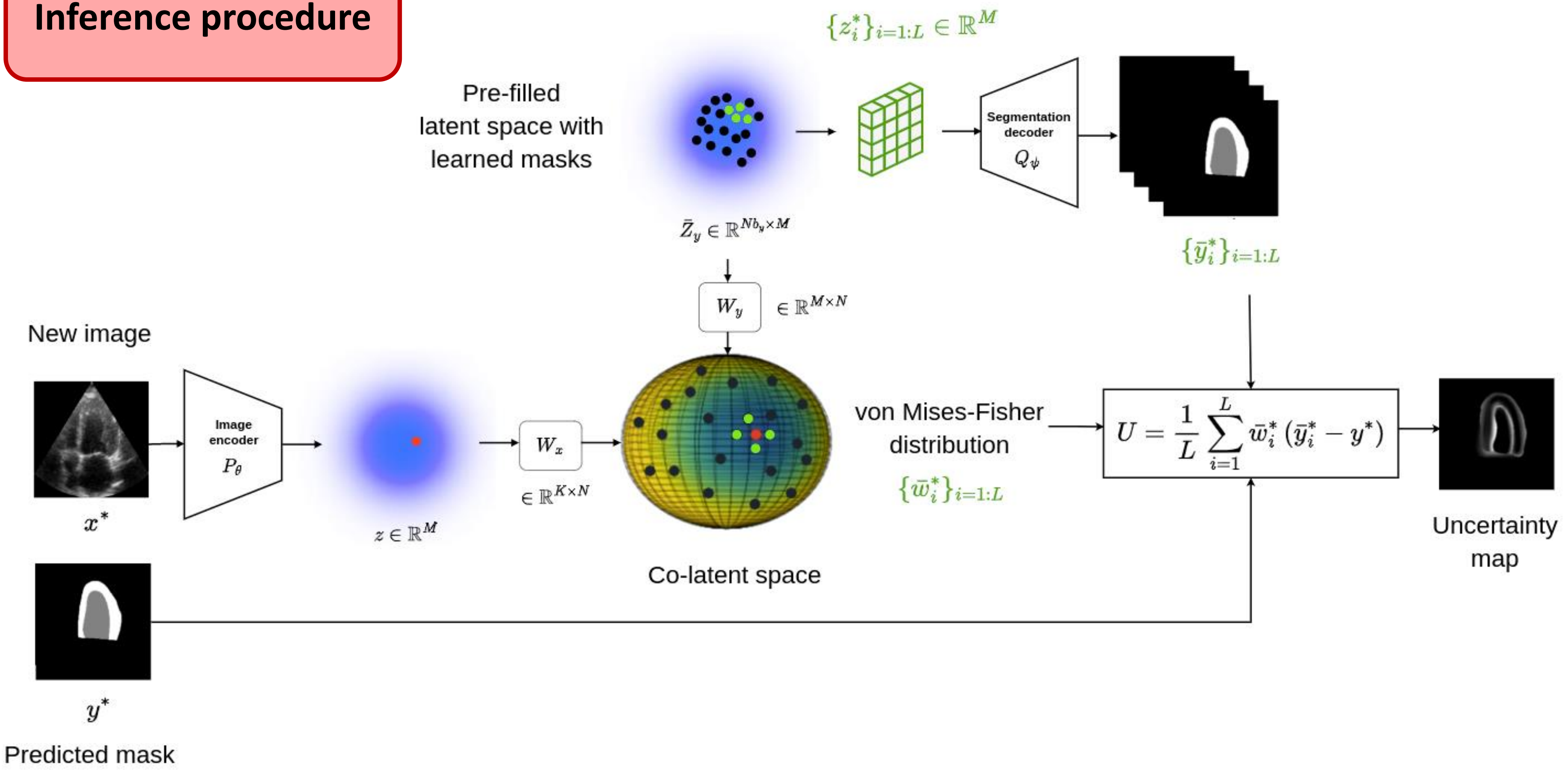
Training procedure



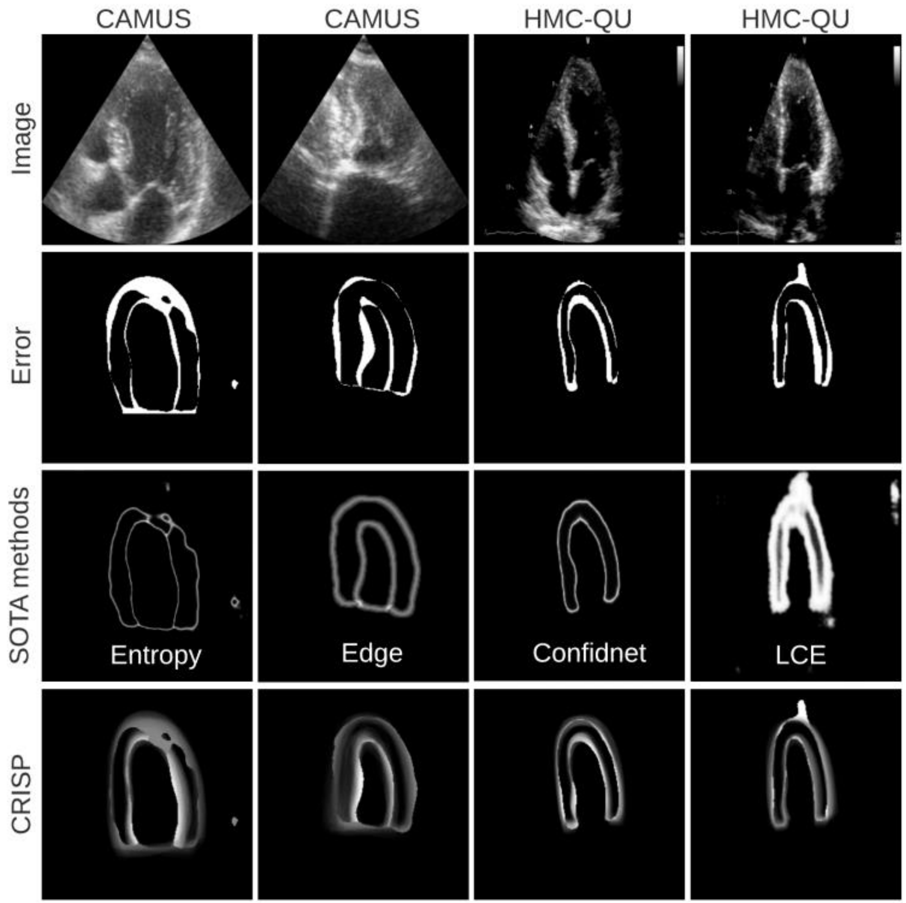
Uncertainty estimation for image segmentation

[MICCAI, 2022]

Inference procedure



- **Uncertainty results**

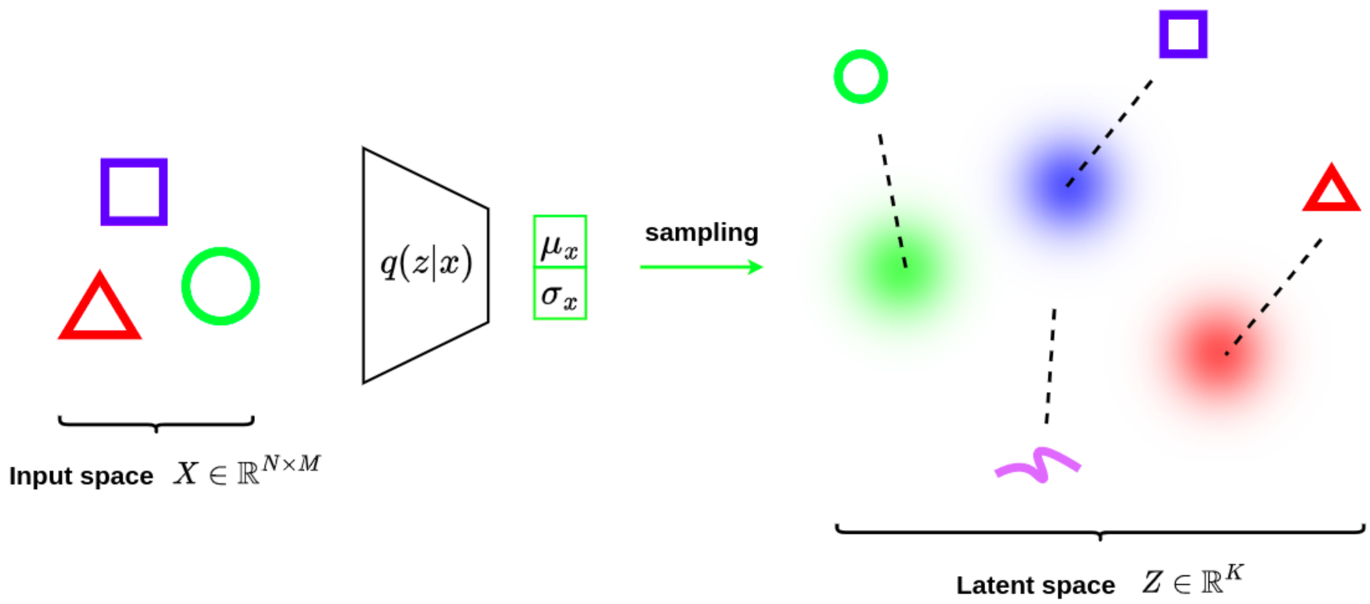


Misc

Probabilistic framework

- **Continuity**

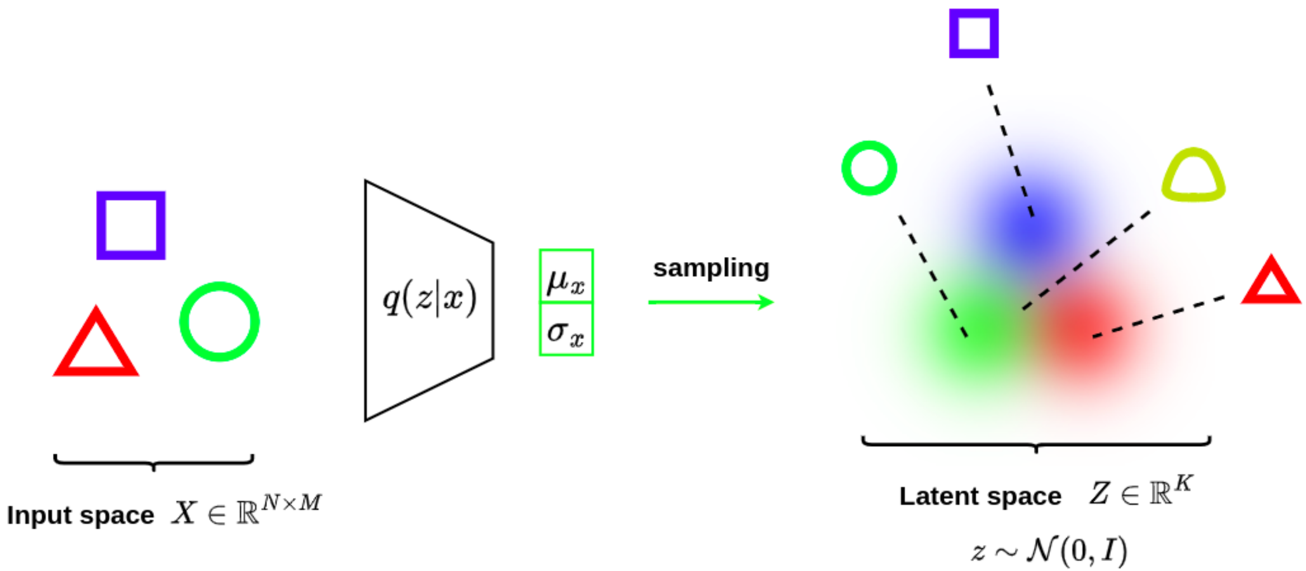
$$\mathcal{N}\left(g(x), h(x)\right)$$



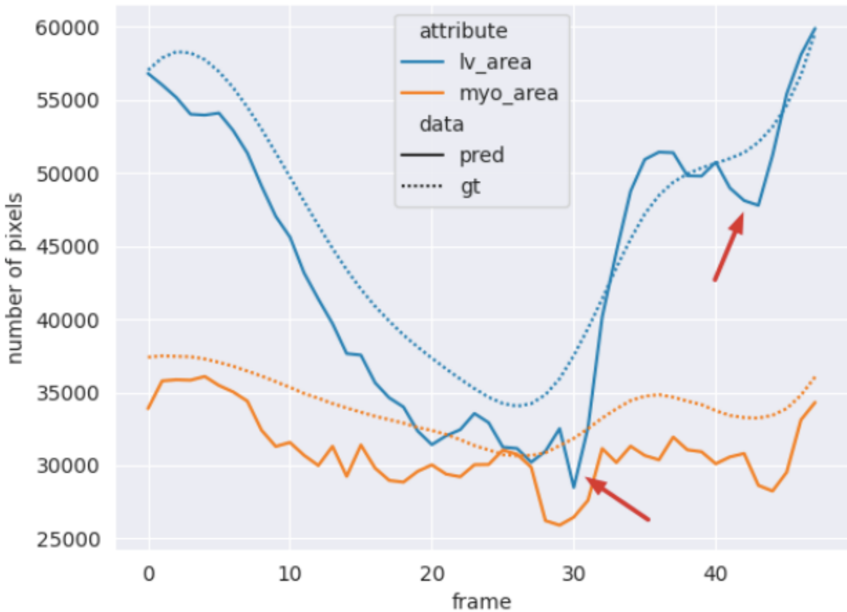
Probabilistic framework

- **Completeness**

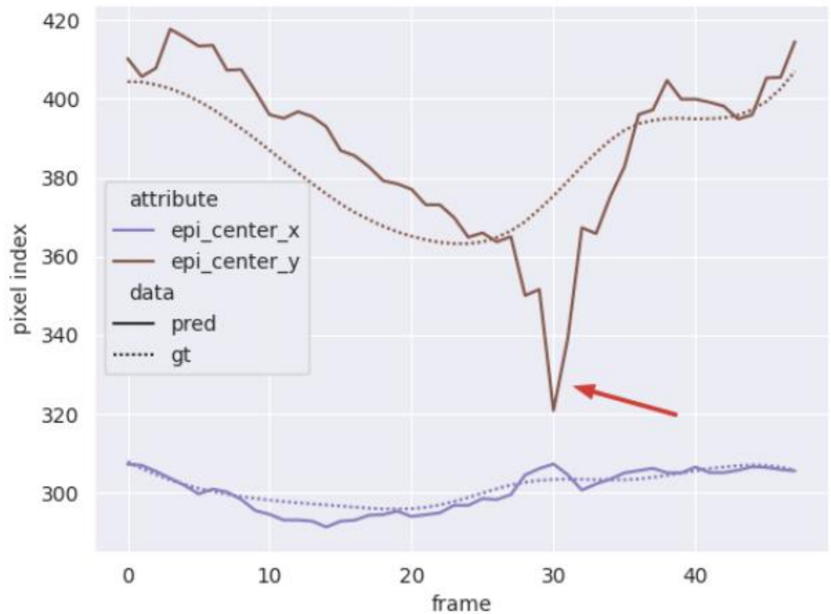
$$\mathcal{N}\left(\cdot, \mathcal{N}(0, I)\right)$$



▶ Temporal inconsistency detection from the latent space



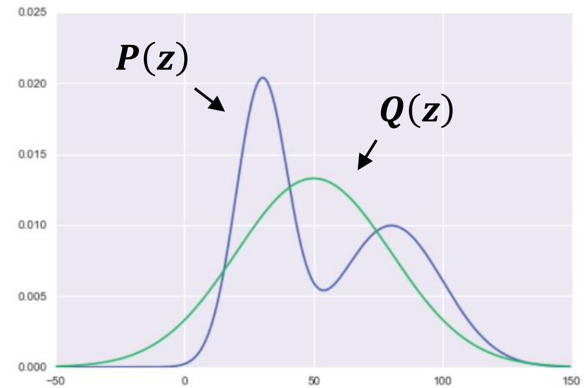
Choppy contraction/dilation of the LV cavity



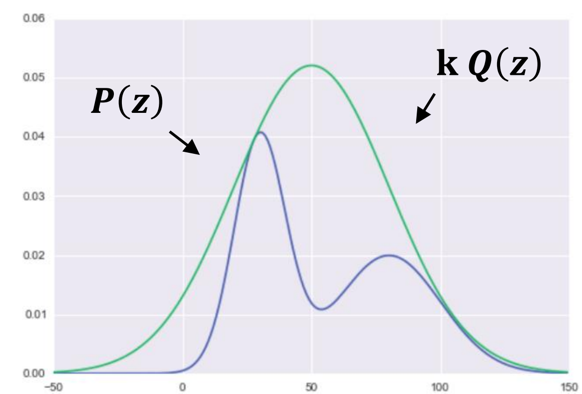
Abrupt vertical shifts of the cardiac shape

Rejection sampling

- **Targeted distribution $P(z)$**
  - ➔ **Parzen window technique**
- **Proposed distribution $Q(z)$**



- **Constrain $kQ(z) > P(z)$**
  - ➔ **Automatic choice of k**



Rejection sampling

- $z \sim Q(z)$
- $u \sim Unif(0, kQ(z))$

- **Computation of $P(z)$**
  - ➔ **If $u \leq P(z)$ then keep z**
  - ➔ **If $u > P(z)$ then reject z**

Populated space