

Variational Auto-Encoders for cardiac shape modeling

Projet IMA 206

Yanis Aït El Cadi,
Zakaria Chahboune,
Clément Leroy,
Tanguy Schmidtlin

Télécom Paris

2024

- 1 Description du modèle
- 2 Evaluation du modèle
- 3 Analyse de l'espace latent
- 4 Conclusion

Données disponibles

Les données sont un ensemble d'images 3D cardiaques prises à plusieurs instants sur les coeurs de 150 patients différents.

Les images sont segmentées en quatres zones:

- le fond,
- le myocarde,
- les ventricules droit et gauche.

Les images sont associées à des métadonnées qui contiennent des informations sur les patients, et une pathologie cardiaque :

- DCM (dilated cardiomyopathy)
- HCM (hypertrophic cardiomyopathy)
- MINF (previous myocardial infarction)
- NOR (normal subjects)
- RV (abnormal right ventricle)

Deux tiers du dataset est utilisé pour la phase d'entraînement.

Pré-traitement du dataset

On ne conserve que les images correspondant au maximum de dilatation du coeur et au maximum de contraction.

On traite les images selon deux transformations :

- recadrage : les images sont recadrées afin d'éliminer une partie du fond (les tailles relatives sont préservées),
- redimensionnement : les images sont redimensionnées en images de 128 par 128 pixels.

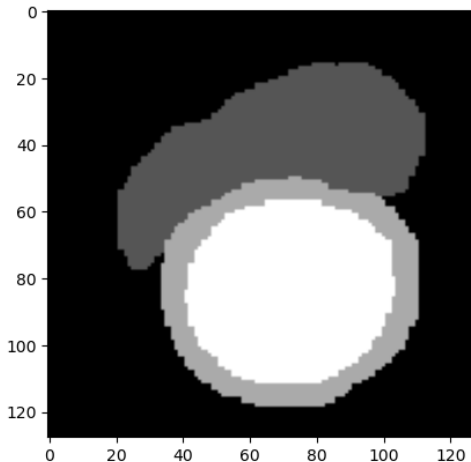
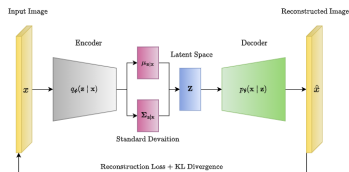


Figure 1: Exemple de slice d'une IRM après pré-traitement

Architecture du modèle

Implémentation d'un β -VAE:



- **Encodeur :**
 - Couches convolutives suivies de couches entièrement connectées
 - Blocs convolutifs successifs : (4, 48), (48, 96), (96, 192), (192, 384)
 - Activation ELU
- **Décodeur :**
 - Couche entièrement connectée produisant $384 \times 8 \times 8$
 - Blocs de convolution transposée successifs : (384, 192), (192, 96), (96, 48), (48, 4)
 - Activation ELU

Fonction de coût

Pour l'entraînement, on définit les fonctions de coût suivante :

$$\mathcal{L} = \mathbb{E}_{q_{\theta_e}(\vec{z}|\mathbf{x})}[-\log q_{\theta_d}(\mathbf{x}|\vec{z})] + \text{KL}(p_{\theta_e}(\vec{z}|\mathbf{x}) \parallel p(\mathbf{z})), \quad (1)$$

$$\mathcal{L} = \mathbb{E}_{q_{\theta_e}(\vec{z}|\mathbf{x})}[-\log q_{\theta_d}(\mathbf{x}|\vec{z})] + \beta \text{KL}(p_{\theta_e}(\vec{z}|\mathbf{x}) \parallel p(\mathbf{z})). \quad (2)$$

- 1 Description du modèle
- 2 Evaluation du modèle**
- 3 Analyse de l'espace latent
- 4 Conclusion

Evaluation de la reconstruction

On utilise l'indice de Sørensen-Dice pour évaluer les performances en reconstruction.

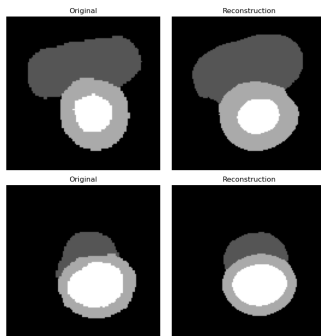


Figure 2: Images originales et images reconstruites par le VAE

Evaluation de la génération

On utilise la Fréchet Inception Distance (FID) pour évaluer les performances en génération d'image.

FID: distance de Fréchet entre les distributions de caractéristiques réelles et de caractéristiques reconstruites. Les caractéristiques sont extraites par un réseau de neurones pré-entraîné (Inception v3).

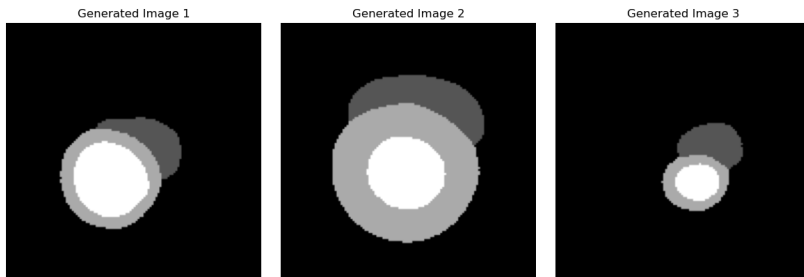
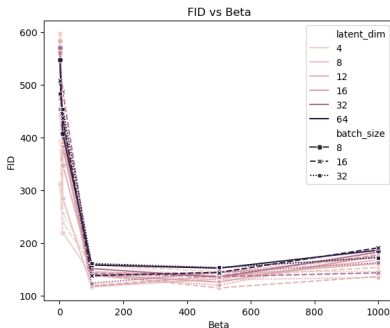
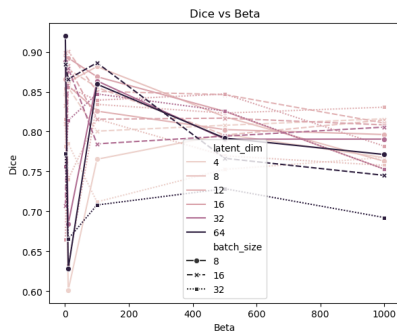


Figure 3: Exemple d'images générées par le modèle

Etude de l'impact des différents paramètres



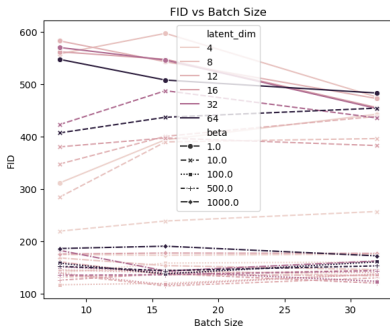
(a) FID vs Beta



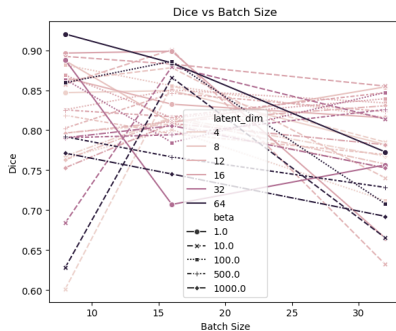
(b) Dice vs Beta

Figure 4: Impact de β sur la FID et l'indice de S rensen-Dice.

Etude de l'impact des différents paramètres



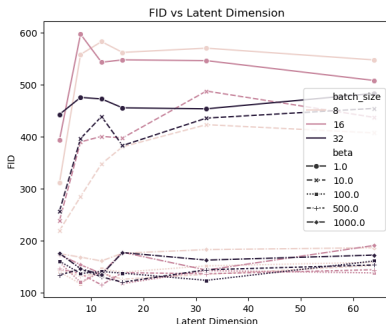
(a) FID vs Batch Size



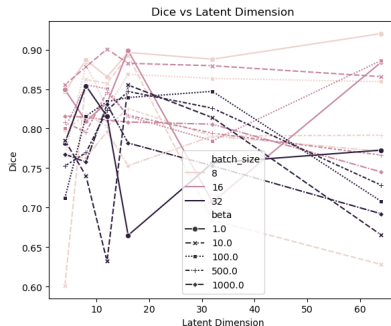
(b) Dice vs Batch Size

Figure 5: Impact de la taille du batch sur la FID et l'indice de S rensen-Dice.

Etude de l'impact des différents paramètres



(a) FID vs Latent Dimension



(b) Dice vs Latent Dimension

Figure 6: Impact de la dimension latente sur la FID et l'indice de Sørensen-Dice.

Paramètres optimaux

Les paramètres optimaux sont les suivants : une dimension latente égale à 8, une taille de batch égale à 8, et un β égal à 100.

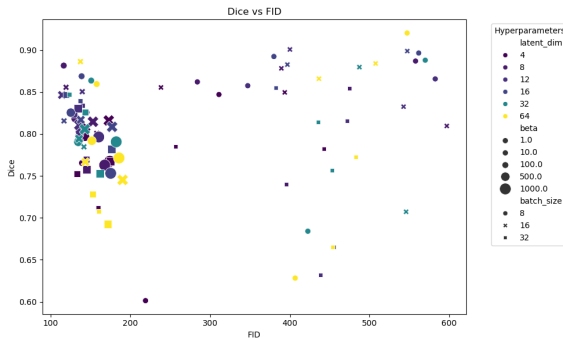
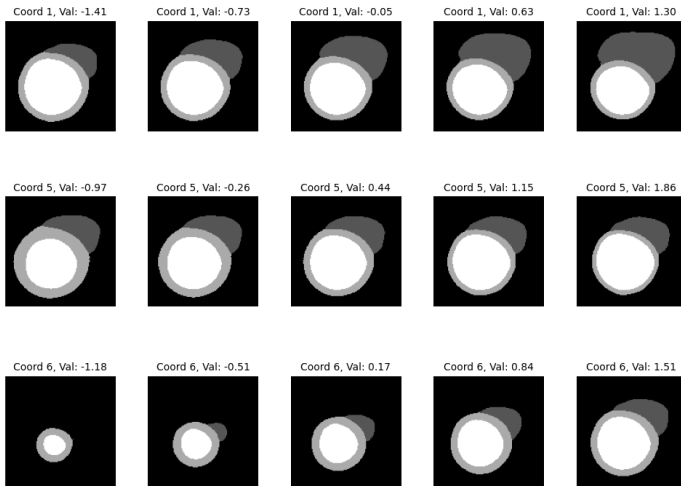


Figure 7: FID et indice de Sørensen-Dice pour des modèles avec différents paramètres de taille de batch, de dimension latente et de β .

- 1 Description du modèle
- 2 Evaluation du modèle
- 3 Analyse de l'espace latent**
- 4 Conclusion

Analyse de l'espace latent



UMAP: Vérification de l'overfit

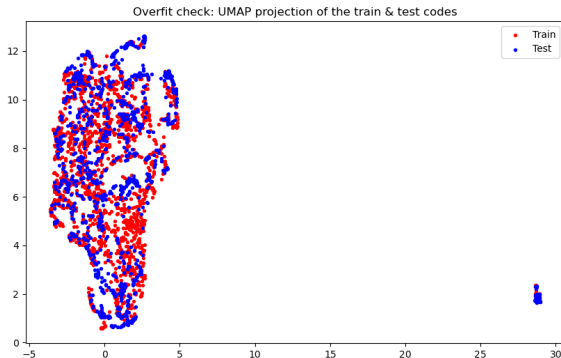


Figure 8: Projection des codes d'entraînement et de test.

UMAP: Outliers

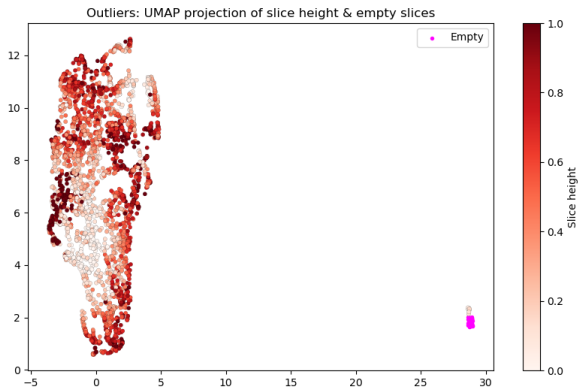


Figure 9: Projection des coupes selon la hauteur relative + coupes vides.

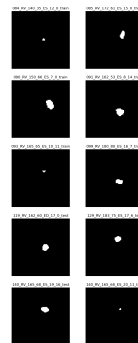


Figure 10:
Coupes non vides
parmi les outliers

UMAP: Groupes

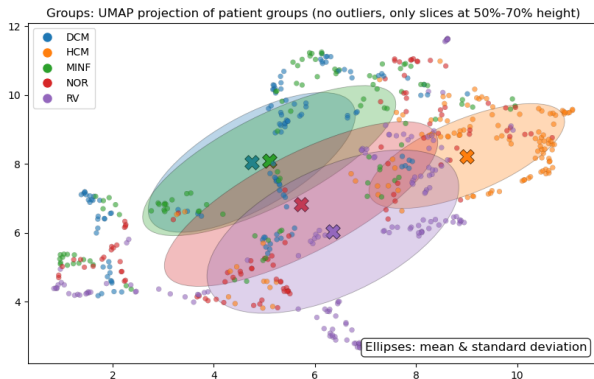


Figure 11: Projection des groupes de patients (sans outliers), coupes entre 50 et 70% de hauteur).

- 1 Description du modèle
- 2 Evaluation du modèle
- 3 Analyse de l'espace latent
- 4 Conclusion**

Conclusion

Les pistes d'amélioration:

- retirer les coupes vides des données d'entraînement,
- vérifier la cohérence des segmentations générées,
- entraîner un modèle pour tenir en compte la cohérence 3D des IRMs.