



UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG
Année : 2021

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

**Application de l'intelligence artificielle à la segmentation de lésions
ischémiques : comparaison aux méthodes manuelles et semi-
automatiques**

Présentée et soutenue publiquement le 1^{er} octobre 2021 à 16h
au Pôle Formation
par **Sevak YEPREMIAN**

JURY

Président :

Monsieur le Professeur Jean-Pierre PRUVO

Assesseurs :

Monsieur le Professeur Xavier LECLERC

Monsieur le Docteur Marco PASI

Monsieur Renaud LOPES

Directeur de thèse :

Monsieur le Docteur Grégory KUCHCINSKI

Avertissement

La Faculté n'entend donner aucune approbation aux opinions émises dans les thèses : celles-ci sont propres à leurs auteurs.

TABLE DES MATIERES

I. INTRODUCTION	1
A. L'accident vasculaire cérébral ischémique	2
1. Epidémiologie	2
2. Définition et physiopathologie	2
3. Œdème cytotoxique et vasogénique	4
B. La diffusion comme technique d'évaluation du cœur de l'infarctus	5
1. Phénomène de diffusion	5
2. L'œdème cytotoxique et vasogénique en diffusion	5
3. L'imagerie de diffusion	6
a) Principe général de l'acquisition IRM	6
b) Acquisition de la séquence de diffusion	7
4. La diffusion : séquence indispensable dans l'ischémie cérébrale	10
a) Diagnostic positif	10
b) Datation d'une lésion ischémique	11
C. Importance du volume ischémique dans la stratégie thérapeutique et le pronostic	12
1. Stratégie thérapeutique	12
2. Impact du volume dans la prise en charge	14
3. Pronostic en fonction du volume ischémique	15
D. Méthodes disponibles en clinique pour quantifier le volume ischémique	16
E. L'intelligence artificielle dans les tâches de segmentation	19
1. Intelligence artificielle : définitions	19
2. Développement d'une méthode d'IA supervisée	21
a) Les données	21
b) L'entraînement	21
c) La fonction de coût	23
3. Généralisabilité d'une application d'IA	25
4. L'IA appliquée aux tâches de segmentation	26
F. Facteurs pouvant altérer la performance des méthodes de segmentation	32
1. Maladie des petits vaisseaux	32
2. Volume lésionnel	33
3. Qualité d'examen	34
G. Objectif de l'étude	34
1. Objectif principal	34
2. Objectifs secondaires	35
II. METHODES	36
A. Population d'étude	37
B. Recueil des données cliniques	38
C. Acquisitions IRM	38
D. Segmentations	39
1. Segmentation semi-automatique à seuillage	39
2. Segmentation par un algorithme d'IA	40

3.	Evaluation des méthodes de segmentation	42
E.	Paramètres pouvant influencer les performances de segmentation.....	43
1.	Trophicité cérébrale et hypersignaux de la substance blanche	43
2.	Volume lésionnel	43
3.	Qualité de la séquence IRM.....	44
F.	Statistiques	44
III.	RESULTATS	46
A.	Population d'étude	47
B.	Comparaison au standard de référence	48
C.	Corrélations aux différents paramètres d'évaluation	53
1.	Scores cliniques	53
2.	Volume lésionnel.....	54
3.	Maladie des petits vaisseaux	56
4.	Qualité d'examen	58
IV.	DISCUSSION	60
A.	Rappel des résultats principaux.....	61
1.	Méthodes de segmentation	61
a)	Deep Learning.....	61
b)	Méthode manuelle	62
c)	Méthode semi-automatique	62
2.	Paramètres pouvant altérer les performances	63
B.	Segmentation de lésions ischémiques dans la littérature	64
C.	Maladie des petits vaisseaux et diffusion	68
D.	Forces des méthodes (semi-)automatiques	69
E.	Limites des méthodes (semi-)automatiques.....	70
V.	CONCLUSION.....	73
VI.	BIBLIOGRAPHIE	75

SOMMAIRE DES FIGURES

Figure 1. Zones d'oligémie bénigne, de nécrose et de pénombre en fonction du temps et du débit sanguin dans l'AVC aigu.....	4
Figure 2. Séquence de diffusion selon Stejskal et Tanner	8
Figure 3. Exemple d'une séquence de diffusion.....	9
Figure 4. Les quatre séquences d'un protocole IRM d'une suspicion d'AVC aigu.....	10
Figure 5. Evolution selon le temps du signal en diffusion et ADC d'une lésion ischémique. .	12
Figure 6. L'imagerie de perfusion	14
Figure 7. Score ASPECTS en scanner.	17
Figure 8. Sous-types d'intelligence artificielle	20
Figure 9. Principe de la cross-validation.	23
Figure 10. La descente de gradient.	24
Figure 11. La descente de gradient et les minima locaux.	24
Figure 12. Représentation graphique de l'optimisation d'un algorithme d'IA.	26
Figure 13. Volume d'acquisition d'une séquence IRM	27
Figure 14. Machine Learning et Deep Learning.	28
Figure 15. Représentation numérique d'une image.	29
Figure 16. Exemple d'opération de convolution.	30
Figure 17. Réseau U-net.	31
Figure 18. Développement du modèle d'IA par apprentissage profond.....	41
Figure 19. Diagramme de flux de la population d'étude	47
Figure 20. Graphiques de Bland-Altman pour chaque méthode de segmentation.....	50
Figure 21. Graphiques de Bland-Altman pour chaque méthode de segmentation, centrés sur la majorité des points.	51
Figure 22. Moyennes de Dice et de distances des barycentres entre méthodes.....	55
Figure 23. Exemple d'une bonne concordance entre les segmentations	59
Figure 24. Exemple d'une mauvaise concordance entre les segmentations	59

SOMMAIRE DES TABLEAUX

Tableau 1. Signal en diffusion d'une lésion ischémique	11
Tableau 2. Paramètres de perfusion et altérations	13
Tableau 3. Distribution en terciles des volumes mesurés.....	45
Tableau 4. Volumes médians segmentés par les différentes méthodes.....	48
Tableau 5. Corrélation de Spearman, coefficient de Dice et distances entre barycentres des volumes segmentés par les différentes méthodes par rapport au Gold Standard	48
Tableau 6. Corrélations entre le volume segmenté et le score NIHSS initial	53
Tableau 7. Corrélations entre le volume segmenté et le score mRS à 6 mois.....	54
Tableau 8. Corrélation aux hypersignaux de la substance blanche selon le tertile	56
Tableau 9. Corrélation à la trophicité cérébrale selon le tertile	57
Tableau 10. Corrélation à la qualité d'examen	58
Tableau 11. L'IA dans les tâches de segmentation de lésions ischémiques cérébrales dans la littérature récente.	65
Tableau 12. Classement SISS en août 2021, trié par Dice décroissant	67
Tableau 13. Classement SPES en août 2021, trié par Dice décroissant	67

LISTE DES ABREVIATIONS

ADC	Coefficient de diffusion apparent (Apparent Diffusion Coefficient)
ASPECTS	Alberta Stroke Program Early CT Score
ATLAS	Anatomical Tracings of Lesions After Stroke
AVC	Accident vasculaire cérébral
BET	Brain extraction tool
CBV	Volume sanguin cérébral (Cerebral Blood Volume)
CNN	Réseau de neurones convolutifs (Convolutional Neural Network)
CTP	Imagerie de perfusion par tomodensitométrie
DEFUSE	Diffusion and Perfusion Imaging Evaluation for Understanding Stroke Evolution
DL	Deep Learning ou apprentissage profond
DSC	Débit sanguin cérébral
DWI	Imagerie de diffusion (Diffusion weighted imaging)
fIRM	Imagerie fonctionnelle par résonance magnétique
FLAIR	Séquence d'inversion-récupération avec atténuation du signal de l'eau (Fluid Attenuation Inversion Recovery)
FOV	Champs de vue (Field Of View)
GMM	Modèle de mélange gaussien (Gaussian Mixture Model)
IA	Intelligence artificielle
IQCODE	Informant Questionnaire on Cognitive Decline in the Elderly
IQR	Intervalle interquartile
IRM	Imagerie par résonance magnétique
IRMp	Imagerie de perfusion par résonance magnétique
ISLES	Ischemic Stroke Lesion Segmentation
LADIS	Leucoaraiosis and Disability
LINDA	Lesion Identification with Neighborhood Data Analysis
ML	Machine Learning ou apprentissage automatique
mRS	Echelle de Rankin modifiée (modified Rankin Scale)
MTT	Temps de transit moyen (Mean Transit Time)
NIHSS	National Institute of Health Stroke Score
NN	Réseau de neurones artificiels (Neural Network)
rCBF	Débit sanguin cérébral relatif (relative Cerebral Blood Flow)
RF	Radiofréquence
SISS	Sub-acute Ischemic stroke lesion segmentation
SPES	Acute stroke outcome/penumbra estimation
SVM	Support Vector Machine
SWI	Séquence de susceptibilité magnétique (Susceptibility Weighted Imaging)
TDM	Tomodensitométrie
TE	Temps d'écho
TOF	Séquence en temps de vol (Time Of Flight)
TR	Temps de répétition
TTP	Temps au pic (Time To Peak)

Résumé

Contexte : La quantification du volume d'un infarctus cérébral est un élément important dans la prise en charge thérapeutique et l'évaluation pronostique d'un accident vasculaire cérébral ischémique. L'objectif de notre étude est d'analyser les performances d'un algorithme d'intelligence artificielle (IA) dans la segmentation de ces lésions par rapport aux méthodes manuelles et semi-automatiques, ainsi que d'évaluer certains paramètres pouvant impacter leur performance.

Méthode : Les analyses portent sur 164 patients présentant un AVC ischémique de moins de 72h, issus de la cohorte Strokedem. Cinq méthodes de segmentation sont évaluées par rapport à une méthode de référence : une méthode manuelle, trois méthodes semi-automatiques (par défaut (A), avec sélection des coupes d'intérêt (B), avec sélection des coupes d'intérêt et correction des lésions (C)) et un algorithme d'IA développé pour l'étude. Leur performance était évaluée par le calcul du coefficient de Spearman, le coefficient de Dice et la distance des barycentres lésionnels. Les paramètres pouvant impacter les performances évaluées sont la maladie des petits vaisseaux, le volume lésionnel et la qualité d'examen.

Résultats : Les volumes médians segmentés sont de 1,83 [0,40-9,41] mL par le Gold Standard, 1,80 [0,43-8,16] mL par la méthode manuelle, 2,86 mL [0,50-10,84] par la méthode d'IA, 22,06 [10,06-37,03] mL par la méthode semi-automatique A, 2,55 [0,70-7,73] mL pour la méthode semi-automatique B et 1,49 [0,24-4,83] mL pour la méthode semi-automatique C. Les corrélations des volumes segmentés par rapport au Gold Standard sont élevées pour toutes les méthodes sauf pour la méthode semi-

automatique A : méthode manuelle 0,954 ($p < 0,001$) ; IA 0,894 ($p < 0,001$) ; méthode semi-automatique A 0,213 ($p = 0,007$), méthode semi-automatique B 0,853 ($p < 0,001$), méthode semi-automatique C 0,896 ($p < 0,001$). Un impact significatif du volume et de la qualité d'examen a été retrouvé pour la segmentation par l'IA, avec une baisse des performances pour les petits volumes et pour les examens de qualité altérée.

Conclusion : Les performances de segmentation observées étaient proches entre la méthode manuelle, l'IA et la méthode semi-automatique C. La méthode manuelle reste globalement supérieure en terme de reproductibilité et de performance pour toute taille de lésion.

I. INTRODUCTION

A. L'ACCIDENT VASCULAIRE CEREBRAL ISCHEMIQUE

1. Epidémiologie

Les accidents vasculaires cérébraux (AVC) touchent plus de 140.000 personnes par an en France et on estime qu'une personne sur six sera victime d'un AVC dans sa vie. Dans 80% des cas, ils sont d'origine ischémique et dans 20% des cas hémorragique. L'AVC est une des principales causes de mortalité en France et dans le monde. Chez la femme, c'est devenu la première cause de mortalité, avant le cancer du sein. Chez l'homme, il est au troisième rang des causes de mortalité mais est responsable de deux fois plus de décès que la femme à âge égal. Près de 90% des cas d'AVC touchent les personnes de plus de 50 ans. Les principaux facteurs de risques chez l'adulte sont l'hypertension artérielle, le tabagisme, la dyslipidémie, le diabète, l'obésité et l'arythmie cardiaque.(1,2)

Les conséquences d'un AVC sont généralement lourdes, avec une mortalité de 30% à 30 jours et seulement 40% des patients reprennent une activité professionnelle à plein temps par la suite. L'AVC est la première cause de handicap acquis et la deuxième cause de démence. 75% des survivants présentent un handicap physique et un AVC sur deux entraîne un handicap moteur ou un trouble du langage.(1,2)

2. Définition et physiopathologie

La Haute Autorité de Santé définit l'AVC par l'arrêt brutal de la circulation sanguine au niveau d'une partie du cerveau. L'AVC ischémique est secondaire à une artère intracrânienne occluse classiquement par un embolie de caillot sanguin ou de l'athérome. Les autres causes d'AVC ischémique sont la maladie des petits vaisseaux,

la dissection artérielle et d'étiologie indéterminée. L'AVC hémorragique est secondaire à une rupture d'un vaisseau intracrânien, sur une lésion vasculaire telle qu'un anévrisme ou une malformation vasculaire(3). Nous développerons dans cet exposé uniquement l'AVC ischémique.

L'interruption de la circulation sanguine brutale est responsable d'un déficit clinique aigu. Le déficit persiste si la circulation sanguine n'est pas rétablie (infarctus constitué), alors qu'il régresse le cas contraire (accident ischémique transitoire). L'infarctus se constitue progressivement avec une zone de nécrose centrale où l'ischémie est irréversible, entourée d'une zone de pénombre ischémique où le débit sanguin cérébral chute suffisamment pour occasionner une souffrance cérébrale avec une traduction clinique mais reste réversible si la circulation est rétablie à temps.(4) Cette zone de pénombre est la cible principale des traitements que constituent la thrombolyse (administration intraveineuse d'un agent thrombolytique) et/ou la thrombectomie (geste endovasculaire visant à rétablir la circulation en retirant mécaniquement le thrombus)(5). La prise en charge thérapeutique doit être la plus précoce possible, d'où l'adage « *Le temps, c'est du cerveau* » (« *Time is brain* » en anglais).

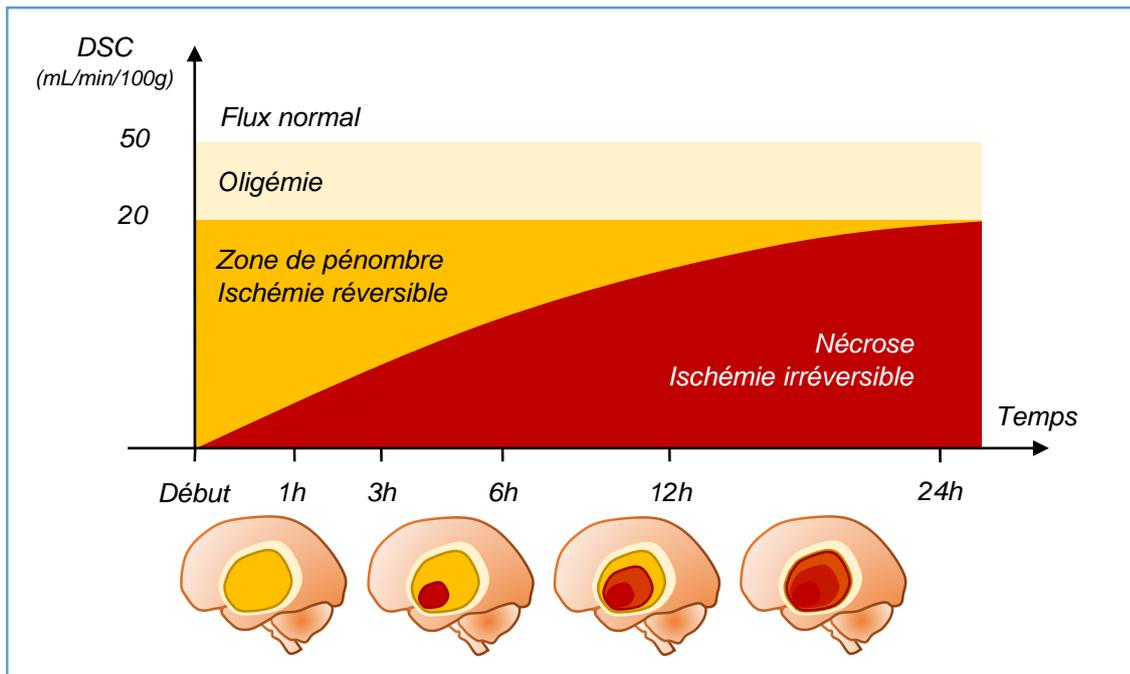


FIGURE 1 . Zones d'oligémie bénigne, de nécrose et de pénombre en fonction du temps et du débit sanguin dans l'AVC aigu. DSC : Débit sanguin cérébral.

3. Œdème cytotoxique et vasogénique

Un œdème cytotoxique s'installe au sein du foyer ischémique en lien avec la nécrose progressive du parenchyme cérébral. L'œdème cytotoxique correspond à l'augmentation intracellulaire du contenu en eau et en sodium dû à la défaillance des canaux ioniques cellulaires suite au déficit énergétique causé par l'ischémie cérébrale. Ce phénomène apparaît précocement après l'arrêt de la circulation sanguine. S'en suit un œdème vasogénique qui correspond à l'accumulation de liquide plasmatique riche en protéines secondaire à l'altération de la barrière hémato-encéphalique. Ce phénomène est maximal à 48-96h du début de l'ischémie.(6)

En IRM, les séquences de diffusion vont permettre de visualiser précocement et de façon efficace l'œdème cytotoxique afin de poser le diagnostic d'ischémie aiguë.(7) Les propriétés techniques de ces séquences seront abordées ci-dessous.

B. LA DIFFUSION COMME TECHNIQUE D’EVALUATION DU CŒUR DE L’INFARCTUS

1. Phénomène de diffusion

L'imagerie de diffusion a comme particularité d'étudier le mouvement aléatoire des molécules d'eau extracellulaires lié à l'agitation thermique (mouvement brownien). Le contraste généré par cette technique est influencé par la plus ou moins grande diffusivité des molécules d'eau dans les régions étudiées.(8) Les structures hydriques tel que le liquide cébrospinal ne présentent pas d'obstacle à la libre diffusion des particules et apparaissent en hyposignal : on parle de diffusion isotrope avec des vecteurs identiques dans toutes les directions de l'espace. Certaines structures normales du parenchyme telles que les faisceaux de substance blanche constituent des barrières à la libre diffusion des molécules : on parle de diffusion anisotrope avec une diffusion restreinte dans certaines directions de l'espace. Cette anisotropie du parenchyme cérébral peut être observée sur les images de diffusion de chaque gradient appliqué sélectivement, avant d'être moyennés pour obtenir l'image « trace » isotrope, comme expliqué par la suite.(9)

2. L'œdème cytotoxique et vasogénique en diffusion

Au sein d'un foyer de nécrose ischémique, le mouvement moléculaire est restreint. Cette restriction de diffusion serait due à la diminution de l'espace extracellulaire suite à l'œdème intracellulaire des cellules infarctées (œdème cytotoxique), freinant la libre circulation des molécules d'eau extracellulaires. Au contraire, l'œdème extracellulaire (vasogénique) est responsable d'une augmentation

de la diffusivité des molécules d'eau, se traduisant par une augmentation du signal en ADC et un iso- ou léger hypersignal en séquence de diffusion. Cet œdème est réversible contrairement à l'œdème cytotoxique.(9)

Le seuil en ADC de $0,615 \times 10^{-3} \text{ mm}^2/\text{s}$ a été proposé pour distinguer les lésions réversibles (> seuil) des lésions irréversibles (< seuil) après thrombolyse précoce.(10) Cependant, la seule évaluation de la valeur d'ADC initiale semble ne pas pouvoir prédire de façon fiable l'évolution du parenchyme, dépendant également du temps et de la perfusion.(11) Une étude a par ailleurs montrée dans le cadre des accidents ischémiques transitoires que les anomalies en diffusion et ADC pouvaient être selon les cas totalement ou partiellement réversibles, notamment en périphérie des hypersignaux diffusion.(12)

3. L'imagerie de diffusion

a) Principe général de l'acquisition IRM

L'imagerie par résonance magnétique est basée sur les propriétés magnétiques des atomes d'hydrogène présentes dans le corps humain. L'atome d'hydrogène est constitué d'un seul proton, élément subatomique de charge positive. Le patient est introduit dans le tunnel de la machine d'IRM où règne un champ magnétique intense (nommé B_0) permettant d'aligner tous les protons contenus dans le champ dans un système en équilibre. On aboutit aux différentes séquences d'IRM par une combinaison d'ondes de radiofréquence et de gradients qui vont déséquilibrer le système et permettre aux noyaux d'hydrogène d'entrer en résonance. Leur retour à

l'état normal d'équilibre produit une variation du champ magnétique qui est captée par l'antenne placée sur la partie du corps étudiée et traduite en signal électrique. (8,13,14)

b) Acquisition de la séquence de diffusion

Grâce à l'imagerie de diffusion, l'atténuation du signal est le reflet du mouvement des molécules d'eau. Elles dérivent de séquences pondérées en T2, le plus souvent en écho de spin - écho planar ultra-rapide (SE-EPI). La séquence de diffusion a été décrite en 1965 par Stejskal et Tanner(15) : un premier gradient de diffusion (gradient de déphasage) est appliqué après l'émission de l'onde de radiofréquence de 90° d'une séquence de spin écho, puis un second gradient de diffusion identique (gradient de rephasage) est appliqué après une onde de radiofréquence de 180° inversant le déphasage. Si les molécules circulent librement, les atomes d'hydrogène déphasés pendant l'application du premier gradient ne subiront pas de rephasage pendant l'application du second gradient, et le signal sera atténué. Si le mouvement moléculaire est restreint, les atomes d'hydrogène déphasés pendant le premier gradient subiront un rephasage pendant le second, et le signal sera préservé.(9)

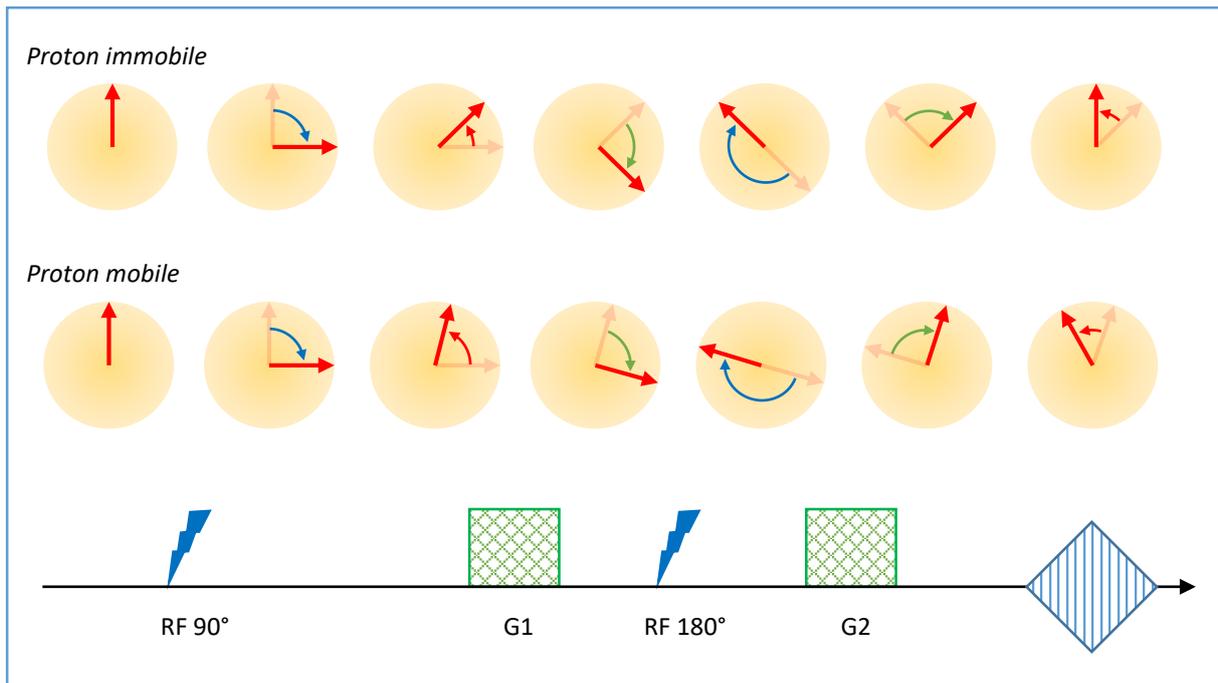


FIGURE 2. Séquence de diffusion selon Stejskal et Tanner

Onde de radiofréquence initiale (RF 90°), 1^{er} gradient de déphasage (G1), onde de radiofréquence d'inversion (RF 180°), 2nd gradient de rephasage (G2) et écho.

En raison de l'anisotropie de diffusion du cerveau, au minimum trois gradients de diffusion dans trois directions de l'espace doivent être appliqués pour obtenir une séquence permettant d'apprécier la diffusivité des molécules. Ainsi on obtient trois images de diffusion, une pour chaque direction (supéro-inférieure, droite-gauche, postéro-antérieure), qui sont combinées pour calculer l'image de diffusion globale appelée « image trace ». Plus il y a de gradients dans des directions différentes, plus l'évaluation de la diffusion sera précise. Les séquences en tenseur de diffusion utilisant au moins six directions permettent ainsi de cartographier le trajet des fibres nerveuses, la diffusion des molécules d'eau étant privilégiée dans le sens des fibres (concept d'anisotropie).(8,9)

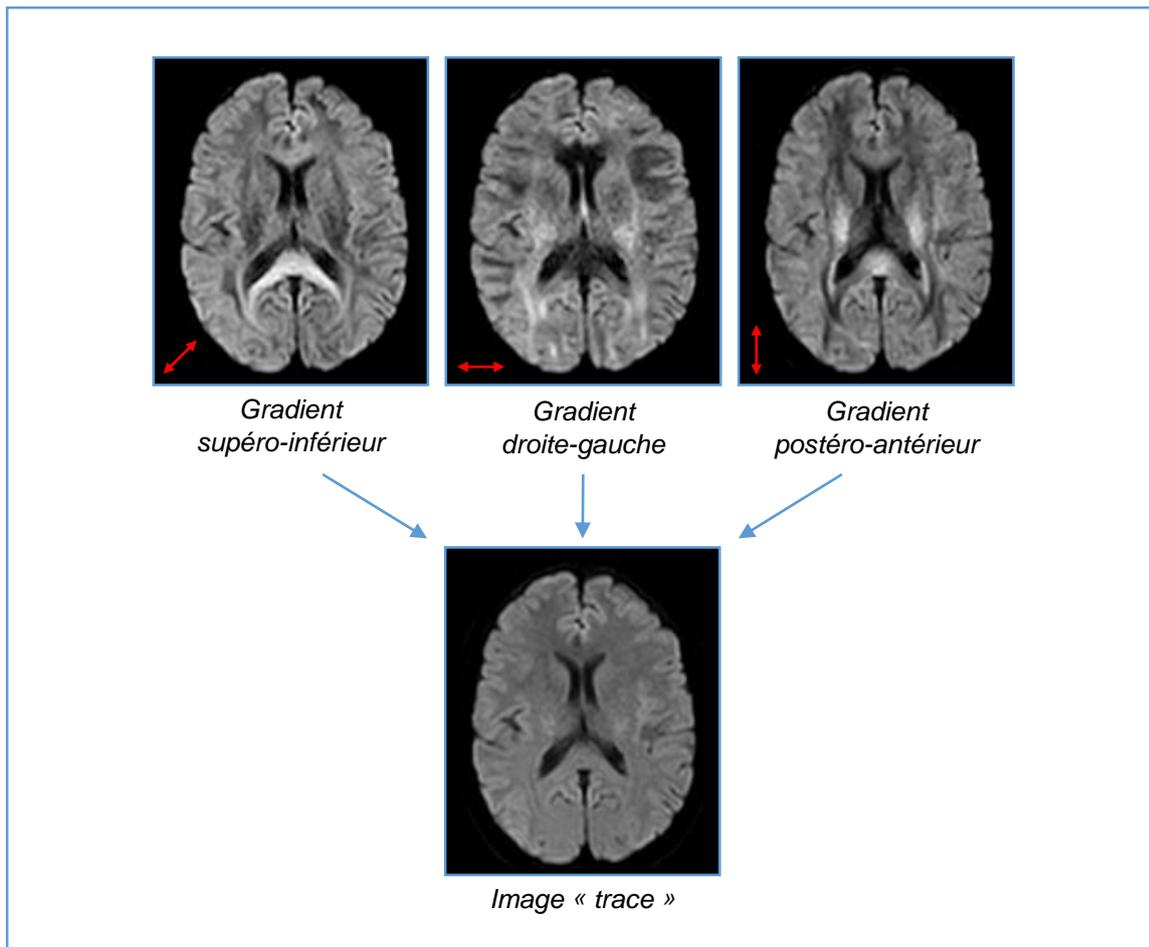


FIGURE 3. Exemple d'une séquence de diffusion montrant les acquisitions séparées des 3 gradients et l'image trace finale.

Le facteur b (appelé *facteur d'atténuation*, en s/mm^2) est une constante qui exprime le degré de pondération en diffusion de la séquence, en fonction des caractéristiques des gradients de diffusion selon la formule $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$ où γ représente le rapport gyromagnétique, G et δ l'amplitude et la durée du gradient, et Δ le délai entre les gradients.(15) L'évaluation quantitative de la diffusivité moléculaire est possible grâce à la cartographie du coefficient de diffusion apparent (ADC), calculé à partir de plusieurs séquences de diffusion à b différent et qui correspond à la pente de la courbe d'atténuation du signal entre différents b . En pratique courante sont utilisés un facteur $b = 0 s/mm^2$ (image T2 sans pondération de diffusion) et $b = 1000 s/mm^2$ (image avec pondération en diffusion).(8,9)

4. La diffusion : séquence indispensable dans l'ischémie cérébrale

a) Diagnostic positif

L'imagerie par résonance magnétique (IRM) cérébrale est l'imagerie de référence en France devant toute suspicion d'AVC. La séquence de diffusion avec cartographie du coefficient d'atténuation fait partie des quatre séquences classiques du protocole IRM réalisé en cas de suspicion d'AVC aigu, incluant également une séquence d'inversion-récupération avec atténuation du signal de l'eau (FLAIR, *fluid-attenuated inversion recovery*), une séquence en écho de gradient pondérée T2 (T2*) ou une séquence de susceptibilité magnétique (SWI, *susceptibility weighted imaging*) et une séquence d'angiographie en temps de vol (TOF, *time of flight*)⁽¹⁶⁾.

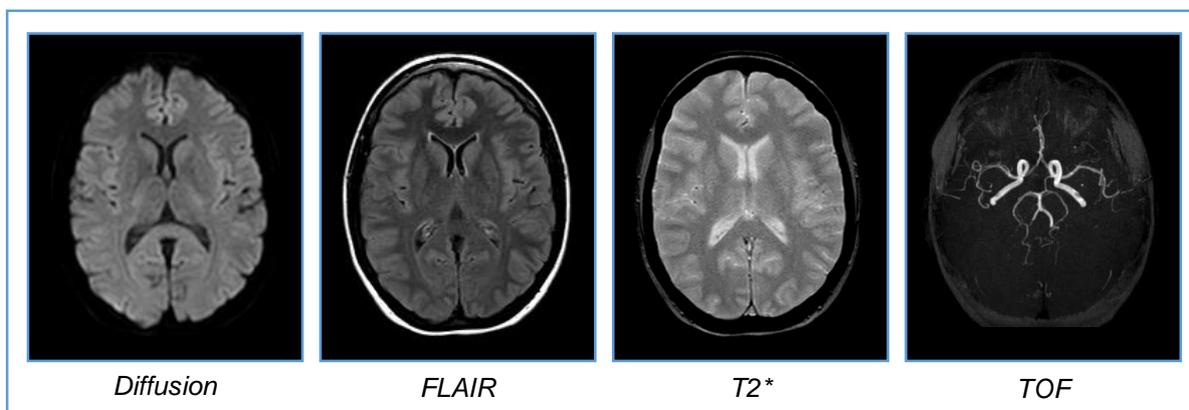


FIGURE 4. Les quatre séquences d'un protocole IRM d'une suspicion d'AVC aigu.

Une lésion ischémique aiguë apparaît précocement en hypersignal diffusion à b élevé et en hyposignal sur la cartographie ADC : on parle de *restriction de la diffusion*. Ces anomalies en diffusion apparaissent avant l'hypersignal en séquence FLAIR. Cette caractéristique fait que la séquence de diffusion est la séquence la plus performante pour le diagnostic positif devant une suspicion d'AVC ischémique.

L'interprétation d'une suspicion de lésion ischémique cérébrale se fait sur la cartographie ADC, l'image « trace » étant surtout utilisée pour la détection des lésions. Certaines lésions apparaissent en hypersignal à b élevé et en hypersignal sur la cartographie ADC, sans restriction de la diffusion : on parle d'*effet T2 (T2 shine-through)*.(17)

b) Datation d'une lésion ischémique

Le couple séquence de diffusion / cartographie ADC permet de séparer les lésions ischémiques en lésions aiguë / subaiguë / chronique selon le signal des lésions sur ces images (Tableau 1).

	<i>Diffusion à b élevé</i>	<i>ADC</i>
Lésion aiguë < 10 jours	hypersignal	hyposignal
Lésion subaiguë 10 jours à 2-3 mois	hypersignal	hypersignal
Lésion chronique > 2-3 mois	iso-hyposignal	hypersignal

TABLEAU 1. *Signal en diffusion d'une lésion ischémique*

Une lésion aiguë date classiquement de < 10 jours, une lésion subaiguë entre 10 jours et 2-3 mois et une lésion chronique au-delà.(18,19) L'évolution du signal d'une lésion ischémique est schématisée dans la Figure 5.

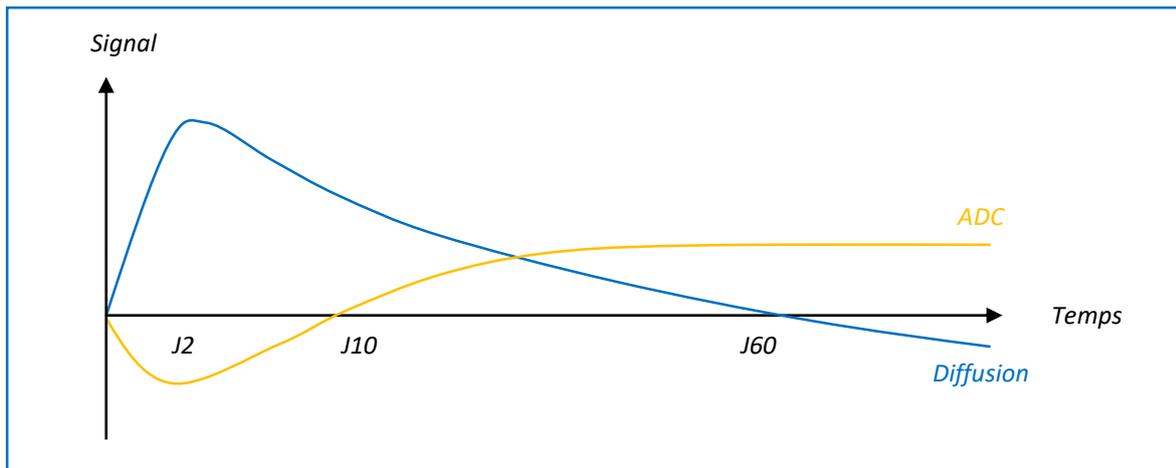


FIGURE 5. Evolution selon le temps du signal en diffusion et ADC d'une lésion ischémique.

Couplé à la séquence FLAIR et à l'imagerie de perfusion, la diffusion permet d'orienter les patients vers une prise en charge thérapeutique si l'heure de début n'est pas connue ou si le délai de 6h est dépassé. Ces points seront abordés dans le chapitre qui suit.

C. IMPORTANCE DU VOLUME ISCHEMIQUE DANS LA STRATEGIE THERAPEUTIQUE ET LE PRONOSTIC

1. Stratégie thérapeutique

En phase suraiguë d'un AVC ischémique (< 6h du début du déficit clinique), la thrombolyse intraveineuse peut être administrée jusqu'à un délai de 4h30 après le début des signes et une thrombectomie mécanique peut être réalisée jusqu'à 6h du début si l'AVC est causé par l'occlusion artérielle proximale de la circulation antérieure.(20,21)

Si l'heure de début est inconnue, par exemple lors d'un AVC du réveil, ou lorsque le délai de 6h est dépassé, le couple diffusion / FLAIR peut orienter la prise en charge. La restriction en diffusion d'une lésion associée à l'absence d'hypersignal FLAIR encourage la thrombolyse.(22–24) La superposition de l'hypersignal FLAIR aux anomalies de signal en diffusion par contre est associée à un risque hémorragique accru après thrombolyse.(25)

Au-delà des 6h, l'imagerie de perfusion est particulièrement pertinente. Les défauts perfusionnels peuvent être étudiés grâce à l'imagerie de perfusion en IRM (IRMp) ou en scanner (CTP). Ces types d'imagerie permettent de quantifier la perfusion cérébrale à l'aide de différents paramètres :

- *TTP (time to peak)* : temps au pic du bolus de produit de contraste ;
- *MTT (mean transit time)* : temps de transit moyen du bolus de produit de contraste dans le réseau capillaire cérébral ;
- *rCBF (relative cerebral blood flow)* : débit sanguin cérébral relatif ;
- *CBV (cerebral blood volume)* : volume sanguin cérébral ;
- *T-max* : temps au maximum de la fonction résiduelle.

Ces paramètres sont traités par un logiciel qui va reconstruire des cartes de perfusion, permettant de distinguer les zones d'oligémie, de pénombre et de nécrose au sein du parenchyme en fonction de ces paramètres (Tableau 2).

	<i>TTP</i>	<i>MTT</i>	<i>rCBF</i>	<i>CBV</i>	<i>T-max</i>
Oligémie	↗	↗	=	=	< 6 s
Pénombre	↗	↗	↘	↗	≥ 6 s
Nécrose	↗	↗	↘↘	↘↘	≥ 6 s

TABLEAU 2. Paramètres de perfusion et altérations selon le défaut perfusionnel

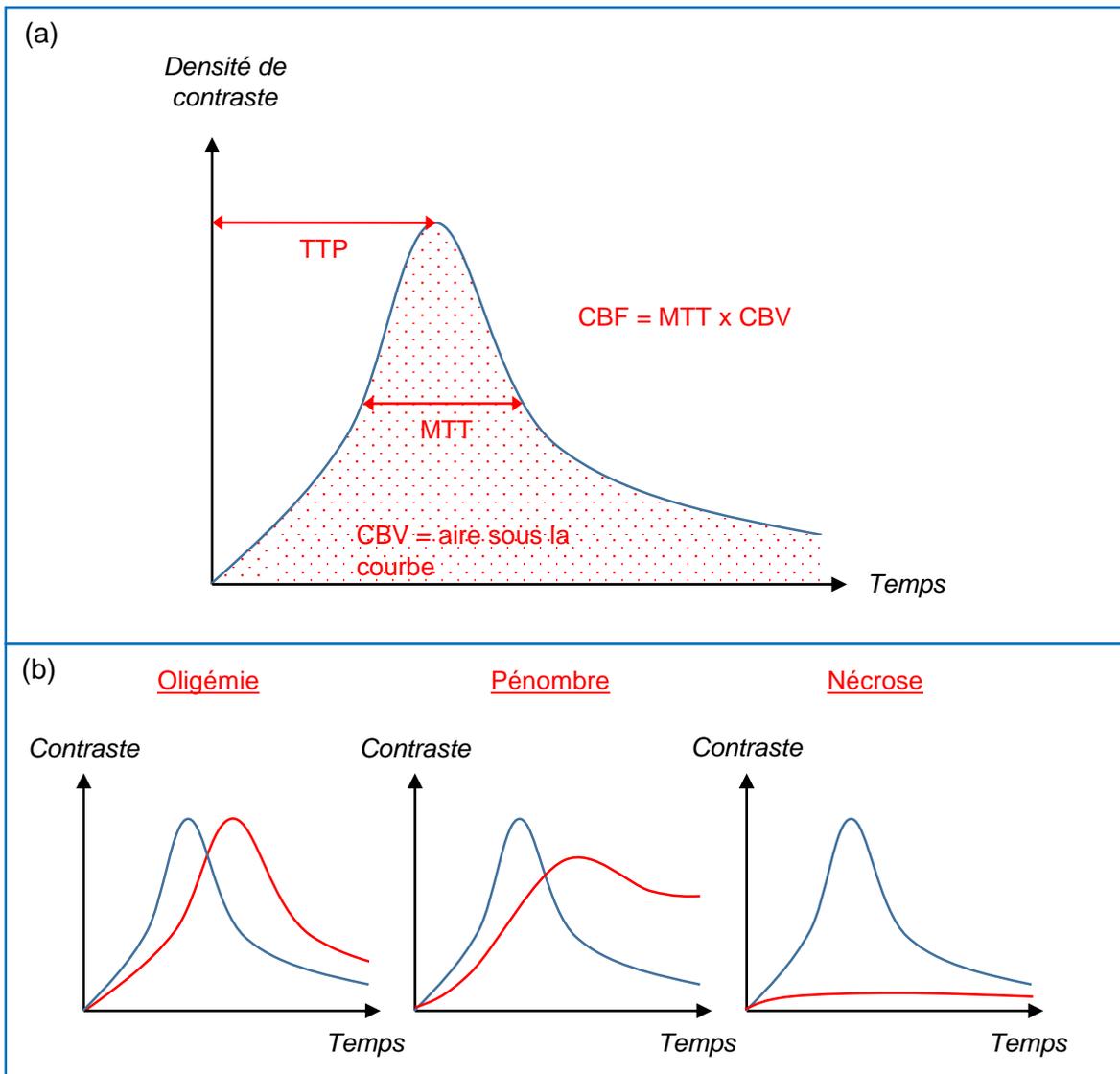


FIGURE 6. L'imagerie de perfusion

(a) Paramètres représentés sur une courbe du bolus de produit de contraste en fonction du temps. (b) Courbes de perfusion selon les altérations observables.

2. Impact du volume dans la prise en charge

L'étude DEFUSE 2 en 2012 montre que le parenchyme cérébral dont le T-max est supérieur ou égal à 6 secondes correspond le plus précisément au parenchyme à risque (zone de pénombre).(26) A l'aide de la cartographie du T-max et de la séquence de diffusion, on peut quantifier le *mismatch*. Ce terme désigne la discordance entre le

volume de parenchyme à risque potentiellement récupérable (pénombre, correspondant à un T-max ≥ 6 secondes) et le volume de parenchyme considéré comme irrécupérable (nécrose, correspondant à la zone lésionnelle en restriction de diffusion).

Si l'AVC est causé par une occlusion artérielle proximale de la circulation antérieure, que le début des signes cliniques date de plus de 6h et moins de 24h et qu'il existe une zone de pénombre significative (mismatch $> 1,8$) avec un volume infarcté < 70 ml en IRM de perfusion, alors une thrombectomie mécanique peut être indiquée.(27,28)

Un volume de nécrose important constitue par contre une mauvaise indication à la thrombectomie. On parle de « profil radiologique malin ». Ce seuil volumique est de l'ordre de 55 ml en CTP et de 70 ml en IRMp.(20,29,30)

Le volume d'infarctus cérébral guide également l'indication d'une éventuelle hémicraniectomie décompressive, dans le cadre de lésions ischémiques de grand volume. L'hémicraniectomie décompressive est recommandée dans les 48h du début des signes cliniques dans les AVC causés par une occlusion de l'artère cérébrale moyenne présentant un déficit neurologique sévère (NIHSS > 15) et un infarctus de grand volume (> 145 mL en séquence de diffusion).(31–34)

3. Pronostic en fonction du volume ischémique

Le volume du cœur ischémique est un biomarqueur en IRM corrélé au pronostic fonctionnel et au risque hémorragique. Un volume > 100 ml en séquence de diffusion ou > 70 mL en IRM de perfusion est associé à un risque hémorragique accru et un mauvais pronostic fonctionnel. Il a été démontré par ailleurs que le bénéfice clinique

de la thrombectomie mécanique en plus de la thrombolyse diminue avec l'augmentation du volume lésionnel.(27,28,35–37) Un volume d'infarctus de > 274 mL en diffusion b1000 et > 244 mL en ADC prédit un pronostic catastrophique à 1 an avec une spécificité de 90%.(31)

Concernant le risque de démence, il est difficile de décrire l'impact de la localisation et du volume lésionnel. Les patients présentant des lésions ischémiques sont associés à un risque accru de développer une démence, que ce soit des AVC déficitaires uniques ou multiples, ou des infarctus cérébraux cliniquement silencieux.(38,39)

D. METHODES DISPONIBLES EN CLINIQUE POUR QUANTIFIER LE VOLUME ISCHEMIQUE

En pratique clinique courante, le volume ischémique peut être quantifié de différentes manières.

Une volumétrie manuelle sur la séquence de diffusion peut être réalisée si aucun logiciel de post-traitement n'est disponible. Les séquences de diffusion étant constituées de coupes 2D, on peut contourner les zones lésionnelles sur chaque coupe, additionner les surfaces et multiplier le tout par l'épaisseur de coupe. Cette méthode apparaît naturellement fastidieuse dès lors qu'il y a plusieurs lésions ou que la lésion est à contours irréguliers, et est difficile à appliquer dans un contexte d'urgence.

Comme discuté plus haut, l'imagerie de perfusion permet de mesurer les zones d'oligémie, de pénombre et de nécrose. Ce type d'imagerie n'est cependant pas systématique.

Le score ASPECTS permet une évaluation rapide de l'étendue de l'ischémie dans le territoire de l'artère cérébrale moyenne. Il est constitué de 10 points, en rapport avec 10 secteurs du cerveau (voir figure 8).(40) Initialement créé pour une évaluation scanographique, ce score a également été adapté à l'IRM.(37)

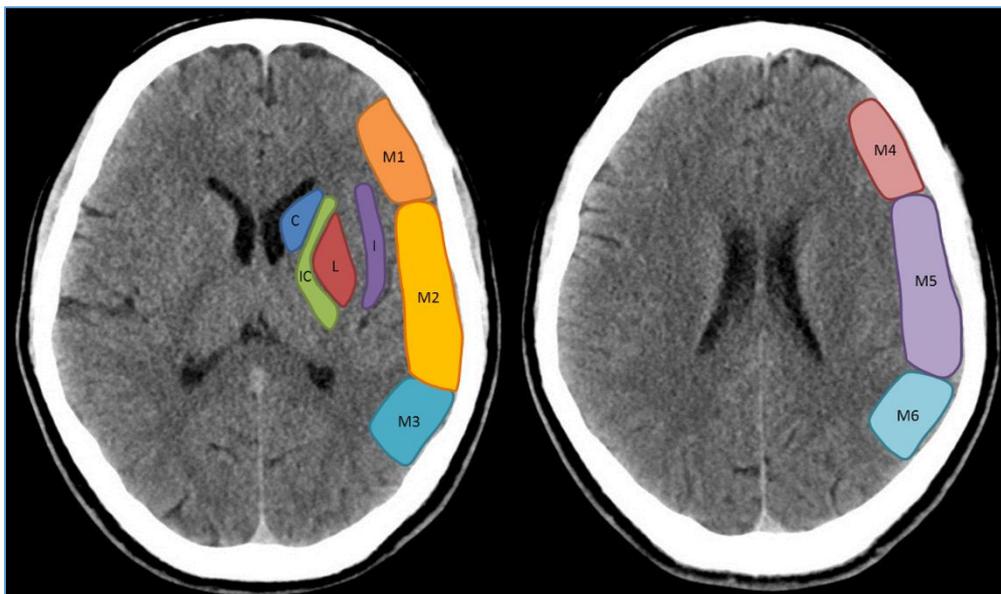


FIGURE 7. Score ASPECTS en scanner.

C : noyau caudé, L : noyau lenticulaire, IC : capsule interne, M1-6 : cortex dans le territoire de l'artère cérébrale moyenne.

Adapté d'un cas du Dr Osamah A. A. Alwalid, issu de Radiopaedia.org, rID: 72706

Plusieurs logiciels sont disponibles dans le commerce dans le cadre de l'imagerie neurovasculaire. Certains logiciels se basent sur les séquences classiques d'IRM ou l'IRM de perfusion et d'autres sur la tomodensitométrie en contraste spontané, l'angioscanner ou le scanner de perfusion. Le logiciel Oléa Médical (Oléa Médical, La Ciotat, France) utilisé dans notre étude permet d'exploiter des séquences de diffusion, mais également de perfusion IRM. Nous pouvons citer également les

logiciels Aidoc, Avicenna.AI, Brainomix, RapidAI et Viz.ai, qui ont été résumés par Soun et al. dans leur article paru en 2021(41) :

- *Aidoc* (Aidoc Medical, Tel-Aviv, Israël) est un logiciel permettant d'identifier une occlusion artérielle proximale en angioscanner et de donner des cartes de perfusion en scanner de perfusion, à l'aide de méthodes d'IA d'apprentissage profond notamment.
- *Avicenna.AI* (Avicenna AI, La Ciotat, France) est un logiciel permettant d'identifier un score ASPECTS sur examen scanographique, basé sur une méthode d'IA d'apprentissage profond.
- *Brainomix* (Brainomix Limited, Oxford, England) est un logiciel permettant d'identifier un score ASPECTS et de localiser une occlusion proximale dans le réseau de l'artère cérébrale moyenne avec une cartographie du déficit en collatérales sur un angioscanner cérébral, ainsi que d'identifier un mismatch perfusionnel en IRM ou scanner de perfusion, basé sur des méthodes d'IA d'apprentissage automatique et profond.
- *RapidAI* (iSchemaView Inc., Menlo Park, United States) est un outil permettant de donner un score ASPECTS sur examen scanographique sans injection, de détecter une occlusion artérielle proximale en angioscanner et d'identifier un mismatch perfusionnel en IRM et scanner de perfusion, basée sur des méthodes d'IA notamment d'apprentissage profond et de forêts d'arbres décisionnels.
- *Viz.ai* (Viz.ai Inc., San Francisco, United States) est un outil mobile permettant de détecter une occlusion artérielle proximale en angioscanner et d'identifier un mismatch perfusionnel en scanner de perfusion, basé sur des méthodes d'IA d'apprentissage profond.

La recherche en intelligence artificielle dans le domaine médical et notamment dans les tâches de segmentation est d'ailleurs en plein expansion, sujet que nous allons aborder maintenant.

E. L'INTELLIGENCE ARTIFICIELLE DANS LES TACHES DE SEGMENTATION

1. Intelligence artificielle : définitions

L'intelligence artificielle (IA) consiste en le développement de technologies visant à imiter une forme d'intelligence réelle. La notion d'IA a vu le jour en 1956 lors de la conférence de Dartmouth aux Etats-Unis.(42) Alain Turing, mathématicien britannique, a proposé dans les années 1950 un test (Test de Turing) pour déterminer si une machine est « intelligente » ou non : un humain est confronté à une conversation à l'aveugle avec un autre humain ou avec une machine ; s'il n'est pas capable de différencier l'interlocuteur humain de l'ordinateur, alors l'ordinateur a passé le test avec succès.(43) Lorsqu'on parle « d'apprentissage » par un ordinateur, on peut citer Arthur Samuel en 1959 qui définit l'apprentissage automatique par le « domaine d'études qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés ».(44)

Le terme « *Intelligence artificielle* » regroupe plusieurs concepts, des plus simples aux plus compliqués, et constitue un champ très large de recherche et développement. Les concepts d'IA, Machine Learning et Deep Learning ne sont pas des notions indépendantes mais s'emboîtent : le Deep Learning est une méthode de Machine Learning et le Machine Learning est une méthode d'IA. Le terme « Machine

Learning » et « apprentissage automatique » sont synonymes, le deuxième terme étant la traduction française du premier. Le Deep Learning (ou apprentissage profond) est une méthode d'apprentissage automatique permettant à l'ordinateur de résoudre des problèmes en apprenant par expérience et de comprendre le monde en terme de hiérarchie de concepts.(45) Il utilise ce qu'on appelle un réseau de neurones artificiels.

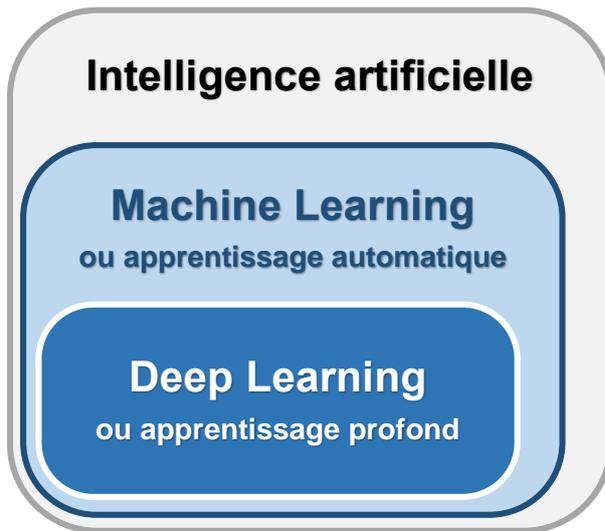


FIGURE 8. Sous-types d'intelligence artificielle

L'utilisation des réseaux de neurones artificiels a déjà été proposé par Yann LeCun en 1990 avec notamment la reconnaissance de codes postaux écrits à la main aux Etats-Unis(46), mais ce n'est qu'à partir de 2012 que l'apprentissage profond se développe de façon exponentielle grâce notamment aux évolutions technologiques avec des machines suffisamment puissantes capables de travailler sur de grandes bases de données qui elles-mêmes se développent avec notamment la base ImageNet(47).

Dans l'imaginaire symbolique actuel, l'IA serait capable de déduire et raisonner avec une logique propre et des connaissances exactes. Or, cette « intelligence » artificielle est en réalité un apprentissage à partir de données, amenant la machine à

une connaissance approximative, statistiquement optimisée. L'apprentissage peut être supervisé ou non supervisé.(45) On parle d'apprentissage supervisé lorsque les données sont labellisées, c'est-à-dire que la base de données utilisée pour développer l'algorithme contient l'objectif à atteindre. Dans le cas contraire, on parlera d'apprentissage non supervisé, lorsque les labels ne sont pas fournis. L'apprentissage non supervisé a peu d'intérêt à l'heure actuelle dans le domaine médical. Nous parlerons de l'apprentissage supervisé dans la suite.

2. Développement d'une méthode d'IA supervisée

a) Les données

Les données sont l'élément central de l'apprentissage automatique. Une base de données importante et de qualité est essentielle à un algorithme performant. La qualité des données en imagerie médicale dépend de la qualité des images, de la représentativité des patients pour une question donnée et de la généralisabilité du modèle d'étude par l'utilisation notamment de données multicentriques. L'apprentissage supervisé utilise des données annotées, c'est-à-dire que pour chaque problème est associée une solution (par exemple, la segmentation d'une lésion sur une image IRM).(48)

b) L'entraînement

Ces données annotées (ou labellisées) sont partagées en trois échantillons pour l'étape d'apprentissage en tant que tel du modèle d'IA : un premier jeu d'entraînement, un second jeu de test et un troisième jeu de validation. L'algorithme

est entraîné sur le premier échantillon de données puis ses hyperparamètres sont optimisés avec le deuxième. Une fois calibré, ses performances sont évaluées avec le dernier échantillon de validation.(48) Ceci nécessite généralement une base de données de grande taille pouvant aller au-delà de dizaines de milliers de sujets à inclure dans la base selon la problématique. Les plus grandes bases de données disponibles contiennent d'ailleurs plusieurs millions de données, comme la base ImageNet qui contenait en août 2021 14197122 images. Cependant, lorsqu'une base de données est de petite taille, l'incertitude statistique se majore. Ces étapes peuvent alors être réalisées grâce à des méthodes complémentaires telle que la cross-validation, ayant comme contrepartie un coût accru en termes de calcul.(45)

La cross-validation permet d'utiliser tous les sujets d'une base pour l'entraînement et le test de l'algorithme d'IA. Cette procédure consiste dans le fait d'utiliser plusieurs sous-ensembles choisis au hasard dans l'ensemble de la base de données. La plus courante est la k-fold cross-validation, où l'ensemble est divisé en un nombre k de sous-ensembles qui ne se chevauchent pas. A chaque essai, un sous-ensemble de test est choisi au hasard, le reste des données servant de sous-ensemble d'entraînement. L'erreur globale sera estimée par une moyenne des erreurs sur un nombre k d'essais.(45)

Une itération d'une cross-validation à 5 niveaux :

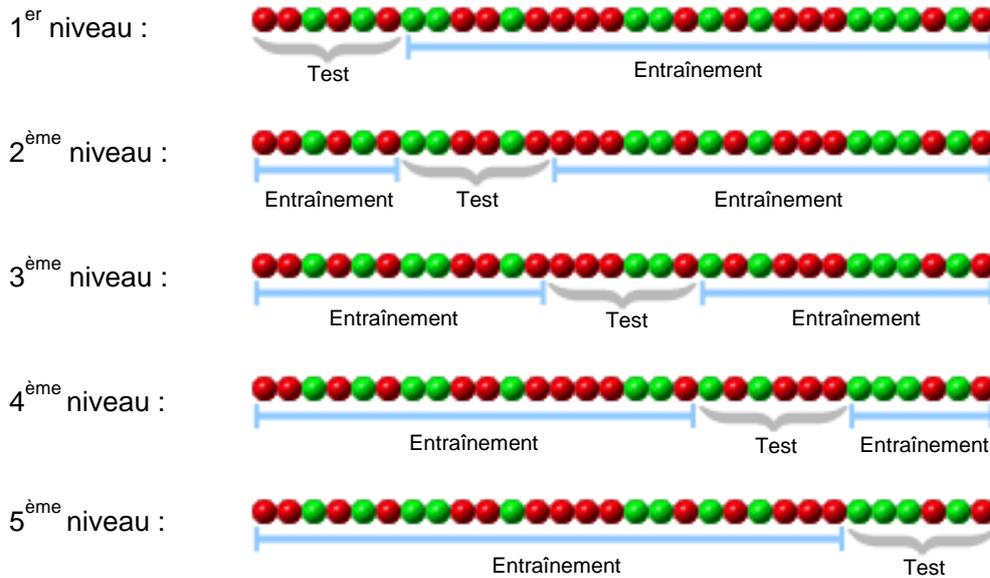


FIGURE 9. Principe de la cross-validation.

Exemple d'une cross-validation à 5 niveaux : l'ensemble est divisé en 5 sous-ensembles sans chevauchement. A chaque niveau, un sous-ensemble de test est choisi au hasard, le reste des données servant de sous-ensemble d'entraînement.

Schéma traduit et adapté de ProClassify User's Guide, Institute for Genomics and Bioinformatics - Graz University of Technology, consultable en ligne sur <https://genome.tugraz.at/proclassify/help/pages/XV.html>

c) La fonction de coût

L'optimisation d'un algorithme d'IA passe par une fonction que l'on cherche à maximiser ou à minimiser. Dans le cadre de l'apprentissage automatique, on cherche généralement à minimiser une fonction, appelée fonction de coût. Cette fonction mesure la pénalité entre une donnée de sortie prédite par l'algorithme (prédiction) et une donnée de sortie attendue (vérité). Les réseaux de neurones multicouches peuvent utiliser un algorithme de rétropropagation pour calculer le gradient de la fonction de coût, donnée indispensable à l'état de l'art des algorithmes d'optimisation qui servent à corriger automatiquement les poids du réseau. L'optimisation de

l'algorithme passe souvent par une méthode de descente de gradient afin de minimiser la fonction de coût au fur et à mesure des itérations. La descente de gradient stochastique est un algorithme d'optimisation de la fonction de coût décrite comme une somme de fonctions dérivables, largement utilisée dans le cadre du Deep Learning. Elle est utile dans le cas où la fonction de coût est une somme de fonctions dérivables. On parle de fonction de coût empirique.

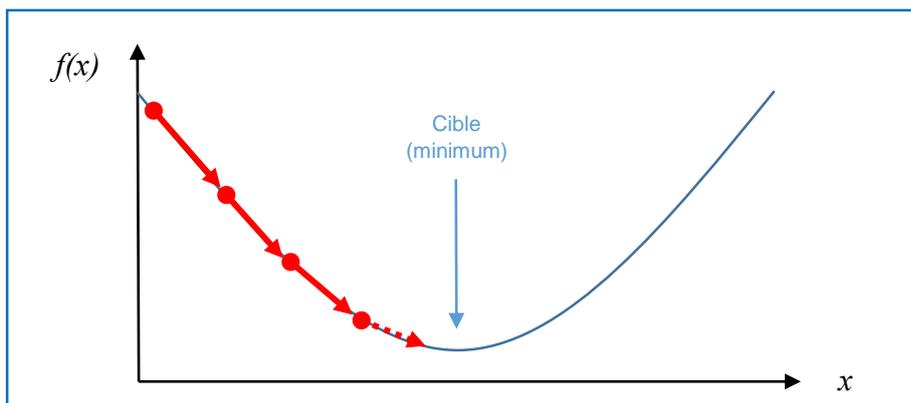


FIGURE 10. La descente de gradient.

L'algorithme de descente de gradient démarre à une coordonnée initiale arbitraire et converge vers le minimum de façon itérative.

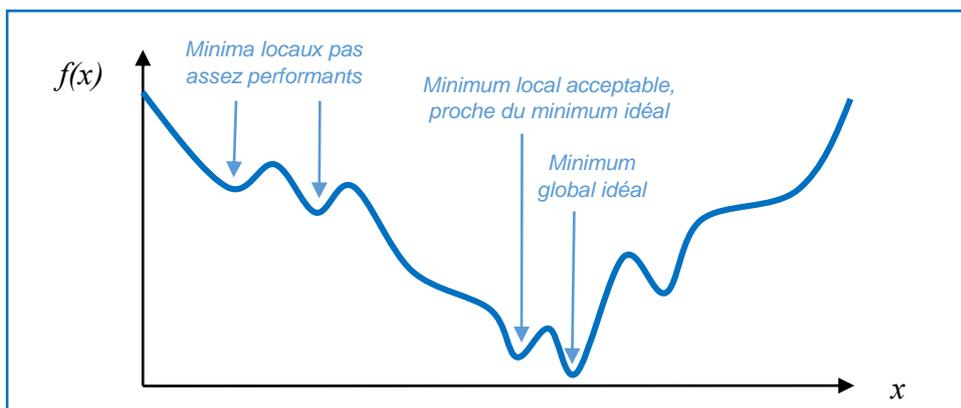


FIGURE 11. La descente de gradient et les minima locaux.

La minimisation selon la descente de gradient peut se heurter à des difficultés en cas de plusieurs minima locaux ou à des phases de plateau. Un minimum local peut être accepté si la fonction de coût présente des valeurs significativement basses.

3. Généralisabilité d'une application d'IA

Un des aspects d'une solution d'IA est la question de sa généralisabilité. Un algorithme est développé grâce à une base de données contenant un certain nombre de sujets ou d'objets, de taille plus ou moins grande. Un intérêt majeur pour cet algorithme est de pouvoir être utilisé de façon fiable en dehors de cette base, sur des données non connues. Il est alors généralisable. Cette caractéristique sera déterminée par deux facteurs, une erreur la plus faible possible pendant la phase d'entraînement et un écart le plus faible possible entre cette erreur et l'erreur lors de la phase de test. Ces deux facteurs correspondent respectivement à ce que l'on appelle le biais et la variance.

Ces deux éléments vont conduire respectivement au sous-entraînement (*underfitting*) ou au surentraînement (*overfitting*) lorsqu'ils sont trop élevés. L'*underfitting* correspond à un algorithme dont l'erreur dans la prédiction des résultats (le biais) est trop élevée suite à la phase d'entraînement : l'algorithme n'est pas suffisamment entraîné ou le modèle n'est pas suffisamment expressif. L'*overfitting* correspond à un entraînement trop poussé ou à un trop grand nombre de paramètres, réalisant des résultats excessivement bons sur l'échantillon d'entraînement mais avec des résultats mauvais sur la phase de test : l'algorithme a appris à reconnaître uniquement les sujets de l'échantillon d'entraînement.(45)

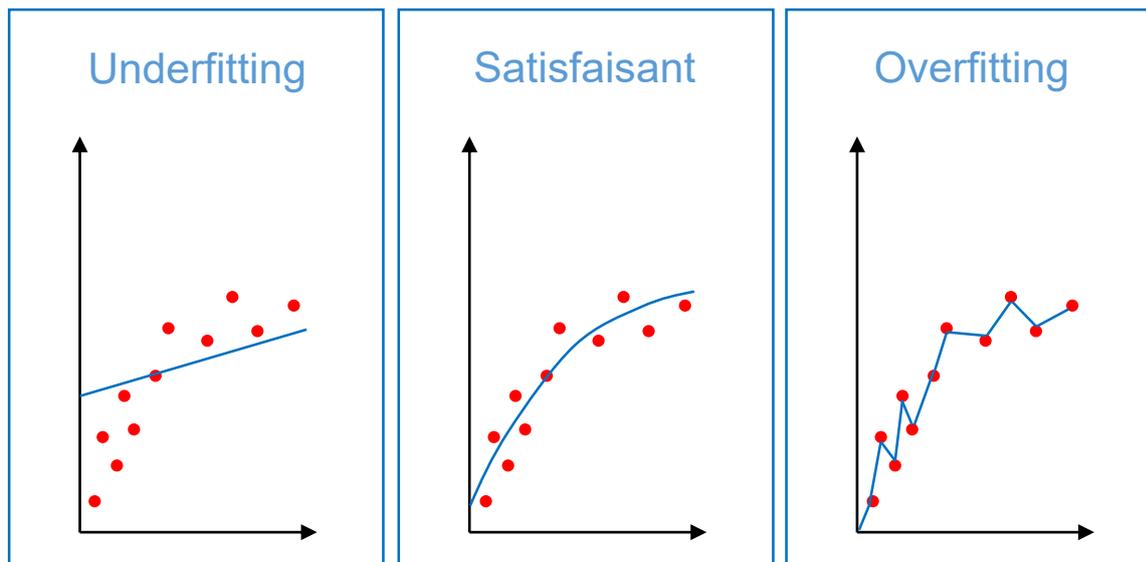


FIGURE 12. Représentation graphique de l'optimisation d'un algorithme d'IA.

L'underfitting correspond à un algorithme non suffisamment entraîné et l'overfitting à un entraînement trop poussé. La courbe du milieu correspond à un entraînement correct, ni trop ni pas assez entraîné.

4. L'IA appliquée aux tâches de segmentation

L'essor de l'intelligence artificielle ces dernières années a généré un nombre croissant de publications dans le domaine de la médecine et en particulier de la neuroradiologie.(49) Une étude questionnant l'intégration de l'IA dans le quotidien du radiologue auprès de 200 radiologues séniors et 70 internes de radiologie dans la région du Nord-Pas-de-Calais en France en 2020 a d'ailleurs montré une perception optimiste de ces outils sur leur impact dans l'exercice futur.(50)

Différents modèles de Machine Learning ont vu le jour afin d'automatiser les tâches de segmentation(51). Les méthodes d'apprentissage automatique de haut niveau ont conduit à des résultats de segmentation nettement meilleurs par rapport aux modèles plus simples, amenant vers une problématique complexe de ces

tâches(52). Les méthodes de Deep Learning ont montré des résultats particulièrement convaincants dans cette application.

Un algorithme de traitement d'images utilise une donnée d'entrée introduite dans le système qui va proposer une donnée de sortie après extraction des caractéristiques les plus informatives (par exemple, position et forme des yeux, du nez, de la bouche pour un visage humain) puis classification (par exemple, visage humain ou non) ou segmentation (par exemple, délimitation d'une lésion ischémique ou d'une tumeur). La donnée d'entrée peut correspondre à des coupes d'IRM : un certain volume d'intérêt est acquis (champ de vue) et présenté sous forme de tranches d'épaisseur égale. Une coupe d'IRM sera définie par une matrice d'une taille donnée (256x256 ; 512x512...) en deux dimensions et d'une épaisseur de coupe permettant d'individualiser des voxels en trois dimensions. Le volume d'un voxel dépend donc intimement de ces trois paramètres que sont le champ de vue, la matrice et l'épaisseur de coupe. Plus le voxel sera petit, plus la résolution spatiale sera bonne. Le contraste dépendra des paramètres d'acquisition et donc du type de séquence ainsi que de l'utilisation éventuelle de produits de contraste.

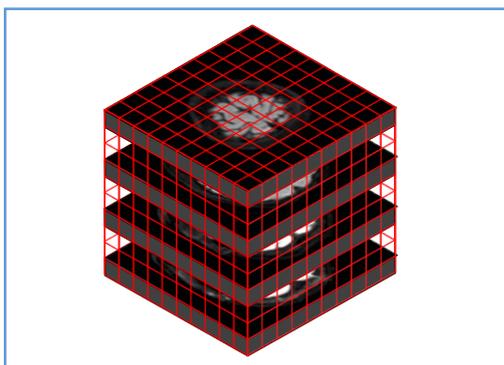


FIGURE 13. Volume d'acquisition d'une séquence IRM

Une séquence IRM 2D est constituée de coupes bidimensionnelles avec une matrice et une épaisseur de coupe. Chaque pixel en 2D d'une coupe est en réalité un voxel en 3D.

Les algorithmes de Deep Learning utilisant une architecture avec des réseaux de neurones convolutifs sont particulièrement efficaces et populaires dans le traitement d'images. Le réseau de neurones convolutifs est conçu pour apprendre

automatiquement et de manière adaptative les hiérarchies spatiales des caractéristiques via un algorithme de rétropropagation.(53) Les deux étapes d'extraction et de classification se font automatiquement sans intervention humaine, contrairement aux méthodes traditionnelles de Machine Learning notamment de Computer Vision (traitement d'images par l'IA).(41)

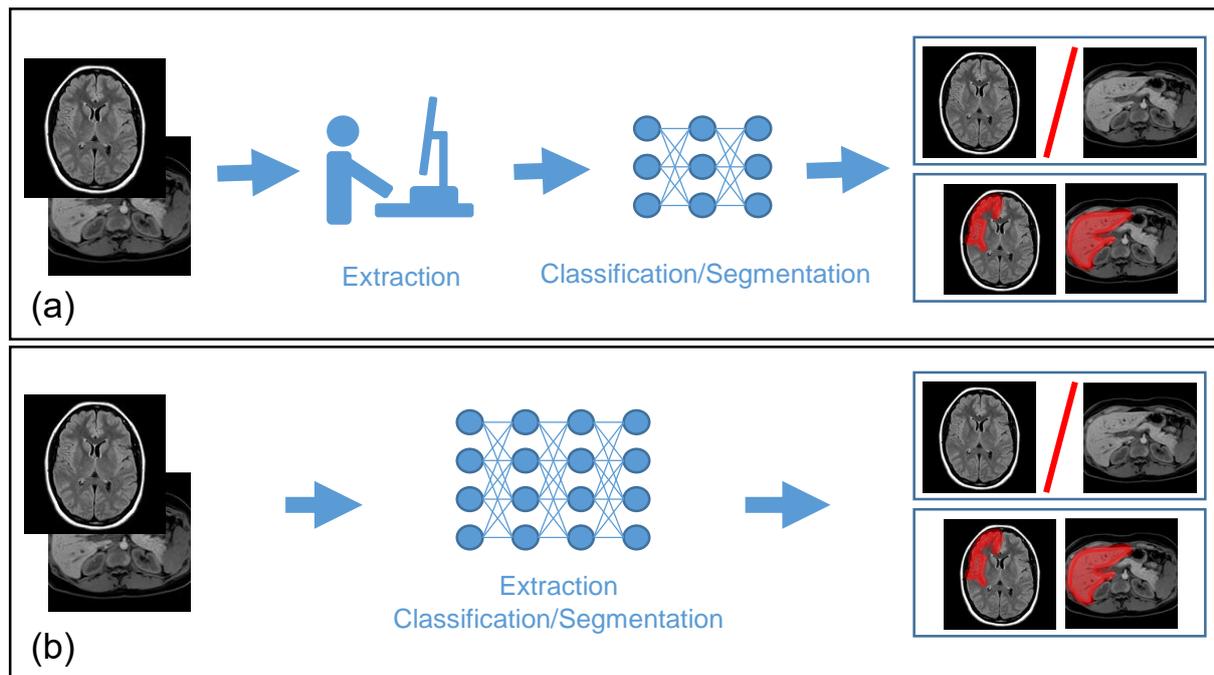


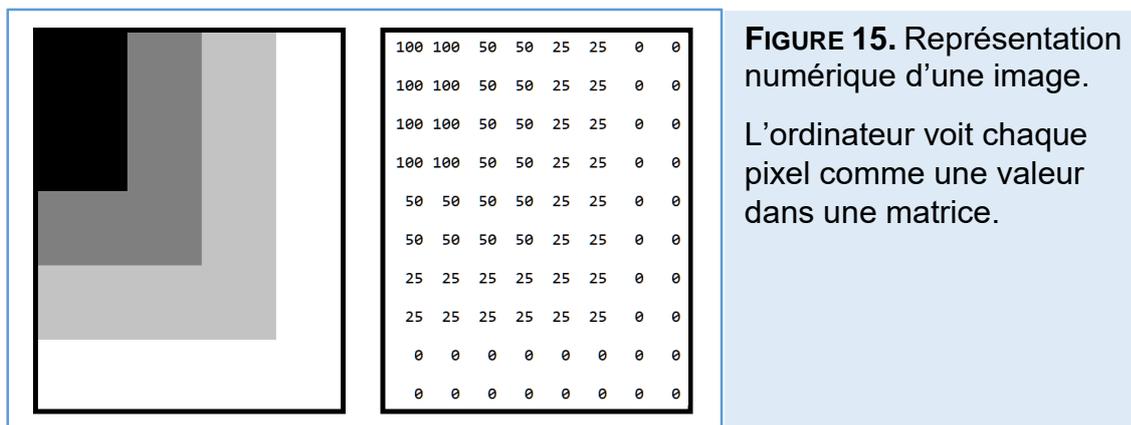
FIGURE 14. Machine Learning et Deep Learning.

Différence schématique entre des méthodes classiques de Machine Learning (a) avec une étape humaine préalable et du Deep Learning (b) où tout est automatisé.

Le radiologue lors de l'interprétation d'une imagerie extrait un ensemble d'éléments pertinent qui vont lui permettre de proposer un diagnostic. De façon similaire, un réseau de neurones artificiels va extraire les données qui lui semblent intéressantes pour une prédiction optimale en modifiant les pondérations des variables de chaque neurone au fil des itérations. Une des règles les plus anciennes sur le fonctionnement de réseaux de neurones a été proposée par Donald Hebb en 1949, avec une fameuse phrase « Cells that fire together, wire together » (en anglais, « des

cellules qui tirent ensemble, se connectent ensemble »). Elle stipule que si deux neurones interconnectés sont activés au même moment alors le poids entre eux devrait être augmenté, ou plus généralement que deux cellules d'un système qui s'activent en même temps de façon répétée, tendent à s'associer de sorte à ce que l'activité de l'une facilite l'activité de l'autre.(54,55)

Au lieu de travailler sur des images qualitatives comme le radiologue, la machine va travailler sur une matrice composée de pixels ou de voxels.(53) Sélectionner les variables pertinentes est essentiel pour l'explicabilité, l'optimisation et la rentabilité d'un modèle.(41,56) Les réseaux de neurones convolutifs permettent cette extraction de données avec une hiérarchie spatiale. A chaque convolution, l'algorithme passe au crible la matrice pour extraire une certaine donnée : une caractéristique pour chaque convolution. Pour faire ceci, plusieurs noyaux (« kernels ») fonctionnent comme extracteurs de caractéristiques, tels qu'un détecteur de bord horizontal ou de bord vertical ou un détecteur de contour.(53)



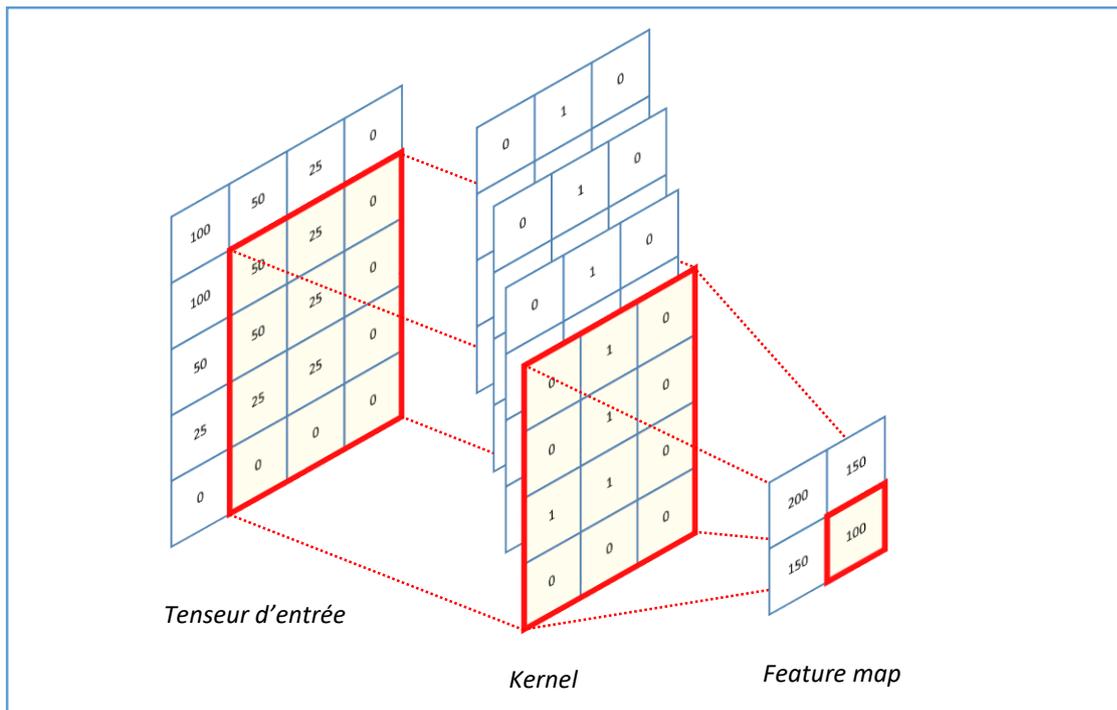


FIGURE 16. Exemple d'opération de convolution.

Un *kernel* (noyau) de 3×4 , sans *padding* (le noyau ne dépasse pas le tenseur d'entrée) et un pas de 1, parcourt le tenseur d'entrée, et réalise un produit élément par élément puis additionne ces valeurs pour donner une valeur de sortie dans la position correspondante dans le tenseur de sortie, appelée *feature map*.(43)

Un type particulier de réseau de neurones convolutif est l'architecture U-net, adapté aux tâches de segmentation, notamment lorsque la taille de l'échantillon de données d'entraînement est modérée. Il se décompose en deux grandes parties :

- La première moitié de la structure est un réseau de convolution avec des couches qui réduisent la résolution de l'image d'entrée (sous-échantillonnage). Comme tout autre réseau convolutif, sa structure hiérarchique permet d'apprendre des caractéristiques à différents niveaux de résolution.
- La seconde partie du réseau se compose de couches de convolution de sur-échantillonnage. Cette partie apprend la carte de segmentation en commençant par la résolution la plus faible puis en augmentant en résolution à chaque couche de convolution jusqu'à atteindre la résolution d'origine.

Dans le but d'avoir suffisamment de données à chaque niveau de résolution, le modèle U-net possède par ailleurs des skip-connexions qui permettent aux cartes de segmentation de chaque niveau de résolution de récupérer les détails de l'image réduite au même niveau de résolution. L'ensemble du réseau forme schématiquement un circuit symétrique descendant puis ascendant en « U », d'où son nom.(57)

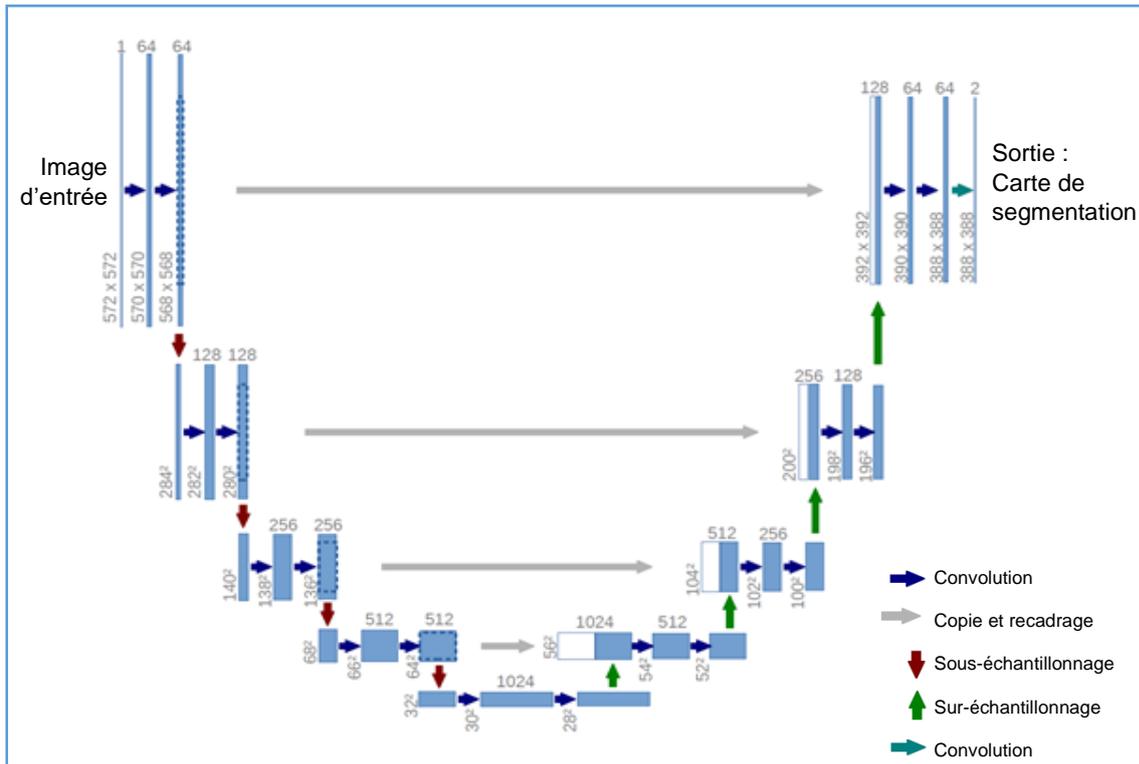


FIGURE 17. Réseau U-net.

Réseau de neurones convolutif de type U-net tel que décrit par O. Ronneberger, P. Fischer et T. Brox (52) : L'image d'entrée subit plusieurs opérations de convolution et de sous-échantillonnage puis de sur-échantillonnage. De plus, à chaque niveau de résolution, une copie est introduite du circuit descendant vers le circuit ascendant. En sortie, l'algorithme propose une carte de segmentation.

Figure traduite et adaptée de l'article original disponible en ligne sur <http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a>

F. FACTEURS POUVANT ALTERER LA PERFORMANCE DES METHODES DE SEGMENTATION

1. Maladie des petits vaisseaux

La maladie des petits vaisseaux cérébraux regroupe des processus pathologiques d'étiologies variées touchant les petites artères, artéριοles, capillaires et veinules du cerveau. Elle est fréquente chez les patients âgés et est responsable d'une multitude de manifestations cliniques et radiologiques. Plusieurs types de lésions cérébrales liées à la maladie des petits vaisseaux peuvent être identifiés en neuro-imagerie, tels que les petits infarctus sous-corticaux, les lacunes, les hypersignaux de la substance blanche, l'élargissement des espaces périvasculaires, l'atrophie cérébrale et les microhémorragies.(58) Elle est responsable au niveau cérébral d'un dommage vasculaire et d'une dérégulation de la circulation sanguine. Une ischémie diffuse chronique infraclinique s'installe avec des lésions axonales, oligodendrogiales et de démyélinisation, visible en IRM sous forme d'hypersignaux diffus en pondération T2 de la substance blanche sus-tentorielle et du pont (leucoaraïose). Une nécrose focale peut se constituer dans la substance blanche ou grise, responsable d'infarctus lacunaires visibles sous formes d'images cavitaires en pondération T1 ou FLAIR. La fragilité vasculaire est responsable de petits saignements focaux (microsaignements ou « microbleeds » en anglais), mais également d'hématomes plus larges.(59,60)

Dans le cadre de la maladie des petits vaisseaux, l'imagerie de diffusion permet d'identifier de petites lésions ischémiques récentes, cliniquement silencieuses. Ces lésions peuvent au cours du suivi laisser place à des anomalies de type lacunaire ou à des hypersignaux en pondération T2 au niveau sous-cortical pouvant fusionner par

la suite aux hypersignaux diffus de la substance blanche, contribuant à la progression de la charge lésionnelle(61,62).

L'imagerie de diffusion a permis de mettre en évidence une modification microstructurale du parenchyme cérébral dans la maladie des petits vaisseaux avec une altération des paramètres de diffusion de la substance blanche. Dans cette maladie, on note une augmentation de la diffusivité et une diminution de l'anisotropie au sein des anomalies de signal de la substance blanche.(58)

Une modification de la cartographie ADC est également décrite dans le cadre de la maladie des petits vaisseaux. La valeur moyenne globale de l'ADC cérébrale serait positivement corrélée à des scores semi-quantitatifs d'hypersignaux de la substance blanche.(63,64) Les mesures d'ADC au sein des anomalies de signal de la substance blanche étaient associées à la charge lésionnelle dans l'étude LADIS (Leucoaraiosis and Disability) et davantage augmentées en présence d'infarctus silencieux. Une augmentation de la diffusion a par ailleurs été décrite, en relation avec la sévérité de l'atrophie cérébrale.(65)

La maladie des petits vaisseaux sera donc évaluée grâce aux paramètres correspondant aux hypersignaux de la substance blanche et la trophicité cérébrale.

2. Volume lésionnel

Un AVC ischémique peut se présenter de différentes manières, par exemple sous forme d'une lésion unique de taille variable ou des lésions multiples, dans un seul territoire ou des territoires vasculaires multiples. Cette hétérogénéité de présentation pourrait impacter les performances de leur segmentation.

3. Qualité d'examen

Une qualité satisfaisante d'un examen d'imagerie est indispensable pour sa bonne analyse. Contrairement au scanner dont l'acquisition au niveau cérébral est rapide (quelques secondes), une acquisition de séquence d'IRM prend plus de temps. Une séquence de diffusion classique avec deux valeurs de b est une des séquences IRM les plus rapides à réaliser (moins d'une minute en général), mais nécessite la parfaite immobilité du patient pendant toute la durée de l'acquisition pour réussir des images de bonne qualité.

La séquence de diffusion peut donc être altérée par la présence d'artéfacts dus à l'acquisition de la séquence (inhomogénéités de champ magnétique, présence d'objets métalliques à proximité...) ou aux mouvements du patient lors de l'acquisition. Ces images dégradées pourraient avoir un impact sur l'efficacité de la segmentation qu'il est intéressant d'étudier.

G. OBJECTIF DE L'ETUDE

1. Objectif principal

L'évaluation du volume lésionnel est un élément important de la prise en charge d'un AVC ischémique, que ce soit en phase aiguë ou chronique, ayant d'importantes implications dans la prise en charge thérapeutique et dans le pronostic fonctionnel et cognitif des patients. La segmentation manuelle est cependant laborieuse, chronophage et sujette à l'erreur humaine. Le but de notre étude est donc de comparer la segmentation de lésions d'AVC ischémiques aiguës par méthode manuelle aux

méthodes semi-automatiques et automatiques, en utilisant respectivement une méthode à seuillage et une méthode d'intelligence artificielle à réseau de neurones.

2. Objectifs secondaires

Notre objectif secondaire était d'identifier les facteurs pouvant altérer les performances des méthodes de segmentation. Le parenchyme cérébral des patients adultes ayant subi un AVC présente généralement en IRM des marqueurs de maladie des petits vaisseaux, principalement des hypersignaux de la substance blanche et une atrophie cérébrale. L'IA et notamment les méthodes de Machine Learning peuvent être utilisées pour segmenter et différencier les hypersignaux de la substance blanche et les lésions d'AVC(66), et évaluer l'atrophie cérébrale. Comme discuté précédemment, la maladie des petits vaisseaux est associée à des modifications parenchymateuses diffuses sur l'imagerie de diffusion et pourrait influencer les performances de segmentation. Il est donc intéressant d'évaluer l'impact de la maladie des petits vaisseaux à travers l'étude des hypersignaux de la substance blanche et de la trophicité cérébrale sur la qualité de segmentation.

Les présentations cliniques et en imagerie des AVC ischémiques pouvant être très variable et possiblement associés à des examens de qualité sous-optimale, il est intéressant également d'étudier ces paramètres par rapport aux segmentations proposés par les différentes méthodes.

II. METHODES

A. POPULATION D'ETUDE

La population d'étude est constituée de 202 patients issus de la cohorte Strokdem inclus entre 2010 et 2015 (investigateur principal : Professeur Régis BORDET, Lille).

La cohorte Strokdem est une étude de cohorte prospective multicentrique observationnelle en milieu hospitalier. Son objectif principal est d'identifier les facteurs pronostiques de l'apparition d'une démence ou d'un déclin cognitif après un AVC aigu. L'étude a été approuvée par le comité d'éthique local et enregistrée sur clinicaltrials.gov (NCT01330160). Les patients ont rencontré un attaché de recherche clinique lors d'un entretien individuel et ont signé un accord écrit et éclairé après avoir reçu des explications sur les objectifs de l'étude. Sont inclus les patients âgés de plus de 18 ans sans démence et admis pour un AVC ischémique ou hémorragique dans les 72 heures suivant l'apparition des symptômes, avec une lésion cérébrale à l'IRM d'admission. Les critères d'exclusion sont : la démence avant l'AVC (définie par un score IQCODE (Informant Questionnaire on Cognitive Decline in the Elderly(67)) supérieur à 64) ; une hémorragie d'origine malformative vasculaire, traumatique, méningée pure ou intraventriculaire ; un patient sous tutelle ; la présence d'une contre-indication à l'IRM ; l'incapacité de parler et de comprendre la langue française ; les déficits neurologiques, y compris l'aphasie, suffisamment sévères pour affecter la compréhension des questionnaires et des tests.

B. RECUEIL DES DONNEES CLINIQUES

Lors de l'inclusion dans la cohorte Strokdem, un questionnaire recense les données sociodémographiques, les facteurs de risque cardiovasculaire et les antécédents médicaux. La gravité clinique de l'AVC est évaluée selon l'échelle du National Institute of Health Stroke Scale (NIHSS)(68) en phase aiguë et par l'échelle de Rankin modifiée(69,70) à 6 mois.

C. ACQUISITIONS IRM

Les IRM dans le cadre de cette étude ont été réalisés dans les 24 à 72h depuis le début des symptômes, sur une machine IRM Philips© Achieva 3 Tesla (Philips Healthcare, Best, The Netherlands) dédiée à la recherche, sur la plateforme d'imagerie du vivant et fonctions de Lille (LIIFE, Pr Jean-Pierre PRUVO, Renaud Lopes).

Les séquences de diffusion ont été réalisées avec les paramètres suivants :

- Durée d'acquisition : 37 s
- TR = 3092 ms, TE = 58 ms
- Deux valeurs de b : b0 et b1000
- Couverture de l'ensemble de l'encéphale dans le plan commissure antérieure-commissure postérieure en 36 coupes
- Champs de vue (FOV, en mm) : 230 (droite-gauche) x 230 (antéro-postérieur) x 144 (hauteur) avec une matrice 256 x 256, une épaisseur de coupe de 4 mm et un intervalle intercoupe de 0 mm
- Taille de voxel acquise (mm) : 2,0 x 2,5 et reconstruite à 0,9 x 0,9

D. SEGMENTATIONS

1. Segmentation semi-automatique à seuillage

Le volume des lésions d'AVC est évalué semi-automatiquement à l'aide de l'application Olea (Olea Sphere 3.0, Olea Médical SAS, La Ciotat, France)(31), sur le fichier dicom de la séquence de diffusion de chaque patient. Olea utilise une méthode à seuil sur la cartographie ADC pour établir une segmentation des lésions ischémiques. Une lésion cérébrale est détectée comme un accident vasculaire cérébral ischémique si sa valeur en ADC est inférieure à $0,6 \times 10^{-3} \text{ mm}^2/\text{s}$.

Nous appelons :

- *Méthode semi-automatique à seuillage par défaut* : les segmentations proposées par Olea après vérification uniquement du côté de l'AVC ;
- *Méthode semi-automatique à seuillage avec sélection des coupes d'intérêt* : les segmentations proposées par Olea après vérification du côté de l'AVC et après avoir manuellement choisi les bonnes coupes IRM sièges de lésions ;
- *Méthode semi-automatique à seuillage avec sélection des coupes d'intérêt et correction des lésions* : les segmentations proposées par Olea après vérification du côté de l'AVC, après avoir manuellement choisi les bonnes coupes IRM sièges de lésions puis appliqué des corrections de la segmentation si nécessaire.

2. Segmentation par un algorithme d'IA

Un algorithme d'IA basé sur l'apprentissage profond utilisant un réseau de neurones artificiels a été développé pour évaluer l'efficacité d'un programme de Deep Learning dans une tâche de segmentation automatique de lésions d'AVC ischémique. Le développement s'est déroulé en plusieurs étapes :

1. *Création de masques* : les lésions d'AVC ont été segmentées manuellement par un radiologue sur les coupes IRM de diffusion b1000 de chaque patient à l'aide d'un outil semi-automatisé du logiciel Olea, basé sur la variation d'intensité et la détection des contours.
2. *Préparation des données* : la boîte crânienne a été supprimée sur des images de diffusion b1000 à l'aide d'un outil d'extraction cérébrale (BET, Brain Extraction Tool (71,72)). Les données ont été normalisées avec des valeurs minimales/maximales comprises entre 0 et 1. Les images ont été découpées en patchs matriciels de taille (32, 32, 16) pour réduire la puissance de calcul requise. Les images prétraitées ont ensuite été distribuées dans des jeux de données d'entraînement, de test et de validation.
3. *Création du modèle d'apprentissage profond* : Un réseau de neurones convolutifs utilisant une architecture de type U-net a été développé en définissant des couches de convolution ainsi que les hyperparamètres associés. Un coefficient Dice est ajouté en tant que fonction de coût. Des générateurs d'augmentation de données sont mis en place pour améliorer la robustesse du modèle.
4. *Entraînement* : L'algorithme est entraîné avec les données prétraitées des jeux d'entraînement et de test avec une validation croisée à 5 niveaux. Les données sont séparées en 5 blocs de données d'entraînement et 5 blocs de données de

test. Chaque sujet est utilisé dans 4 blocs de données d'entraînement et 1 de test, avec 131 sujets pour l'entraînement et 33 pour le test. 40 des 131 sujets de l'entraînement sont sélectionnés aléatoirement pour l'étape de validation.

5. *Prédiction de l'image globale* : une fonction de prédiction est créée, qui découpe les images en patches et fait une prédiction sur chaque patch, puis regroupe les patches dans une prédiction d'image entière finale.
6. *Évaluation et rétropropagation de gradient* : La précision est évaluée avec le score Dice et des améliorations sont apportées en revenant aux étapes 2, 3 ou 4.

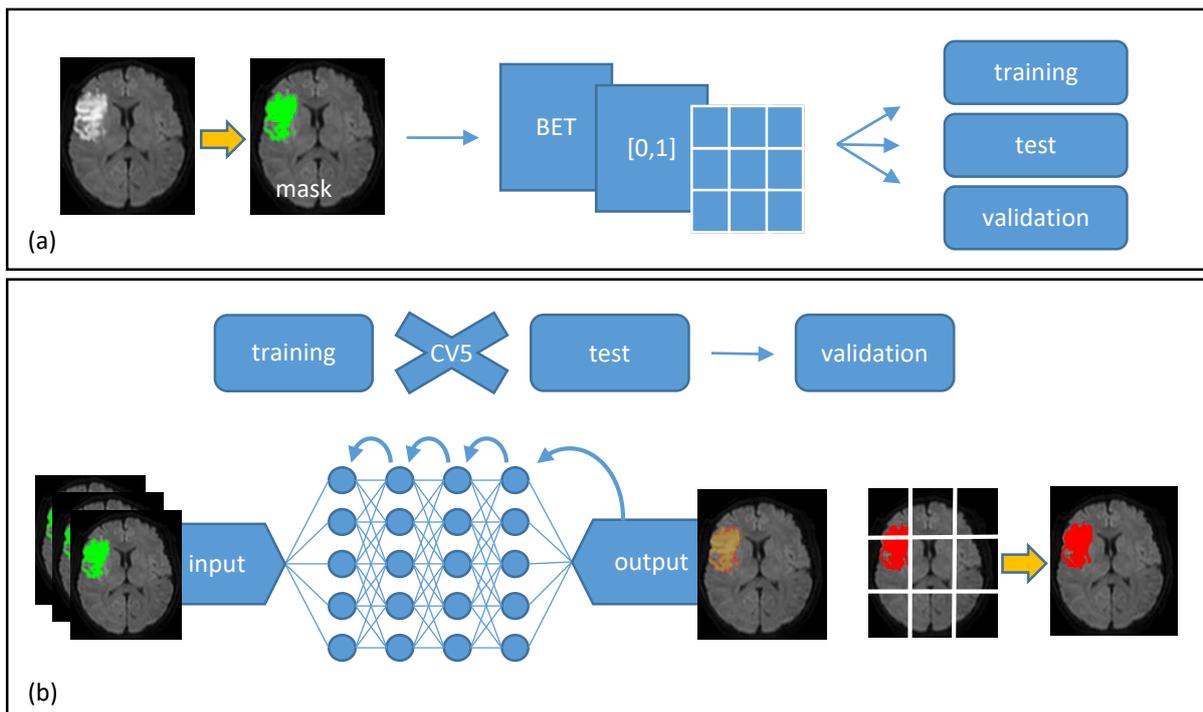


FIGURE 18. Développement du modèle d'IA par apprentissage profond.

- (a) Prétraitement des données avec création des masques de segmentation, extraction du cerveau, normalisation et découpage en patches puis distribution en trois jeux de données d'entraînement, test et validation.
- (b) Création du réseau de neurones artificiels et entraînement avec cross-validation à 5 niveaux puis validation ; prédiction sur patches regroupés en prédiction finale sur image globale.

3. Evaluation des méthodes de segmentation

Le Gold Standard pour les segmentations des lésions ischémiques correspond à la segmentation manuelle des lésions par un radiologue expert référent, appelé observateur 1.

Les méthodes évaluées sont (cf paragraphes précédents) :

- *Méthode manuelle* : segmentations manuelles par un radiologue expert (différent de l'observateur 1), appelé observateur 2, sur les coupes IRM de diffusion b1000 à l'aide d'un outil semi-automatisé du logiciel Olea, basé sur la variation d'intensité et la détection des contours.
- *Méthode semi-automatique à seuil* :
 - 1) Par défaut ;
 - 2) Avec sélection des coupes d'intérêt ;
 - 3) Avec sélection des coupes d'intérêt et correction des lésions.
- *Méthode automatique à l'aide d'un algorithme d'intelligence artificielle par apprentissage profond* (Deep Learning).

E. PARAMETRES POUVANT INFLUENCER LES PERFORMANCES DE SEGMENTATION

1. Trophicité cérébrale et hypersignaux de la substance blanche

La trophicité cérébrale et les hypersignaux de la substance blanche ont été mesurés sur chaque patient inclus pour la segmentation.

Les volumes cérébraux et intracrâniens ont été évalués sur des images cérébrales tridimensionnelles pondérées en T1 à l'aide de l'outil volBrain 1.0 disponible sur [https://volbrain.upv.es/\(73\)](https://volbrain.upv.es/(73)). Le volume cérébral a été normalisé par rapport au volume intracrânien.

Les hypersignaux de la substance blanche ont été évaluées avec des images cérébrales tridimensionnelles pondérées en T1 et des images FLAIR pondérées en T2 en utilisant l'outil lesionBrain 1.0 disponible également sur [https://volbrain.upv.es/\(73\)](https://volbrain.upv.es/(73)).

2. Volume lésionnel

Les volumes lésionnels sont recueillis pour chaque méthode de segmentation, décrites précédemment. Les segmentations réalisées par Deep Learning ont également été revue individuellement par un radiologue, les causes d'échec ont été recueillies qualitativement.

3. Qualité de la séquence IRM

Les séquences de diffusion ont été évalués visuellement par un radiologue. Un score de qualité est attribué à chaque séquence :

- 0 : séquence inexploitable
- 1 : séquence comportant des déformations et des artéfacts de mouvement
- 2 : séquence comportant des déformations
- 3 : séquence comportant des artéfacts de mouvement
- 4 : séquence de très bonne qualité

F. STATISTIQUES

Les analyses statistiques ont été réalisées à l'aide du logiciel SPSS version 26 (IBM corp., Armonk, NY, USA). La distribution des variables quantitatives était appréciée visuellement. Les données quantitatives sont présentées sous la forme médiane [intervalle interquartile IQR].

La performance des méthodes évaluées était tout d'abord appréciée selon plusieurs critères :

- Chevauchement avec la lésion segmentée par la méthode de référence, évaluée par le coefficient de Dice. Le coefficient de Dice va de 0 (aucun chevauchement) à 1 (chevauchement parfait). Il se définit comme le double de l'intersection de deux groupes de valeurs divisé par l'union de ces deux groupes. Il peut être exprimé selon la formule suivante :

$$Dice = \frac{2 \cdot |X \cap Y|}{|X| \cup |Y|}$$

- Distance du barycentre de la segmentation réalisée par rapport au barycentre de la lésion segmentée par la méthode de référence.
- Corrélation avec le volume de la lésion segmentée par la méthode de référence, évaluée par le calcul du coefficient de corrélation non paramétrique de Spearman.
- Corrélation avec la sévérité clinique (score NIHSS initial) et le pronostic fonctionnel (mRS à 6 mois), évaluée par le calcul du coefficient de corrélation non paramétrique de Spearman.

Puis nous avons évalué l'impact du volume d'infarctus, du volume des hypersignaux de la substance blanche et de la trophicité cérébrale et de la qualité d'examen sur la performance de ces méthodes.

Pour étudier l'effet du volume lésionnel sur l'efficacité de la segmentation, la population a été divisée en 3 terciles selon un volume croissant des lésions (1^{er} tercile : petites lésions, 2^{ème} tercile : lésions moyennes, 3^{ème} tercile : grandes lésions), comme indiqué dans le Tableau 3.

	1^{er} tercile	2^{ème} tercile	3^{ème} tercile
Volume ischémique (mL)	0,1 – 0,84	0,87 – 5,18	5,55 – 140,92
Volume des hypersignaux de la substance blanche (mL)	0,01 – 2,58	2,67 -7,93	8,11 – 94,59
Volume cérébral (mL)	0,631 – 0,756	0,757 – 0,808	0,810 – 0,891

TABLEAU 3. *Distribution en terciles des volumes mesurés.*

Pour étudier l'effet de la qualité d'examen sur les performances de segmentation, la population est divisée en examens de très bonne qualité (score qualité 4) et examens avec artéfacts et/ou déformations (score qualité < 4).

Une p value < 0.05 était considérée comme étant statistiquement significative.

III. RESULTATS

A. POPULATION D'ÉTUDE

Sur un échantillon de 202 patients de la cohorte Strokdem inclus entre 2010 et 2015, 164 patients ont été inclus dans notre étude (voir Figure 19).

La population étudiée présentait un âge médian de 65 ans [56-75]. Le score NIHSS médian à l'inclusion était de 1 [0-2] et le mRS médian à 6 mois était de 1 [0-2]. Les volumes segmentés par les différentes méthodes sont disponibles dans le Tableau 4, en comparaison au Gold Standard. Le volume cérébral normalisé médian était de 0,78 [0,74-0,82]. Le volume médian des hypersignaux de la substance blanche était de 4,6 cm³ [1,51-11,29].

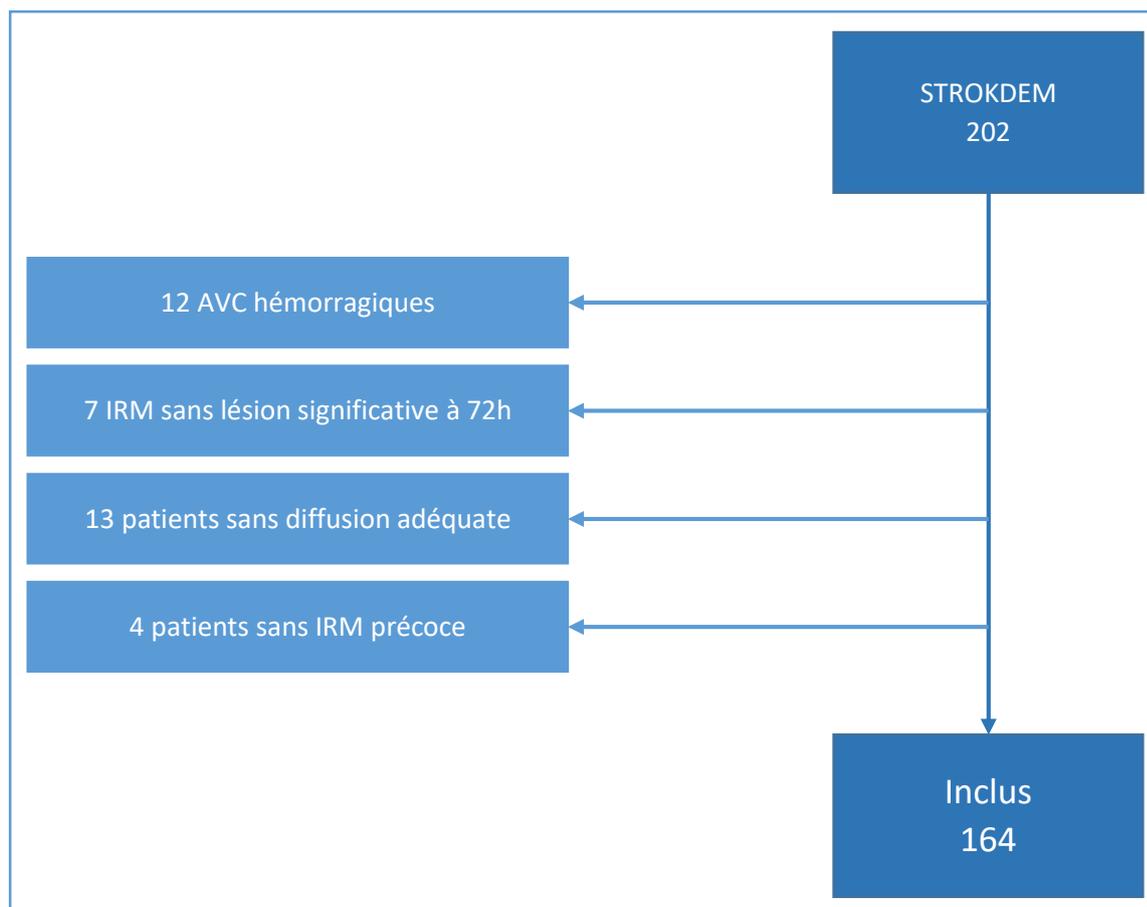


FIGURE 19. Diagramme de flux de la population d'étude

Méthode de segmentation	Médiane (mL)	[Q1-Q3]
Gold Standard (Observateur 1)	1,83	[0,4-9,41]
Méthode manuelle (Observateur 2)	1,80	[0,43-8,16]
Deep Learning	2,86	[0,50-10,84]
Méthode semi-automatique à seuil par défaut	22,06	[10,06-37,03]
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt	2,55	[0,70-7,73]
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt et correction des lésions	1,49	[0,24-4,83]

TABLEAU 4. Volumes médians segmentés par les différentes méthodes

B. COMPARAISON AU STANDARD DE REFERENCE

Les coefficients de corrélation, Dice et distances entre barycentres des volumes segmentés par les différentes méthodes par rapport au Gold Standard sont résumés dans le Tableau 5.

	Rho (ρ)	Dice médian [Q1-Q3]	Distance médiane entre barycentres en mm [Q1-Q3]
Méthode manuelle (Observateur 2)	0,954 (<0,001)	0,69 [0,51-0,80]	1,00 [1,84-3,61]
Deep Learning	0,894 (<0,001)	0,63 [0,35-0,75]	1,51 [2,96-10,81]
Méthode semi-automatique à seuil par défaut	0,213 (0,007)	0,07 [0,00-0,29]	33,12 [52,22-70,57]
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt	0,853 (<0,001)	0,36 [0,11-0,56]	5,68 [9,30-18,35]
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt et correction des lésions	0,896 (<0,001)	0,46 [0,24-0,63]	1,84 [3,99-7,16]

TABLEAU 5. Corrélation de Spearman (Rho), coefficient de Dice et distances entre barycentres des volumes segmentés par les différentes méthodes par rapport au Gold Standard

Un graphique représentant la différence entre chaque méthode et le Gold Standard par rapport à la moyenne des différences (graphique de Bland-Altman) est représenté dans la Figure 20 et la Figure 21.

Les volumes segmentés manuellement par les deux observateurs 1 et 2 sont très proches, avec respectivement un volume médian segmenté de 1,83 ml et 1,80 ml.

La segmentation semi-automatique à seuil par défaut semble largement surestimer le volume lésionnel avec un volume médian segmenté de 22,06 ml, rattrapé par la sélection des coupes d'intérêt (volume médian de 2,55 ml) qui semble toujours surestimer le volume lésionnel, alors que la sélection des coupes avec correction des lésions semble sous-estimer légèrement le volume lésionnel avec un volume médian à 1,49 ml.

Le volume médian segmenté par Deep Learning est de 2,86 ml, en faveur d'une légère surestimation.

La corrélation entre volumes de segmentation par rapport au Gold Standard est élevée pour la segmentation manuelle par l'observateur 2, les segmentations semi-automatiques hors par défaut et la segmentation par Deep Learning, avec une corrélation la plus forte pour la segmentation manuelle (coefficient de Spearman à 0,954). Ceci serait en accord avec une segmentation volumétrique meilleure pour la segmentation manuelle, suivie de près par les segmentations semi-automatiques (hormis celle par défaut) et la segmentation par Deep Learning.

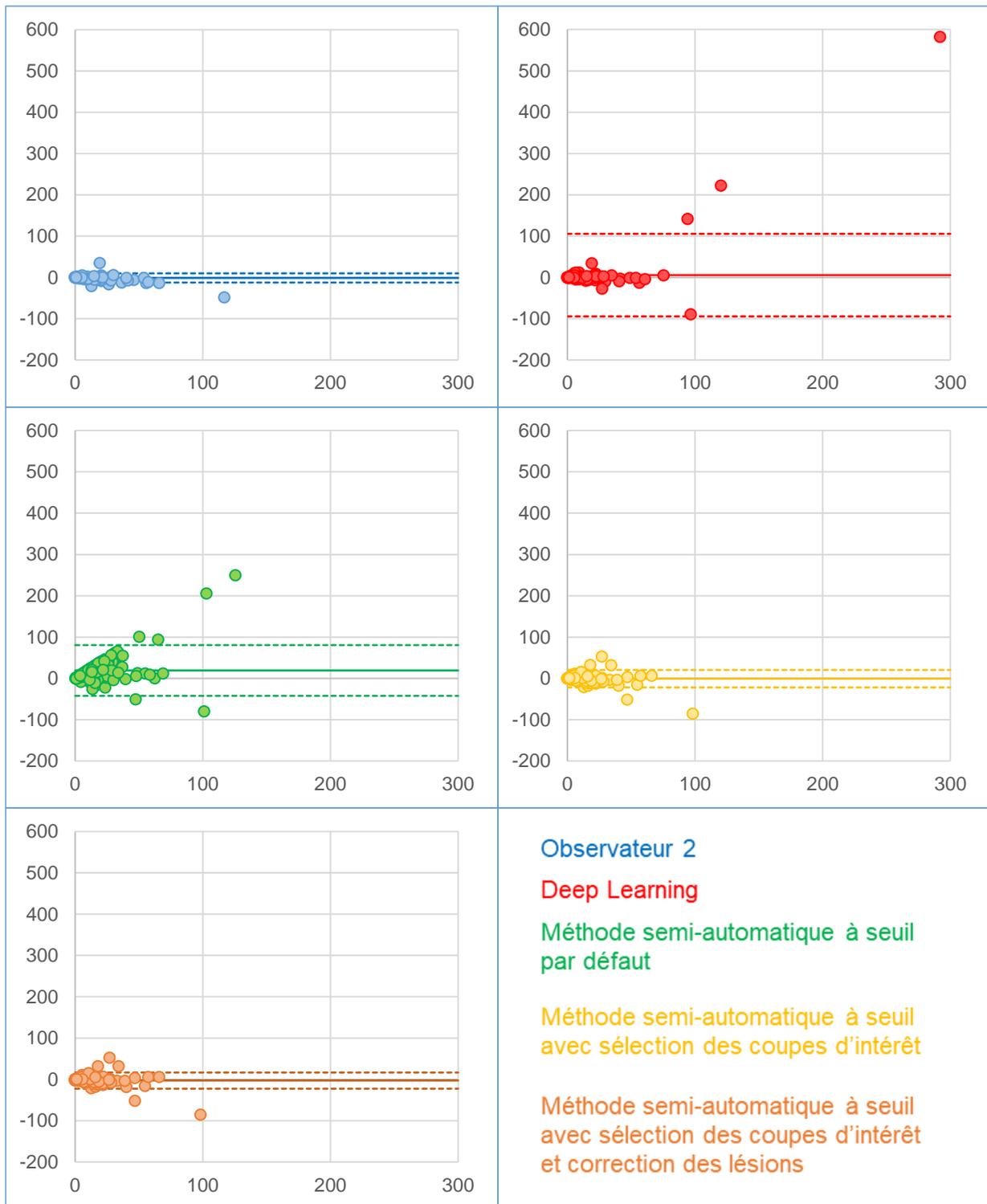


FIGURE 20. Graphiques de Bland-Altman pour chaque méthode de segmentation.

- En abscisse : moyenne entre la méthode étudiée et le Gold Standard
- En ordonnée : différence par rapport au Gold Standard
- Ligne pleine : différence moyenne par rapport au Gold Standard
- Lignes pointillée : intervalles de confiance à 95%

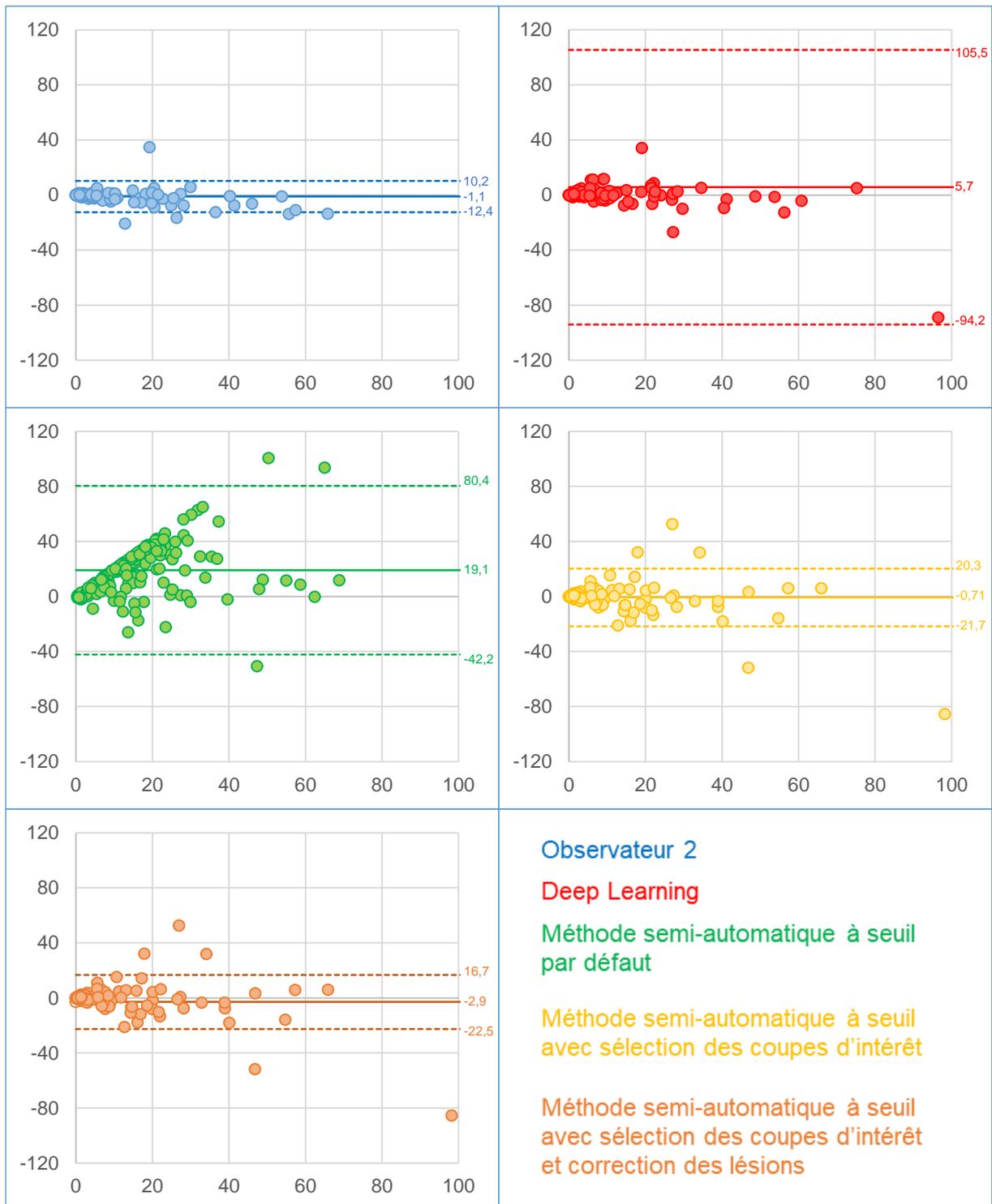


FIGURE 21. Graphiques de Bland-Altman pour chaque méthode de segmentation, centrés sur la majorité des points.

La segmentation semi-automatique par défaut présente un coefficient Rho très bas (0,213), alors que ce coefficient remonte drastiquement dès la sélection des coupes présentant des lésions (0,853) sans augmenter de façon importante après ajout de corrections (0,896). Ceci s'expliquerait par le fait que le logiciel réaliserait une mauvaise sélection des coupes lésionnelles en incluant dans la sélection lésionnelle des artéfacts ou de fausses plages lésionnelles, alors que la segmentation est globalement bien réalisée sur les coupes où l'on trouve effectivement des lésions.

La segmentation par Deep Learning présente une forte corrélation (0,894), similaire à celle de la segmentation semi-automatique avec sélection des coupes et correction des lésions. Elle présente cependant une plus grande dispersion des volumes avec un intervalle de confiance plus large sur le modèle de Bland-Altman, probablement surestimé par des valeurs extrêmes et aberrantes.

Le coefficient de Dice permet d'évaluer le chevauchement des segmentations (overlap). Les segmentations, manuelle et par Deep Learning, présentent des résultats proches avec une médiane de Dice à respectivement 0,69 et 0,63 et un bon overlap, meilleur pour la segmentation manuelle. Les segmentations semi-automatiques (hormis celles par défaut) présentent un overlap un peu moins bon (0,46 après sélection et correction et 0,36 après sélection). Les résultats pour la segmentation semi-automatique par défaut sont très bas (0,07), en rapport vraisemblablement avec la mauvaise sélection et segmentation des lésions.

En dehors des lésions correctement segmentées, des segmentations hors cibles ont été proposés pour 71/164 (43%) patients par l'algorithme de Deep Learning. Pour 32/71 (45%) de ces patients, ces hors cibles correspondaient à des artéfacts de la séquence de diffusion, notamment des artéfacts de bord. Pour 18/71 (25%) de ces patients, il s'agissait de localisations aberrantes à type de faux positifs lésionnels au

sein des hémisphères cérébraux, du cervelet, de la moelle ou des plexus choroïdes. Enfin, 21/71 (30%) de ces patients présentaient ces deux types de segmentations hors cibles. Ces erreurs de segmentations étaient plus ou moins importantes en volume et en impact sur les scores de performance.

C. CORRELATIONS AUX DIFFERENTS PARAMETRES D'EVALUATION

1. Scores cliniques

Les corrélations entre volumes et scores de prédiction clinique (NIHSS initial, mRS à 6 mois) sont modérées pour l'ensemble des méthodes, mais significatives. La faible variabilité des scores de sévérité clinique et pronostique ne permettent pas de discriminer les différentes méthodes. Les coefficients de corrélation sont disponibles dans le Tableau 6 et le Tableau 7.

	Spearman Rho	p
Volume Observateur 1 (Gold Standard)	0.300	<0.001
Volume Observateur 2	0.316	<0.001
Volume Deep Learning	0.302	<0.001
Méthode semi-automatique à seuil par défaut	0.047	0.548
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt	0.258	0.001
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt et correction des lésions	0.331	<0.001

TABLEAU 6. *Corrélations entre le volume segmenté et le score NIHSS initial*

	Spearman Rho	p
Volume Observateur 1 (Gold Standard)	0.288	<0.001
Volume Observateur 2	0.299	<0.001
Volume Deep Learning	0.293	<0.001
Méthode semi-automatique à seuil par défaut	0.186	0.146
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt	0.297	0.001
Méthode semi-automatique à seuil avec sélection des coupes d'intérêt et correction des lésions	0.331	<0.001

TABLEAU 7. *Corrélations entre le volume segmenté et le score mRS à 6 mois*

2. Volume lésionnel

Pour les petites lésions, on observe un meilleur overlap avec la méthode manuelle par rapport aux méthodes d'IA. Pour les lésions de moyenne et grande taille, les performances sont similaires pour les méthodes manuelles et Deep Learning, tous les deux supérieurs à la segmentation semi-automatique.

Sur les calculs de distances entre barycentres on voit que le Deep Learning montre de mauvais résultats pour les petites lésions, en faveur d'une mauvaise détection et segmentation de ces lésions. Des résultats similaires sont retrouvés avec la segmentation semi-automatique avec sélection seule des lésions. Par ailleurs, la variabilité n'est pas significative entre les méthodes manuelle, semi-automatique avec sélection et correction des lésions et le Deep Learning.

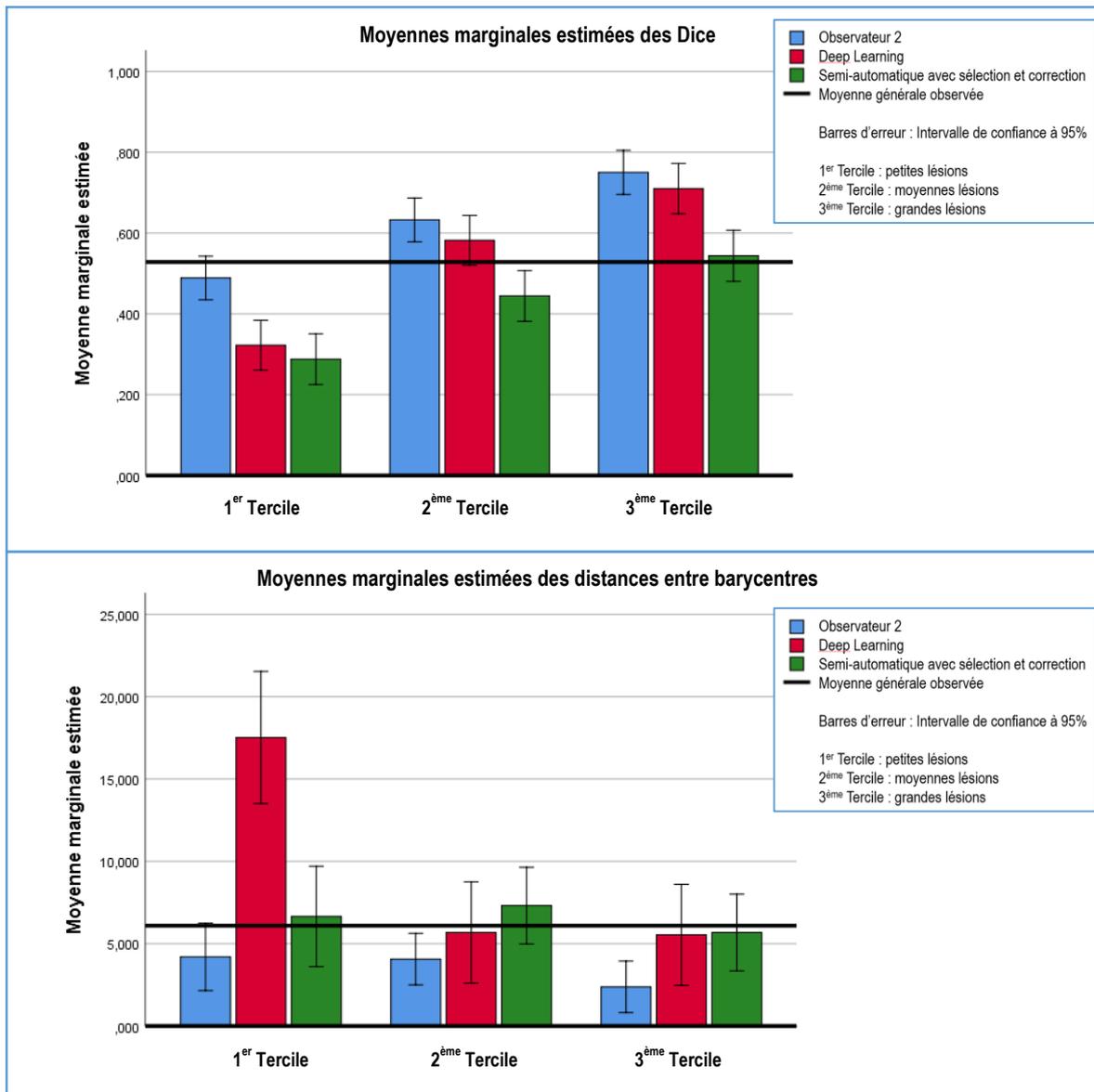


FIGURE 22. Moyennes de Dice et de distances des barycentres entre méthodes.

Moyennes estimées du Dice et de la distance des barycentres des volumes segmentés par l'observateur 2, le Deep Learning et la méthode semi-automatique à seuil avec sélection des coupes d'intérêt et correction des lésions.

3. Maladie des petits vaisseaux

Les études de l'effet de la trophicité cérébrale et de la leucopathie sur les performances de segmentation n'ont pas montré de résultats significatifs. Ces données sont résumées dans le Tableau 8 et le Tableau 9. Les résultats sont présentés selon le 1^{er}, 2^{ème} et 3^{ème} tercile de la population d'étude représentant respectivement les petites, moyennes et grandes lésions.

Dice par rapport au Gold Standard	Tercile	Médiane	[Q1-Q3]
Observateur 2	1	0,69	[0,46-0,80]
	2	0,70	[0,57-0,79]
	3	0,70	[0,49-0,79]
Deep Learning	1	0,63	[0,23-0,74]
	2	0,64	[0,45-0,75]
	3	0,63	[0,35-0,76]
Méthode semi-automatique à seuil (par défaut)	1	0,06	[0,00-0,29]
	2	0,13	[0,01-0,30]
	3	0,06	[0,00-0,32]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	1	0,29	[0,09-0,52]
	2	0,44	[0,19-0,63]
	3	0,37	[0,13-0,55]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	1	0,42	[0,15-0,61]
	2	0,53	[0,30-0,70]
	3	0,44	[0,28-0,63]
Distance des barycentres par rapport au Gold Standard	Tercile	Médiane	[Q1-Q3]
Observateur 2	1	1,75	[0,99-3,77]
	2	1,71	[0,82-2,91]
	3	1,99	[1,13-5,17]
Deep Learning	1	3,19	[1,57-8,62]
	2	3,38	[1,55-18,26]
	3	2,77	[1,31-13,41]
Méthode semi-automatique à seuil (par défaut)	1	43,34	[26,51-69,16]
	2	54,45	[37,98-70,74]
	3	52,49	[37,32-70,66]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	1	11,68	[5,98-22,04]
	2	7,96	[4,46-15,98]
	3	9,40	[6,57-17,09]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	1	4,33	[1,84-11,00]
	2	3,06	[1,73-5,45]
	3	4,25	[1,84-7,84]

TABLEAU 8. Relation entre la qualité de segmentation et le volume des hypersignaux de la substance blanche selon le tercile (1 : 1^{er} tercile, 2 : 2^{ème} tercile, 3 : 3^{ème} tercile)

Dice par rapport au Gold Standard	Tercile	Médiane	[Q1-Q3]
Observateur 2	1	0,67	[0,57-0,78]
	2	0,64	[0,43-0,78]
	3	0,74	[0,57-0,80]
Deep Learning	1	0,62	[0,38-0,73]
	2	0,57	[0,21-0,76]
	3	0,70	[0,42-0,76]
Méthode semi-automatique à seuil (par défaut)	1	0,062	[0,00-0,32]
	2	0,10	[0,00-0,38]
	3	0,10	[0,00-0,29]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	1	0,42	[0,15-0,56]
	2	0,34	[0,10-0,56]
	3	0,37	[0,13-0,58]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	1	0,46	[0,30-0,65]
	2	0,38	[0,12-0,61]
	3	0,52	[0,27-0,66]
<hr/>			
Distance des barycentres par rapport au Gold Standard	Tercile	Médiane	[Q1-Q3]
Observateur 2	1	1,99	[1,22-3,28]
	2	1,92	[1,20-5,11]
	3	1,58	[0,86-3,51]
Deep Learning	1	2,83	[1,33-14,94]
	2	3,42	[1,40-9,59]
	3	2,76	[1,60-10,61]
Méthode semi-automatique à seuil (par défaut)	1	53,91	[37,98-71,60]
	2	57,01	[32,31-70,88]
	3	46,35	[29,36-64,99]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	1	9,40	[5,40-17,27]
	2	8,843	[6,19-13,60]
	3	11,21	[4,99-20,03]
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	1	3,89	[1,74-7,20]
	2	4,61	[2,33-7,95]
	3	3,33	[1,67-6,46]

TABLEAU 9. Relation entre la qualité de segmentation et la trophicité cérébrale selon le tercile (1 : 1er tercile, 2 : 2^{ème} tercile, 3 : 3^{ème} tercile)

4. Qualité d'examen

On observe de meilleurs résultats en terme de distance de barycentres avec des examens de bonne qualité (score qualité 4) par rapport aux examens de qualité altérée (score qualité < 4) pour les segmentations par Deep Learning et la segmentation semi-automatique par sélection des lésions. Les autres données n'ont pas montré de significativité sur ce point, disponibles dans le Tableau 10.

Dice par rapport au Gold Standard	Qualité	Médiane [Q1-Q3]	p
Observateur 2	0	0,68 [0,40-0,80]	0,568
	1	0,71 [0,53-0,78]	
Deep Learning	0	0,63 [0,13-0,74]	0,267
	1	0,64 [0,42-0,75]	
Méthode semi-automatique à seuil (par défaut)	0	0,10 [0,05-0,28]	0,920
	1	0,07 [0,00-0,31]	
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	0	0,27 [0,11-0,49]	0,168
	1	0,41 [0,14-0,58]	
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	0	0,39 [0,13-0,63]	0,359
	1	0,48 [0,27-0,65]	
Distance des barycentres par rapport au Gold Standard	Qualité	Médiane [Q1-Q3]	P
Observateur 2	0	2,29 [0,97-4,49]	0,365
	1	1,77 [1,02-3,30]	
Deep Learning	0	5,59 [1,97-21,97]	0,029
	1	2,61 [1,46-8,07]	
Méthode semi-automatique à seuil (par défaut)	0	52,77 [34,70-65,31]	0,846
	1	52,22 [30,17-70,93]	
Méthode semi-automatique à seuil (sélection des coupes d'intérêt)	0	15,08 [7,78-27,82]	0,003
	1	8,37 [5,37-15,48]	
Méthode semi-automatique à seuil (sélection des coupes d'intérêt et correction des lésions)	0	4,19 [1,94-10,61]	0,399
	1	3,82 [1,76-7,02]	

TABLEAU 10. Relation entre la qualité de segmentation et la qualité d'examen (0 : examens de bonne qualité avec score qualité à 4, 1 : examens de qualité altérée avec score qualité < 4)

Exemples illustratifs de segmentations :

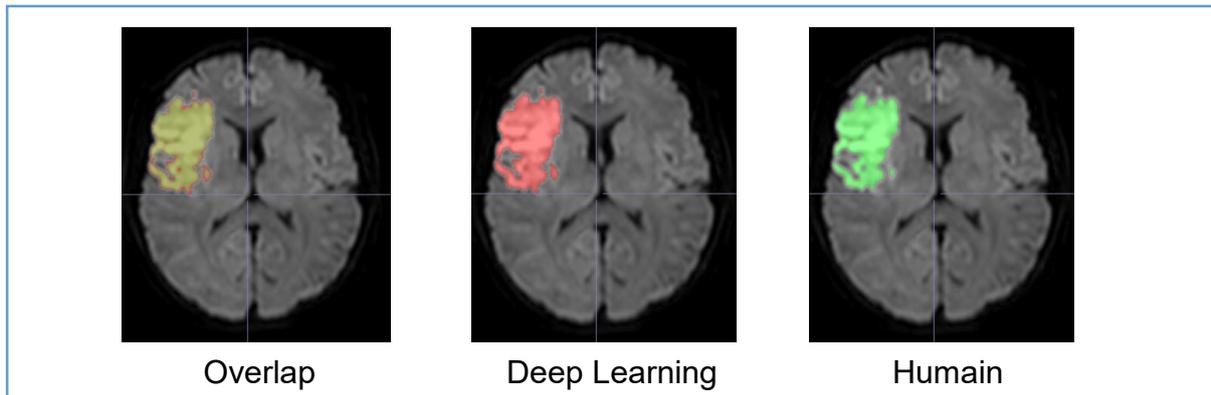


FIGURE 23. Exemple d'une bonne concordance entre les segmentations par méthode manuelle et par Deep Learning. Les deux segmentations se chevauchent correctement, avec un Dice à 0,79 entre la segmentation humaine par l'observateur 2 et le Deep Learning. On note une très légère surestimation par l'algorithme de Deep Learning par rapport à l'observateur 2.

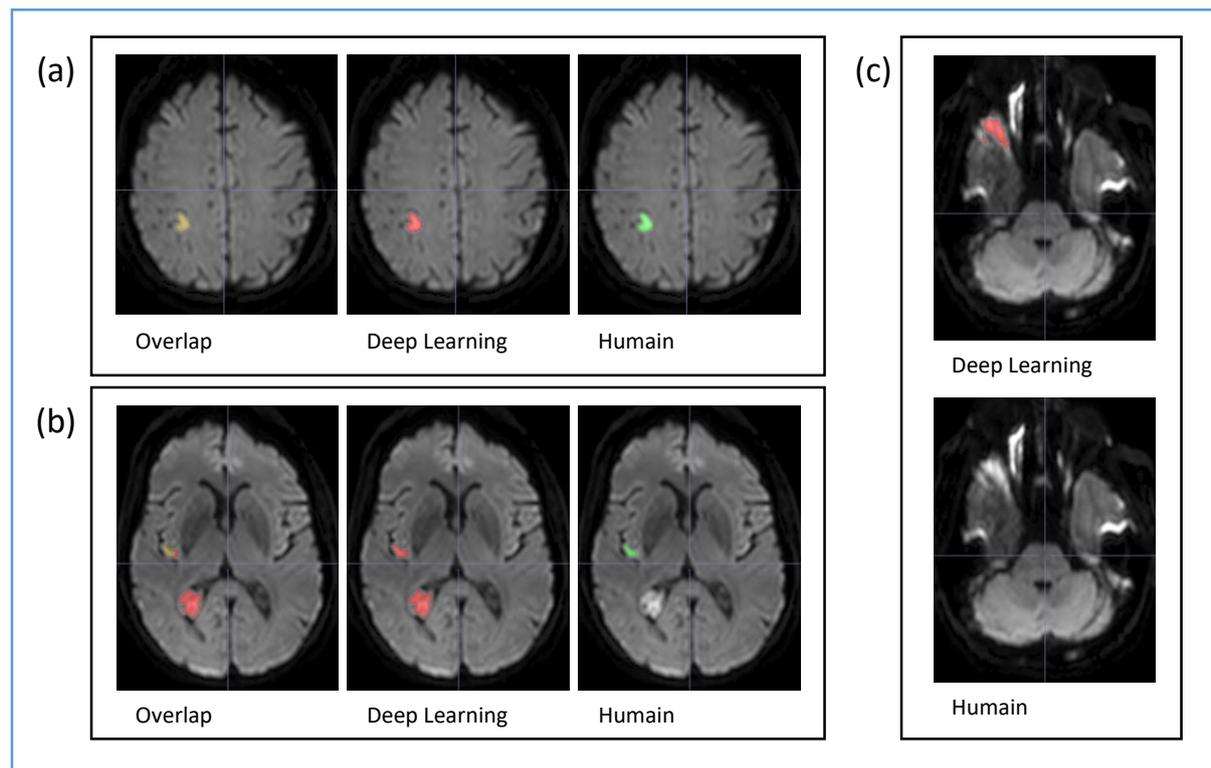


FIGURE 24. Exemple d'une mauvaise concordance entre les segmentations par méthode manuelle et par Deep Learning.

Les segmentations des lésions précentrale droite (a) et insulaire droite (b) se chevauchent correctement mais on observe des faux positifs de segmentation de la part de l'algorithme de Deep Learning avec une segmentation du plexus choroïde droit (b) et d'un artefact en avant du pôle temporal droit (c), faisant chuter le Dice à 0,53.

IV. DISCUSSION

A. RAPPEL DES RESULTATS PRINCIPAUX

1. Méthodes de segmentation

a) Deep Learning

La segmentation lésionnelle par méthode de Deep Learning a montré de bonnes performances avec une forte corrélation entre volumes segmentés par rapport au Gold Standard (coefficient de Spearman à 0,894 avec $p < 0,001$), un bon chevauchement entre segmentations (Dice médian à 0,63) et une faible distance entre barycentres (médiane de 1,51 mm).

Cette méthode présente des résultats similaires à la segmentation semi-automatique avec sélection des coupes d'intérêt et correction des lésions en terme de corrélation des volumes segmentés par rapport à la référence (respectivement 0,894 et 0,896), mais présente de meilleurs résultats sur le chevauchement des lésions et la distance des barycentres lésionnels (respectivement Dice à 0,63 et 0,46 ; distances entre barycentres de 1,51 mm et 1,84 mm). Notre algorithme de Deep Learning présente néanmoins quelques points faibles et reste moins bon que la méthode manuelle qui présente les meilleurs coefficients de corrélation (0,954), Dice (0,69) et distance entre barycentres (1,00 mm).

Notre algorithme de Deep Learning a une tendance à surestimer les lésions par rapport au Gold Standard, avec un volume médian segmenté respectivement de 2,86 ml contre 1,83 ml. Sur les graphiques de Bland-Altman, on observe une forte dispersion des valeurs par rapport au biais moyen, avec l'intervalle de confiance le plus large parmi les différentes méthodes, vraisemblablement surestimé en raison de valeurs aberrantes ou extrêmes. 43% des segmentations incluaient par ailleurs des

segmentations hors cibles avec notamment des artéfacts de la séquence de diffusion et de faux positifs parenchymateux.

Avec l'étude des volumes ischémiques par terciles, on observe une amélioration des performances du Deep Learning avec l'augmentation du volume des lésions. On note particulièrement de mauvais résultats pour les petites lésions, avec une distance entre barycentres et un Dice significativement altérés (voir Figure 18). Une bonne qualité d'examen semble améliorer légèrement les distances entre barycentres (on passe de 1,97 mm pour les examens de qualité altérée à 1,46 mm pour les examens de bonne qualité, $p=0,029$) même si l'impact sur le chevauchement n'est pas significatif.

b) Méthode manuelle

Notre étude a montré une forte corrélation entre les segmentations manuelles réalisées par les observateurs 1 et 2, en faveur d'une bonne reproductibilité inter-observateurs sur une segmentation de lésions d'AVC. L'étude de l'overlap et de la distance entre barycentres lésionnels ont montré les meilleurs résultats pour la segmentation manuelle avec un coefficient Dice le plus élevé et un décalage des barycentres le plus faible, par rapport aux autres méthodes de segmentation.

c) Méthode semi-automatique

La segmentation semi-automatique par défaut utilisant une méthode à seuil a présenté les moins bonnes performances avec une détection de faux positifs en rapport avec des artéfacts ou de fausses plages lésionnelles parenchymateuses. Les performances s'améliorent drastiquement dès lors qu'une sélection manuelle des

coupes lésionnelles est effectuée, correspondant à la méthode semi-automatique avec sélection des coupes. Ceci serait lié au fait que le logiciel segmente de façon automatique globalement correctement les lésions, mais voit ses performances chuter en raison des faux positifs. L'amélioration des performances est très faible d'ailleurs entre la méthode semi-automatique avec sélection des coupes lésionnelles et l'ajout de corrections manuelles dans la segmentation, allant dans le sens d'une bonne segmentation des lésions lorsqu'elles sont détectées mais désavantagée par la présence de faux positifs.

2. Paramètres pouvant altérer les performances

Plusieurs paramètres pouvant influencer l'efficacité des tâches de segmentation ont été étudiés.

L'étude de l'impact du volume lésionnel a démontré une amélioration du score de Dice et de la distance entre barycentres avec l'augmentation du volume lésionnel, pour toutes les méthodes. La méthode manuelle reste supérieure de façon significative aux autres méthodes automatiques et semi-automatiques pour les petites lésions. Pour les moyennes et grandes lésions, l'algorithme de Deep Learning rejoint les performances humaines. L'algorithme voit par contre ses performances chuter pour les petites lésions avec une mauvaise détection et segmentation de celles-ci. La méthode de segmentation à seuil reste inférieure à ces deux précédentes techniques.

La qualité de l'examen semble influencer l'efficacité de la segmentation par Deep Learning, avec de meilleures performances pour les examens de bonne qualité. Ospel et al. ont montré en 2020 que l'effet du volume partiel sur les séquences de diffusion pouvait être une source importante d'erreur dans la mesure du volume

lésionnel dans l'AVC aigu.(74) En observant les marges d'erreurs pour un volume lésionnel de 70 ml, la plage de volumes potentiellement mal classés s'étendant entre 46,5 ml et 94,1 ml pour une taille de voxels de 2 x 2 x 4 mm, passait à un intervalle situé entre 64,4 ml et 80,1 ml pour une taille de voxels de 1 x 1 x 2 mm.(74) En raison des contours généralement complexes des lésions ischémiques de grande taille, une erreur de volume pourrait aboutir à un mauvais choix thérapeutique. L'utilisation de séquences de diffusion haute résolution permettrait de diminuer ces erreurs de mesure.

La trophicité cérébrale et la leucopathie n'ont pas démontré d'effet significatif sur les performances de segmentation des différentes méthodes.

B. SEGMENTATION DE LESIONS ISCHEMIQUES DANS LA LITTERATURE

De nombreuses études de segmentation automatique de lésions d'AVC sont menées pour développer des méthodes capables de délimiter la zone lésionnelle sans ou avec le moins d'intervention humaine possible, en utilisant des examens tomodensitométriques(75–77), ainsi qu'en IRM. Sarmiento et al. fournit une analyse de 1933 travaux qui compare des approches récentes d'intelligence artificielle entre 2015 et 2018 à ce sujet.(51) Le nombre d'études présentant et évaluant des modèles d'intelligence artificielle pour la détection et la segmentation dans le domaine de l'imagerie médicale a considérablement augmenté au cours de la dernière décennie, et concerne tout autant l'imagerie par résonance magnétique que la tomodensitométrie à rayons X.(49)

	Auteurs	Méthode d'IA	Modalité d'imagerie
Réseaux de neurones	Clerigues et al. 2020 (78)	Réseau de neurones convolutifs de type U-Net	IRM (T1, T2, FLAIR, DWI)
	Qiu et al. 2020 (79)	Réseau de neurones convolutifs de type U-Net et Forêt d'arbres décisionnels	TDM sans injection
	Kasasbeh et al. 2019 (80)	Réseau de neurones	TDM de perfusion
	Guerrero et al. 2018 (66)	Réseau de neurones convolutifs (uResNet)	IRM (T1, FLAIR)
	Zhang et al. 2018 (81)	Réseau de neurones	IRM (DWI)
	Chen et al. 2017 (82)	Deux réseaux de neurones convolutifs combinés	IRM (DWI)
Autres techniques	Boldsen et al. 2018 (83)	Forêt d'arbres décisionnels	IRM (DWI)
	Ramamurthy et al. 2017 (84)	Diverses méthodes combinées	IRM (DWI)
	McKinley et al. 2017 (85)	Forêt d'arbres décisionnels	IRM (T1, T2, DWI)
	Nag et al. 2017 (86)	GMM	IRM (DWI)
	Pustina et al. 2016 (87)	LINDA	IRM (T1)
	Forkert et al. 2015 (88)	SVM	IRM (DWI)
	Maier et al. 2015 (89)	Forêt d'arbres décisionnels	IRM (T1, T2, FLAIR, DWI, fIRM)
	Purushotham et al. 2015 (90)	Méthode à seuil ADC	IRM (DWI)

NN : réseau de neurones artificiels ; CNN : réseau de neurones convolutifs ; SVM : support vector machine ; GMM : modèle de mélange gaussien ; LINDA : Lesion Identification with Neighborhood Data Analysis ; fIRM : IRM fonctionnelle

TABLEAU 11. L'IA dans les tâches de segmentation de lésions ischémiques cérébrales dans la littérature récente.

Les méthodes d'apprentissage automatique de haut niveau, telles que les approches de type forêt d'arbres décisionnels aléatoires (*random decision forest*) et de réseau de neurones convolutif, ont conduit à des résultats de segmentation nettement meilleurs par rapport aux méthodes plus simples.(52,81) Notre étude va d'ailleurs dans ce sens, avec globalement de meilleures performances de la méthode de Deep Learning par rapport à la méthode à seuil. Pour exemple :

- Zhang et al. ont entraîné en 2018 un modèle de réseaux de neurones convolutif sur les images de diffusion de lésions ischémiques aiguës variées de 242 sujets (90 pour l'entraînement, 62 pour la validation et 90 pour le test), avec un coefficient Dice de 0,79.
- Maier et al. ont publié une étude comparative en 2016 de différentes méthodes de segmentation de lésions ischémiques subaiguës en FLAIR, avec les meilleurs résultats pour les forêts d'arbres décisionnels et les réseaux de neurones convolutif, avec un Dice de 0,67 pour les deux méthodes.

Des bases de données publiques de lésions cérébrales ischémiques et hémorragiques existent pour faciliter la recherche sur ce sujet. Soun et al. en a résumé les principaux dans son article de 2021, avec notamment les bases ATLAS (Anatomical Tracings of Lesions After Stroke), ISLES (Ischemic Stroke Lesion Segmentation) 2016-2017 et ISLES 2018, contenant des imageries IRM et TDM de lésions ischémiques (ATLAS, ISLES 2016-2017) et hémorragiques (ISLES 2018) et les bases CQ500 et RSNA Brain Hemorrhage CT Dataset, contenant des imageries TDM de lésions hémorragiques.(41,78,91–95)

Des classements des meilleures performances sont disponibles à partir de ces bases de données publiques, sur notamment des lésions ischémiques subaiguës

(SISS : sub-acute ischemic stroke lesion segmentation) et de lésions ischémiques aiguës (SPES : acute stroke perfusion estimation). Le premier concours de segmentation ISLES a eu lieu en 2015. Depuis, un classement régulièrement actualisé est mis en ligne sur <https://www.smir.ch/ISLES/Start2015>.(95) Les classements SISS et SPES en cours des 5 meilleures méthodes sont disponibles dans le Tableau 12 et le Tableau 13 (moyennes \pm écart-type). Pour ces bases de données, le Gold Standard consistait en la segmentation manuelle des lésions ischémiques sur séquence de diffusion (image trace). Parmi ces participants au concours, certains utilisent des réseaux de neurones convolutifs tel que Clèrigues et al. (*clera2*)(78), Kamnitsas et al. (*kamnk1*)(96) et Zhang et al. (*zhanr6*)(81). D'autres utilisent des méthodes différentes de Machine Learning, tel que Mckinley et al. (*mckir1*)(97) qui utilise un algorithme basé sur les forêts aléatoires d'arbres décisionnels. Nous n'avons pas retrouvé d'article commentant la méthodologie des participants *ghosp1*, *lianl1*, *sathr2* et *cheng2*.

Participant	Dice	Distance de Hausdorff	Valeur prédictive positive	Sensibilité
ghosp1	0.63 (\pm 0.25)	34.57 (\pm 24.27)	0.74 (\pm 0.23)	0.62 (\pm 0.26)
clera2	0.59 (\pm 0,31)	34.74 (\pm 28.90)	0.65 (\pm 0.35)	0.60 (\pm 0.30)
kamnk1	0.59 (\pm 0,31)	39.61 (\pm 30.68)	0.68 (\pm 0.33)	0.60 (\pm 0.27)
zhanr6	0.58 (\pm 0,31)	38.89 (\pm 35.27)	0.60 (\pm 0.33)	0.68 (\pm 0.24)
lianl1	0.57 (\pm 0,29)	43.02 (\pm 30.48)	0.58 (\pm 0.30)	0.64 (\pm 0.29)

TABLEAU 12. Classement SISS sur test en août 2021, trié par Dice décroissant

Participant	Dice	Distance de Hausdorff	Valeur prédictive positive	Sensibilité
clera2	0.84 (\pm 0.10)	20.66 (\pm 13.89)	0.82 (\pm 0.15)	0.89 (\pm 0.06)
ghosp1	0.83 (\pm 0.08)	40.74 (\pm 21.34)	0.79 (\pm 0.12)	0.89 (\pm 0.07)
mckir1	0.82 (\pm 0.08)	29.02 (\pm 16.29)	0.83 (\pm 0.10)	0.82 (\pm 0.14)
sathr2	0.82 (\pm 0.09)	45.49 (\pm 31.49)	0.81 (\pm 0.11)	0.87 (\pm 0.14)
cheng2	0.81 (\pm 0.10)	28.12 (\pm 21.72)	0.81 (\pm 0.11)	0.82 (\pm 0.12)

TABLEAU 13. Classement SPES sur test en août 2021, trié par Dice décroissant

C. MALADIE DES PETITS VAISSEAUX ET DIFFUSION

Plusieurs types de lésions cérébrales liées à la maladie des petits vaisseaux peuvent être identifiés en neuro-imagerie, tels que les petits infarctus sous-corticaux, les lacunes, les hypersignaux de la substance blanche, l'élargissement des espaces périvasculaires, l'atrophie cérébrale et les microhémorragies.(58)

Les changements microstructuraux sont fréquents, l'augmentation de la diffusivité et la diminution de l'anisotropie étant les résultats les plus fiables observés, principalement dans la substance blanche. Ces modifications tissulaires sont associées à une sévérité clinique et en particulier à un dysfonctionnement exécutif. Ils peuvent également précéder les marqueurs IRM habituels de la maladie, tels que les hypersignaux de la substance blanche, les microhémorragies et les lacunes.(98)

Les modifications en tenseur de diffusion et la présence d'hypersignaux de la substance blanche seraient deux processus interdépendants mais se produisant dans des fenêtres temporelles différentes. Les changements de diffusivité refléteraient la première phase des changements de la substance blanche et les anomalies de signal de la substance blanche apparaîtraient comme une conséquence de ces changements.(99)

Les hypersignaux de la substance blanche sévères étaient corrélées avec les hypersignaux en diffusion dans la substance blanche profonde chez des patients victimes d'accident vasculaire transitoire et contribueraient à un risque accru de récurrence d'AVC.(100)

Une grande partie des petites lésions ischémiques détectées en diffusion se transformeraient au fil du temps en lacunes et infarctus sous-corticaux et les lésions

ischémiques focales asymptomatiques pourraient conduire en phase chronique à des hypersignaux T2, contribuant à la progression des anomalies de signal de la substance blanche dans la maladie des petits vaisseaux(61,62).

Les valeurs en ADC de la substance blanche subissent également des modifications chez les patients atteints d'une maladie des petits vaisseaux, comme vu dans l'introduction, avec une augmentation des valeurs ADC moyennes du cerveau et au sein des anomalies de signal de la substance blanche(64,65).

L'imagerie de diffusion peut donc varier avec la maladie des petits vaisseaux sous-jacente et avec la sévérité de l'atrophie cérébrale. Ces variations pourraient avoir une répercussion sur l'efficacité de la segmentation de lésions ischémiques par des méthodes utilisant un seuil pour délimiter les éléments pathologiques sur des séquences de diffusion sans autre traitement d'image que le calcul des cartes ADC et b1000, tel que le logiciel étudié dans notre travail, mais notre étude n'a pas retrouvé d'impact significatif. Les méthodes d'IA d'apprentissage profond utilisent des masques conçus par un prétraitement des images qui normalisent les valeurs quantitatives de chaque voxel, ce qui pourrait rendre l'algorithme moins sensible à ces variations. Notre étude n'a d'ailleurs pas retrouvé d'effet significatif concernant ce point.

D. FORCES DES METHODES (SEMI-)AUTOMATIQUES

Les méthodes d'IA qu'elles soient totalement automatiques telle qu'un algorithme de Deep Learning ou semi-automatiques, permettent un gain de temps considérable par rapport à une segmentation manuelle fastidieuse par un observateur

humain. De plus, les performances de méthodes optimisées arrivent au même niveau que les performances humaines, rendant ces outils très intéressants en pratique.

L'intelligence artificielle a par ailleurs l'avantage, par rapport à l'analyse humaine, d'une évaluation objective et quantitative, d'une capacité à détecter des patterns subtils à l'échelle du voxel, de résultats rapides et d'une possibilité de mise en œuvre à grande échelle.(41)

Le pronostic fonctionnel après un AVC est profondément lié à la localisation et à l'extension de la lésion ischémique. Des modèles d'IA d'apprentissage automatique de haut niveau peuvent être utilisées pour prédire les résultats fonctionnels selon le volume et la localisation d'un infarctus cérébral.(88) Les scores cliniques de notre étude n'étaient cependant pas adaptés pour une analyse correcte de cette problématique.

E. LIMITES DES METHODES (SEMI-)AUTOMATIQUES

Sur les segmentations semi-automatiques prédites par notre méthode à seuil, des corrections manuelles ont dû être apportées pour quasiment tous les patients. La majorité des volumes lésionnels prédits ont été sous-estimés par rapport aux volumes évalués manuellement.

La base de données monocentrique et la méthode de validation croisée pour le développement de l'algorithme de Deep Learning diminuent la généralisabilité de la segmentation basée sur l'apprentissage automatique. La réutilisation des patients dans les jeux d'entraînement, de test et de validation surestime probablement l'efficacité de l'algorithme. Sachant que le but de l'étude n'était pas de développer une

méthode de segmentation à usage public, ce compromis a été fait concernant la taille relativement modeste de notre base de données de patients.

De plus, il est important de préciser sur quelle séquence exactement est réalisée la segmentation. Une segmentation « sur diffusion » peut par exemple être réalisée sur la carte de diffusion b1000 ou la cartographie ADC. Les volumes segmentés sont généralement proches entre ces deux méthodes mais il a été montré une différence systématique avec une légère sous-estimation des volumes sur la cartographie ADC comparativement à la séquence de diffusion b1000.(31)

Même si l'intelligence artificielle gagne du terrain dans l'automatisation des tâches de segmentation, l'intervention humaine est toujours nécessaire, ne serait-ce que pour la vérification des résultats proposés. Parmi les méthodes de classification comparées par Maier et al., aucune n'a montré de supériorité sur la segmentation manuelle humaine(52).

De plus, il n'existe pas d'explication actuellement sur le fonctionnement exact des réseaux de neurones : on parle de « boîte noire ». La littérature montre que ces algorithmes sont performants, mais ne permettent pas d'expliquer pourquoi, et comment tel algorithme arrive à telle conclusion, même si celle-ci est correcte. Les cartes d'activation sont dans ce cadre un outil intéressant pour expliquer les performances de ces modèles, en montrant les zones jugées comme pertinentes par l'algorithme pour conclure à un résultat.(101)

La comparabilité entre les méthodes de segmentation automatique ou semi-automatique décrites dans la littérature est limitée par des bases de données et des schémas de validation non publics ainsi que par l'utilisation de différentes vérités terrain(51).

Une segmentation automatique précise et reproductible est une tâche difficile en raison de divers facteurs tels que différentes régions d'une lésion qui peuvent avoir des propriétés de contraste différentes dans différents types d'imagerie médicale(51) et une grande variation dans la forme, la localisation et l'apparence de la lésion(52). D'ailleurs, malgré de bonnes performances dans notre étude, l'algorithme de Deep Learning présentait, comme vu plus haut, un pourcentage assez élevé de faux positifs avec l'inclusion dans la segmentation lésionnelle, d'artéfacts et de fausses lésions parenchymateuses.

Comme discuté précédemment, différents paramètres tel que la maladie des petits vaisseaux, largement présente chez les patients victimes d'AVC, est associée à des modifications diffuses du parenchyme dans l'imagerie, ce qui ajoute de la complexité aux tâches de segmentation en raison des valeurs quantitatives non homogènes du parenchyme cérébral, en particulier dans la substance blanche.

Travailler sur une population plus homogène en terme de séquences d'acquisition et l'utilisation de la data-augmentation sont des pistes intéressantes à explorer pour augmenter les performances d'un algorithme d'IA pour les petites bases de données, aux dépens cependant de sa généralisabilité.

Parallèlement, l'utilisation de multiples séquences et de l'imagerie multimodale incluant des acquisitions classiques et de l'imagerie de perfusion permettent d'améliorer les performances des modèles d'IA par rapport à un algorithme qui travaille sur un seul type d'images.(41,102,103)

V. CONCLUSION

Notre étude a montré des performances de segmentation de lésions ischémiques aiguës proches entre une méthode manuelle par un radiologue et une méthode automatique par un algorithme d'intelligence artificielle à réseaux de neurones convolutif. La méthode manuelle reste globalement supérieure en terme de reproductibilité et de performance pour toute taille de lésion. Notre algorithme de Deep Learning a l'avantage d'être automatique, rapide et performant, mais manque de fiabilité pour les lésions de petit volume. Ses performances s'améliorent avec l'augmentation en taille des lésions et la qualité d'examen.

Le logiciel de détection et segmentation à seuillage sur cartographie d'ADC réalisait une bonne segmentation des lésions ischémiques mais présentait trop de faux positifs pour être utilisée de façon totalement automatique, une correction humaine étant toujours nécessaire. Cette méthode semi-automatique présente néanmoins un gain de temps considérable en pratique par rapport à une segmentation manuelle.

Les marqueurs de la maladie des petits vaisseaux n'ont pas montré d'impact sur les performances des différentes méthodes de segmentation étudiées. Les performances de segmentation ont par contre été influencées par le volume lésionnel et la qualité de l'examen.

VI. BIBLIOGRAPHIE

1. Lecoffre C. L'accident vasculaire cérébral en France : patients hospitalisés pour AVC en 2014 et évolutions 2008-2014. :11.
2. Lecoffre C. Mortalité par accident vasculaire cérébral en France en 2013 et évolutions 2008-2013. :6.
3. Maïer B, Boursin P, Mazighi M. Phase aiguë des accidents vasculaires cérébraux, chaque minute compte. *Soins*. 2018;63(828):28–31. doi: 10.1016/j.soins.2018.06.009.
4. Baron JC. Mapping the ischaemic penumbra with PET: implications for acute stroke treatment. *Cerebrovasc Dis*. 1999;9(4):193–201. doi: 10.1159/000015955.
5. Berkhemer OA, Fransen PSS, Beumer D, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. *N Engl J Med*. 2015;372(1):11–20. doi: 10.1056/NEJMoa1411587.
6. Rosenberg GA. Ischemic brain edema. *Prog Cardiovasc Dis*. 1999;42(3):209–216. doi: 10.1016/s0033-0620(99)70003-4.
7. Oppenheim C, Samson Y, Manai R, et al. Prediction of malignant middle cerebral artery infarction by diffusion-weighted imaging. *Stroke*. 2000;31(9):2175–2181. doi: 10.1161/01.str.31.9.2175.
8. Hoa D, Micheau A, Gahide G, et al. INF-WP-5 L'IRM pas a pas : un module interactif d'enseignement des bases physiques de l'IRM. *Journal de Radiologie*. 2007;88(10):1545–1546. doi: 10.1016/S0221-0363(07)81764-1.
9. Oppenheim C, Naggara O, Hamon M, et al. Imagerie par résonance magnétique de diffusion de l'encéphale chez l'adulte : technique, résultats normaux et pathologiques. *EMC - Radiologie*. 2005;2(2):133–164. doi: 10.1016/j.emcrad.2005.01.004.
10. Purushotham A, Mlynash M, Olivot J, et al. Apparent Diffusion Coefficient Distinguishes Ischemic Core from Reversible Diffusion Lesions. 2009. p. E115–E116.
11. An H, Ford AL, Vo K, Powers WJ, Lee J-M, Lin W. Signal Evolution and Infarction Risk for Apparent Diffusion Coefficient Lesions in Acute Ischemic Stroke Are Both Time- and Perfusion-Dependent. *Stroke*. American Heart Association; 2011;42(5):1276–1281. doi: 10.1161/STROKEAHA.110.610501.
12. Kidwell CS, Alger JR, Di Salle F, et al. Diffusion MRI in patients with transient ischemic attacks. *Stroke*. 1999;30(6):1174–1180. doi: 10.1161/01.str.30.6.1174.
13. Comprendre l'IRM. . <https://www.elsevier-masson.fr/comprendre-lirm-9782294753718.html>. Accessed June 20, 2021.
14. Bitar R, Leung G, Perng R, et al. MR Pulse Sequences: What Every Radiologist Wants to Know but Is Afraid to Ask. *RadioGraphics*. 2006;26(2):513–537. doi: 10.1148/rg.262055063.
15. Stejskal EO, Tanner JE. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J Chem Phys*. American Institute of Physics; 1965;42(1):288–292. doi: 10.1063/1.1695690.
16. Jauch EC, Saver JL, Adams HP, et al. Guidelines for the Early Management of Patients With Acute Ischemic Stroke. *Stroke*. American Heart Association; 2013;44(3):870–947. doi: 10.1161/STR.0b013e318284056a.

17. Labriffe M, Leiber L-M. Diffusion, cartographie ADC, effet T2 en IRM cérébrale : kézaco ? *Journal d'imagerie diagnostique et interventionnelle*. 2019;2(6):289–293. doi: 10.1016/j.jidi.2019.07.011.
18. Allen LM, Hasso AN, Handwerker J, Farid H. Sequence-specific MR Imaging Findings That Are Useful in Dating Ischemic Stroke. *RadioGraphics*. 2012;32(5):1285–1297. doi: 10.1148/rg.325115760.
19. Lansberg MG, Thijs VN, O'Brien MW, et al. Evolution of apparent diffusion coefficient, diffusion-weighted, and T2-weighted signal intensity of acute stroke. *AJNR Am J Neuroradiol*. 2001;22(4):637–644.
20. Saver JL, Goyal M, Bonafe A, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *N Engl J Med*. 2015;372(24):2285–2295. doi: 10.1056/NEJMoa1415061.
21. Emberson J, Lees KR, Lyden P, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. *Lancet*. 2014;384(9958):1929–1935. doi: 10.1016/S0140-6736(14)60584-5.
22. Ebinger M, Galinovic I, Rozanski M, Brunecker P, Endres M, Fiebach JB. Fluid-attenuated inversion recovery evolution within 12 hours from stroke onset: a reliable tissue clock? *Stroke*. 2010;41(2):250–255. doi: 10.1161/STROKEAHA.109.568410.
23. Thomalla G, Cheng B, Ebinger M, et al. DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4-5 h of symptom onset (PRE-FLAIR): a multicentre observational study. *Lancet Neurol*. 2011;10(11):978–986. doi: 10.1016/S1474-4422(11)70192-2.
24. Oppenheim C, Logak M, Dormont D, et al. Diagnosis of acute ischaemic stroke with fluid-attenuated inversion recovery and diffusion-weighted sequences. *Neuroradiology*. 2000;42(8):602–607. doi: 10.1007/s002340000356.
25. Cho A-H, Kim JS, Kim S-J, et al. Focal fluid-attenuated inversion recovery hyperintensity within acute diffusion-weighted imaging lesions is associated with symptomatic intracerebral hemorrhage after thrombolysis. *Stroke*. 2008;39(12):3424–3426. doi: 10.1161/STROKEAHA.108.516740.
26. Lansberg MG, Straka M, Kemp S, et al. MRI profile and response to endovascular reperfusion after stroke (DEFUSE 2): a prospective cohort study. *Lancet Neurol*. 2012;11(10):860–867. doi: 10.1016/S1474-4422(12)70203-X.
27. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. *N Engl J Med*. 2018;378(1):11–21. doi: 10.1056/NEJMoa1706442.
28. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. *N Engl J Med*. 2018;378(8):708–718. doi: 10.1056/NEJMoa1713973.
29. Inoue M, Mlynash M, Straka M, et al. Patients with the malignant profile within 3 hours of symptom onset have very poor outcomes after intravenous tissue-type plasminogen activator therapy. *Stroke*. 2012;43(9):2494–2496. doi: 10.1161/STROKEAHA.112.653329.
30. Campbell BCV, Mitchell PJ, Kleinig TJ, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med*. 2015;372(11):1009–1018. doi: 10.1056/NEJMoa1414792.

31. Casolla B, Kuchcinski G, Kyheng M, et al. Infarct Volume Before Hemispherectomy in Large Middle Cerebral Artery Infarcts Poorly Predicts Catastrophic Outcome. *Stroke*. American Heart Association; 2020;51(8):2404–2410. doi: 10.1161/STROKEAHA.120.029920.
32. Vahedi K, Hofmeijer J, Juettler E, et al. Early decompressive surgery in malignant infarction of the middle cerebral artery: a pooled analysis of three randomised controlled trials. *The Lancet Neurology*. Elsevier; 2007;6(3):215–222. doi: 10.1016/S1474-4422(07)70036-4.
33. Committee TESO (ESO) EC and the EW. Guidelines for Management of Ischaemic Stroke and Transient Ischaemic Attack 2008. CED. Karger Publishers; 2008;25(5):457–507. doi: 10.1159/000131083.
34. Powers WJ, Rabinstein AA, Ackerson T, et al. Guidelines for the Early Management of Patients With Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*. American Heart Association; 2019;50(12):e344–e418. doi: 10.1161/STR.0000000000000211.
35. Xie Y, Oppenheim C, Guillemin F, et al. Pretreatment lesional volume impacts clinical outcome and thrombectomy efficacy. *Ann Neurol*. 2018;83(1):178–185. doi: 10.1002/ana.25133.
36. Ozbek D, Ozturk Tan O, Ekinci G, Midi I. Risk of hemorrhage in ischemic stroke and its relationship with cerebral microbleeds. *Clin Neurol Neurosurg*. 2018;168:112–117. doi: 10.1016/j.clineuro.2018.03.007.
37. Singer OC, Kurre W, Humpich MC, et al. Risk assessment of symptomatic intracerebral hemorrhage after thrombolysis using DWI-ASPECTS. *Stroke*. 2009;40(8):2743–2748. doi: 10.1161/STROKEAHA.109.550111.
38. Vermeer SE, Prins ND, den Heijer T, Hofman A, Koudstaal PJ, Breteler MMB. Silent Brain Infarcts and the Risk of Dementia and Cognitive Decline. *New England Journal of Medicine*. Massachusetts Medical Society; 2003;348(13):1215–1222. doi: 10.1056/NEJMoa022066.
39. Kalaria RN, Akinyemi R, Ihara M. Stroke injury, cognitive impairment and vascular dementia. *Biochim Biophys Acta*. 2016;1862(5):915–925. doi: 10.1016/j.bbadis.2016.01.015.
40. Pexman JH, Barber PA, Hill MD, et al. Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke. *AJNR Am J Neuroradiol*. 2001;22(8):1534–1542.
41. Soun JE, Chow DS, Nagamine M, et al. Artificial Intelligence and Acute Stroke Imaging. *AJNR Am J Neuroradiol*. 2021;42(1):2–11. doi: 10.3174/ajnr.A6883.
42. Julia L, Khayat O, Gassée J-L. L'intelligence artificielle n'existe pas. 2019.
43. TURING AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*. 1950;LIX(236):433–460. doi: 10.1093/mind/LIX.236.433.
44. Samuel A. Some Studies in Machine Learning Using the Game of Checkers. *IBM J Res Dev*. 1959; doi: 10.1147/rd.33.0210.
45. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA, USA: MIT Press; 2016.
46. Le Cun Y, Matan O, Boser B, et al. Handwritten zip code recognition with multilayer networks. 10th International Conference on Pattern Recognition [1990] Proceedings. 1990. p. 35–40 vol.2. doi: 10.1109/ICPR.1990.119325.

47. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*. Curran Associates, Inc.; 2012. <https://papers.nips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>. Accessed August 17, 2021.
48. Gautherot M, Yepremian S, Bretzner M, et al. 15 minutes pour comprendre et évaluer un logiciel d'intelligence artificielle appliquée à l'imagerie médicale. *Journal d'imagerie diagnostique et interventionnelle*. 2021;4(3):167–171. doi: 10.1016/j.jidi.2020.12.001.
49. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp*. 2018;2(1):35. doi: 10.1186/s41747-018-0061-6.
50. Waymel Q, Badr S, Demondion X, Cotten A, Jacques T. Impact of the rise of artificial intelligence in radiology: What do radiologists think? *Diagnostic and Interventional Imaging*. 2019;100(6):327–336. doi: 10.1016/j.diii.2019.03.015.
51. Sarmiento RM, Vasconcelos FFX, Reboucas Filho PP, Wu W, De Albuquerque VHC. Automatic Neuroimage Processing and Analysis in Stroke - A Systematic Review. *IEEE Rev Biomed Eng*. 2019;1–1. doi: 10.1109/RBME.2019.2934500.
52. Maier O, Schröder C, Forkert ND, Martinetz T, Handels H. Classifiers for Ischemic Stroke Lesion Segmentation: A Comparison Study. Hu D, editor. *PLoS ONE*. 2015;10(12):e0145118. doi: 10.1371/journal.pone.0145118.
53. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. SpringerOpen; 2018;9(4):611–629. doi: 10.1007/s13244-018-0639-9.
54. Chakraverty S, Sahoo DM, Mahato NR. Hebbian Learning Rule. In: Chakraverty S, Sahoo DM, Mahato NR, editors. *Concepts of Soft Computing: Fuzzy and ANN with Programming*. Singapore: Springer; 2019. p. 175–182. doi: 10.1007/978-981-13-7430-2_12.
55. Keyzers C, Gazzola V. Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos Trans R Soc Lond B Biol Sci*. 2014;369(1644):20130175. doi: 10.1098/rstb.2013.0175.
56. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23(19):2507–2517. doi: 10.1093/bioinformatics/btm344.
57. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing; 2015. p. 234–241. doi: 10.1007/978-3-319-24574-4_28.
58. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *The Lancet Neurology*. 2013;12(8):822–838. doi: 10.1016/S1474-4422(13)70124-8.
59. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *The Lancet Neurology*. 2010;9(7):689–701. doi: 10.1016/S1474-4422(10)70104-6.
60. Greenberg SM, Nandigam RNK, Delgado P, et al. Microbleeds versus macrobleeds: evidence for distinct entities. *Stroke*. 2009;40(7):2382–2386. doi: 10.1161/STROKEAHA.109.548974.

61. Koch S, McClendon MS, Bhatia R. Imaging evolution of acute lacunar infarction: leukoaraiosis or lacune? *Neurology*. 2011;77(11):1091–1095. doi: 10.1212/WNL.0b013e31822e1470.
62. Conklin J, Silver FL, Mikulis DJ, Mandell DM. Are acute infarcts the cause of leukoaraiosis? Brain mapping for 16 consecutive weeks. *Ann Neurol*. 2014;76(6):899–904. doi: 10.1002/ana.24285.
63. Mascalchi M, Moretti M, Nave RD, et al. Longitudinal evaluation of leukoaraiosis with whole brain ADC histograms. *Neurology*. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology; 2002;59(6):938–940. doi: 10.1212/WNL.59.6.938.
64. Helenius J, Soine L, Salonen O, Kaste M, Tatlisumak T. Leukoaraiosis, ischemic stroke, and normal white matter on diffusion-weighted MRI. *Stroke*. 2002;33(1):45–50. doi: 10.1161/hs0102.101228.
65. Jokinen H, Schmidt R, Ropele S, et al. Diffusion changes predict cognitive and functional outcome: the LADIS study. *Ann Neurol*. 2013;73(5):576–583. doi: 10.1002/ana.23802.
66. Guerrero R, Qin C, Oktay O, et al. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *NeuroImage: Clinical*. 2018;17:918–934. doi: 10.1016/j.nicl.2017.12.022.
67. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine*. Cambridge University Press; 1989;19(4):1015–1022. doi: 10.1017/S0033291700005742.
68. Brott T, Adams HP, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*. 1989;20(7):864–870. doi: 10.1161/01.str.20.7.864.
69. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott Med J*. 1957;2(5):200–215. doi: 10.1177/003693305700200504.
70. Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. *Stroke*. 1999;30(8):1538–1541. doi: 10.1161/01.str.30.8.1538.
71. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143–155. doi: 10.1002/hbm.10062.
72. Jenkinson M, Pechaud M, Smith S. BET2 - MR-Based Estimation of Brain, Skull and Scalp Surfaces. :1.
73. Manjón JV, Coupé P. volBrain: An Online MRI Brain Volumetry System. *Front Neuroinform*. Frontiers; 2016;10. doi: 10.3389/fninf.2016.00030.
74. Ospel JM, Jaffray A, Schulze-Zachau V, Kozerke S, Federau C. Spatial Resolution and the Magnitude of Infarct Volume Measurement Error in DWI in Acute Ischemic Stroke. *AJNR Am J Neuroradiol*. 2020;41(5):792–797. doi: 10.3174/ajnr.A6520.
75. Bhadauria HS, Dewal ML. Efficient denoising technique for CT images to enhance brain hemorrhage segmentation. *J Digit Imaging*. 2012;25(6):782–791. doi: 10.1007/s10278-012-9453-y.
76. Hajimani E, Ruano MG, Ruano AE. An intelligent support system for automatic detection of cerebral vascular accidents from brain CT images. *Comput Methods Programs Biomed*. 2017;146:109–123. doi: 10.1016/j.cmpb.2017.05.005.

77. de Haan B, Clas P, Juenger H, Wilke M, Karnath H-O. Fast semi-automated lesion demarcation in stroke. *Neuroimage Clin.* 2015;9:69–74. doi: 10.1016/j.nicl.2015.06.013.
78. Clèrigues A, Valverde S, Bernal J, Freixenet J, Oliver A, Lladó X. Acute and sub-acute stroke lesion segmentation from multimodal MRI. *Computer Methods and Programs in Biomedicine.* 2020;194:105521. doi: 10.1016/j.cmpb.2020.105521.
79. Qiu W, Kuang H, Teleg E, et al. Machine Learning for Detecting Early Infarction in Acute Stroke with Non-Contrast-enhanced CT. *Radiology. Radiological Society of North America;* 2020;294(3):638–644. doi: 10.1148/radiol.2020191193.
80. Kasasbeh A, Christensen S, Parsons M, Campbell B, Albers G, Lansberg M. Artificial Neural Network Computer Tomography Perfusion Prediction of Ischemic Core. *Stroke.* 2019;50. doi: 10.1161/STROKEAHA.118.022649.
81. Zhang R, Zhao L, Lou W, et al. Automatic Segmentation of Acute Ischemic Stroke From DWI Using 3-D Fully Convolutional DenseNets. *IEEE Trans Med Imaging.* 2018;37(9):2149–2160. doi: 10.1109/TMI.2018.2821244.
82. Chen L, Bentley P, Rueckert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. *Neuroimage Clin.* 2017;15:633–643. doi: 10.1016/j.nicl.2017.06.016.
83. Boldsen JK, Engedal TS, Pedraza S, et al. Better Diffusion Segmentation in Acute Ischemic Stroke Through Automatic Tree Learning Anomaly Segmentation. *Frontiers in Neuroinformatics.* 2018;12:21. doi: 10.3389/fninf.2018.00021.
84. Ramamurthy K, Menaka R. Delineation of ischemic lesion from brain MRI using symmetric bit plane pattern and curvelet co-occurrence matrix. *International Journal of Innovative Technology and Exploring Engineering.* 2019;8:201–206.
85. McKinley R, Häni L, Gralla J, et al. Fully automated stroke tissue estimation using random forest classifiers (FASTER). *J Cereb Blood Flow Metab.* 2017;37(8):2728–2741. doi: 10.1177/0271678X16674221.
86. Nag MK, Koley S, China D, et al. Computer-assisted delineation of cerebral infarct from diffusion-weighted MRI using Gaussian mixture model. *Int J Comput Assist Radiol Surg.* 2017;12(4):539–552. doi: 10.1007/s11548-017-1520-x.
87. Pustina D, Coslett H, Turkeltaub P, Tustison N, Schwartz M, Avants B. Automated segmentation of chronic stroke lesions using LINDA: Lesion Identification with Neighborhood Data Analysis. *Human brain mapping.* 2016;37. doi: 10.1002/hbm.23110.
88. Forkert ND, Verleger T, Cheng B, Thomalla G, Hilgetag CC, Fiehler J. Multiclass Support Vector Machine-Based Lesion Mapping Predicts Functional Outcome in Ischemic Stroke Patients. Baron J-C, editor. *PLoS ONE.* 2015;10(6):e0129569. doi: 10.1371/journal.pone.0129569.
89. Maier O, Wilms M, von der Gablentz J, Krämer UM, Münte TF, Handels H. Extra Tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences. *Journal of Neuroscience Methods.* 2015;240:89–100. doi: 10.1016/j.jneumeth.2014.11.011.
90. Purushotham A, Campbell B, Straka M, et al. Apparent diffusion coefficient threshold for delineation of ischemic core: Research. *International journal of stroke : official journal of the International Stroke Society.* 2013;10. doi: 10.1111/ijss.12068.

91. Liew S-L, Anglin JM, Banks NW, et al. A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. *Sci Data*. 2018;5:180011. doi: 10.1038/sdata.2018.11.
92. Chilamkurthy S, Ghosh R, Tanamala S, et al. Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans. arXiv:180305854 [cs]. 2018; <http://arxiv.org/abs/1803.05854>. Accessed August 15, 2021.
93. Flanders AE, Prevedello LM, Shih G, et al. Construction of a Machine Learning Dataset through Collaboration: The RSNA 2019 Brain CT Hemorrhage Challenge. *Radiology: Artificial Intelligence*. Radiological Society of North America; 2020;2(3):e190211. doi: 10.1148/ryai.2020190211.
94. Winzeck S, Hakim A, McKinley R, et al. ISLES 2016 and 2017-Benchmarking Ischemic Stroke Lesion Outcome Prediction Based on Multispectral MRI. *Frontiers in Neurology*. 2018;9:679. doi: 10.3389/fneur.2018.00679.
95. Maier O, Menze BH, von der Gablentz J, et al. ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical Image Analysis*. 2017;35:250–269. doi: 10.1016/j.media.2016.07.009.
96. Kamnitsas K, Chen L, Ledig C, Rueckert D, Glocker B. Multi-Scale 3D Convolutional Neural Networks for Lesion Segmentation in Brain MRI. :4.
97. McKinley R, Häni L, Wiest R, Reyes M. Segmenting the Ischemic Penumbra: A Decision Forest Approach with Automatic Threshold Finding. 2016. p. 275–283. doi: 10.1007/978-3-319-30858-6_24.
98. Lyoubi-Idrissi AL, Jouvent E, Poupon C, Chabriat H. Diffusion magnetic resonance imaging in cerebral small vessel disease. *Revue Neurologique*. 2017;173(4):201–210. doi: 10.1016/j.neurol.2017.03.005.
99. Pelletier A, Periot O, Dilharreguy B, et al. Age-Related Modifications of Diffusion Tensor Imaging Parameters and White Matter Hyperintensities as Inter-Dependent Processes. *Front Aging Neurosci*. *Frontiers*; 2016;7. doi: 10.3389/fnagi.2015.00255.
100. Fang H, Zhao L, Pei L, et al. Severity of White Matter Lesions Correlates with Subcortical Diffusion-Weighted Imaging Abnormalities and Predicts Stroke Risk. *Journal of Stroke and Cerebrovascular Diseases*. Elsevier; 2017;26(12):2964–2970. doi: 10.1016/j.jstrokecerebrovasdis.2017.07.025.
101. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning Deep Features for Discriminative Localization. arXiv:151204150 [cs]. 2015; <http://arxiv.org/abs/1512.04150>. Accessed September 1, 2021.
102. Stier N, Vincent N, Liebeskind D, Scalzo F. Deep learning of tissue fate features in acute ischemic stroke. 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2015. p. 1316–1321. doi: 10.1109/BIBM.2015.7359869.
103. Nielsen A, Hansen MB, Tietze A, Mouridsen K. Prediction of Tissue Outcome and Assessment of Treatment Effect in Acute Ischemic Stroke Using Deep Learning. *Stroke*. American Heart Association; 2018;49(6):1394–1401. doi: 10.1161/STROKEAHA.117.019740.

AUTEUR : Nom : YEPREMIAN Prénom : Sevak

Date de soutenance : 01/10/2021

Titre de la thèse : Application de l'intelligence artificielle à la segmentation de lésions ischémiques : comparaison aux méthodes manuelles et semi-automatiques

Thèse - Médecine - Lille 2021

Cadre de classement : Radiologie et imagerie médicale

DES + spécialité : Radiologie, Neuroradiologie

Mots-clés : neuroradiologie, accident vasculaire cérébral, intelligence artificielle

Résumé :

Contexte : La quantification du volume d'un infarctus cérébral est un élément important dans la prise en charge thérapeutique et l'évaluation pronostique d'un accident vasculaire cérébral ischémique. L'objectif de notre étude est d'analyser les performances d'un algorithme d'intelligence artificielle (IA) dans la segmentation de ces lésions par rapport aux méthodes manuelles et semi-automatiques, ainsi que d'évaluer certains paramètres pouvant impacter leur performance.

Méthode : Les analyses portent sur 164 patients présentant un AVC ischémique de moins de 72h, issus de la cohorte Strokedem. Cinq méthodes de segmentation sont évaluées par rapport à une méthode de référence : une méthode manuelle, trois méthodes semi-automatiques (par défaut (A), avec sélection des coupes d'intérêt (B), avec sélection des coupes d'intérêt et correction des lésions (C)) et un algorithme d'IA développé pour l'étude. Leur performance était évaluée par le calcul du coefficient de Spearman, le coefficient de Dice et la distance des barycentres lésionnels. Les paramètres pouvant impacter les performances évaluées sont la maladie des petits vaisseaux, le volume lésionnel et la qualité d'examen.

Résultats : Les volumes médians segmentés sont de 1,83 [0,40-9,41] mL par le Gold Standard, 1,80 [0,43-8,16] mL par la méthode manuelle, 2,86 mL [0,50-10,84] par la méthode d'IA, 22,06 [10,06-37,03] mL par la méthode semi-automatique A, 2,55 [0,70-7,73] mL pour la méthode semi-automatique B et 1,49 [0,24-4,83] mL pour la méthode semi-automatique C. Les corrélations des volumes segmentés par rapport au Gold Standard sont élevées pour toutes les méthodes sauf pour la méthode semi-automatique A : méthode manuelle 0,954 ($p < 0,001$) ; IA 0,894 ($p < 0,001$) ; méthode semi-automatique A 0,213 ($p = 0,007$), méthode semi-automatique B 0,853 ($p < 0,001$), méthode semi-automatique C 0,896 ($p < 0,001$). Un impact significatif du volume et de la qualité d'examen a été retrouvé pour la segmentation par l'IA, avec une baisse des performances pour les petits volumes et pour les examens de qualité altérée.

Conclusion : Les performances de segmentation observées étaient proches entre la méthode manuelle, l'IA et la méthode semi-automatique C. La méthode manuelle reste globalement supérieure en terme de reproductibilité et de performance pour toute taille de lésion.

Composition du Jury :

Président : Professeur Jean-Pierre PRUVO

Asseseurs : Professeur Xavier LECLERC, Docteur Marco PASI, Monsieur Renaud LOPES

Directeur de thèse : Docteur Grégory KUCHCINSKI