



HAL
open science

U-Net 3D par patches et régularisation spatiale avec CRF pour l'extraction de cerveau sur IRM de porcelet

Patty Coupeau, Jean-Baptiste Fasquel, Edouard Mazerand, Claudia
Montero-Menei, Philippe Menei, Mickael Dinomais

► To cite this version:

Patty Coupeau, Jean-Baptiste Fasquel, Edouard Mazerand, Claudia Montero-Menei, Philippe Menei, et al.. U-Net 3D par patches et régularisation spatiale avec CRF pour l'extraction de cerveau sur IRM de porcelet. ORASIS 2021, Centre National de la Recherche Scientifique [CNRS], Sep 2021, Saint Ferréol, France. hal-03339669

HAL Id: hal-03339669

<https://hal.science/hal-03339669>

Submitted on 9 Sep 2021

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.

U-Net 3D par patches et régularisation spatiale avec CRF pour l'extraction de cerveau sur IRM de porcelet

Patch-based 3D U-Net and CRF-based spatial regularization for piglet MRI brain extraction

P. Coupeau¹ J.-B. Fasquel¹ E. Mazerand² C. Montero-Menei² P.Menei^{2,3} M. Dinomais^{1,4}

¹ LARIS, SFR MATHSTIC, Université d'Angers, France

² CRCINA, INSERM, Université de Nantes, Université d'Angers, France

³ Département de neurochirurgie, Centre Hospitalier Universitaire d'Angers, France

⁴ Département de médecine physique et de réadaptation, Centre Hospitalier Universitaire d'Angers, France

patty.coupeau@etud.univ-angers.fr

Résumé

Du fait de ses similarités morphologiques avec le cerveau de l'enfant, le porcelet est un modèle prisé. Il est notamment utilisé dans l'étude de la paralysie cérébrale afin de quantifier l'impact de la lésion précoce sur le neurodéveloppement en particulier sur des IRM que nous considérons dans ces travaux. Pour disposer d'une évaluation objective, il est nécessaire de développer des algorithmes automatiques de segmentation des différentes structures et tissus. L'étape préliminaire cruciale consiste à extraire le cerveau. Nous proposons ici une méthode d'extraction entièrement automatique du cerveau du porcelet combinant un réseau U-Net 3D par patches et l'utilisation de champs aléatoires conditionnels (CRF) pour la régularisation spatiale finale. Notre méthode est entraînée et évaluée à partir de 15 IRM T1 de porcelets âgés de 15 jours et comparée à l'outil BET, standard pour le cerveau humain. L'influence de la taille et de la distribution des patches utilisés lors de l'apprentissage profond est étudiée tout comme les bénéfices de la régularisation spatiale. Les premiers résultats montrent que notre approche est prometteuse en terme de DICE moyen (0.954) et de distance d'Hausdorff (7.73), surpassant les performances de BET (DICE de 0.754 et distance d'Hausdorff de 24.31).

Mots Clef

Extraction de cerveau, U-Net, Patches, CRF, IRM du porcelet, Paralysie cérébrale.

Abstract

Because of its morphological similarities with the child brain, piglet is a popular model. It is used in particular in the study of cerebral palsy in order to quantify the impact of the early lesion on neurodevelopment from MRI as

we consider in this work. For an objective evaluation, it is necessary to develop automatic algorithms for the segmentation of the different structures and tissues. The crucial preliminary step consists in extracting the brain. We propose here a fully automatic piglet brain extraction method combining a 3D patch-based U-Net and conditional random fields (CRF) for final spatial regularization. Our method is trained and evaluated from 15 T1 MRI of 15-day-old piglets and compared to the BET tool, standard for the human brain. The influence of the size and distribution of the patches used during deep learning is studied as well as the benefits of spatial regularization. The first results show that our approach is promising in terms of average DICE (0.954) and Hausdorff distance (7.73), outperforming BET (DICE of 0.754 and Hausdorff distance of 24.31).

Keywords

Brain extraction, U-Net, Patches, CRF, Piglet MRI, Cerebral palsy.

1 Introduction

L'utilisation du porcelet comme modèle animal pour l'étude du développement cérébral de l'enfant s'accroît [13]. En effet, les caractéristiques anatomiques du cerveau du porcelet, incluant la distribution de matière blanche et de matière grise, sont très similaires à celles du nouveau-né [5]. Ainsi, le modèle du porcelet permet de mieux comprendre le processus de maturation du cerveau et d'anticiper de futurs troubles neurodéveloppementaux [9]. Depuis peu, des recherches relatives à la paralysie cérébrale tentent de caractériser l'impact de la lésion ischémique (à l'origine du trouble moteur) sur le neurodéveloppement du nouveau-né. Pour cela, elles s'appuient sur le modèle du porcelet reproduisant les lésions[12].

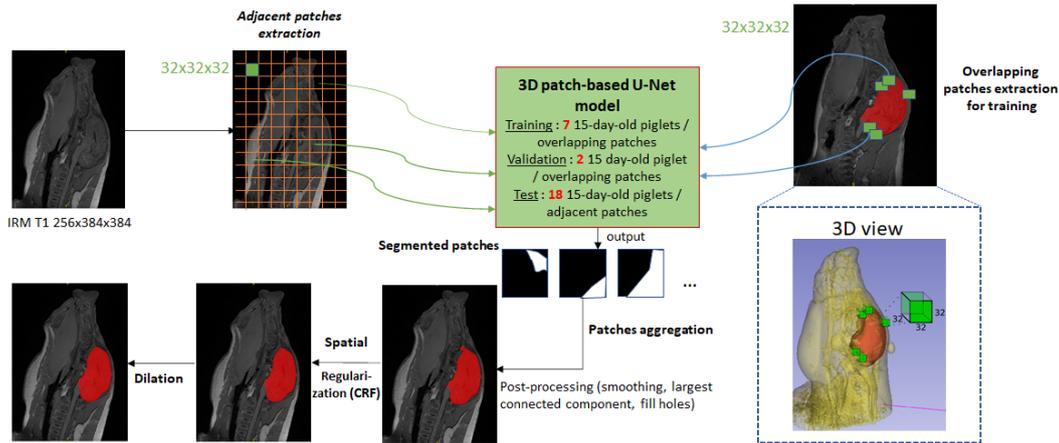


FIGURE 1 – Vue d’ensemble de la méthode d’extraction automatique du cerveau de porcelet sur IRM T1 proposée

Dans l’objectif de quantifier à l’échelle macroscopique les conséquences de la lésion sur le développement cérébral, l’imagerie par résonance magnétique (IRM) est un outil adapté [13]. Néanmoins, il est nécessaire de développer des algorithmes d’analyse d’images IRM tels que la segmentation des structures et tissus cérébraux ou des analyses morphométriques qui permettront un suivi longitudinal quantitatif du développement cérébral.

Bien que plusieurs travaux aient été réalisés sur le cerveau de l’enfant [3], il en existe peu relatif au porcelet. Seules deux études s’attellent à la création d’un atlas du cerveau du porcelet apparaissent dans la littérature [13, 6]. Leur objectif est d’automatiser la segmentation de structures cérébrales grâce à une approche par atlas comme cela se fait beaucoup sur le cerveau humain. Cependant, une étape préliminaire à ces analyses volumétriques entraîne des difficultés : l’extraction du cerveau du porcelet.

Actuellement, le masque du cerveau est dessiné entièrement à la main par un opérateur à l’aide de logiciels dédiés (ITKSnap¹, MITK²). Il s’agit d’une opération longue (environ 1h30 par IRM) et inappropriée à la construction de grandes bases de données de cerveaux de porcelets segmentés. Du fait de la petite taille du cerveau porcine et de l’abondance de muscles et de graisses [6], les techniques habituellement considérées pour l’extraction de cerveau humain tel que le célèbre outil BET (Brain Extraction Tool) [16] fonctionnent mal.

C’est pourquoi, il est nécessaire de trouver une alternative à ces algorithmes afin d’automatiser l’extraction du cerveau du porcelet sur IRM. Ce travail propose une solution en combinant un réseau de neurones profond de segmentation sémantique (U-Net [27]) et une régularisation spatiale finale avec champs aléatoires conditionnels (CRF) [14]. La combinaison des deux outils constitue une approche originale capable de corriger la segmentation produite par le réseau de neurones en exploitant l’information de voisinage

à travers du CRF. Par ailleurs, afin d’associer contexte générale et information locale durant l’apprentissage profond, nous proposons une approche par patches 3D, comme récemment proposé pour l’analyse d’IRM cérébrale chez l’homme [8, 17], ceci constituant la seconde originalité de cette contribution.

La méthode proposée est détaillée dans le Section 2. Les résultats des premières expériences illustrant le potentiel de celle-ci sont présentés dans la Section 3 et discutés dans la Section 4. La Section 5 conclut cet article.

2 Méthode

La Fig.1 donne un aperçu de la méthode proposée utilisant un exemple d’image IRM T1 de porcelet considérée dans l’étude. A partir d’une base de données d’entraînement annotée, le réseau de neurones est entraîné à extraire le cerveau du porcelet de l’IRM.

Le modèle considéré est un réseau U-Net recevant en entrée des patches 3D de dimension 32^3 voxels de l’IRM T1 et produisant en sortie des patches de même taille contenant la segmentation du cerveau. Ainsi, chaque IRM de porcelet est découpée en patches de cette dimension qui vont passer par le réseau U-Net afin d’être segmentés avant d’être agrégés pour reconstituer l’image du masque.

Après quelques opérations élémentaires indiquées dans la figure 1 (conservation de la composante connexe principale, lissage de celle-ci et élimination des éventuelles cavités), une régularisation spatiale est mise en oeuvre en s’appuyant sur des champs aléatoires conditionnels [14]. Une dilatation morphologique est appliquée afin de garantir l’obtention de l’intégralité du cerveau dans la segmentation finale (marge de précaution pour les étapes ultérieures de segmentation des structures cérébrales internes).

Nous détaillons par la suite l’architecture du réseau considéré (Section 2.1) puis la stratégie d’extraction de patches choisie pour l’entraînement et l’inférence (Section 2.2).

1. <http://www.itknap.org/pmwiki/pmwiki.php>
 2. [https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_\(MITK\)](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK))

2.1 Architecture du réseau U-Net

Le réseau U-Net constitue une architecture standard pour la segmentation d'images médicales [27]. Il constitue une extension des réseaux entièrement convolutifs (FCN) [18]. L'architecture FCN se compose d'un chemin dit d'encodage où l'entrée est sous-échantillonnée à l'aide d'opérations successives de convolution et de maxpooling. La carte de caractéristiques qui en découle est introduite dans une carte d'activation pour prédire la classe de chacun des pixels. Le réseau U-Net (Fig.2) possède également ce chemin d'encodage mais ce dernier est suivi d'un autre chemin dit de décodage presque similaire au chemin de l'encodeur, donnant ainsi l'architecture en forme de U. Lors du décodage, les opérations de pooling sont remplacées par des opérations de sur-échantillonnage.

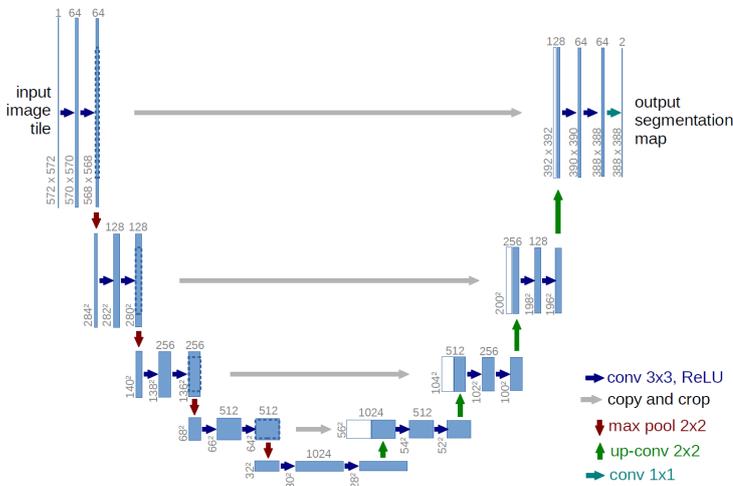


FIGURE 2 – Architecture du réseau U-Net standard (image extraite de [27])

Les paramètres du réseau U-Net implémenté seront présentés dans la Section 3.

2.2 Approche par patches 3D

L'architecture U-net a montré de très bonnes performances pour la segmentation d'images [8]. Cependant, plus la taille de l'image d'entrée est grande, plus la capacité de mémoire GPU nécessaire pour l'entraînement est importante. De plus, lorsque l'architecture prend l'IRM entière ou une coupe de l'IRM comme entrée, le modèle a tendance à manquer des détails dans certaines régions de l'image [8]. Pour surmonter ce problème, nous proposons une approche par patches qui présente l'avantage d'être plus précise (le réseau peut se focaliser sur des informations locales au patch) tout en nécessitant moins de mémoire pour l'entraînement et l'inférence, réduisant ainsi les temps de calcul.

Dans le but d'enrichir l'information de contexte contenue dans le patch, une bonne solution consiste à utiliser des patches 3D qui bénéficient du contexte dans les 3 directions (axiale, coronale et sagittale) [17].

Nous présentons par la suite la stratégie d'extraction de patches 3D pour les phases d'entraînement et d'inférence.

Entraînement. Pour l'entraînement, nous avons décidé de travailler avec des patches pouvant se chevaucher (Fig.1) car cette propriété aide le réseau convolutif à voir à la fois le contexte local et global [17]. En effet, la prédiction de chaque voxel provient de la décision de l'ensemble des patches contenant ce voxel (Eq.2).

Pour chaque voxel de chaque patch, notre modèle U-Net fournit une probabilité d'appartenance au cerveau. Les paramètres du réseau sont adaptés afin de minimiser la fonction de coût cross-entropy binaire définie par l'Eq.1 [25].

$$L(Y, \hat{y}) = -(Y \log(\hat{y}) + (1 - Y) \log(1 - \hat{y})) \quad (1)$$

où Y_i est la vraie classe du voxel x_i (0 pour le fond, 1 pour le cerveau) et \hat{y}_i est la probabilité d'appartenance au cerveau du voxel x_i . En cas de recouvrement de patches, \hat{y}_i est moyennée sur l'ensemble des patches concernés selon la relation :

$$\hat{y}_i = \frac{\sum_{p \in P_i} \text{Softmax}_p(y_i)}{|P_i|} \quad (2)$$

avec P_i l'ensemble des patches contenant x_i et $\text{Softmax}_p(y_i)$ la prédiction d'appartenance au cerveau du voxel x_i dans le patch p définie comme tel :

$$\text{Softmax}_p(y_i) = \frac{\exp^{y_i}}{\exp^{(1-y_i)} + \exp^{y_i}} \quad (3)$$

y_i signifiant classer le voxel x_i comme cerveau et $1 - y_i$ comme appartenant au fond.

Par ailleurs, la région du cerveau occupant un espace relativement petit de l'IRM (problématique de la représentativité des classes soulignées dans [17]), nous avons imposé la répartition suivante des patches d'entraînement (Fig.3) :

- 85% contiennent du cerveau
- 15% contiennent des parties de la tête externes au cerveau

Dans le but d'apprendre avec davantage de précision les limites du cerveau, 50% des patches du cerveau sont localisés en bordure de celui-ci.

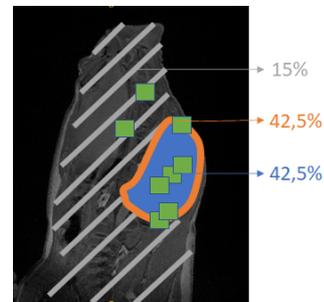


FIGURE 3 – Répartition des patches d'entraînement

La sélection des patches de chaque région s'est faite selon la stratégie suivante :

- **Patchs externes au cerveau** : tirage aléatoire de patchs ne contenant aucun voxel du cerveau, en ignorant la région du fond dont l'intensité est nulle (i.e. extérieure à la tête du porcelet).
- **Patchs internes au cerveau** : tirage aléatoire de distribution uniforme de patchs dont le voxel central se situe dans un rayon de 30 voxels du barycentre du cerveau.
- **Patchs des contours** : tirage aléatoire de distribution uniforme de patchs dont le voxel central se situe sur le contour du cerveau.

Inférence. Pour l'inférence, nous avons privilégié l'utilisation de patchs adjacents afin de simplifier les calculs. Beaucoup d'études ont montré l'efficacité de la technique des patchs adjacents [8, 10] pour la segmentation (Fig.1).

Chaque IRM de test est divisée en patchs adjacents 3D. Pour chaque voxel x_i de chaque patch p , le réseau U-Net retourne un vecteur de probabilité v_i d'appartenance à chaque classe (fond ou cerveau). Chaque voxel est attribué à la classe de plus grande probabilité selon l'Eq.4 :

$$\hat{y}_i = \operatorname{argmax} v_i \quad (4)$$

Les patchs adjacents 3D sont ensuite agrégés afin de reformer l'image entière.

3 Expériences

Nous présentons dans cette section les données utilisées, le protocole d'évaluation ainsi que les résultats de nos expériences.

3.1 Base de données

Notre méthode a été évaluée à partir de 15 images IRM T1 3D de 15 porcelets différents (14 mâles et 1 femelle) âgés de 15 (± 3) jours. Le poids moyen et le volume cérébral moyen des porcelets considérés à deux semaines étaient de 4.29 (± 1.96) kg et de 51.31 (± 5.26) cm^3 respectivement. Les images 3D T1 acquises sont de taille 256x384x384 voxels avec une résolution spatiale de 0.5x0.5x0.5 mm^3 .

3.2 Protocole d'évaluation

L'ensemble des expériences ont été réalisées dans un environnement Python sur Windows 64 bits avec un processeur Intel Core i7 @ 2.70 GHz avec 32Go de RAM et un processeur graphique Nvidia Quadro RTX 3000.

Pour implémenter l'architecture U-Net et réaliser les expériences d'apprentissage profond, nous avons utilisé la librairie PyTorch [22]. Les paramètres détaillés du réseau considéré sont donnés par la Table 1. La régularisation spatiale a été réalisée à l'aide de la librairie *pydensecrf*³.

Nous avons découpé notre base de données de la façon suivante : 7 porcelets (et donc 7 IRM) sont utilisés pour l'entraînement, 1 porcelet pour la validation et 7 porcelets

Paramètres	Overlapping Patch-based 3D U-Net
Taille d'entrée	Patchs de taille 32^3
Convolutions	3^3 convolutions (pas = 1, padding)
Fonction d'activation	ReLU
maxPooling	2^3 maxPooling (pas = 2)
Sortie	Softmax
Nb d'opérations	11 conv., 2 maxPooling, 2 up-convolution

TABLE 1 – Résumé des paramètres de l'architecture U-Net 3D avec patchs proposée

pour la phase de test. Le modèle a été entraîné avec "Stochastic Gradient Descent (SGD)" sur 75 epochs à partir de mini-batchs de taille 8. Une stratégie de réduction du taux d'apprentissage par plateau a été exploitée avec un taux d'apprentissage initial $Lr_0 = 0.1$ et un facteur de réduction $\sigma = 0.01$. La fonction de coût cross-entropy binaire mentionnée dans la Section 2.2 (Eq.1) a été utilisée.

Pour entraîner le réseau U-Net et évaluer la qualité des segmentations obtenues, nous avons considéré comme référence une segmentation manuelle du cerveau réalisée avec le logiciel ITKSnap. Pour l'évaluation quantitative des segmentations, nous avons calculé le DICE [19] défini par Eq.5 calculant la similarité entre la vérité terrain I et la segmentation prédite I' :

$$DICE(I, I') = \frac{2|I \cap I'|}{|I| + |I'|} \quad (5)$$

où $|\cdot|$ représente la cardinalité de l'ensemble.

Nous avons également décomposé le DICE en précision et rappel pour certaines expériences. En effet, dans l'objectif de segmenter par la suite les structures et tissus cérébraux, le critère principal consiste à ne pas perdre de partie du cerveau. Pour cela, nous veillons à maximiser le rappel tout en assurant une précision acceptable.

Par ailleurs, la distance d'Hausdorff (HD) [20] et la distance de surface symétrique moyenne (ASSD) [15] entre la vérité terrain et le masque prédit ont été évaluées. La distance d'Hausdorff entre deux ensembles I et I' est la distance maximale entre un point de I et le point le plus proche de I' . Elle est définie par :

$$HD(I, I') = \max(\max_{a \in I} \min_{b \in I'} |b - a|, \max_{b \in I'} \min_{a \in I} |a - b|) \quad (6)$$

La distance de surface symétrique moyenne entre deux ensembles I et I' est la moyenne de toutes les distances entre les points de I et le point le plus proche de I' et vice versa (de I' vers I) comme indiqué par l'Eq.7. La distance d'un voxel a à la surface de I' est donnée par $D_{I'}(a)$.

$$ASSD(I, I') = \frac{1}{|I| + |I'|} \times \left(\sum_{a \in I} D_{I'}(a) + \sum_{b \in I'} D_I(b) \right) \quad (7)$$

3. <https://github.com/lucasb-eyer/pydensecrf>

La taille des patchs utilisés pour l’entraînement est importante car elle fait varier la quantité d’information locale représentée [17]. Ainsi, nous avons étudié son impact en comparant l’utilisation de patchs de taille 32^3 et 64^3 . Par ailleurs, la quantité de patchs extraits influence la fréquence de chevauchement de ces derniers. C’est pourquoi, nous avons aussi étudié ce paramètre.

De la même manière, nous avons comparé notre stratégie de sélection des patchs d’entraînement avec d’autres afin de définir l’importance de la répartition des patchs ainsi que l’impact d’un surentraînement spécifique au niveau des contours. Trois stratégies ont été considérées :

- **Stratégie 1** : Distribution entièrement aléatoire des patchs dans l’IRM
- **Stratégie 2** : 15% des patchs externes au cerveau / 85% des patchs distribués uniformément dans la région du cerveau autour du barycentre
- **Stratégie 3** : Notre stratégie (détaillée dans la Section 2.2) soit 15% des patchs externes au cerveau / 42,5% des patchs distribués uniformément dans la région du cerveau autour du barycentre / 42,5% des patchs focalisés sur les contours du cerveau.

Dans un second temps, nous avons mis en évidence les effets de la régularisation spatiale avec CRF en comparant les segmentations obtenues avec et sans régularisation. Nous avons également étudié les conséquences de la dilatation.

Enfin, nous avons comparé notre méthode avec l’outil Brain Extraction Tool (BET) de FSL⁴. Le choix de BET se justifie par le fait qu’il s’agit d’un standard pour l’extraction automatique du cerveau humain sur IRM offrant de très bons résultats [16]. Par ailleurs, l’outil s’appuie sur des modèles déformables et ne dépend donc pas du nombre de données à disposition contrairement à d’autres méthodes basées sur un atlas [7]. Pour la comparaison, nous avons calculé les métriques présentées précédemment ainsi que le volume cérébral segmenté afin de mettre en évidence la méthode perdant le moins de volume de cerveau.

3.3 Résultats

Dans cette section, nous présentons les résultats obtenus pour les diverses expériences décrites précédemment.

Impact de la taille et de la distribution des patchs d’entraînement. L’utilisation de petits patchs (32^3) permet de réduire considérablement les temps d’entraînement par rapport à des patchs plus grands (64^3) comme illustré dans la Table 2. Néanmoins, ils contiennent moins d’information sur le contexte, ce qui dégrade nettement les résultats quantitatifs (Table 2) et qualitatifs (Fig.4).

La Table 2 montre aussi que les résultats sont meilleurs lorsque l’on augmente le nombre de petits patchs par IRM à 700 (probablement du fait d’un chevauchement plus fréquent). Les patchs de taille 32^3 permettent d’atteindre un

Patchs	DICE	HD	ASSD	Entraînement
300×64^3	0.930	26.79	2.07	4hrs30
300×32^3	0.881	21.34	3.56	1hrs10
700×64^3	NaN	NaN	Nan	more than 7hrs
700×32^3	0.950	28.85	1.68	2hrs30

TABLE 2 – Comparaison de la durée d’entraînement et des performances du réseau U-Net en fonction du nombre et de la taille des patchs utilisés.

DICE moyen de 0.950 contre 0.930 avec les patchs de taille 64^3 tout en réduisant presque de moitié le temps d’entraînement. Par ailleurs, bien que la distance d’Hausdorff moyenne soit dégradée, la distance de surface moyenne est améliorée (1.68 avec les patchs de taille 32^3 contre 2.07 avec ceux de taille 64^3). Ces résultats sont illustrés par l’image de droite sur la Fig.4 segmentant avec plus de précision le cerveau porcine.

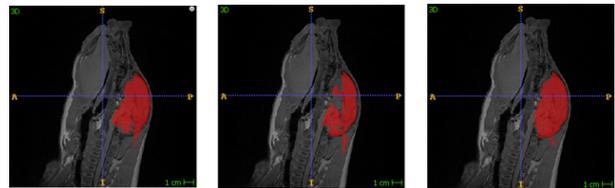


FIGURE 4 – Exemples de segmentations obtenues en fonction de la taille et du nombre de patchs utilisés (de gauche à droite 300×64^3 , 300×32^3 , 700×32^3)

La Table 3 compare les performances du réseau en fonction de la répartition des patchs d’entraînement de taille 32^3 .

	Precision	Rappel	DICE	HD	ASSD
Stratégie 1	0.849	0.607	0.728	26.21	6.46
Stratégie 2	0.893	0.915	0.904	50.29	5.27
Stratégie 3	0.952	0.948	0.950	28.85	1.68

TABLE 3 – Comparaison des performances du réseau U-Net en fonction de la répartition des patchs d’entraînement

On constate qu’en forçant 85% des patchs d’entraînement à contenir des parties du cerveau (stratégie 2), on parvient à considérablement améliorer les performances du réseau (+ 0.176 pour le DICE, -1.19 pour ASSD). Le rappel (puisque notre objectif principal est de le maximiser) est augmenté de 0.308. La distance d’Hausdorff moyenne est quant à elle très largement dégradée du fait d’une erreur grossière sur l’un des porcelets. Par ailleurs, l’approche par contours (stratégie 3) permet d’améliorer encore davantage les performances du réseau (+0.046 pour le DICE, -3.59 pour ASSD, -21.44 pour HD par rapport à la stratégie 2). Bien que le HD soit un peu moins bon qu’avec la stratégie 1, la moyenne des distances à la surface (ASSD) est bien meilleure. On peut noter le bon équilibre entre précision et rappel avec la stratégie 3 qui est celle de notre méthode.

4. <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET/UserGuide>

Effets de la régularisation spatiale. Sur la Fig.5, on observe que la régularisation spatiale à l'aide des champs aléatoires conditionnels (CRF) a permis de corriger les irrégularités de la segmentation. Par ailleurs, en comparant les deux premières lignes de la Table 4, on remarque d'une part, que l'utilisation du CRF a permis de réduire considérablement la distance d'Hausdorff et la distance de surface symétrique moyenne. D'autre part, l'ajout de cette opération augmente raisonnablement le temps de traitement par IRM. Néanmoins, elle a également dégradé le rappel.

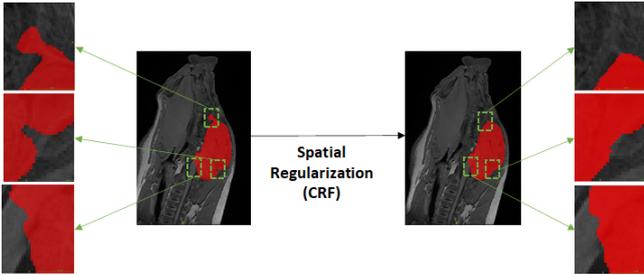


FIGURE 5 – Impact de l'utilisation de champs aléatoires conditionnels sur le masque du cerveau extrait

En pratique, on essaie d'augmenter le rappel en appliquant une dilatation morphologique en fin de traitement. Ainsi, la méthode proposée intègre cette dernière étape d'ajustement (avec un élément structurant sous forme de boule 3D de rayon unitaire). Dans la Table 4, on remarque que l'ajout de la dilatation permet d'augmenter le rappel et de réduire les distances HD et ASSD. Cela traduit une récupération de certaines zones du cerveau qui avaient été perdues.

Comparaison avec BET. La Table 5 donne les résultats obtenus avec notre méthode et avec l'outil BET pour les 7 porclets test. On constate que notre méthode surpasse BET pour l'ensemble des porclets avec une amélioration de toutes les métriques considérées (DICE, HD et ASSD). Par ailleurs, excepté pour le porclet 3, le volume cérébral estimé se rapproche davantage du volume cérébral réel. On observe une tendance à surestimer le volume cérébral avec notre méthode, ceci résultant du choix de considérer une marge (dilatation appliquée) pour la segmentation ultérieure des tissus et structures.

La Fig.6 donne des exemples de segmentation obtenus. On constate que notre méthode parvient à extraire le cerveau du porclet avec plus de précision et de constance que BET qui a tendance à perdre des parties ou à incorporer dans son masque des zones externes (porclets 3 et 4).

4 Discussion

Ce travail illustre la capacité du réseau U-Net 3D avec patchs d'extraire avec précision le cerveau du porclet sur IRM T1 grâce notamment à une focalisation des patchs d'entraînement sur les régions de bord. Par ailleurs, la méthode proposée présente l'avantage de fonctionner à partir d'une petite base de données d'entraînement (seulement 7

porcelets). Nous avons également trouvé un équilibre dans la taille des patchs afin de combiner des informations relatives au contexte globale et aux caractéristiques locales dans des temps de calcul raisonnables. Le nombre conséquent de patchs d'entraînement extraits par IRM (700) nous permet de compenser la modeste taille de données tout en assurant des chevauchements plus systématiques rendant la prédiction plus précise. De plus, notre méthode est capable de compenser les erreurs de segmentation grâce à une chaîne de post-traitement constituée de traitement d'image et d'une régularisation spatiale par champs aléatoires conditionnels (CRF). On soulignera que les CRF pourraient être entièrement intégrés dans le modèle U-Net comme récemment proposé [24]. Par ailleurs, l'utilisation de patchs se chevauchant lors de l'inférence pourrait compenser l'incertitude de prédiction liée aux effets de bords et ainsi réduire les irrégularités dans la segmentation [17]. Cela constituerait une alternative au post-traitement.

De par l'ensemble de ces propriétés, nous avons vu que notre solution dépasse incontestablement, dans le cas du porclet, des outils d'extraction de cerveau spécifiques à l'Homme comme BET.

Bien que notre approche soit prometteuse, elle souffre de certaines limitations. Premièrement, les performances ont été évaluées sur une base de données test relativement petite (seulement 7 porclets). De nouvelles données IRM seront bientôt prêtes afin d'évaluer sur une population plus grande l'efficacité de la solution.

Par ailleurs, nous avons choisi une stratégie de sélection de patchs, certes guidée, mais aléatoire, rendant inconstant le nombre de patchs impliqués dans une décision. Il serait intéressant d'étudier si une approche tenant compte de tous les patchs recouvrant l'image originale telle que présentée dans [17] offrirait de meilleurs résultats.

Une autre limitation probable est celle du stade de développement. En effet, durant les premiers mois de vie, le cerveau du porclet change très vite. Les contrastes entre les différents tissus cérébraux s'accroissent et la taille du cerveau augmente. Ainsi, les caractéristiques de texture et d'intensité apprises à 15 jours risquent de ne plus être totalement valables quelques semaines plus tard. Cela se traduirait par une meilleure précision sur les porclets ayant un cerveau à un stade de développement plus proche des données d'entraînement (volume cérébral moyen de 52.73cm^3). Cette dynamique du développement cérébral s'observe déjà avec les 7 porclets considérés dont les âges fluctuent de 13 à 16 jours avec des variations de volume de 46.05cm^3 à 54.94cm^3 (voir tableau 5). Ces variations de développement sont également propres à chaque individu. Évidemment, les changements autour de 15 jours restent minimes et ne dégradent pas vraiment l'efficacité du réseau. Néanmoins, l'application de ce même réseau un mois plus tard (âge où le volume cérébral est en moyenne de $70,5\text{cm}^3$ selon des études préliminaires) risque d'offrir des résultats moins bons. Pour pallier ce problème, il serait possible d'entraîner un réseau de neurones à différents stade de

	Precision	Rappel	DICE	HD	ASSD	Temps de traitement
U-Net + PP	0.952	0.948	0.950	28.85	1.68	1min
U-Net + PP + CRF	0.975	0.927	0.951	8.52	1.32	1min46s
U-Net + PP + CRF + Dilatation	0.948	0.960	0.954	7.73	1.28	1min48s

TABLE 4 – Comparaison des extractions de cerveau obtenues avec régularisation spatiale et dilatation (PP : post-traitement, impliquant la conservation de la composante connexe principale, le lissage de celle-ci et l’élimination des éventuelles cavités)

	Age	Volume cérébral	Volume cérébral estimé		DICE		HD		ASSD	
			BET	Notre méthode	BET	Notre méthode	BET	Notre méthode	BET	Notre méthode
Porcelet 1	16J	53.45cm ³	45.20cm ³	55.03cm³	0.890	0.959	13.64	7.28	2.91	1.16
Porcelet 2	13J	48.81cm ³	46.04cm ³	49.57cm³	0.870	0.955	14.18	9.43	3.44	1.25
Porcelet 3	13J	46.05cm ³	45.83cm³	44.76cm ³	0.680	0.956	29.07	8.48	8.52	1.21
Porcelet 4	15J	52.40cm ³	38.76cm ³	54.70cm³	0.409	0.959	49.16	5.38	16.49	1.19
Porcelet 5	15J	50.82cm ³	32.94cm ³	51.99cm³	0.675	0.948	37.01	8.60	8.33	1.45
Porcelet 6	16J	54.94cm ³	48.54cm ³	52.89cm³	0.892	0.956	11.04	6.40	2.90	1.25
Porcelet 7	15J	48.72cm ³	43.14cm ³	49.61cm³	0.861	0.946	16.03	8.54	3.55	1.48

TABLE 5 – Résultats obtenus avec BET et notre méthode sur les 7 porcelets test (les âges exacts sont donnés à titre indicatif).

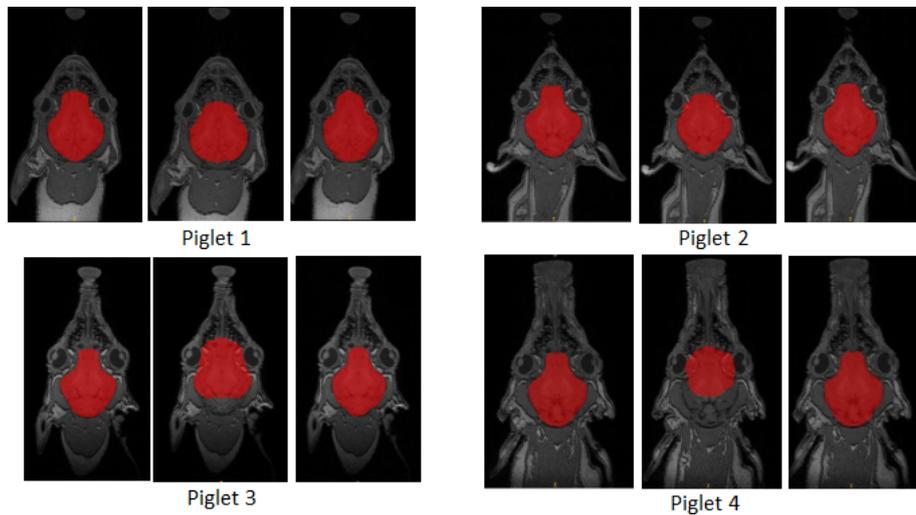


FIGURE 6 – Exemples de segmentations de 4 porcelets. De gauche à droite : vérité terrain, résultat avec BET, notre méthode.

développement (séparés d’un mois environ) jusqu’à ce que le cerveau se stabilise. Une autre solution serait d’entraîner un unique réseau U-Net avec des données de porcelets de tout âge mais cela supposerait beaucoup plus de données et une augmentation considérable des temps d’entraînement, avec le risque d’une perte de précision à un âge donné.

Enfin, nous avons testé notre méthode uniquement sur le porcelet. Il serait intéressant de l’exploiter sur d’autres modèles animaux de plus en plus exploités comme les rongeurs [26], les macaques [4], le furet [21] ou le chien [2].

5 Conclusions

Nous avons proposé une méthode originale pour l’extraction du cerveau du porcelet sur IRM T1 combinant apprentissage profond et régularisation spatiale. Les premiers résultats montrent que cette approche permet d’obtenir une

bonne qualité de segmentation (DICE moyen de 0.954, HD de 7.73 et ASSD de 1.28), surpassant des outils spécifiques à l’Homme comme BET, à partir d’une base de données limitée. Une possible limitation de cette méthode concerne sa robustesse vis à vis du développement cérébral. Il s’agira d’un aspect à considérer à travers la question ouverte de savoir si l’on peut utiliser un seul et unique réseau (entraîné sur des exemples à différents stades de développement) ou bien un réseau spécifique à chaque stade.

Les perspectives à ce travail sont nombreuses. En effet, à partir du cerveau extrait, on peut envisager la segmentation des structures et tissus cérébraux du porcelet afin de mettre en place diverses études (analyses morphométriques, analyses des informations structurales par graphe [1, 23]) sur cerveaux sains et lésés qui permettront de caractériser l’impact de la lésion ischémique liée à la paralysie cérébrale.

Une autre perspective consiste à mettre en place une méthodologie générale transposable à différents animaux.

Références

- [1] J.-B. Fasquel, and N. Delanoue "A graph based image interpretation method using a priori qualitative inclusion and photometric relationships", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 41, pp. 1043-1055, 2019.
- [2] X. Liu, R. Tian, Z. Zuo, H. Zhao, L. Wu, Y. Zhuo, Y.Q. Zhang, and L. Chen "A high-resolution MRI brain template for adult Beagle", *Magnetic Resonance Imaging*, Vol. 68, pp. 148-157, 2020.
- [3] A. Makropoulos, S. J. Counsell, and D. Rueckert, "A review on automatic fetal and neonatal brain MRI segmentation", *NeuroImage*, Vol. 170, pp. 231-248, 2018.
- [4] Y. Balbastre, D. Rivière, N. Souedet, C. Fischer, A.-S. Herard, S. Williams, M. Vandenberghe, J. Flament, R. Aron-Badin, P. Hantraye, J.-F. Mangin, and T. Delzescaux "A validation dataset for Macaque brain MRI segmentation", *Data in Brief*, Vol. 16, pp. 37-42, 2018.
- [5] K.L. Thibault, and S.S. Margulies, "Age-dependant material properties of the porcine cerebrum : effect on pediatric inertial head injury criteria", *Journal of Biomechanics*, Vol. 31, pp. 1119-1126, 1998.
- [6] M.S. Conrad, B.P. Sutton, R.N. Dilger, and R.W. Johnson, "An In Vivo Three-Dimensional Magnetic Resonance Imaging-Based Averaged Brain Collection of the Neonatal Piglet (*Sus scrofa*)", *PLoS One*, Vol. 9, pp. 1396-1402, 2014.
- [7] J. Lohmeier, T. Kaneko, B. Hamm, M.R. Makowski, and H. Okano "atlasBREX : Automated template-derived brain extraction in animal MRI", *Scientific Reports*, Vol. 9, 2019.
- [8] B. Lee, N. Yamanakkanavar, and J.Y. Choi, "Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture", *PLoS One*, Vol. 15, 2020.
- [9] C.N. Devi, A. Chandrasekharan, V.K. Sundararaman, and Z.C. Alex, "Automatic segmentation of infant brain MR images : With special reference to myelinated white matter", *Biocybernetics and Biomedical Engineering*, Vol. 37, pp. 143-158, 2017.
- [10] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle "Brain tumor segmentation with Deep Neural Networks", *Medical Image Analysis*, Vol. 35, pp. 18-31, 2017.
- [11] J. Lafferty, A. McCallum, and F. Pereira "Conditional random fields : Probabilistic models for segmenting and labeling sequence data", *Proceedings of the Eighteenth International Conference on Machine Learning*, pp. 282-289, 2001.
- [12] E. Mazerand, C. Gallet, L. Le Fournier, M. Dinomais, C. Montero-Menei, and P. Menei, "Development of a porcine model of cerebral palsy : first step before a study on the efficacy of a stereotactic stem cell therapy", *European Academy of Childhood Disability*, 2019.
- [13] H. Gan, Q. Zhang, H. Zhang, Y. Chen, J. Lin, T. Kang, J. Zhang, F.A. Troy, and B. Wang, "Development of new population-averaged standard templates for spatial normalization and segmentation of MR images for postnatal piglet brains", *Magnetic Resonance Imaging*, Vol. 32, pp. 1396-1402, 2014.
- [14] P. Krähenbühl, and V. Koltun, "Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials", *Advances in Neural Information Processing Systems*, Vol. 24, pp. 109-117, 2011.
- [15] V. Yeghiazaryan, and I. Voiculescu "Family of boundary overlap metrics for the evaluation of medical image segmentation", *Journal of Medical Imaging*, Vol. 5, 2018.
- [16] S.M. Smith, "Fast robust automated brain extraction", *Human Brain Mapping*, Vol. 17, pp. 143-155, 2002.
- [17] M. Ben naceur, M. Akil, R. Saouli, and R. Kachouri "Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy", *Medical Image Analysis*, Vol. 63, 2020.
- [18] E. Shelhamer, J. Long, and T. Darrell "Fully Convolutional Networks for Semantic Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 640-651, 2017.
- [19] A.P Zijdenbos, B.M. Dawant, R.A. Margolin, and A.C. Palmer "Morphometric analysis of white matter lesions in MR images : method and validation", *IEEE Transactions on Medical Imaging*, Vol. 13, pp. 716-724, 1994.
- [20] M. Beauchemin, K. P. B. Thomson, and G. Edwards "On the Hausdorff distance used for the evaluation of segmentation results", *Canadian Journal of Remote Sensing*, Vol. 24, pp. 3-8, 1998.
- [21] E.B. Hutchinson, S.C. Schwerin, K.L. Radomski, N. Sadeghi, J. Jenkins, M.E. Komlos, M.O. Irfanoglu, S.L. Juliano, and C. Pierpaoli "Population based MRI and DTI templates of the adult ferret brain and tools for voxelwise analysis", *NeuroImage*, Vol. 152, pp. 575-589, 2017.
- [22] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala "PyTorch : An Imperative Style, High-Performance Deep Learning Library", *Advances in Neural Information Processing Systems*, Vol. 32, 2019.
- [23] J. Chopin, J.-B. Fasquel, H. Mouchère, R. Dahyot, and I. Bloch "Semantic image segmentation based on spatial relationships and inexact graph matching", *International Conference on Image Processing Theory, Tools and Applications*, 2020.
- [24] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille "DeepLab : Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, pp. 834-848, 2018.
- [25] S. Mannor, D. Peleg, and R. Rubinstein "The cross entropy method for classification", *In Proceedings of the 22nd international conference on Machine learning*, 2005.
- [26] R. Feo, and F. Giove "Towards an efficient segmentation of small rodents brain : A short critical review", *Journal of Neuroscience Methods*, Vol. 323, pp. 82-89, 2019.
- [27] O. Ronneberger, P. Fischer, and T. Brox "U-Net : Convolutional networks for biomedical image segmentation", *Lecture Notes in Computer Science*, Vol. 9351, pp. 234-241, 2015.