

SegSRGAN : Segmentation et Super-Résolution Generative Adversarial Networks d'IRM néonatales

Quentin Delannoy - Ingénieur de recherche

Lundi 8 mars 2021



ANR MAIA
(ANR-15-CE23-0009)

Les enfants prématurés peuvent garder des séquelles neurologiques de leur prématurité :

Chez les grands prématurés 10% présentent des handicaps moteurs et 15% des déficiences intellectuelles (source : <https://epipage2.inserm.fr/index.php/fr/prematurite-fr/etat-des-connaissances>)

La compréhension des anomalies cérébrales en lien avec ces séquelles reste limitée.

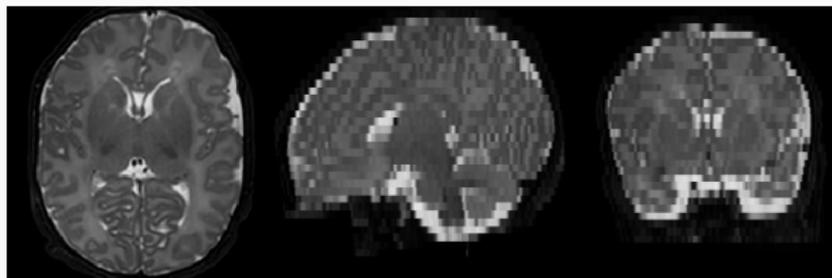
Données : 1500 images IRM de 400 enfants prématurés (**Epirmex**). Les images IRM ont une résolution de $0.5 \times 0.5 \times 3 \text{ mm}^3$ (temps d'acquisition faible)

Objectif : Étudier certaines zones du cerveau pour améliorer la compréhension du phénomène sous-jacent à ces séquelles.

Moyen : Segmenter certaines zones du cerveau pour pouvoir les analyser.

Segmentation d'images anisotropes

Exemple d'images IRM (Epirmex) : résolution : $0.5 \times 0.5 \times 3 \text{ mm}^3$



Pipeline usuel de segmentation d'images anisotropes :

1. Interpolation
2. Segmentation

Problème : Présence d'approximations liées à l'interpolation.

Idée : Essayer de rendre la segmentation robuste à ces approximations

En pratique : Fusionner super-résolution et segmentation.

- ① Présentation générale de SegSRGAN
- ② Bases de données utilisées
- ③ SegSRGAN : Segmentation corticale
- ④ Étude des résultats sur la base Epirmex
- ⑤ Dernières extensions
- ⑥ Conclusion

Présentation générale de SegSRGAN

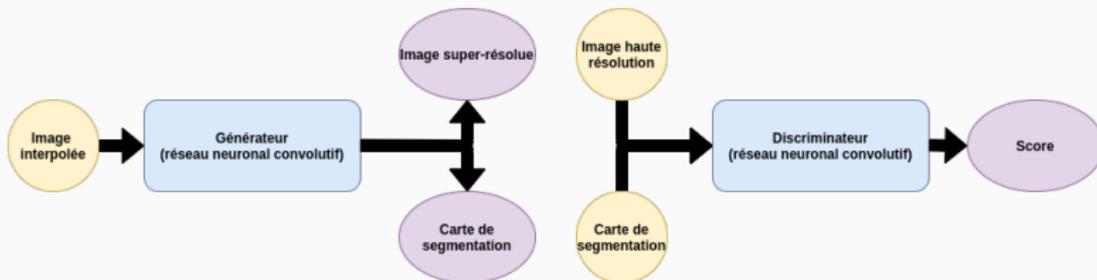
Architecture basée sur le principe des réseaux GAN (Generative Adversarial Network).

Objectif : Générer des données qui ressemblent à celles de la base de données d'entraînement.

Deux réseaux antagonistes :

- **Discriminateur** : Reconnaître les données générées des vraies données
- **Générateur** : Générer des données qui ressemblent aux vraies données

Image interpolée (isotrope) : image anisotrope interpolée avec des splines cubiques



Problèmes d'optimisation :

Entraînement du discriminateur (D) :

$$\min_D \lambda_{gp} \mathcal{L}_{gp} - \mathcal{L}_{adv}$$

- Maximiser le score calculé sur les "vraies" images (\mathcal{L}_{adv})
- Minimiser le score calculé sur les images générées par le générateur (\mathcal{L}_{adv})
- Hypothèse WGAN-GP : Rester proche d'une fonction 1-lipschitzienne (\mathcal{L}_{gp})

Entraînement du générateur (G) :

$$\min_G \mathcal{L}_{rec} + \lambda_{adv} \mathcal{L}_{adv}$$

- Minimiser la distance voxel par voxel entre les vraies données et les données générées (\mathcal{L}_{rec})
- Maximiser le score discriminateur calculé sur les images générées (\mathcal{L}_{adv})

Relation antagoniste entre les deux réseaux :

$$\mathcal{L}_{adv} = \mathbb{E}_{\mathbf{X} \sim \mathbb{P}_{\mathbf{X}}, \mathbf{S}_{\mathbf{X}} \sim \mathbb{P}_{\mathbf{S}_{\mathbf{X}}}} [D((\mathbf{X}, \mathbf{S}_{\mathbf{X}}))] - \mathbb{E}_{\mathbf{Z} \sim \mathbb{P}_{\mathbf{Z}}} [D(G(\mathbf{Z}))]$$

$\mathbf{X}, \mathbf{S}_{\mathbf{X}}$: IRM haute résolution et vraie segmentation, \mathbf{Z} : images interpolées

Bases de données utilisées

Rappel :

L'entraînement nécessite une base de données contenant :

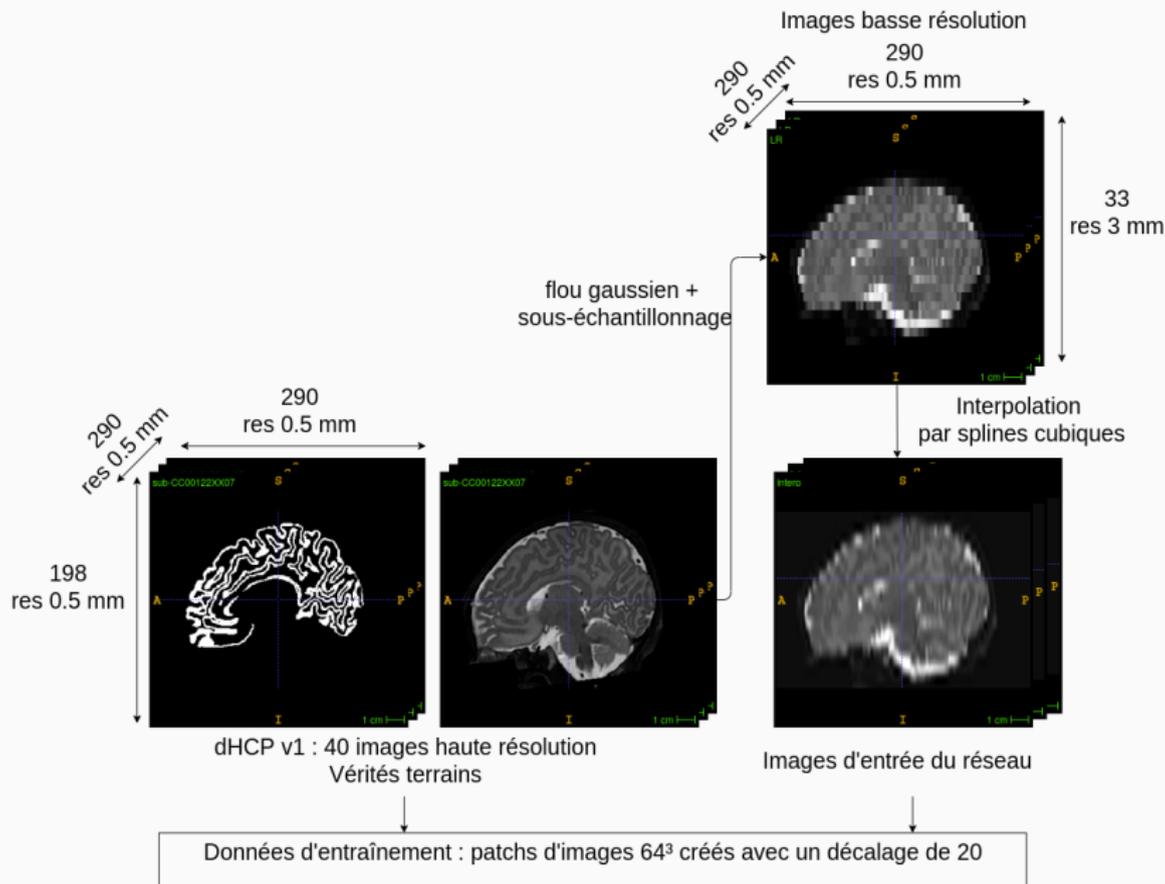
- Des images IRM anisotropes (basse résolution) qui seront données en entrée à la méthode.
- Des images IRM et de segmentation isotropes (haute résolution).

Objectif de la méthode : Obtenir une carte de segmentation et une image IRM de haute résolution à partir de l'image IRM de basse résolution.

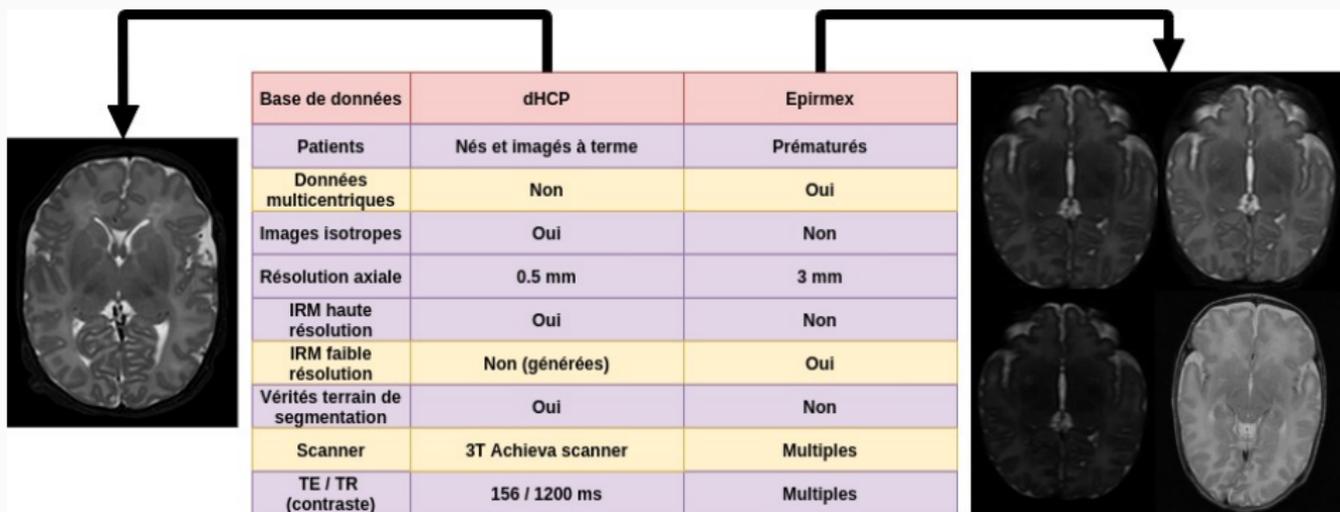
Problème : Nous n'avons pas accès à une base de données qui contient la même image IRM en basse et haute résolution ainsi qu'une carte de segmentation

Solution : Utilisation d'une base de données (dHCP) contenant les segmentations et des images IRM (haute résolution) puis simulation d'images IRM basse résolution

Génération des données d'entraînement



Comparaison entre les bases de données dHCP et Epirmex



Base de données	dHCP	Epirmex
Patients	Nés et imagés à terme	Prématurés
Données multicentriques	Non	Oui
Images isotropes	Oui	Non
Résolution axiale	0.5 mm	3 mm
IRM haute résolution	Oui	Non
IRM faible résolution	Non (générées)	Oui
Vérités terrain de segmentation	Oui	Non
Scanner	3T Achieva scanner	Multiples
TE / TR (contraste)	156 / 1200 ms	Multiples

Augmentation de la variabilité de la base dHCP : Contraste et ajout de bruit.

Analyse des deux bases complémentaires :

dHCP :

Vérités terrains (analyse quantitative)

IRM basse résolution générées

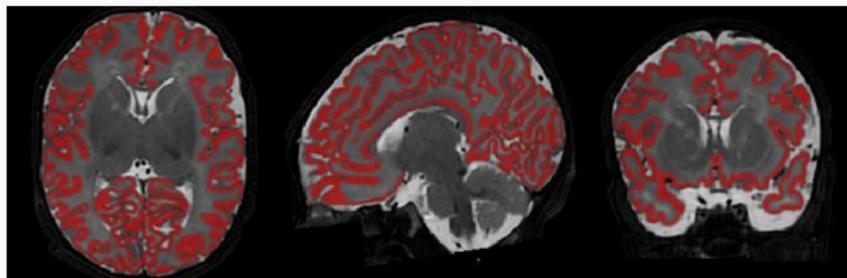
Epirmex :

Pas de vérités terrains (analyse qualitative)

Vraies IRM basse résolution

SegSRGAN : Segmentation corticale

Objectif : Segmentation du cortex (rouge)



Indices utilisés :

Reconstruction haute résolution :

$$\text{PSNR}(X, Y) = 10 \log_{10} \frac{(\max_i X_i)^2}{\frac{1}{|X|} \sum_i (X_i - Y_i)^2}$$

X et Y image haute résolution réelle et estimée

Segmentation :

$$\text{Dice}(S, G) = \frac{2|S \cap G|}{|S| + |G|}$$

G et S segmentation (binaire) réelle et estimée

Valeurs moyennes obtenues sur 8 images tests (patch 128^3 avec un decalage de 30) :

PSNR :

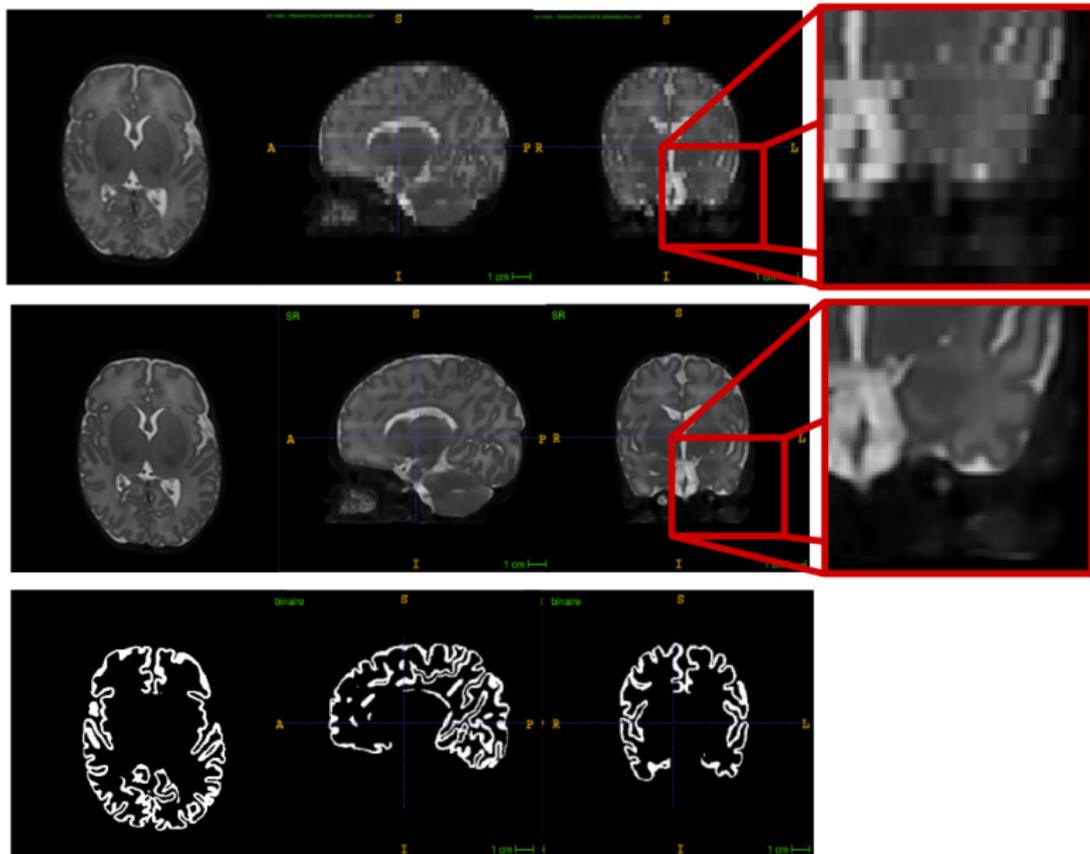
SegSRGAN : 26.96

Interpolation par spline cubique : 24.22

Dice :

SegSRGAN : 0.855

Évaluation qualitative des résultats (Epirmex)



Étude des résultats sur la base Epirmex

Analyse de la segmentation basée sur le volume du cortex.

Objectif final : corréler ce volume avec le devenir de l'enfant.

Plusieurs facteurs peuvent influencer sur les volumes obtenus :

- **L'âge du patient** au moment où il est imagé
- **Les paramètres d'acquisition** des images

Données utilisées pour étudier l'impact de ces facteurs : 43 patients ont 4 acquisitions effectuées le même jour avec des contrastes différents (contrastés identiques entre les patients)

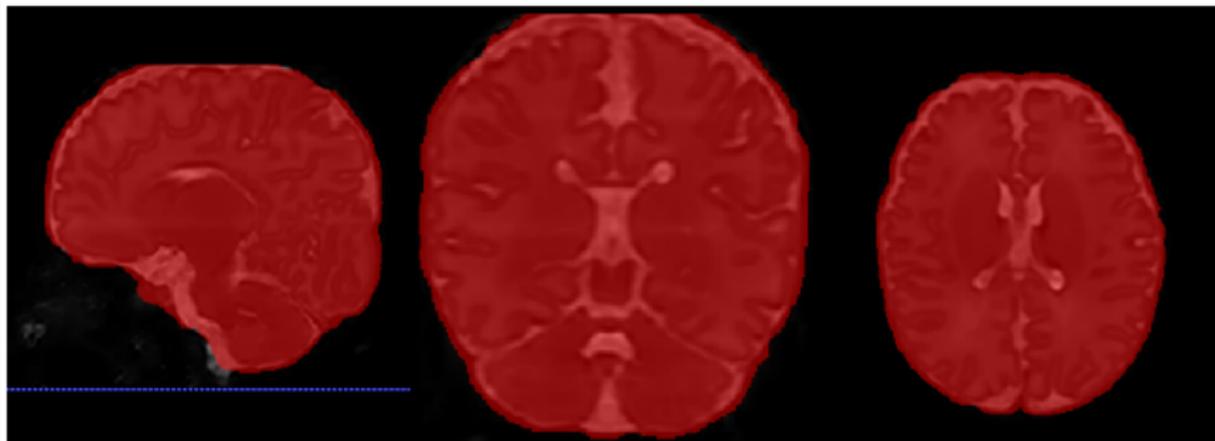
L'âge du patient 1/2

Afin de rendre comparable le volume cortical sur des enfants qui ont un développement différent :

Diviser le volume cortical par le volume intracrânien.

Moyen : estimation d'un masque intracrânien.

Par exemple en utilisant : Brain Extraction Tool (BET).



Valeur de TE (ms)		120	200	280	360
Ecart-type divisé par moyenne	Volume cortex	0.0939	0.0942	0.0960	0.1053
	Ratio volume cortex volume intracrânien	0.0543	0.0396	0.0443	0.0875
Distance inter-quartile divisée par médiane	Volume cortex	0.0956	0.1018	0.0852	0.0960
	Ratio volume cortex volume intracrânien	0.0471	0.0465	0.0534	0.0764

⇒ Suite à la division du volume cortical par le volume intracrânien, nous observons une **réduction de la variabilité** des résultats **entre les différents patients**.

Comparaison basée sur la présence d'IRM acquises avec des contrastes différents le même jour sur un même patient (43 patients).

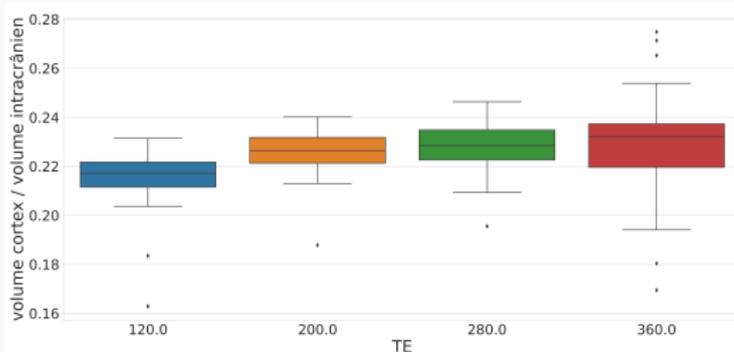
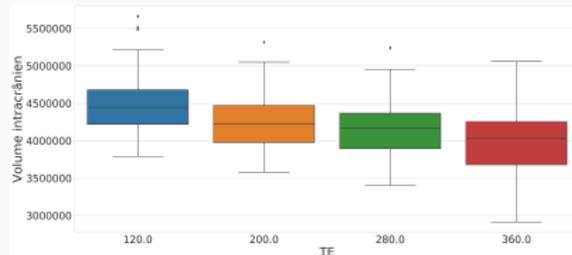
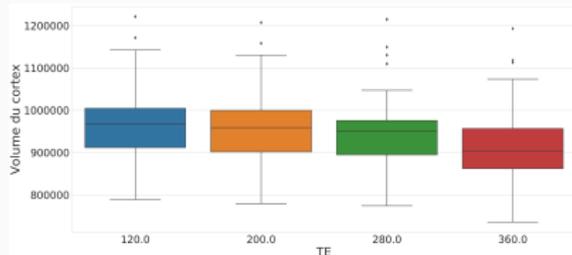
Par exemple ici avec le sujet Abvg4 du centre A :



Segmentation corticale obtenue avec TE=120ms en blanc et TE=280ms en rouge

Impact des paramètres d'acquisition (TE) 2/2

On retrouve cet impact avec plusieurs patients :



⇒ Le paramètre TE a une influence sur les résultats obtenus et nécessite donc d'être pris en compte pour comparer les volumétries obtenues sur différentes images

Dernières extensions

Dernières extensions :

- Segmentation du cerveau en plusieurs classes
- Estimation d'un masque qui sépare la partie droite et gauche du cerveau. Cela permet également d'obtenir un masque intracrânien

Démonstration

Conclusion

Même si certaines différences existent entre les données dHCP et Epirmex. Les résultats observés semblent pouvoir donner lieu à une analyse sur epirmex.

Une attention doit être portée aux facteurs pouvant influencer sur les résultats.

Implémentation en python / Keras.

Disponibilité d'un package sur Pipy¹.

Perspectives :

- Validation définitive des segmentations multi-classes sur la base Epirmex
- Étude des facteurs influant sur les résultats en segmentation multi-classe

1. https://github.com/Qdelannoy/SegSRGAN_multi_labels/tree/develop

SegSRGAN : Segmentation et Super-Résolution Generative Adversarial Networks d'IRM néonatales

Quentin Delannoy - Ingénieur de recherche

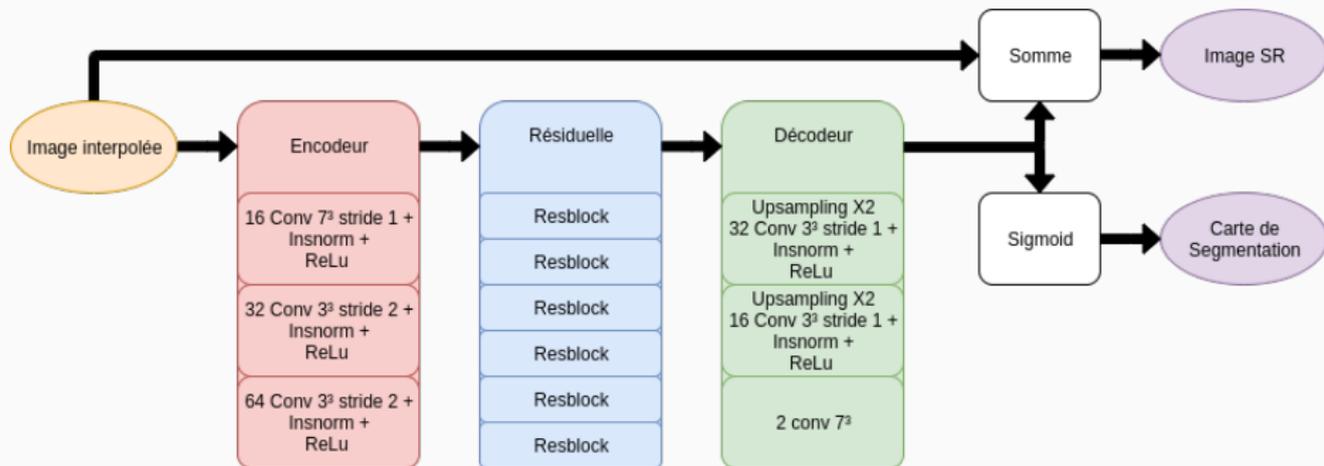
Lundi 8 mars 2021



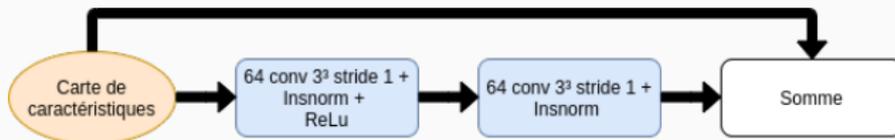
ANR MAIA
(ANR-15-CE23-0009)

Architecture du générateur (segmentation binaire)

Architecture du Générateur :



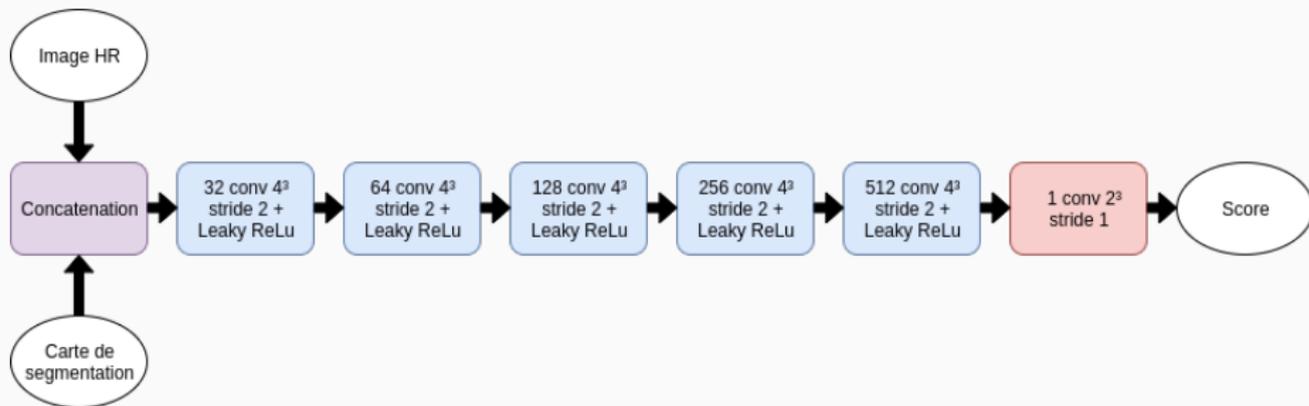
Resblock :



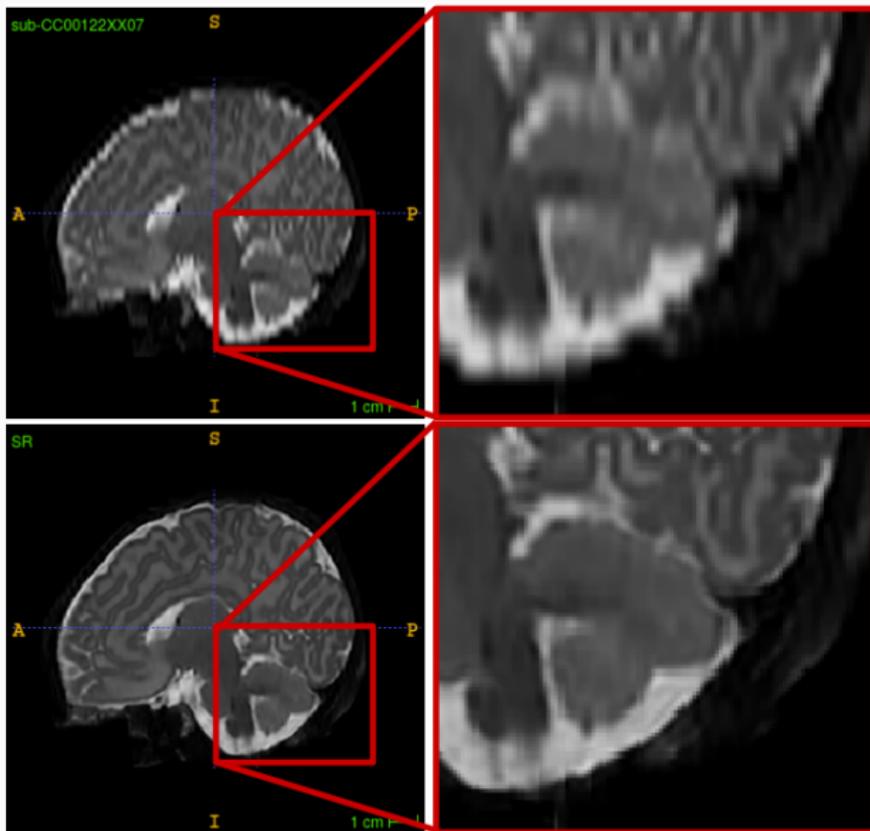
Architecture du discriminateur (segmentation binaire)

Données d'entraînement : patchs d'image de taille 64^3

Architecture du Discriminateur :



Comparaison visuelle de l'image interpolée et SR (dHCP). Segmentation binaire



Comparaison avec d'autres méthodes, toutes appliquées après interpolation de l'image basse résolution :

	SegSRGAN	IMAPA	DrawEM
Dice	0.855(\pm 0.014)	0.786(\pm 0.023)	0.730(\pm 0.010)

Table 1: Dice scores mean value (and standard deviation) for segmentation map. Calculated over 8 dHCP images

Fonction de coût (1/3)

D : discriminateur , **G** : générateur.

Problèmes d'optimisation :

$$\min_D \lambda_{gp} \mathcal{L}_{gp} - \mathcal{L}_{adv}$$

$$\min_G \mathcal{L}_{rec} + \lambda_{adv} \mathcal{L}_{adv}$$

Le jeu "concurrentiel" \mathcal{L}_{adv} :

$$\mathcal{L}_{adv} = \mathbb{E}_{\mathbf{X} \sim \mathbb{P}_{\mathbf{X}}, \mathbf{S}_{\mathbf{X}} \sim \mathbb{P}_{\mathbf{S}_{\mathbf{X}}}} [D((\mathbf{X}, \mathbf{S}_{\mathbf{X}}))] - \mathbb{E}_{\mathbf{Z} \sim \mathbb{P}_{\mathbf{Z}}} [D(G(\mathbf{Z}))]$$

X, **S_X** : IRM haute résolution et vraie segmentation, **Z** : images interpolées
 \mathcal{L}_{rec} , comparaison voxel par voxel entre vérités terrains et images générées.

Le type de GAN implémenté est un WGAN-GP (Gulrajani et al., 2017) :

\mathcal{L}_{gp} , terme de pénalisation pour le discriminateur (distance à une fonction 1-lipschitzienne) . Hypothèse WGAN

Fonction de coût (2/3)

Rappel :

$$\min_D \lambda_{gp} \mathcal{L}_{gp} - \mathcal{L}_{adv}$$

$$\min_G \mathcal{L}_{rec} + \lambda_{adv} \mathcal{L}_{adv}$$

en notant $(\mathbf{X}, \mathbf{S}_\mathbf{X})$ les "vraies" segmentations et IRM haute résolution et \mathbf{Z} les images interpolées.

\mathcal{L}_{gp} est définie comme :

$$\mathcal{L}_{gp} = \mathbb{E}_{\widehat{\mathbf{XS}}} [(\|\nabla_{\widehat{\mathbf{XS}}} D(\widehat{\mathbf{XS}})\|_2 - 1)^2]$$

$$\text{Avec } \widehat{\mathbf{XS}} = (1 - \varepsilon)(\mathbf{X}, \mathbf{S}_\mathbf{X}) + \varepsilon G(\mathbf{Z}) \text{ et } \varepsilon \sim U[0, 1]$$

Et \mathcal{L}_{rec} comme :

$$\mathcal{L}_{rec} = \mathbb{E}_{\mathbf{X} \sim \mathbb{P}_{\mathbf{X}}, \mathbf{S}_{\mathbf{X}} \sim \mathbb{P}_{\mathbf{S}_{\mathbf{X}}}} [\rho((\mathbf{X}, \mathbf{S}_{\mathbf{X}}) - G(\mathbf{Z}))]$$

Où ρ (robust Charbonnier loss (Charbonnier et al., 1997)) est calculé pixel à pixel :

$$\rho(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i^2 + 10^{-6})}$$

Bibliographie

- Charbonnier, P., L. Blanc-Féraud, G. Aubert, and M. Barlaud (1997). Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing* 6(2), 298–311.
- Gulrajani, I., F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville (2017). Improved training of wasserstein gans. *CoRR abs/1704.00028*.