Réseaux antagonistes génératifs pour la reconstruction super-résolution et la segmentation en IRM

Quentin Delannoy, Chi-Hieu Pham, Clément Cazorla, Carlos Tor-Díez, Guillaume Dollé, Hélène Meunier, Nathalie Bednarek, Ronan Fablet, Nicolas Passat, François Rousseau

▶ **To cite this version:**

Quentin Delannoy, Chi-Hieu Pham, Clément Cazorla, Carlos Tor-Díez, Guillaume Dollé, et al.. Réseaux antagonistes génératifs pour la reconstruction super-résolution et la segmentation en IRM. Extraction et Gestion des Connaissances - Atelier Apprentissage Profond : Théorie et Applications (APTA@EGC), 2020, Bruxelles, Belgique. hal-02387754

HAL Id: hal-02387754
https://hal.science/hal-02387754

Submitted on 15 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Réseaux antagonistes génératifs pour la reconstruction super-résolution et la segmentation en IRM

Quentin Delannoy*, Chi-Hieu Pham**, Clément Cazorla*, Carlos Tor-Díez**, Guillaume Dollé***, Hélène Meunier****, Nathalie Bednarek*,****, Ronan Fablet♯, Nicolas Passat*, François Rousseau***

* Université de Reims Champagne Ardenne, CReSTIC, 51097 Reims, France
** IMT Atlantique, LaTIM U1101 INSERM, UBL, Brest, France
*** Université de Reims Champagne Ardenne, CNRS, LMR UMR 9008, 51097 Reims, France
**** Service de médecine néonatale et réanimation pédiatrique, CHU de Reims, France
♯ IMT Atlantique, Lab-STICC UMR CNRS 6285, Brest, France

Résumé. La faible résolution et l'anisotropie des données induisent des difficultés pour l'analyse cérébrale néonatale en IRM (imagerie par résonance magnétique). Dans la plupart des pipelines d'analyse IRM, les données sont d'abord rééchantillonnées, puis segmentées par des approches (semi-)automatiques. En d'autres termes, la reconstruction et la segmentation de l'image sont effectuées séparément. Nous proposons une méthodologie permettant d'effectuer simultanément la reconstruction haute résolution et la segmentation des données IRM du cerveau néonatal. Notre stratégie s'appuie principalement sur des réseaux antagonistes génératifs (GAN). Nous décrivons et discutons cette architecture et les résultats qu'elle permet d'obtenir.

1 Introduction

Des études à long terme sur les conséquences de la prématurité ont démontré que la majorité des nouveau-nés prématurés peuvent présenter des déficits moteurs, cognitifs et comportementaux importants (Hack et Fanaroff, 2000; Marlow et al., 2005). Au demeurant, notre compréhension de la nature des anomalies cérébrales sous-jacentes à ces séquelles neurologiques reste limitée. Dans ce contexte, l'imagerie par résonance magnétique (IRM) offre des possibilités uniques d'investigation in vivo du cerveau humain. En raison de leur faible résolution et de leur anisotropie, l'analyse de ces images cérébrales néonatales reste difficile. Ainsi, l'amélioration de la résolution des images et la segmentation du cerveau à partir de celles-ci est une condition indispensable pour permettre des analyses morphométriques robustes.

Lorsqu'il s'agit d'images anisotropes à faible résolution, l'une des premières composantes clés d'un pipeline de traitement des données IRM cliniques est l'estimation d'images isotropes, par suréchantillonnage. La super-résolution (SR) (Greenspan, 2008) est une technique qui vise à améliorer la résolution d'une image après son acquisition. Cependant, la SR constitue un problème inverse difficile à résoudre ; en particulier, l'estimation de la texture et des détails est délicate. Dans ce contexte, les techniques d'apprentissage profond supervisé ont permis de

réelles améliorations par rapport aux approches basées sur des modèles. Ainsi, l'application de réseaux neuronaux convolutionnels 3D (CNNs) aboutit à des résultats prometteurs pour les données IRM (Pham et al., 2017; Chen et al., 2018).

Cependant, l'utilisation d'une fonction de coût basée uniquement sur une comparaison point à point de l'image estimée peut conduire à un lissage excessif des images haute résolution obtenues (Johnson et al., 2016). Dans l'optique de rendre ces images plus « réalistes », une composante perceptuelle (Ledig et al., 2017) peut être ajoutée à la fonction de coût. Dans notre cas, cette dernière sera basée sur la fonction de coût d'un GAN (*generative adversarial network*). Il est à noter que l'usage de GAN a été récemment proposé pour la segmentation d'IRM cérébrales (Moeskops et al., 2017).

Parmi les structures cérébrales d'intérêt, les structures minces et en particulier la matière grise corticale, restent difficiles à analyser, principalement à cause de leur forte dégradation dans les images basse résolution. Néanmoins, le cortex est une région d'intérêt, comme le soulignent des travaux récents portant par exemple sur la gyrification (Dubois et al., 2008; Lefèvre et al., 2016; Orasanu et al., 2016), la connectivité corticale (Ball et al., 2013a) ou le développement cortical (Ball et al., 2013b; Yu et al., 2016).

Notre objectif est de proposer une méthodologie dédiée à l'analyse d'images IRM anisotropes basse résolution. En particulier, nous visons à traiter des structures anatomiques complexes, dont le cortex. Pour ce faire, nous proposons une approche basée sur les GAN, nommée SegSRGAN, qui génère à la fois une image super-résolue et une carte de segmentation corticale à partir d'une seule image basse résolution.

2 Super-resolution et segmentation d'image : formulation

2.1 Super-résolution

Le but d'une méthode de SR est d'estimer une image haute résolution (HR) $\mathbf{X} \in \mathbb{R}^m$ à partir d'une image basse résolution (LR) observée $\mathbf{Y} \in \mathbb{R}^n$, avec $m > n$. Un tel problème peut être formulé à l'aide du modèle d'observation linéaire :

$$\mathbf{Y} = H_{\downarrow} B \mathbf{X} + N = \Theta \mathbf{X} + N \tag{1}$$

où $N \in \mathbb{R}^n$ est un bruit additif, $B \in \mathbb{R}^{m \times m}$ est une matrice de dispersion (« flou »), $H_{\downarrow} \in \mathbb{R}^{n \times m}$ est une matrice de décimation (sous-échantillonnage) et $\Theta = H_{\downarrow} B \in \mathbb{R}^{n \times m}$.

Une façon usuelle de résoudre ce problème de SR consiste à définir la matrice Θ^{-1} comme la combinaison d'un opérateur de restauration $F \in \mathbb{R}^{m \times m}$ et d'un opérateur d'interpolation $S^{\uparrow} \in \mathbb{R}^{m \times n}$ qui calcule l'image LR interpolée $\mathbf{Z} \in \mathbb{R}^m$ associée à $\mathbf{Y}$ (i.e. $\mathbf{Z} = S^{\uparrow}\mathbf{Y}$). Dans le contexte d'un apprentissage supervisé, étant donné un ensemble d'images HR $\mathbf{X}_i$ et leurs images LR correspondantes $\mathbf{Y}_i$, l'opérateur de restauration F peut être estimé tel que :

$$\widehat{F} = \arg\min_{F} \sum_{i} d(\mathbf{X}_i - F(\mathbf{Z}_i)) \tag{2}$$

où d peut être, par exemple, une norme ℓ_2, une norme ℓ_1 ou une variante dérivable de la norme ℓ_1 telle que définie par Charbonnier et al. (1997). À l'instar de Lai et al. (2017), notre fonction de coût sera de type « Charbonnier ». Dans (Pham et al., 2017; Chen et al., 2018), il a été démontré que les CNNs 3D pouvaient être utilisés pour estimer avec précision la fonction de restauration $\widehat{F}$ pour les images IRM du cerveau.

2.2 Segmentation

Afin d'équilibrer les contributions de l'image SR et de la segmentation dans la fonction de coût, la segmentation de l'image est considérée comme un problème de régression supervisée :

$$\mathbf{S_X} = R\left(\mathbf{X}\right) \tag{3}$$

où R désigne une fonction non linéaire de l'image interpolée $\widehat{\mathbf{X}}$ vers la carte de segmentation $\mathbf{S_X}$. Comme pour le problème de SR, en supposant que nous avons un ensemble d'images interpolées $\widehat{\mathbf{X}}$ et leurs cartes de segmentation correspondantes $\mathbf{S}_{\mathbf{X}_i}$, une approche générale pour résoudre un tel problème de segmentation consiste à trouver la correspondance R tel que :

$$\widehat{R} = \arg\min_{R} \sum_{i} d(\mathbf{X}_i - R(\mathbf{X}_i)) \tag{4}$$

3 Description de la méthode

Les approches basées sur les GAN consistent à former un premier réseau G, dit générateur, qui estime pour une image d'entrée interpolée donnée, l'image HR et la carte de segmentation correspondante. Un second réseau D, dit discriminateur, est conçu pour différencier les couples d'images HR et de segmentations réels des couples générés d'images SR et de segmentations.

3.1 Fonction de coût

Afin d'éviter de potentiels problèmes de saturation du gradient, que nous pouvons rencontrer avec la fonction de coût traditionnelle [1] « minimax » des GAN, la fonction de coût appelée WGAN-GP (Gulrajani et al., 2017) est utilisée. Ce type de GAN vise à minimiser la distance de Wasserstein entre les deux distributions $\mathbb{P}_r$ et $\mathbb{P}_g$ (resp. données réelles et générées) :

$$W(\mathbb{P}_r, \mathbb{P}_g) = \inf_{\gamma \in \Pi(\mathbb{P}_r, \mathbb{P}_g)} \mathbb{E}_{(x,y)\sim\gamma}\left[||x-y||\right] \tag{5}$$

$$= \sup_{||f||_L \leq 1} \mathbb{E}_{x\sim\mathbb{P}_r}[f(x)] - \mathbb{E}_{x\sim\mathbb{P}_g}[f(x)] \tag{6}$$

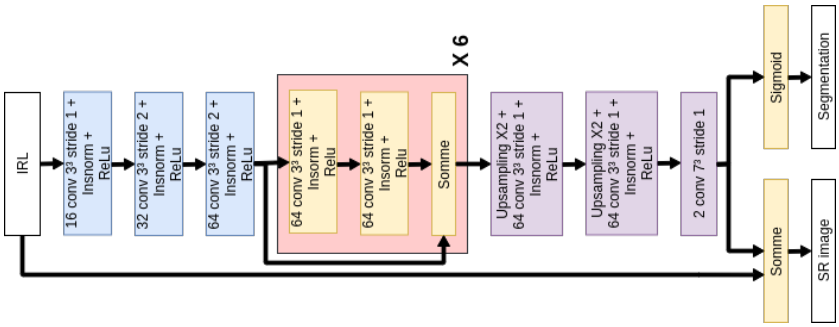
où, $\Pi(\mathbb{P}_r, \mathbb{P}_g)$ est l'ensemble de toutes les distributions dont les marginales sont respectivement $\mathbb{P}_r$ et $\mathbb{P}_g$, et le supremum est calculé sur toutes les fonctions 1-lipschitziennes f.

Dans ce GAN, le discriminateur apprend la fonction paramétrée f tandis que le générateur vise à minimiser cette distance. Ainsi, la partie antagoniste de la fonction de coût est :

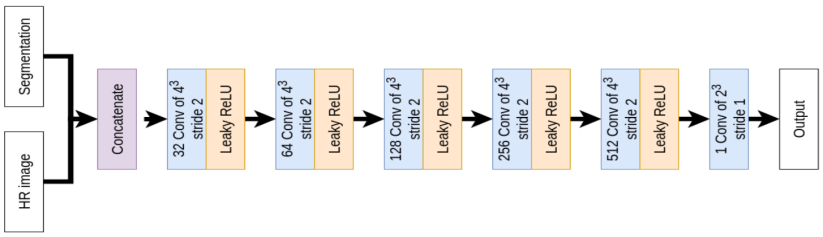
$$\mathcal{L}_{adv} = \mathbb{E}_{\mathbf{X}\sim\mathbb{P}_\mathbf{X}, \mathbf{S_X}\sim\mathbb{P}_{\mathbf{S_X}}}[D((\mathbf{X}, \mathbf{S_X}))] - \mathbb{E}_{\mathbf{Z}\sim\mathbb{P}_\mathbf{Z}}[D(G\left(\mathbf{Z}\right))] \tag{7}$$

où $\mathbf{X}$ et $\mathbf{S_X}$ sont les vraies image HR et carte de segmentation, respectivement, D est le discriminateur, G le générateur et $\mathbf{Z}$ l'image interpolée. C'est à travers ce terme que sera exprimé le jeu entre le générateur (minimiser $\mathcal{L}_{adv}$) et le discriminateur (maximiser $\mathcal{L}_{adv}$).

1. C'est-à-dire en résolvant $\min_G \max_D V(D,G) = \mathbb{E}_{x\in p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\in p_z(z)}[\log(1-D(G(z)))]$.



(a) Architecture du générateur.



(b) Architecture du discriminateur.

FIG. 1 – *Architecture des réseaux de neurones. (a) Générateur. (b) Discriminateur.*

Rappelons que le discriminateur vise à estimer une fonction 1-lipschitzienne ; un terme de régularisation sur la valeur de son gradient est donc ajouté. Finalement, la fonction de coût (à minimiser) du discriminateur est :

$$\mathcal{L}_{dis} = \lambda_{gp}\mathbb{E}_{\widehat{\mathbf{XS}}}[(\|(\nabla_{\widehat{\mathbf{XS}}}D(\widehat{\mathbf{XS}})\|_2 - 1)^2] - \mathcal{L}_{adv} \tag{8}$$
$$\text{avec } \widehat{\mathbf{XS}} = (1-\varepsilon)(\mathbf{X}, \mathbf{S_X}) + \varepsilon G(\mathbf{Z}) \text{ et } \varepsilon \sim U[0,1]$$

où $\lambda_{gp} > 0$ et ∇ désignent respectivement le coefficient de pénalité du gradient et l'opérateur de gradient.

La fonction de coût du générateur est construite en ajoutant un terme de comparaison point par point ρ (Charbonnier et al., 1997) des images cibles en sortie et des images estimées :

$$\mathcal{L}_{gen} = \lambda_{adv}\mathcal{L}_{adv} + \mathbb{E}_{\mathbf{X}\sim\mathbb{P}_{\mathbf{X}},\mathbf{S_X}\sim\mathbb{P}_{\mathbf{S_X}}}\left[\rho\left((\mathbf{X}, \mathbf{S_X}) - G(\mathbf{Z})\right)\right] \tag{9}$$
$$\text{avec } \rho(x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(x_i^2 + \nu^2)} \text{ et } \nu = 10^{-3}$$

3.2 Architecture du réseau

Architecture du générateur Le réseau du générateur est un réseau basé sur la convolution, avec des blocs résiduels. Il prend en entrée l'image LR interpolée. Il est composé de 18 couches de convolution : trois pour la partie encodage, deux fois six pour la partie résiduelle et trois

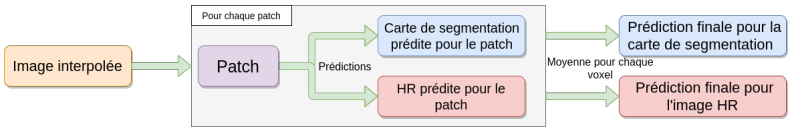


FIG. 2 – *Fabrication de la prédiction durant la phase de test.*

pour la partie décodage (voir Figure 1(a)). Afin d'améliorer les performances de la procédure d'entraînement, des couches de normalisation d'instance sont utilisées sur le résultat de chaque convolution, avant application de la fonction d'activation.

Lors de l'encodage, la taille des cartes de caractéristiques est divisée par 4 ; leur nombre passe successivement de 1 (entrée) à 64. La partie résiduelle est composée de six blocs. Chaque bloc transforme l'entrée grâce à deux couches de convolution puis somme le résultat avec l'entrée. Enfin, la partie décodeur double la taille de chaque carte de caractéristiques puis les transforme par le biais d'une couche de convolution. La dernière couche de convolution produit deux images 3D : la première sera transformée en carte de probabilité de classe (en utilisant une activation sigmoïde) ; la seconde sera additionnée avec l'image interpolée originale.

Architecture du discriminateur Le réseau du discriminateur est entièrement convolutif. Il prend en entrée une image HR et une carte de segmentation. Le discriminateur contient cinq couches de convolution avec un nombre croissant de noyaux, augmentant par un facteur 2 de 32 à 512 noyaux (voir Figure 1(b)). Chaque convolution (sauf la dernière) est effectuée avec un pas de deux, ce qui permet de réduire successivement la taille des cartes de caractéristiques par deux jusqu'à arriver à une représentation sous forme d'un scalaire (lorsque l'entrée est de taille 64^3).

3.3 Entraînement et évaluation

Avant chaque application du réseau, chaque image LR est normalisée (division par la plus grande intensité de l'image) et interpolée via des splines cubiques. Ces deux étapes sont effectuées lors de l'entraînement et de l'évaluation de la méthode. Avant d'appliquer le réseau, les images sont également divisées en patchs d'une taille donnée.

Entraînement Les données d'entraînement sont constituées de patchs d'images interpolées (entrée du réseau) et de patchs de cartes de segmentation et d'images HR. À partir de ces données d'entraînement, le discriminateur et le générateur sont entraînés de la manière suivante. Pour chaque mise à jour des poids du générateur, les poids du discriminateur sont mis à jour cinq fois. La méthode d'optimisation choisie est Adam avec les mêmes paramètres que ceux de l'article original (Kingma et Ba, 2014). La taille de batch est de 32 patchs qui sont de taille 64^3, créés avec un décalage de 20 entre chacun. Enfin λ_{gp} et λ_{adv} sont respectivement fixés à 100 et 0.001. Pour chaque entraînement effectué, le nombre maximal d'époques est fixé à 200 et les poids finaux (parmi les poids de chaque époque) sont ceux qui maximisent la performance sur un ensemble de données de test.

Evaluation La figure 2 illustre le processus d'évaluation de la méthode.

4 Resultats

4.1 Données

Nous travaillons sur la base de données IRM dHCP [2]. Cette base contient 40 IRM cérébrales de nouveau-nés imagés à terme. En plus de ces acquisitions, la base dHCP fournit, pour chaque sujet, des cartes de segmentation du cerveau. Nous nous intéressons ici à la carte de segmentation (binaire) du cortex cérébral.

4.2 Pré-traitement : génération d'images LR

Toutes les images de la base dHCP sont isotropes et de résolution suffisante ($0.5 \times 0.5 \times 0.5$ mm^3) pour être considérées de haute résolution. Par conséquent, afin de former un modèle de super-résolution, nous devons déterminer des images LR correspondantes pour ces couples images HR / cartes de segmentation. Les images LR sont générées en utilisant le modèle proposé par Greenspan (2008). En notant X l'image HR et X_{LR} l'image LR associée :

$$X_{LR} = H_{\downarrow} B X \tag{10}$$

où B est une matrice floue et $H_{\downarrow}$ est une décimation par sous-échantillonnage. En particulier, nous considérons un filtre gaussien B avec un écart-type :

$$\sigma = \frac{\text{res}}{2\sqrt{2\log 2}} \tag{11}$$

où res est la résolution de l'image LR, ici fixé à $0.5 \times 0.5 \times 3$ mm^3 (avec une forte anisotropie destinée à prendre en compte les conditions réelles d'acquisition des données cliniques).

4.3 Métriques utilisées

Segmentation Afin d'évaluer la qualité des résultats de segmentation, nous considérons le score de Dice, qui est une mesure standard (comprise entre $[0, 1]$) pour évaluer une segmentation S par rapport à une vérité-terrain G :

$$\text{Dice}(S, G) = \frac{2|S \cap G|}{|S| + |G|} \tag{12}$$

En particulier, plus le Dice est proche de 1, plus S et G sont semblables.

Super-résolution Mesurer la performance des algorithmes SR est moins simple. En effet, pour mesurer la similitude d'aspect visuel entre deux images, une distance entre l'intensité des voxels SR et HR peut ne pas être suffisante. La performance de la reconstruction SR est alors mesurée par deux indices : le PSNR et le SSIM (Wang et al., 2004), définis comme :

$$\text{PSNR}(X, Y) = 10 \log_{10} \frac{\left(\max_i X_i\right)^2}{\frac{1}{|X|}\sum_i |X_i - Y_i|^2} \tag{13}$$

2. http://www.developingconnectome.org

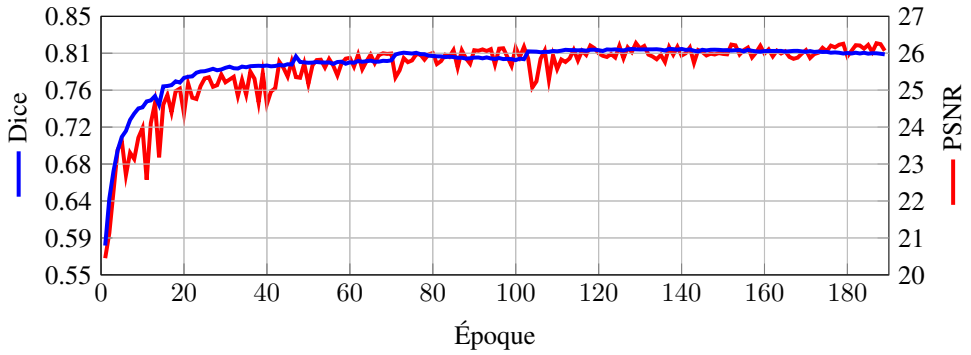


FIG. 3 – *Dice et PSNR moyens à chaque époque, mesurés sur les huit images test de dHCP.*

où X_i et Y_i sont les valeurs de X et Y au point i, respectivement, et :

$$\text{SSIM}(X,Y) = \frac{(2\mu_X\mu_Y + c_1) + (2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \tag{14}$$

où μ_X (resp. μ_Y) est la moyenne des valeurs X (resp. Y), σ_X (resp. σ_Y) est l'écart-type des valeurs X (resp. Y), σ_{XY} est la covariance entre les valeurs X et Y, et c_1, c_2 sont des stabilisateurs numériques liés quadratiquement à la dynamique de l'image.

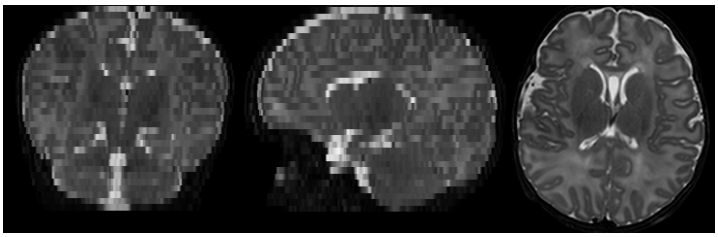
Plus la valeur de ces indices est élevée, meilleure est la similitude entre les deux images.

4.4 Convergence (entraînement)

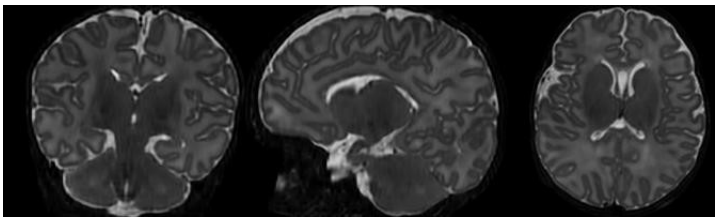
Tout d'abord, nous observons l'évolution des scores Dice et PSNR, tout au long de l'entraînement. Les patchs sur les image tests sont créés de la même façon que ceux de la base d'entraînement. Le PSNR et le Dice sont ensuite calculés pour chaque patch. Les PSNR et Dice finaux sont obtenus en établissant la moyenne des valeurs calculées dans chaque patch. Les résultats sont représentés dans la Figure 3, qui fournit l'évolution des scores de Dice et de PSNR à la fin de chaque époque. Le Dice initial a une valeur très faible, proche de $0,5$, et augmente ensuite jusqu'à $0,8$. Il semble converger aux alentours de l'époque 100. Le PSNR converge également, mais de façon plus bruitée. Cependant, la taille des pics diminue alors que le score tend à se stabiliser, suivant le même comportement que le Dice.

4.5 Résultats obtenus sur une base test

Les résultats présentés dans cette partie ont été obtenus avec des patchs de taille 128^3 voxels et un décalage de 30 voxels entre patchs successifs. La figure 4 fournit un exemple de résultat obtenu sur une image de la base test de dHCP. Les résultats quantitatifs présentés ci-après ont été calculés à partir des 8 images de dHCP utilisées comme base de données test. Le tableau 1 résume les Dices de notre méthode (SegSRGAN), de la méthode IMAPA (Tor-Díez et al., 2018) et de DrawEM (Makropoulos et al., 2014). Comme dans un contexte clinique typique, ces méthodes ont été appliquées sur des images interpolées (en utilisant une



(a) Image LR (initiale)



(b) Image SR (reconstruite)

(c) Carte de segmentation du cortex

FIG. 4 – *Image LR de la base test de dHCP (a) et les résultat obtenus sur celle-ci (b) et (c)*

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

TAB. 1 – *Moyennes (et écarts-type) du Dice pour les cartes de segmentation estimées, calculées sur huit images test.*

spline cubique). On observe que, quantitativement, l'approche proposée conduit aux meilleurs résultats de segmentation corticale avec une amélioration significative par rapport aux deux autres méthodes. De plus, comme mentionné par Tor-Díez et al. (2018), l'utilisation de IMAPA appliquée aux images originales HR de dHCP conduit à un score moyen de 0,887 (écart-type de 0,011). Enfin, le résultat obtenu sur les images interpolées ne diminue que de 3% par rapport à IMAPA appliquée sur les images HR.

Le tableau 2 résume le PSNR et le SSIM de notre méthode versus une interpolation par spline cubique. On peut observer que les deux scores de qualité pour la reconstruction de l'image SR traduisent de meilleurs résultats avec SegSRGAN qu'avec l'interpolation par spline cubique, qui constitue une base standard de comparaison et est l'entrée du réseau.

	SegSRGAN	Interpolation par spline cubique
PSNR	26.96	24.22
SSIM	0.73	0.63

TAB. 2 – *Valeurs moyennes du PSNR et du SSIM, calculées sur huit images test.*

5 Conclusion

Dans cet article, nous avons proposé une nouvelle solution méthodologique pour effectuer la reconstruction et la segmentation d'images IRM 3D complexes. Les résultats obtenus tendent à prouver la pertinence de notre approche, avec des résultats satisfaisants tant en termes de reconstruction SR que de segmentation. Cependant, il est important de garder en tête que la méthode a ici été entraînée et évaluée pour une reconstruction super-résolution d'une résolution axiale initiale fixée (3 mm) à une autre (0.5 mm). Cependant, prendre en compte une variabilité de la résolution initiale peut également être intéressant d'un point de vue clinique. Cet aspect, ainsi que la segmentation d'autres zones du cerveau, constituent les travaux à venir.

Remerciements Ces travaux ont été soutenus par l'Agence Nationale de la Recherche (contrat ANR-15-CE23-0009) ; l'INSERM et l'Institut Mines Télécom Atlantique (Chaire "Imagerie médicale en thérapie interventionnelle") ; la Fondation pour la Recherche Médicale (contrat DIC2016123636453) ; et l'*American Memorial Hospital Foundation*. Nous remercions NVIDIA pour la fourniture de la carte GPU Titan Xp utilisée dans le cadre de ces recherches.

Références

Ball, G., J. P. Boardman, P. Aljabar, A. Pandit, T. Arichi, N. Merchant, D. Rueckert, A. D. Edwards, et S. J. Counsell (2013a). The influence of preterm birth on the developing thalamocortical connectome. *Cortex 49*(6), 1711–1721.

Ball, G., L. Srinivasan, P. Aljabar, S. J. Counsell, G. Durighel, J. V. Hajnal, M. A. Rutherford, et A. D. Edwards (2013b). Development of cortical microstructure in the preterm human brain. *Proceedings of the National Academy of Sciences of the United States of America 110*(23), 9541–9546.

Charbonnier, P., L. Blanc-Féraud, G. Aubert, et M. Barlaud (1997). Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing 6*(2), 298–311.

Chen, Y., Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, et D. Li (2018). Brain MRI super resolution using 3D deep densely connected neural networks. In *ISBI*, pp. 739–742.

Dubois, J., M. Benders, A. Cachia, F. Lazeyras, R. Ha-Vinh Leuchter, S. V. Sizonenko, C. Borradori-Tolsa, J. F. Mangin, et P. S. Hüppi (2008). Mapping the early cortical folding process in the preterm newborn brain. *Cerebral Cortex 18*(6), 1444–1454.

Greenspan, H. (2008). Super-resolution in medical imaging. *The Computer Journal 52*(1), 43–63.

Gulrajani, I., F. Ahmed, M. Arjovsky, V. Dumoulin, et A. C. Courville (2017). Improved training of Wasserstein GANs. In *NIPS*, pp. 5769–5779.

Hack, M. et A. A. Fanaroff (2000). Outcomes of children of extremely low birthweight and gestational age in the 1990s. *Seminars in Neonatology 5*, 89–106.

Johnson, J., A. Alahi, et L. Fei-Fei (2016). Perceptual losses for real-time style transfer and super-resolution. In *ECCV*, pp. 694–711.

Kingma, D. P. et J. Ba (2014). Adam: A method for stochastic optimization. *CoRR abs/1412.6980*.

Lai, W.-S., J.-B. Huang, N. Ahuja, et M.-H. Yang (2017). Deep Laplacian pyramid networks for fast and accurate super-resolution. *CoRR abs/1704.03915*.

Ledig, C., L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et W. Shi (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*, pp. 105–114.

Lefèvre, J., D. Germanaud, J. Dubois, F. Rousseau, I. de Macedo Santos, H. Angleys, J.-F. Mangin, P. S. Hüppi, N. Girard, et F. De Guio (2016). Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns? *Cerebral Cortex 26*(7), 3023–3035.

Makropoulos, A., I. S. Gousias, C. Ledig, P. Aljabar, A. Serag, J. V. Hajnal, A. D. Edwards, S. J. Counsell, et D. Rueckert (2014). Automatic whole brain MRI segmentation of the developing neonatal brain. *IEEE Transactions on Medical Imaging 33*(9), 1818–1831.

Marlow, N., D. Wolke, M. A. Bracewell, M. Samara, et E. S. Group (2005). Neurologic and developmental disability at six years of age after extremely preterm birth. *The New England Journal of Medicine 352*, 9–19.

Moeskops, P., M. Veta, M. W. Lafarge, K. A. J. Eppenhof, et J. P. W. Pluim (2017). Adversarial training and dilated convolutions for brain MRI segmentation. In *DLMIA and ML-CDS*, pp. 56–64.

Orasanu, E., A. Melbourne, M. J. Cardoso, H. Lomabert, G. S. Kendall, N. J. Robertson, N. Marlow, et S. Ourselin (2016). Cortical folding of the preterm brain: A longitudinal analysis of extremely preterm born neonates using spectral matching. *Brain and Behavior 6*(8), e00488.

Pham, C.-H., A. Ducournau, R. Fablet, et F. Rousseau (2017). Brain MRI super-resolution using deep 3D convolutional networks. In *ISBI*, pp. 197–200.

Tor-Díez, C., N. Passat, I. Bloch, S. Faisan, N. Bednarek, et F. Rousseau (2018). An iterative multi-atlas patch-based approach for cortex segmentation from neonatal MRI. *Computerized Medical Imaging and Graphics 70*, 73–82.

Wang, Z., A. C. Bovik, H. R. Sheikh, et E. P. Simoncelli (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing 13*(4), 600–612.

Yu, Q., A. Ouyang, L. Chalak, T. Jeon, J. Chia, V. Mishra, M. Sivarajan, G. Jackson, N. Rollins, S. Liu, et H. Huang (2016). Structural development of Human fetal and preterm brain cortical plate based on population-averaged templates. *Cerebral Cortex 26*(11), 4381–4391.