

Segmentation des hyper-intensités de la matière blanche en quelques secondes à l'aide d'un réseau de neurones convolutif et de transfert d'apprentissage

Élodie Puybareau¹ Yongchao Xu^{1,2,3} Joseph Chazalon^{1,4} Isabelle Bloch² Thierry Géraud¹

¹ Laboratoire de Recherche et Développement de l'EPITA (LRDE), Le Kremlin-Bicêtre, France
elodie.puybareau@lrde.epita.fr

² LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France

³ Huazhong University of Science and Technology, Wuhan, China

⁴ L3i, Université de La Rochelle, France

Résumé

Dans cet article, nous proposons une méthode automatique et rapide pour segmenter les hyper-intensités de la matière blanche (WMH) dans des images IRM cérébrales 3D, en utilisant un réseau de neurones entièrement convolutif (FCN) et du transfert d'apprentissage. Ce FCN est le réseau neuronal du Visual Geometry Group (VGG) pré-entraîné sur la base ImageNet pour la classification des images naturelles, et affiné avec l'ensemble des données d'entraînement du concours MICCAI WMH. Nous considérons trois images pour chaque coupe du volume à segmenter, provenant des acquisitions en T1, en FLAIR, et le résultat d'un opérateur morphologique appliqué sur le FLAIR, le top-hat, qui met en évidence les petites structures de forte intensité. Ces trois images 2D sont assemblées pour former une image 2D-3 canaux interprétée comme une image en couleurs, ensuite passée au FCN pour obtenir la segmentation 2D de la coupe correspondante. Nous traitons ainsi toutes les coupes pour former la segmentation de sortie 3D. Avec une telle technique, la segmentation de WMH sur un volume cérébral 3D prend environ 10 secondes, pré-traitement compris. Notre technique a été classée 6e sur 20 participants au concours MICCAI WMH.

Mots Clef

IRM cérébrale 3D, segmentation de lésions, hyper-intensités de la matière blanche, morphologie mathématique, réseaux de neurones convolutifs.

Abstract

In this paper, we propose a fast automatic method that segments white matter hyper-intensities (WMH) in 3D brain MR images, using a fully convolutional network (FCN) and transfer learning. This FCN is the Visual Geometry Group neural network (VGG) pre-trained on ImageNet for natural image classification, and fine tuned with the training dataset of the MICCAI WMH Challenge. We consider three images for each slice of

the volume to segment : the T1 image, the FLAIR image, and the result of a morphological operator that emphasizes small bright structures. These three 2D images are assembled to form a 3 channels 2D image interpreted as a color image, that inputs the FCN to obtain the 2D segmentation of the corresponding slice. We process all slices to form the 3D output segmentation. With such a technique, the segmentation of WMH on a 3D brain volume takes about 10 seconds including pre-processing. Our technique was ranked 6-th over 20 participants at the MICCAI WMH Challenge.

Keywords

3D brain MRI, lesion segmentation, white matter hyper-intensities, mathematical morphology, fully convolutional network.

1 Introduction

1.1 Contexte

Ce travail a été réalisé dans le contexte du concours MICCAI WMH¹. L'objectif était de fournir une méthode entièrement automatisée pour la segmentation des hyper-intensités de la matière blanche (WMH) d'origine vasculaire. Les WMH sont les conséquences de maladies des petits vaisseaux et sont visibles sur les images IRM du cerveau [25]. Les maladies des petits vaisseaux sont impliquées dans les pathologies cérébro-vasculaires et sont une cause du déclin cognitif et de la perte fonctionnelle au cours du vieillissement [17]. L'étude des paramètres des WMH (volume, forme, etc.) peut donc être une clé pour le diagnostic et le suivi des patients traités pour la démence ou les maladies neurodégénératives.

L'analyse visuelle des images pour détecter les WMH est un processus difficile et les méthodes automatisées pourraient être utiles pour guider le diagnostic [2]. Cependant, l'évaluation et la comparaison des tech-

1. <http://wmh.isi.uu.nl>

Hôpital	Imageur IRM	Ne	Nt
UMC Utrecht	3T Philips Achieva	20	30
NUHS Singapore	3T Siemens TrioTim	20	30
VU Amsterdam (AMS)	3T GE Signa HDxt	20	30
	1.5T GE Signa HDxt	0	10
	3T Philips Ingenuity	0	10

TABLE 1 – Présentation de la base de données du concours, avec Ne et Nt les nombres de volumes respectivement d’entraînement et de test (restés secrets). Chaque volume (image 3D) contient environ 45 coupes qui nous servent à former les images 2D analysées par le réseau.

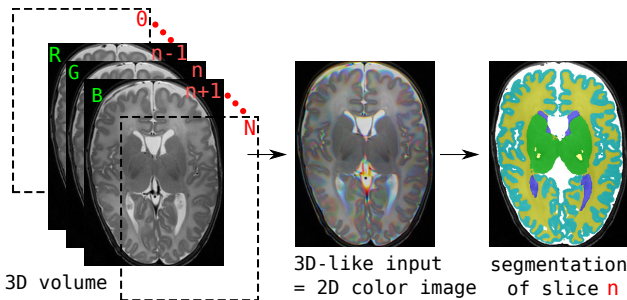


FIGURE 1 – Illustration de l’image en pseudo-couleurs venant du volume 3D et de la segmentation associée (méthode utilisée dans [27]).

niques automatisées de segmentation des WMH restent difficiles en raison de la diversité des ensembles de données et des critères d’évaluation [2]. L’objectif du concours MICCAI WMH était de comparer les techniques automatisées de segmentation des WMH. Cette comparaison donne un classement des techniques appliquées aux données acquises à partir de différents imageurs IRM : origines différentes, résolutions différentes, etc. Les données utilisées pendant le concours provenaient de trois hôpitaux et de cinq machines différentes, comme indiqué dans la table 1. L’organisateur a fourni un ensemble de 60 volumes (images 3D) pour l’entraînement. Les 110 volumes de test pour la phase d’évaluation sont restés secrets.

1.2 État de l’art

La segmentation des WMH a toujours été un problème difficile, et il est très compliqué d’obtenir une méthode fiable entièrement automatisée [2, 5]. En effet, un coefficient de similarité (Dice) ne sera considéré comme bon que s’il est supérieur à 0,7 [2], ce qui peut être difficile à obtenir pour des petites structures comme les WMH. Le problème de la segmentation des WMH peut être considéré avec plusieurs stratégies. L’utilisation de la modalité FLAIR semble être l’une des meilleures pour ce type de segmentation : il est possible de segmenter les hyper-intensités avec un seuil d’intensité en FLAIR optimal, déterminé par analyse des histogrammes comme décrit par Jack et al. [8].

Certaines méthodes reposent sur des champs de Markov aléatoires, soit en utilisant l’image FLAIR [10], soit en utilisant une autre modalité [22]. Dans [16], des opérateurs morphologiques et des représentations de l’image par arbre de coupes (*max-tree*) sont utilisés pour segmenter les WMH dans les images cérébrales T2 du nouveau-né, mais cette méthode n’est pas entièrement automatisée.

L’apprentissage automatique a réellement amélioré certains résultats de segmentation. De nouvelles méthodes ont émergé en utilisant des procédures d’apprentissage automatique (supervisées ou non). Dans [6], la segmentation des WMH est effectuée en utilisant des réseaux de neurones convolutifs (CNN) avec des informations de localisation anatomique, et dans [7], en utilisant l’apprentissage par transfert avec une adaptation de domaine et des correctifs, atteignant un score de Dice de 0,76. Les auteurs ont suggéré l’idée de remplacer leur réseau CNN par un réseau convolutif entièrement connecté, par exemple U-Net [18].

Dans un travail précédent récent [27], les volumes IRM 3D du cerveau ont été segmentés en utilisant un réseau de neurones entièrement convolutif (FCN) et l’apprentissage par transfert. Le réseau utilisé pour l’apprentissage par transfert est VGG (Visual Geometry Group) [23], pré-entraîné sur l’ensemble de données ImageNet. Il prend en entrée une image 2D en « couleurs » (3 canaux RVB) qui est ici une image *pseudo-3D*, composée de 3 coupes consécutives du volume 3D (voir figure 1), qui rappelle le $2D+\epsilon$ [1], mais est utilisable sur toutes les coupes et reforme une image 2D RVB au lieu de 3 images 2D. Cette méthode utilise une seule modalité et fournit de bons résultats pour la segmentation des structures du cerveau.

Pour ce concours, nous avons étendu le travail précédent sur la segmentation d’IRM de cerveau [27] à la segmentation des WMH, en tirant parti de la puissance d’un FCN pré-entraîné sur un grand ensemble de données et en l’ajustant sur l’ensemble des données fournies par le concours. Les principales contributions sont :

1. une technique de pré-traitement par morphologie mathématique qui met l’accent sur les petites lésions pour améliorer leur segmentation ;
2. une technique de fusion qui permet d’exploiter plusieurs modalités (T1, FLAIR et FLAIR après pré-traitement morphologique) dans la chaîne de segmentation pour chaque coupe.

Une présentation de la méthode proposée est donnée dans la figure 2. La méthode est entièrement automatique et utilise les images 3D T1 et FLAIR. Les détails de l’ensemble de la méthode sont donnés ci-après. Cette méthode est très rapide, environ 10 secondes sont nécessaires pour traiter un volume de scan complet, mais est aussi efficace : nous avons atteint la 6^e place lors du concours MICCAI WMH.

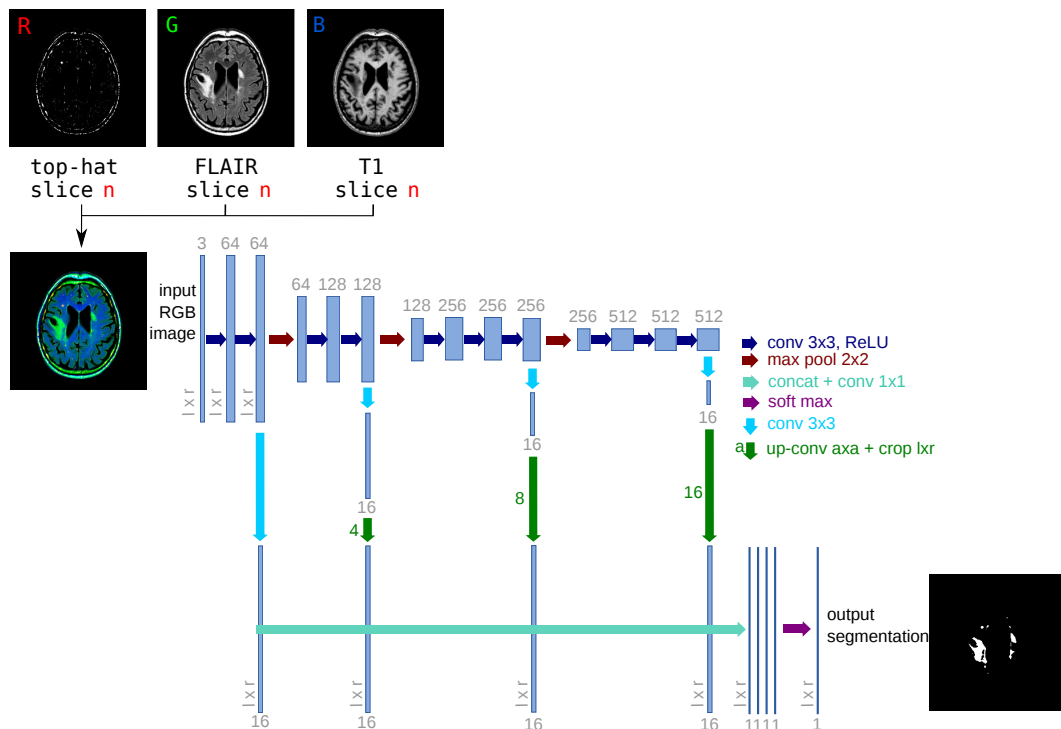


FIGURE 2 – Architecture du réseau proposé. Nous réalisons un réglage fin (*fine tuning*) sur les 4 premières parties convolutives et combinons linéairement les cartes de caractéristiques des plus fines aux grossières du réseau VGG pré-entraîné [23]. Chaque image en « couleurs » (Input) est construite à partir de la coupe n des séquences T1 et FLAIR, et d'un résultat de pré-traitement morphologique (par « *top-hat* »).

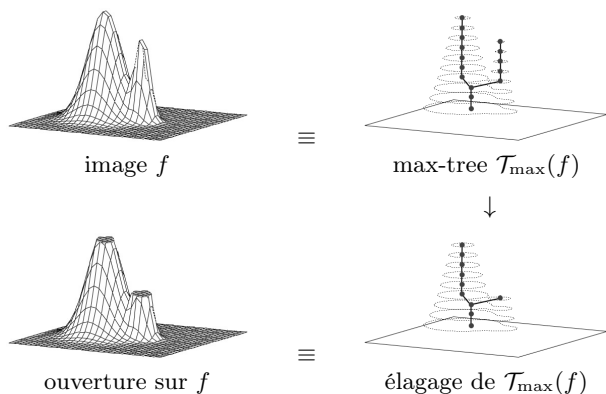


FIGURE 3 – Un opérateur morphologique connexe (ici une ouverture) appliqué à une représentation de l'image par arbre.

2 Description de la méthode

2.1 Prétraitement morphologique

Alors que les opérateurs les plus populaires de la morphologie mathématique sont définis à partir d'éléments structurants, la classe des « opérateurs connexes » ne l'est pas nécessairement [12, 20]. Cette classe est très intéressante car ces opérateurs ne modifient pas les contours des objets (ils ne peuvent pas créer de

nouveaux contours, ils suppriment simplement ceux qui existent). Certains opérateurs connexes peuvent être facilement définis à partir de certaines représentations arborescentes d'une image en niveaux de gris [9, 19, 21]. Ces représentations expriment l'inclusion des composantes connexes obtenues par seuillages successifs de l'image. Le calcul, le stockage et le traitement d'un tel arbre sont très efficaces [3, 15]. Dans ce qui suit, nous nous appuyons sur la représentation sous forme de *max-tree*, dénotée par \mathcal{T}_{\max} , obtenue par seuillage croissant. Un exemple de cet arbre est présenté dans la figure 3 en haut à droite. Si l'on élague cet arbre, comme dans la figure 3 en bas à droite, on peut reconstruire la fonction représentée dans la figure 3 en bas à gauche. Dans ce qui suit, nous filtrons toute composante du *max-tree* dont la taille (ou la surface, c'est-à-dire le nombre de pixels) est inférieure à un seuil N , ce qui conduit à une *ouverture surfacique* [24]. Dans toutes nos expériences, nous prenons $N = 25$ pixels.

Dans la figure 4(b), un opérateur de ce type a été appliqué à l'image de la figure 4(a). Dans ce résultat, quelques petites composantes de forte intensité sont supprimées, sans que le reste de l'image ne soit modifié (c'est un opérateur connexe); notons également que cet opérateur a la capacité de filtrer les objets faiblement contrastés de la même manière que les objets fortement contrastés. Les composantes élimi-

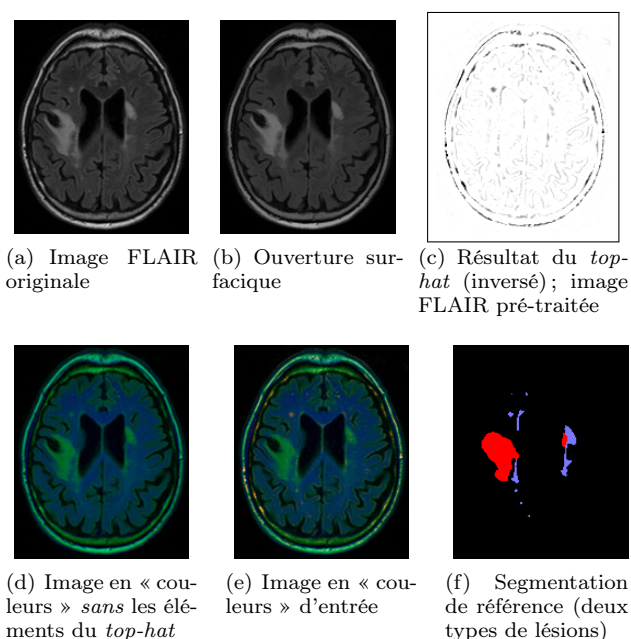


FIGURE 4 – Illustration de la procédure de *top-hat*.

nées correspondent à de petits pics d'intensité dans la matière blanche. Le résidu de cette étape de filtrage, c'est-à-dire le résultat moins l'entrée, est appelé chapeau haut-de-forme (*top-hat*); il est représenté en figure 4(c).

Le principal problème avec la segmentation des WMH est que ce sont de petites lésions. Une piste pour aider le réseau à détecter ces lésions est d'« améliorer » les données en entrée du réseau. Le prétraitement morphologique est appliqué à l'image FLAIR dont les valeurs d'intensité des voxels sont converties sur 8 bits. Dans l'image finale, appelée « FLAIR pré-traitée » dans ce qui suit et illustrée dans la figure 4(c), les petites lésions sont particulièrement visibles et les grandes n'apparaissent pas. Il ne nous est pas nécessaire au vu des résultats d'enlever les pixels de la boîte crânienne.

2.2 Réseaux profonds convolutifs pour la segmentation des WMH

Les réseaux de neurones totalement convolutifs (FCN) et le transfert d'apprentissage ont prouvé leur efficacité pour la segmentation des images naturelles [13], et de plus en plus dans d'autres domaines tels que l'imagerie médicale. Dans un précédent article [27], nous avons proposé d'exploiter cette méthode pour segmenter des images IRM 3D cérébrales, bien que ces images soient très différentes des images naturelles (voir aussi le matériel supplémentaire [26]). Au vu des bons résultats obtenus, nous l'avons adaptée à la segmentation des WMH. Nous nous appuyons sur le réseau VGG à 16 couches [23], qui a été pré-entraîné sur des millions d'images naturelles d'ImageNet pour la classification d'images [11]. Pour notre application,

nous ne conservons que les quatre premières parties convolutives appelées « réseau de base » et rejetons les couches entièrement connectées à la fin du réseau VGG. Ce réseau de base est principalement composé de couches convolutives : $z_i = w_i \times x + b_i$, d'unités linéaires rectifiées (ReLU) pour la fonction d'activation non linéaire : $f(z_i) = \max(0, z_i)$, et de couches de *max pooling* entre deux étapes successives, où x est l'entrée de chaque couche convolutive, w_i est le poids synaptique, et b_i est le terme de biais. Les trois couches de *max pooling* divisent le réseau de base en quatre étapes de cartographies de caractéristiques allant de fines à grossières. Inspirés par les travaux [13, 14], nous ajoutons des couches convolutives spécialisées (avec une taille de noyau 3×3) avec 16 cartes de caractéristiques après les couches convolutives à la fin de chaque étape. Toutes les couches spécialisées sont ensuite redimensionnées à la taille de l'image d'origine et concaténées ensemble. Nous ajoutons à la fin une dernière couche convolutive avec une taille de noyau 1×1 . Cette dernière couche combine linéairement les cartes de caractéristiques fines à grossières dans les couches spécialisées concaténées, et fournit le résultat final de la segmentation. Ce résultat final correspond au label le plus probable pour chaque pixel : fond ou lésion (deux types de lésions sont détectées par le réseau). L'architecture de réseau proposée est schématisée figure 2.

L'architecture décrite ici est très similaire à celle utilisée dans [14] pour l'analyse d'images 2D en couleurs de la rétine. Pour notre application, la question est de savoir comment préparer les entrées appropriées étant donné qu'une image IRM cérébrale est un volume 3D. Pour obtenir des images d'entrée RVB, nous proposons d'utiliser les coupes 2D de différentes modalités. Plus précisément, pour former une image en couleurs artificielles (3 canaux) en entrée pour le réseau pré-entraîné afin de segmenter la coupe n , on utilise les coupes n des images 3D FLAIR, T1 et FLAIR pré-traitée en guise de canaux de couleurs de l'image 2D d'entrée du réseau. Les canaux vert, bleu et rouge portent donc des informations différentes. La correspondance entre un canal et le type de donnée qu'il contient est arbitraire mais cette correspondance doit toujours rester la même entre les phases d'apprentissage et de test. Ce processus est représenté en figure 2 (à gauche). Chaque image 2D en couleurs forme ainsi une représentation d'une partie (une coupe des volumes FLAIR et T1) du volume IRM. L'utilisation d'une telle représentation 2D nous permet de contourner les besoins coûteux en calcul et en mémoire d'un FCN 3D.

Pour la phase d'entraînement, nous utilisons une fonction de perte logistique avec une loi multinomiale pour une tâche de classification un-parmi-plusieurs, en passant des prédictions réelles à travers une couche *softmax* pour obtenir une distribution de probabilité sur les classes. Au cours de l'entraînement, nous utilisons la stratégie classique d'augmentation des données par variation de l'échelle et rotation. Nous affinons le ré-

seau pour les premières 50k itérations en utilisant un *learning rate* l_r de 10^{-8} , et les derniers 100k avec un plus petit *learning rate* de 10^{-9} . Nous minimisons la fonction de perte avec une descente de gradient stochastique de *momentum* de 0,99 pour les premières 50k itérations et de 0,999 pour les 100k suivantes et un *weight_decay* de 0,0005. La fonction de perte est moyennée sur 20 images (valeur par défaut donnant de bons résultats dans notre cas).

Lors de la phase de test, après avoir pré-traité le volume (image 3D), nous préparons l'ensemble des images 2D en couleurs.

L'entraînement et les tests ont été effectués sur une carte GPU *NVIDIA GeForce GTX 1080 Ti* (11 Go de RAM). La phase d'entraînement dure 4 heures tandis que la phase de test dure moins de 10 secondes pour un volume entier. Les 4 heures d'apprentissage restent raisonnables et le réseau peut être réappris si des acquisitions provenant de nouvelles machines sont fournies. Cela pourrait être utile pour améliorer la robustesse du réseau, surtout si les nouvelles acquisitions ont des caractéristiques différentes : résolution, contraste, etc.

La sortie du réseau pour une coupe, pendant la phase d'inférence, est une coupe segmentée 2D. L'ensemble des coupes segmentées constitue un volume 3D ayant la même taille que le volume initial et ne contenant que les lésions segmentées.

Enfin, nous aimerions mettre en avant que, dans le cas de petites lésions, l'utilisation de la multi-modalité peut être une clé importante pour la segmentation. En effet, les petites lésions ne sont pas présentes sur beaucoup de coupes, limitant ainsi l'intérêt des représentations 3D. L'utilisation conjointe des coupes T1 et FLAIR apporte plus d'informations pour la détection des lésions. Comme la proportion des petites lésions est faible par rapport au nombre total de pixels, chaque information pouvant être discriminante pour la détection des lésions est importante.

3 Expériences et résultats

Cette section présente les expériences et résultats obtenus, tout d'abord en utilisant l'ensemble de données fournies pour l'apprentissage, puis les résultats du concours. Les mesures utilisées pour évaluer les résultats sont les suivantes.

Dice : Coefficient de similarité Dice.

H95 : Distance de Hausdorff (modifiée au 95^e percentile, en mm) [4].

AVD : Différence de volume moyen (en pourcentage).

Recall : Sensibilité pour les lésions individuelles (composantes connexes), ou rappel (entre 0 et 1).

F1 : Score $F1 = 2PR/(P + R)$ pour les lésions individuelles (entre 0 et 1), où P et R sont respectivement la précision et le rappel.

Type	Dice ↑	AVD ↓	Recall ↑	F1 ↑
pseudo-3D [27]	0,72	23,90	0,38	0,46
2D sans top-hat	0,72	28,24	0,39	0,48
2D avec top-hat	0,75	22,63	0,61	0,63

TABLE 2 – Comparaison quantitative de l'influence du *top-hat* sur l'ensemble des données d'entraînement. Pour le concours, nous avons soumis la méthode 2D avec *top-hat*. Dans ce tableau et dans les suivants, ↑ (resp. ↓) signifie qu'une valeur plus élevée (resp. plus faible) est meilleure.

Origine	Dice ↑	H95 ↓	AVD ↓	Recall ↑	F1 ↑
UMC Utrecht	0,74	11,22	19,07	0,70	0,66
NUHS Singapore	0,77	8,28	17,64	0,61	0,68
AMS GE 3T	0,75	6,75	21,91	0,62	0,71
AMS GE 1.5T	0,73	10,94	16,66	0,60	0,71
AMS Philips 3T	0,50	70,27	46,33	0,57	0,53
Moyenne pondérée	0,73	14,54	21,71	0,63	0,67
Rang $\in [0, 1]$ ↓	0,122	0,180	0,004	0,352	0,159

TABLE 3 – Résultats de notre méthode sur l'ensemble des données du concours. Les cinq premières lignes sont les résultats pour chaque mesure d'évaluation de chaque groupe de données. La sixième ligne est la moyenne pondérée de chaque mesure. La dernière ligne correspond au rang de chaque mesure ; notre méthode est extrêmement performante du point de vue de la quantification des WMH (mesure AVD).

3.1 Phase de développement

Pour cette partie, nous avons utilisé les 60 volumes de la base d'entraînement du concours pour valider et ajuster notre approche. Chacun de ces volumes contient environ 45 coupes (pour chaque modalité) qui servent à former les images 2D utilisées par notre réseau. Ces données ont été acquises dans trois hôpitaux en utilisant trois machines de différents fournisseurs (voir table 1).

Entraînement. Nous avons entraîné notre modèle sur les images extraites de 30 volumes (10 provenant de chaque hôpital) choisis au hasard. Le modèle a été ré-appris en utilisant les paramètres décrits dans la section précédente.

Test. Nous avons testé notre réseau sur les images extraites des 30 volumes restants. Nous avons évalué les résultats selon quatre mesures, en utilisant le code fourni par les organisateurs : le coefficient de Dice, la différence moyenne du volume (AVD), la sensibilité pour la détection de lésions individuelles et le score F1 pour les lésions individuelles. Les résultats sont présentés dans la table 2, ligne 3.

La figure 5 illustre nos résultats de segmentation, avec une comparaison qualitative avec la segmentation de référence ; on y voit que les faux négatifs et faux positifs sont très réduits.

ÉQUIPE	RANG ↓	Dice ↑	H95 ↓	AVD ↓	REC ↑	F1 ↑
nlp_logix	0,0485	0,77	7,16	18,37	0,73	0,78
k2	0,1368	0,77	9,79	19,08	0,59	0,70
ipmi-bern	0,2498	0,69	9,72	19,92	0,44	0,57
misp	0,1659	0,72	14,88	21,36	0,63	0,68
lrde	0,1635	0,73	14,54	21,71	0,63	0,67
cian	0,0366	0,78	6,82	21,72	0,83	0,70
sysu_media	0,0076	0,80	6,30	21,88	0,84	0,76
achilles	0,2962	0,63	11,82	24,41	0,45	0,52
nic-vicorob	0,0735	0,77	8,28	28,54	0,75	0,71
tig	0,3858	0,60	17,86	34,34	0,38	0,42
scan	0,2762	0,63	14,34	34,67	0,55	0,51
knight	0,4159	0,70	17,03	39,99	0,25	0,35
skkumedneuro	0,3492	0,58	19,02	58,54	0,47	0,51
tignet	0,3802	0,59	21,58	86,22	0,46	0,45
nist	0,4747	0,53	15,91	109,98	0,37	0,25
text_class	0,5725	0,50	28,23	146,64	0,27	0,29
nih_cidi	0,2697	0,68	12,82	196,38	0,59	0,54
upc_dlmi	0,4337	0,53	27,01	208,49	0,57	0,42
neuro.ml	0,5960	0,51	37,36	614,05	0,71	0,21
hadi	0,8886	0,23	52,02	828,61	0,58	0,11

TABLE 4 – Résultats du concours, triés par rapport à la mesure AVD (REC est le rappel).

Validation du choix du *top-hat* 2D. Pour évaluer l’influence du *top-hat* et de l’entrée 2D sur nos résultats, nous avons entraîné et testé notre réseau avec et sans le *top-hat* sur les images 2D (le dernier canal a été remplacé par la même coupe de T1 pour la version sans *top-hat*) et nous avons testé l’approche *pseudo-3D* de notre travail précédent [27] (l’image d’entrée correspondant à la coupe n a pour canaux les coupes $n - 1$, n et $n + 1$ du volume 3D). Les résultats sont présentés dans la table 2. Les résultats de l’approche 3D et de l’approche 2D sans *top-hat* sont assez similaires. La détection des lésions est un peu meilleure pour la version 2D, mais ce n’est pas significatif. L’information 3D n’est donc pas pertinente pour cette application. L’approche 2D incluant le *top-hat* a légèrement amélioré le Dice et l’AVD, mais surtout a permis une amélioration de la détection des lésions. Sur cet ensemble de données, la mesure de la sensibilité pour une lésion individuelle (*Recall*) est passée de 0,38 ou 0,39 à 0,61 grâce à la procédure du *top-hat* 2D (cf. table 2). L’amélioration des petites lésions a en effet aidé à la détection des lésions.

3.2 Résultats du concours

Les données d’évaluation du concours MICCAI WMH sont restées secrètes. Elles étaient composées de 110 acquisitions (volumes) provenant de trois hôpitaux et de cinq machines différentes de trois fournisseurs (voir table 1).

Les mesures utilisées pour l’évaluation sont les cinq mesures décrites au début de cette section. Le rang est compris entre 0 (meilleur résultat) et 1 (pire résultat). Il est calculé par les organisateurs du concours comme suit : les équipes sont triées de la meilleure à la moins bonne pour chaque mesure. La meilleure équipe sera classée 0 et la moins bonne 1. Les autres équipes

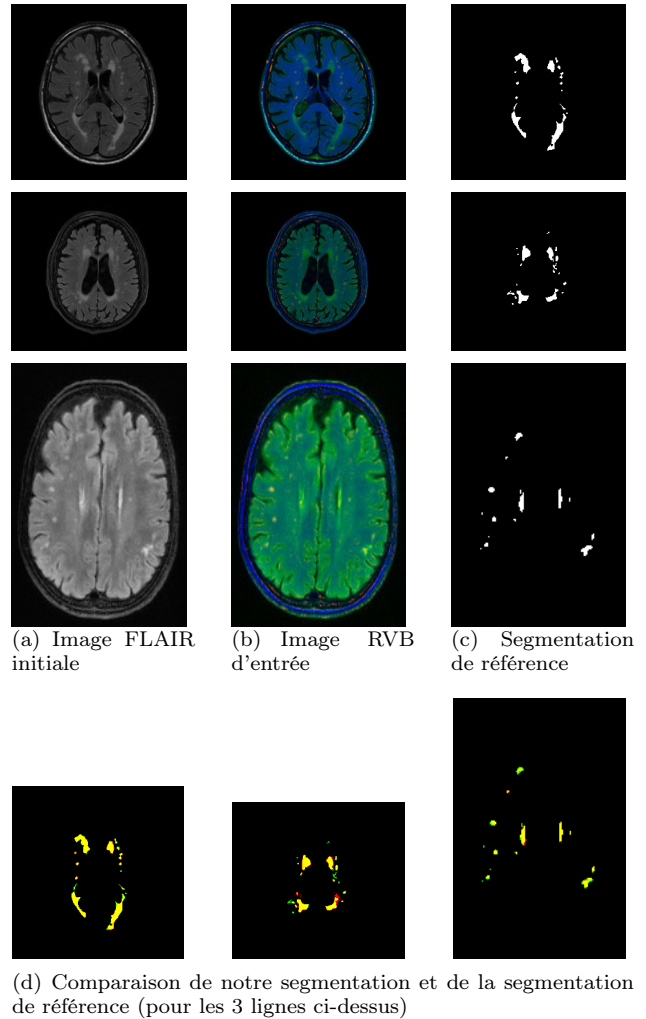


FIGURE 5 – Exemple de résultats sur chaque base de données. Les images d’entrée sont assez différentes car elles proviennent de différents hôpitaux. De haut en bas : Utrecht, Singapour et Amsterdam. L’intersection (les pixels en commun) entre la segmentation de référence et notre résultat est représentée en jaune ; les pixels verts (faux négatifs) sont dans la segmentation de référence mais pas dans notre segmentation ; les rouges (faux positifs) sont dans notre segmentation mais pas dans la référence.

sont classées entre 0 et 1 par rapport à leur performance dans la fourchette de cette mesure. Le classement final est la moyenne des 5 rangs des mesures. La table 3 montre les résultats de notre méthode pour le concours.

Notre meilleure performance concerne la mesure AVD où nous avons un très bon classement comme le montrent les tables 3 et 4, alors que notre moins bon score est pour la détection des lésions (*Recall*). Sur la base de ces mesures, nous pouvons conclure que notre méthode manque encore certaines lésions mais que celles détectées sont proches de la segmentation

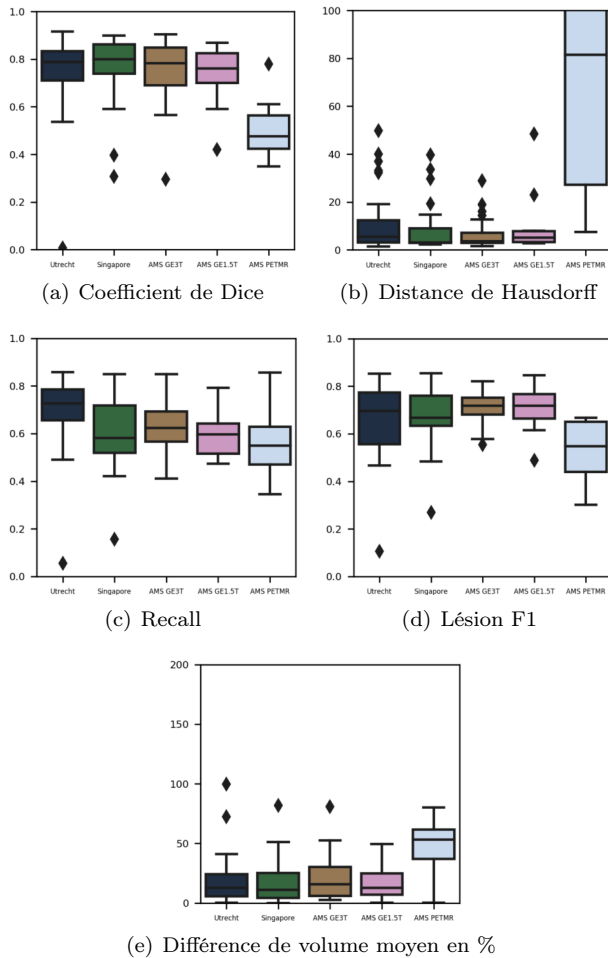


FIGURE 6 – Statistiques des mesures d’évaluation en fonction de l’origine des images. De gauche à droite : UMC Utrecht en bleu foncé, NUHS Singapour en vert, VU Amsterdam (AMS) 3T GE Signa HDxt en marron, AMS 1.5T GE Signa HDxt en rose, et AMS 3T Philips Ingenuity en bleu clair.

de référence. Nous étions classés 0,164 sur la moyenne générale et avons atteint la 6^e place du concours. Les meilleures méthodes du concours utilisent également l’apprentissage en profondeur pour la segmentation. Elles utilisent principalement des patches d’images au lieu d’images entières. Concernant l’architecture utilisée, il est intéressant de noter que le gagnant de ce concours utilise une architecture U-Net avec des étapes de pré-traitement et de post-traitement.

L’analyse des résultats par rapport à chaque source de données révèle que notre méthode fonctionne mal sur les données acquises avec l’imageur Philips, pour lesquelles aucune donnée d’apprentissage n’était disponible. Comme les données de test sont restées secrètes, nous n’avons pas pu approfondir nos investigations. Nous pouvons également noter que même si aucune donnée d’entraînement n’était disponible pour l’imageur GE 1.5T, notre méthode fonctionne encore bien

sur ces données. Les boîtes à moustaches (figure 6) montrent les hétérogénéités de ces résultats. Malgré cela, notre méthode apparaît assez stable selon les différents dispositifs d’acquisition.

4 Conclusion

Dans le cadre du concours MICCAI WMH, nous avons développé une méthode rapide, robuste et automatisée qui segmente les hyper-intensités de la matière blanche (WMH) dans les images cérébrales IRM 3D. Cette méthode étend une autre méthode précédemment proposée pour la segmentation d’images IRM du cerveau. Il s’agit d’un réseau de neurones entièrement convolutif (FCN) affiné avec un apprentissage par transfert pour segmenter les lésions. Il prend en entrée des images 2D en couleurs (3 canaux), correspondant à une coupe d’un volume 3D. Les deux premiers canaux sont composés d’une combinaison de coupes des acquisitions FLAIR et T1. Pour améliorer la détection des lésions, le dernier canal est la coupe de l’image FLAIR pré-traitée, où les petites hyper-intensités sont rehaussées par un opérateur de *top-hat*. Le FCN est le réseau VGG pour la classification des images naturelles, pré-entraîné sur la base de données ImageNet. Nous l’avons affiné sur l’ensemble de données d’entraînement du concours. Lorsque toutes les coupes sont traitées, les résultats sont réunis pour reconstruire le volume segmenté 3D résultant. Au cours de la phase de développement, nous avons validé les avantages des images pré-traitées 2D avec rehaussement des petites lésions sur les images 2D contenant uniquement les coupes T1 et FLAIR, ainsi que sur une image de type pseudo-3D comme dans [27]. Avec cette technique, la segmentation complète des WMH sur un volume cérébral 3D prend environ 10 secondes. Lors du concours, les organisateurs ont exécuté notre méthode sur leur base de données composée de 5 jeux différents. Nos résultats sont assez stables par rapport aux méthodes d’acquisition, sauf pour l’imageur Philips, ce qui démontre la fiabilité de notre méthode. Enfin, nous avons atteint un score élevé sur la mesure AVD avec un rang de 0,004 (le meilleur est 0, le pire est 1) pour une valeur AVD de 21,71. Notre score global est de 0,164 (moyenne des 5 mesures utilisées pour le concours), nous conduisant à la 6^e place.

Grâce à notre méthode, nous pourrions envisager d’autres applications à d’autres fins telles que l’évaluation de la prolifération tumorale (avec les données du concours *TUPAC16 MICCAI Grand Challenge*) afin de prédire le score de prolifération tumorale du cancer du sein, ou la prédiction de la survie des patients à partir de l’étude des lésions cérébrales, avec les données du concours *MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS)*.

Remerciements Les auteurs remercient les organisateurs du concours WMH de MICCAI 2017, ainsi que *NVIDIA Corporation* pour le don de la carte graphique *GeForce GTX 1080 Ti*.

Références

- [1] K. Aderghal, J. Benois-Pineau, and K. Afdel. Classification of sMRI for Alzheimer's disease diagnosis with CNN : Single siamese networks with 2D+ ϵ approach and fusion on ADNI. In *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval*, pages 494–498, 2017.
- [2] M. E. Caligiuri, P. Perrotta, A. Augimeri, F. Rocca, A. Quattrone, and A. Cherubini. Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging : A review. *Neuroinformatics*, 13(3) :261–276, 2015.
- [3] E. Carlinet and T. Géraud. A comparative review of component tree computation algorithms. *IEEE Trans. on Image Processing*, 23(9) :3885–3895, 2014.
- [4] M.-P. Dubuisson and A. K. Jain. A modified Hausdorff distance for object matching. In *Proc. of the Intl. Conf. on Pattern Recognition*, volume 1, pages 566–568. IEEE, 1994.
- [5] D. García-Lorenzo, S. Francis, S. Narayanan, D. L. Arnold, and D. L. Collins. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical Image Analysis*, 17(1) :1–18, 2013.
- [6] M. Ghafoorian et al. Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities. *Scientific Reports*, 7(5110), 2017.
- [7] M. Ghafoorian et al. Transfer learning for domain adaptation in MRI : Application in brain lesion segmentation. In *Proc. of MICCAI, Part III*, volume 10435 of *LNCS*, pages 516–524. Springer, 2017.
- [8] C. R. Jack et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. *Journal of Magnetic Resonance Imaging*, 14(6) :668–676, 2001.
- [9] R. Jones. Component trees for image filtering and segmentation. In E. Coyle, editor, *Proc. of the IEEE Workshop on Nonlinear Signal and Image Processing*, Mackinac Island, 1997.
- [10] R. Khayati, M. Vafadust, F. Towhidkhan, and M. Nabavi. Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and Markov random field model. *Computers in Biology and Medicine*, 38(3) :379–390, 2008.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Image-net classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [12] G. Lazzara, T. Géraud, and R. Levillain. Planting, growing and pruning trees : Connected filters applied to document image analysis. In *Proc. of the IAPR Intl. Workshop on Document Analysis Systems (DAS)*, pages 36–40, Tours, France, 2014.
- [13] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proc. of IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- [14] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. V. Gool. Deep retinal image understanding. In *Proc. of MICCAI, Part II*, volume 9901 of *LNCS*, pages 140–148. Springer, 2016.
- [15] A. Meijster and M. H. F. Wilkinson. A comparison of algorithms for connected set openings and closings. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(4) :484–494, 2002.
- [16] B. Morel, Y. Xu, A. Virzi, T. Géraud, C. Adamsbaum, and I. Bloch. A challenging issue : Detection of white matter hyperintensities on neonatal brain MRI. In *Proc. of the Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 93–96, 2016.
- [17] L. Pantoni. Cerebral small vessel disease : From pathogenesis and clinical characteristics to therapeutic challenges. *The Lancet Neurology*, 9(7) :689–701, 2010.
- [18] O. Ronneberger, P. Fischer, and T. Brox. U-net : Convolutional networks for biomedical image segmentation. In *Proc. of MICCAI, Part III*, volume 9351 of *LNCS*, pages 234–241. Springer, 2015.
- [19] P. Salembier and J. Serra. Flat zones filtering, connected operators and filters by reconstruction. *IEEE Trans. on Image Processing*, 3(8) :1153–1160, 1995.
- [20] P. Salembier and M. H. Wilkinson. Connected operators. *IEEE Signal Processing Magazine*, 26(6) :136–157, 2009.
- [21] P. Salembier, A. Oliveras, and L. Garrido. Antiextensive connected operators for image and sequence processing. *IEEE Trans. on Image Processing*, 7(4) :555–570, 1998.
- [22] C. Schwarz, E. Fletcher, C. DeCarli, and O. Carmichael. Fully-automated white matter hyperintensity detection with anatomical prior knowledge and without FLAIR. *Information Processing in Medical Imaging*, 21 :239–251, 2009.
- [23] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [24] L. Vincent. Grayscale area openings and closings, their efficient implementation and applications. In *Proc. of the EURASIP 1st Workshop on Mathematical Morphology and its Applications to Signal Processing*, pages 22–27, Barcelona, Spain, May 1993.
- [25] J. M. Wardlaw et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *The Lancet Neurology*, 12(8) :822–838, 2013.
- [26] Y. Xu, T. Géraud, and I. Bloch. Extra materials for paper [27]. <http://publications.lrde.epita.fr/xu.17.icip>, 2017.
- [27] Y. Xu, T. Géraud, and I. Bloch. From neonatal to adult brain MR image segmentation in a few seconds using 3D-like fully convolutional network and transfer learning. In *Proc. of the 23rd IEEE Intl. Conf. on Image Processing (ICIP)*, pages 4417–4421, Beijing, China, September 2017.