

ECN 2025

Encadré par :

Diana MATEUS et Oriane THIERY

Fait par :

Ilyes LAMSSALAK et Maria VALLE

Apprentissage de caractéristiques d'images médicales non redondantes pour la fusion multimodale

PETvsCLin

Sommaire

1

Contexte et objectifs

2

Bibliographie

3

Méthodologie

4

Nos Travaux

5

Validation expérimentale

6

Conclusion et perspectives

1

Contexte et objectifs

Patients atteints de cancer lymphatique (DLBCL)

Identification par descripteurs de
texture et forme (Radiomiques)



Images
TEP

Calcul analytique

Apprentissage Machine

Hand Crafted Radiomics (HCR)

Non-redondantes

Deep Learning Radiomics (DLR)

Non-redondance ?

Indépendance statistique

Des variables aléatoires x_1 et x_2 sont dites statistiquement indépendants ssi $P(x_1, x_2) = P(x_1) \cdot P(x_2)$

Mesure :

Divergence de Kullback-Leibler

$$D_{\text{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{Q(x)} \right)$$

Information Mutuelle

$$I(X; Y) = D_{\text{KL}}(P_{(X,Y)} \parallel P_X \otimes P_Y)$$

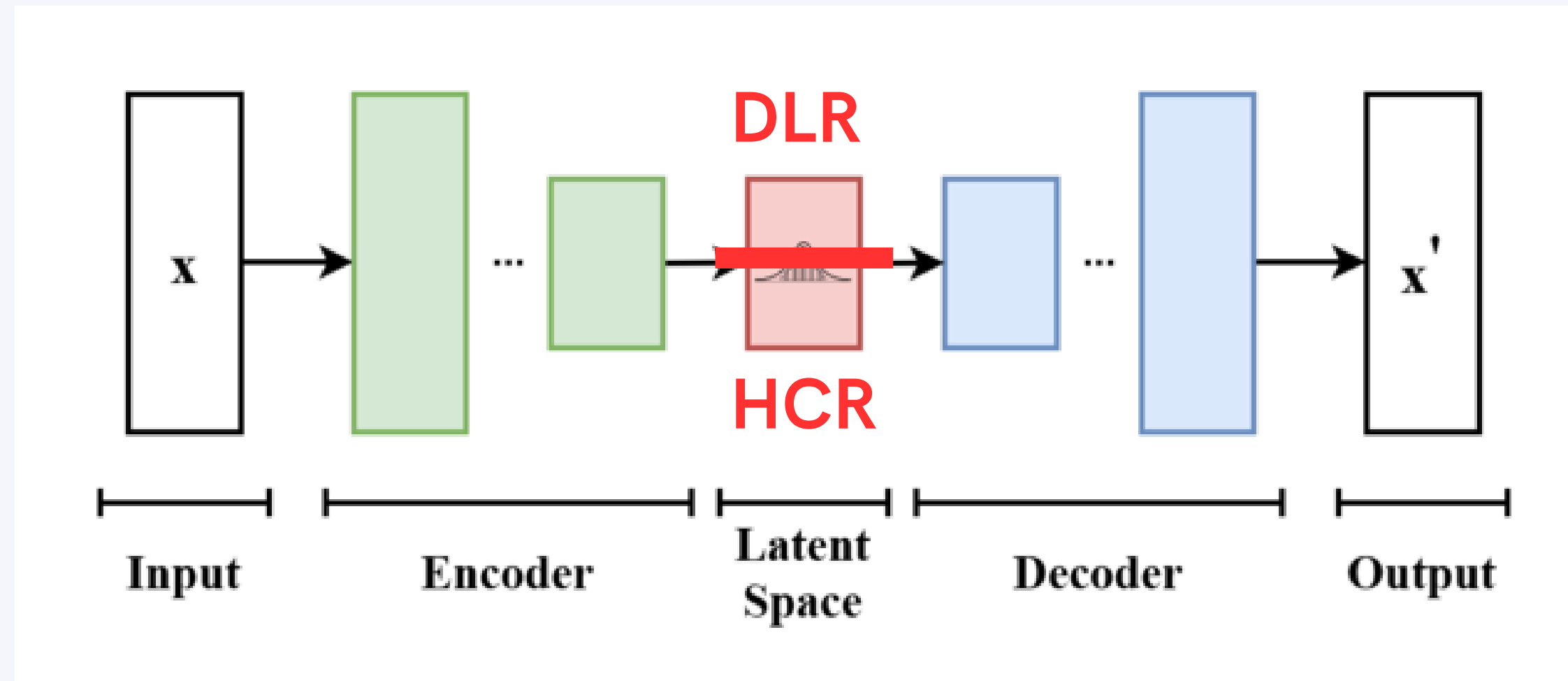
Loi conjointe

Lois marginales

2

Bibliographie

Variational Autoencoder (VAE)



Contrastive Variational Autoencoder (cVAE)

Les articles

Où notre projet s'insère

— Identification de patients de haut risque de cancer

Thiery, Oriane, et al. "Graph-based multimodal multi-lesion DLBCL treatment response prediction from PET images." Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, 2023

D'où vient l'idée

6
6

Vétil, Rebeca, et al. "Non-redundant combination of hand-crafted and deep learning radiomics: Application to the early detection of pancreatic cancer." MICCAI Workshop on Cancer Prevention through Early Detection, 2023

Le point de départ

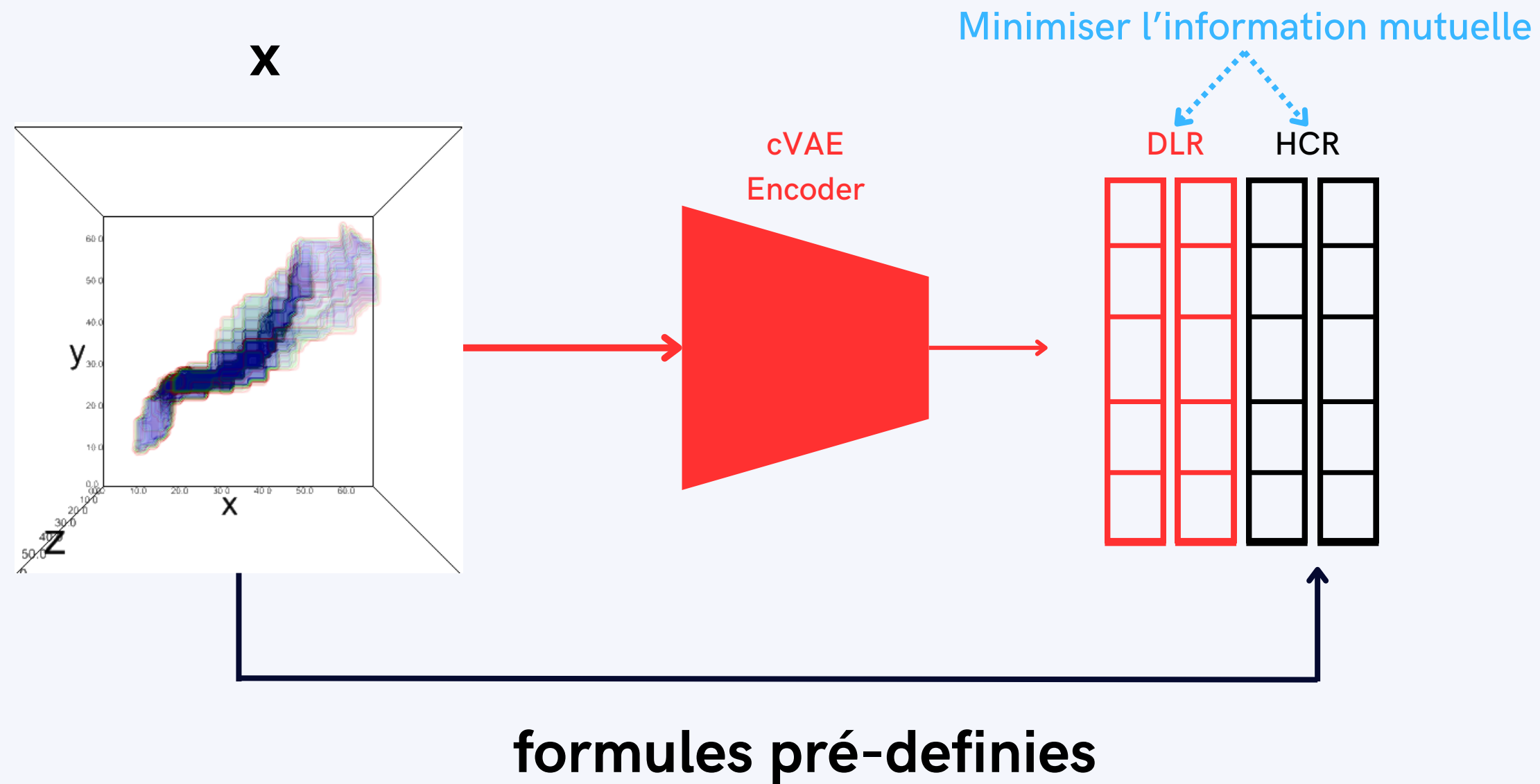
— Codes disponibles

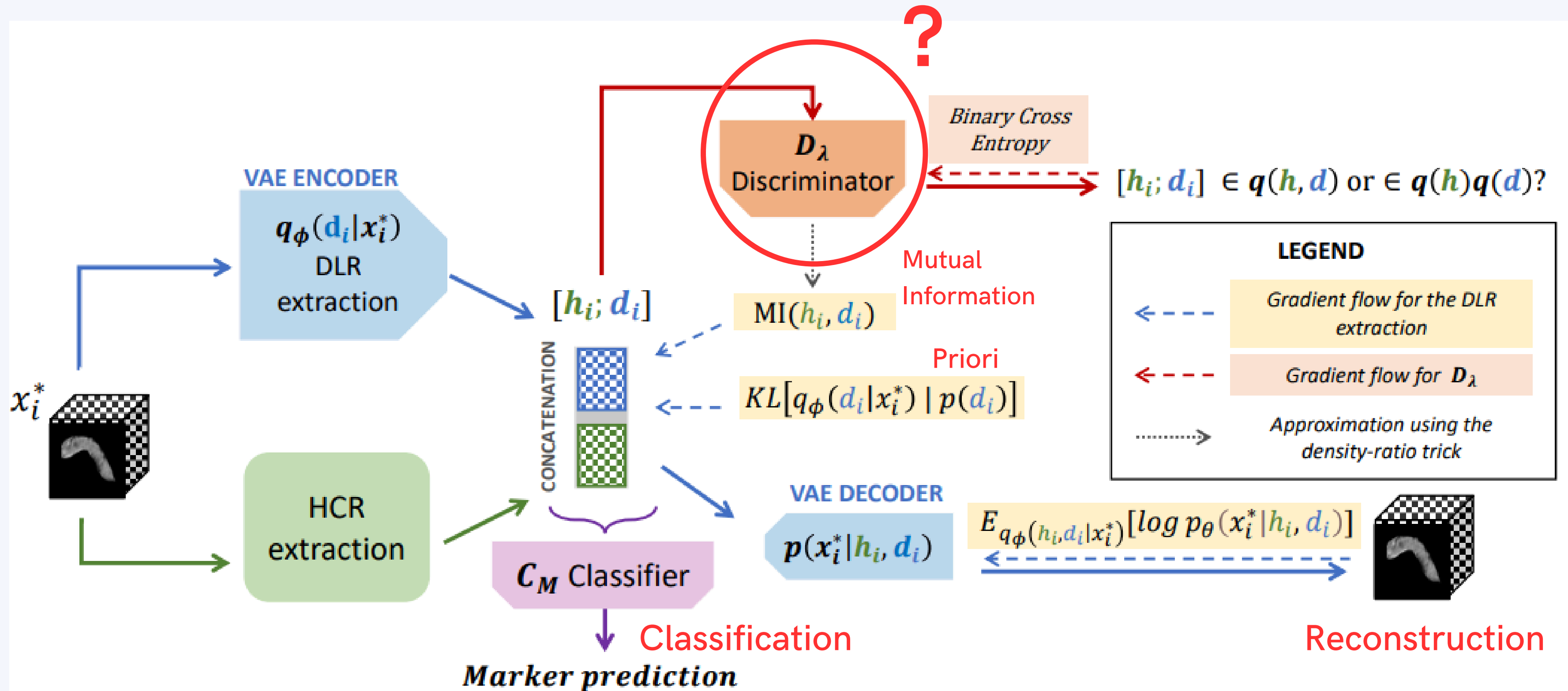
Louiset, Robin, et al. "SepVAE: a contrastive VAE to separate pathological patterns from healthy ones." ICML 3rd Workshop on Interpretable Machine Learning in Healthcare (IMLH), 2023

3

Méthodologie

Extraction des radiomiques



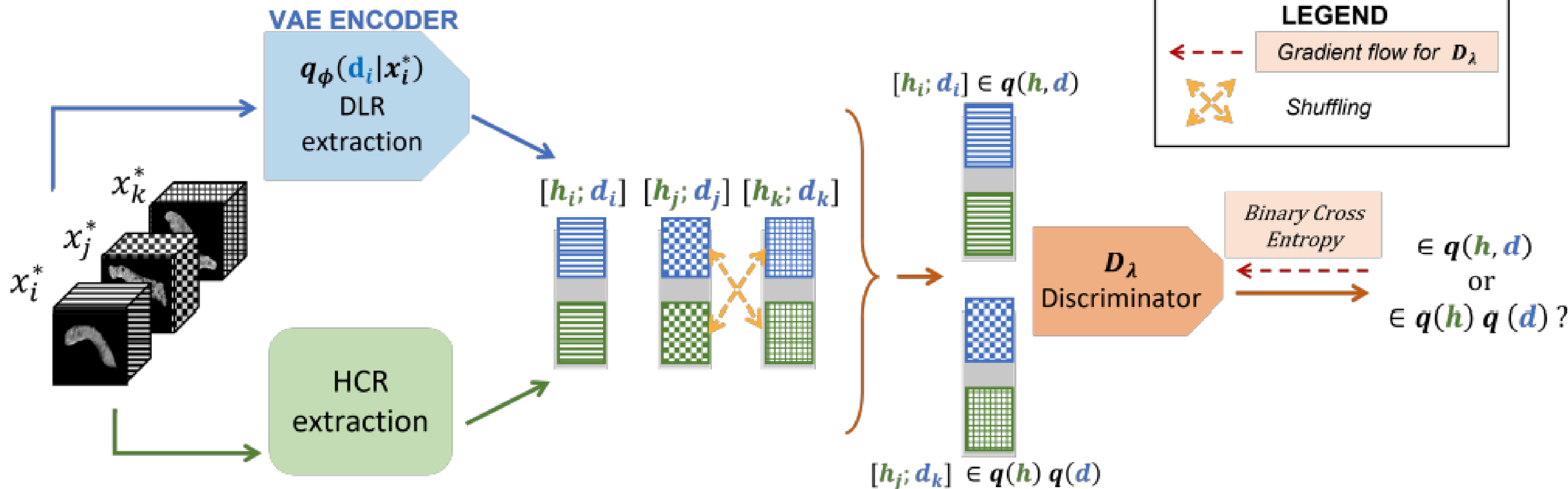


Vétil, Rebeca, et al. "Non-redundant combination of hand-crafted and deep learning radiomics: Application to the early detection of pancreatic cancer." MICCAI Workshop on Cancer Prevention through Early Detection, 2023

Classificateur Indépendant

Density-Ratio Trick

$$KL[q(h, d) \mid q(h)q(d)] = \mathbb{E}_{q(h, d)} \left[\log \frac{q(h, d)}{q(h)q(d)} \right] \approx \sum_i \text{ReLU} \left(\left[\log \frac{\mathcal{D}_\lambda(h_i, d_i)}{1 - \mathcal{D}_\lambda(h_i, d_i)} \right] \right)$$



4

Nos Travaux

Étapes

À partir du code de :
Louiset, Robin, et al. "SepVAE: a contrastive VAE to separate pathological patterns from healthy ones." ICML 3rd Workshop on Interpretable Machine Learning in Healthcare (IMLH), 2023

- 1 Compréhension et démarrage du code
- 2 Recherche des base de donnés
- 3 Adaptations du code (fonction de coût et structure)
- 4 Validation de l'apprentissage
 - Structure du espace latent (PCA, LDA)
 - Classification (Régréssion Logistique, SVM, Random Forest)
 - Reconstruction
 - Visualisation de la fonction de coût

Fonction de Coût du code original

$$\mathcal{LVAE} = -\mathbb{E}_{c,s \sim q_{\phi_c}, q_{\phi_s}(c,s|x,y)} [\log p_{\theta}(x|c, s, y)] + \text{KL}(q_{\phi_c}(c|x) \| p_{\theta}(c))$$

Reconstruction

Commum Prior

$$+ \text{KL}(q(c, s) \| q(c)q(s)) - \mathbb{E}_{s \sim q_{\phi_s}(s|x,y)} [\log p_{\theta}(y|s)] + \text{KL}(q_{\phi_s}(s|x, y) \| p_{\theta}(s|y))$$

Mutual Information

Classification (salient)

Salient Prior

Fonction de Côt

- $x \in \mathbb{R}^V$: Image
- $y \in \{0, 1\}$: Labels binaire
- $x^* = x \times y$: Image masquée
- $h \in \mathbb{R}^{N_h}$: Ensemble des HCR
- $d \in \mathbb{R}^{N_d}$: Ensemble des DLR
- $p_\theta(x^*|y, h, d)$: Modèle génératif supposé pour reconstruire x^*
- $q_\phi(d|x^*)$: Approximation de la distribution a posteriori de d

$$\mathcal{LVAE} = -\mathbb{E}_{q_\phi(h, c|x^*)} [\log p_\theta(x^*|y, h, d)] + \text{KL} [q_\phi(d|x^*) || p(d)]$$

Reconstruction

DLR Prior

$$+ \kappa \text{KL} [q(h, d) || q(h)q(d)] + \mathbb{E}_{q_\theta(d, h|x^*)} [\log p_\theta(y|h, d)]$$

Discriminateur ←

Mutual Information

Classification

4.1

Données

MedMNIST2021

— □ Classssification binaire

2D

- Pneumonia (Rayon X poitrine)
- 5856 échantillons (Tr: 4708 | V: 524 | Te: 624)
- Images 64x64

3D

- AdrenalMNIST (Forme d'abdomen CT)
- 1584 échantillons (Tr: 1188 | V: 98 | Te: 298)
- Images 64x64x64

4.2

Validation Expérimentale

Classification

Space latent

Fonction de coût

Reconstruction

Évaluation Visuelle

Quantification de l'Erreur

Space latent

Influence de la dimension

| HCR seule | | | | HCR + DLR | |
|----------------------|-----|-----|-----|-----------|------|
| H8 | H32 | D16 | D32 | HD24 | HD64 |
| 0 | 0 | 0 | | | |
| DIRE EXPERIENCE FAIT | | | | AUC (%) | |

Space latent | Choix des hyperparamètres

```
# hyper-parameter
hyperparams = {
    "learning_rate" : 8*1e-4,
    "factor_learning_rate" : 1e-4,
    "alpha": 1 , # Paramètres de régularisation L2
    "beta_c": 1e-2, # Poids de la divergence de KL entre la distribution latente des DLR et d'u
    "kappa": 50, # Poids de l'erreur mutuelle information provenant du factor_classifier
    "gamma": 1, # Poids de l'erreur de classification dans l'entraînement du C-VAE
    "fader_param": 1e-2,
    "reconstruction_param" : 1,
    "dlr_size": 18,
    "hcr_size": 10,
    "epochs":10,
}
```

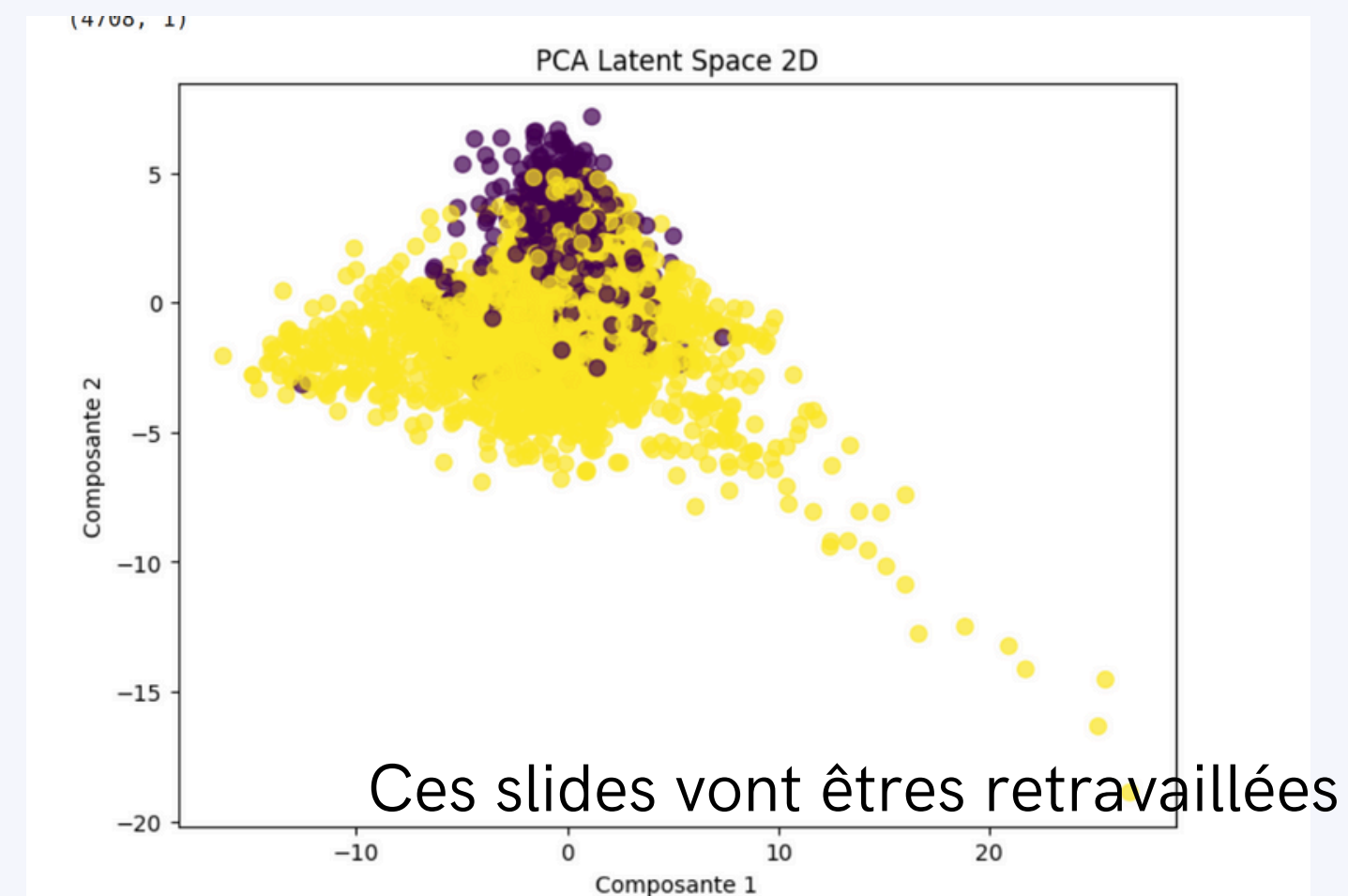
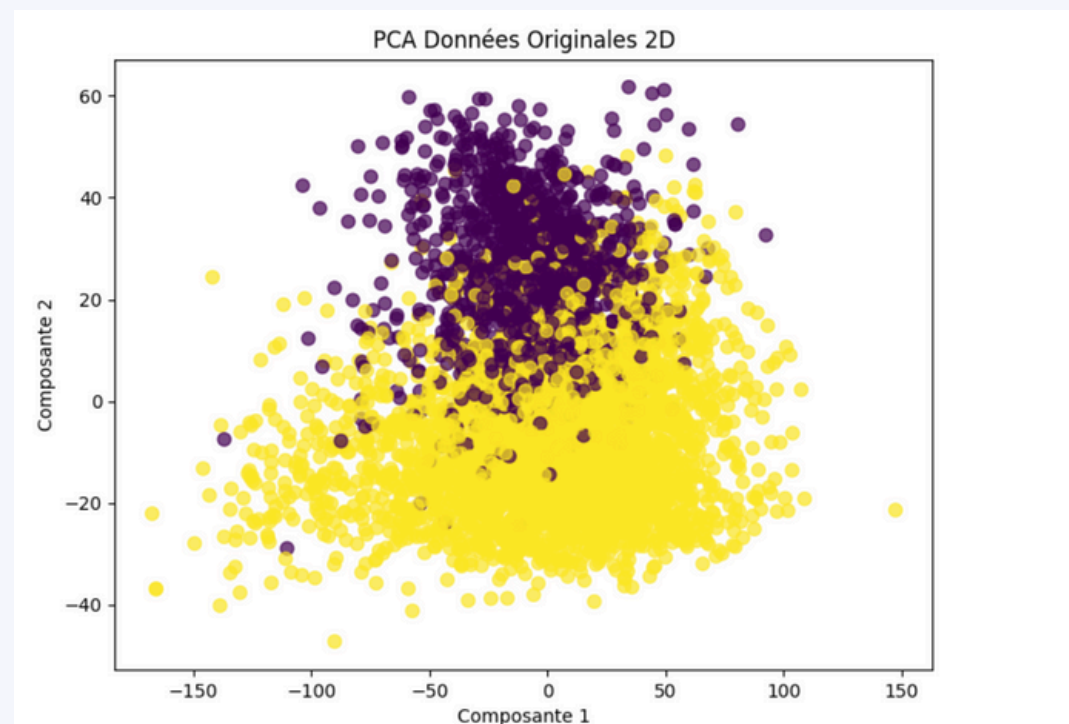
Space latent | Analyse des hyperparamètres

Poids du vecteur de la régression linéaire incoming

Ces slides vont êtres retravaillées

Space latent

Analyse de la Structure



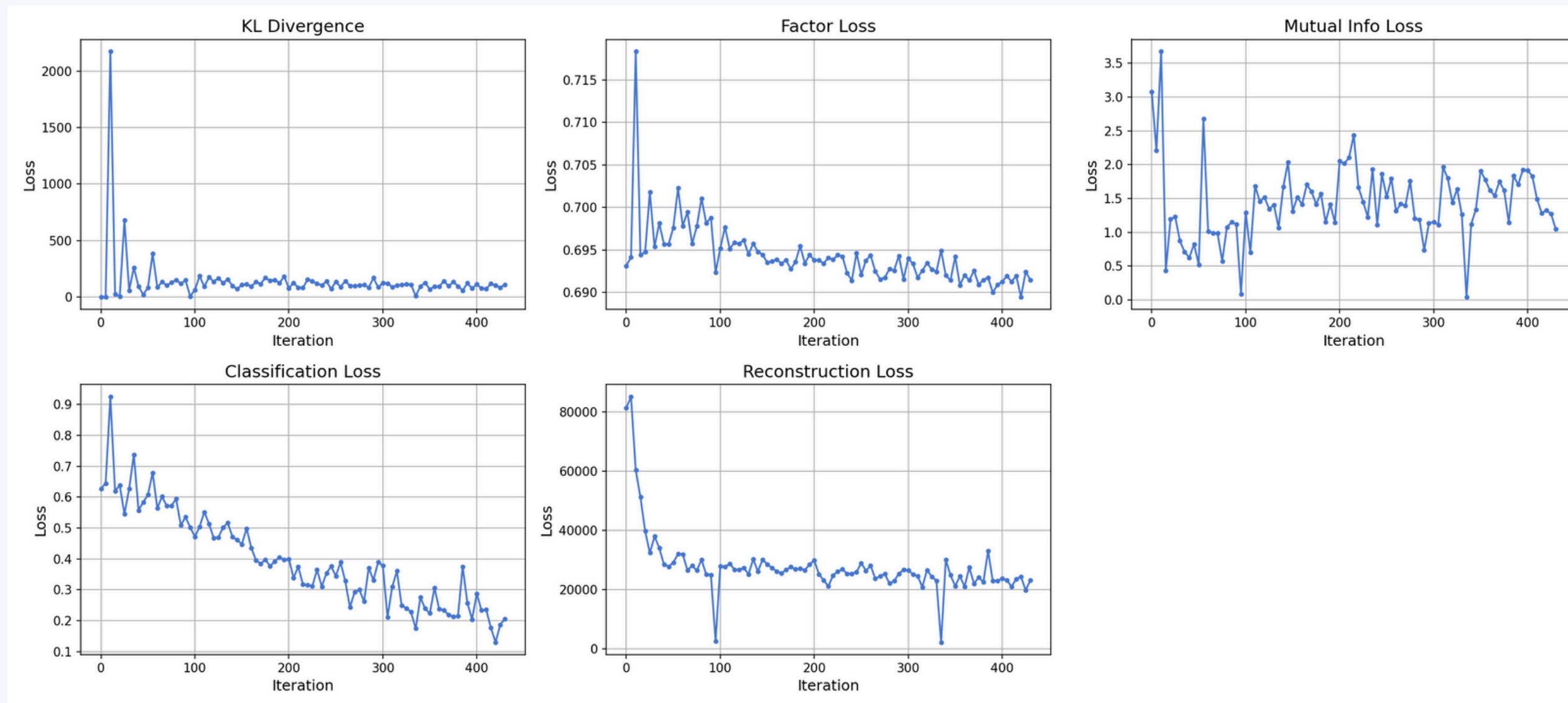
Space latent |

```
➡ TEST B-ACC : 0.8271093578785887
TEST FACTOR B-ACC : 0.5
Linear probe trained space latent, latents HCR + DLR : 0.8173076923076923
Linear probe trained space original, latents HCR + DLR : 0.8333333333333334
Linear probe trained, latents HCR + DLR (SVM) : 0.6442307692307693
/usr/local/lib/python3.11/dist-packages/sklearn/svm/_base.py:305: ConvergenceWarning:
  warnings.warn(
Linear probe trained, latents HCR + DLR (Random forest) : 0.7996794871794872
Variance expliquée du PCA sur les données latentes: [0.37339947 0.20371431]
Variance du PCA du sur les données initiales : [0.38265646 0.10186525]
(4708, 1)
```

AUC (%)

Fonction de coût

Qualité de l'entraînement



5

Conclusion et perspectives

5

Quest-ce que ce qu'on a apporté de concret dans notre travail + idees pour futur

Références

- [1] Thiery, Oriane, et al. "Graph-based multimodal multi-lesion DLBCL treatment response prediction from PET images." Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, 2023
- [2] Vétil, Rebeca, et al. "Non-redundant combination of hand-crafted and deep learning radiomics: Application to the early detection of pancreatic cancer." MICCAI Workshop on Cancer Prevention through Early Detection, 2023
- [3] Louiset, Robin, et al. "SepVAE: a contrastive VAE to separate pathological patterns from healthy ones." ICML 3rd Workshop on Interpretable Machine Learning in Healthcare (IMLH), 2023

Merci pour votre écoute

Élèves:

Ilyes LAMSSALAK : Ilyes.Lamssalak@eleves.ec-nantes.fr

Maria Clara VALLE : Maria-Clara.Valle@eleves.ec-nantes.fr