



Université de Poitiers

Faculté de Médecine et Pharmacie

ANNEE 2021

THESE

POUR LE DIPLOME D'ETAT
DE DOCTEUR EN MEDECINE
(Décret du 25 novembre 2016)

Présentée et soutenue publiquement
Le 18 octobre 2021 à Poitiers
Par Madame Imane M'RABET KHARROUBI

Evaluation de la charge lésionnelle de sclérose en plaques en imagerie par résonance magnétique : segmentation automatisée par intelligence artificielle

COMPOSITION DU JURY

Président : Monsieur le Professeur Rémy GUILLEVIN

Membres :

Monsieur le Professeur Rémy GUILLEVIN
Monsieur le Professeur Jean-Philippe NEAU
Monsieur le Docteur Benoît GIANELLI
Monsieur le Docteur Nikolaos EFTHYMIADIS
Monsieur le Docteur Mathieu NAUDIN

Directeur de thèse : Monsieur le Docteur Benoît GIANELLI



Université de Poitiers

Faculté de Médecine et Pharmacie

ANNEE 2021

THESE

POUR LE DIPLOME D'ETAT
DE DOCTEUR EN MEDECINE
(Décret du 25 novembre 2016)

Présentée et soutenue publiquement
Le 18 octobre 2021 à Poitiers
Par Madame Imane M'RABET KHARROUBI

Evaluation de la charge lésionnelle de sclérose en plaques en imagerie par résonance magnétique : segmentation automatisée par intelligence artificielle

COMPOSITION DU JURY

Président : Monsieur le Professeur Rémy GUILLEVIN

Membres :

Monsieur le Professeur Rémy GUILLEVIN
Monsieur le Professeur Jean-Philippe NEAU
Monsieur le Docteur Benoît GIANELLI
Monsieur le Docteur Nikolaos EFTHYMIADIS
Monsieur le Docteur Mathieu NAUDIN

Directeur de thèse : Monsieur le Docteur Benoît GIANELLI

Le Doyen,

Année universitaire 2020 - 2021

LISTE DES ENSEIGNANTS DE MEDECINE
Professeurs des Universités-Praticiens Hospitaliers

- BOULETI Claire, cardiologie
- BOURMEYSTER Nicolas, biologie cellulaire
- BRIDOUX Frank, néphrologie
- BURUCOA Christophe, bactériologie – virologie
- CHEZE-LE REST Catherine, biophysique et médecine nucléaire
- CHRISTIAENS Luc, cardiologie
- CORBI Pierre, chirurgie thoracique et cardio-vasculaire
- DAHYOT-FIZELIER Claire, anesthésiologie – réanimation
- DEBAENE Bertrand, anesthésiologie réanimation
- DEBIAIS Françoise, rhumatologie
- DROUOT Xavier, physiologie
- DUFOUR Xavier, Oto-Rhino-Laryngologie
- FAURE Jean-Pierre, anatomie
- FRASCA Denis, anesthésiologie-réanimation
- FRITEL Xavier, gynécologie-obstétrique
- GERVAIS Elisabeth, rhumatologie
- GICQUEL Ludovic, pédopsychiatrie
- GILBERT Brigitte, génétique
- GOMBERT Jean-Marc, immunologie
- GOUJON Jean-Michel, anatomie et cytologie pathologiques
- GUILLEVIN Rémy, radiologie et imagerie médicale
- HAUET Thierry, biochimie et biologie moléculaire
- INGRAND Pierre, biostatistiques, informatique médicale
- ISAMBERT Nicolas, cancérologie
- JAAFARI Nematollah, psychiatrie d'adultes
- JABER Mohamed, cytologie et histologie
- JAYLE Christophe, chirurgie thoracique et cardio-vasculaire
- KARAYAN-TAPON Lucie, cancérologie
- KEMOUN Gilles, médecine physique et de réadaptation (**en disponibilité**)
- KRAIMPS Jean-Louis, chirurgie générale
- LECLERE Franck, chirurgie plastique, reconstructrice
- LECRON Jean-Claude, biochimie et biologie moléculaire
- LELEU Xavier, hématologie
- LEVARD Guillaume, chirurgie infantile
- LEVEQUE Nicolas, bactériologie-virologie
- LEVEZIEL Nicolas, ophtalmologie
- MACCHI Laurent, hématologie
- MCHEIK Jiad, chirurgie infantile
- MEURICE Jean-Claude, pneumologie
- MIGEOT Virginie, santé publique
- MILLOT Frédéric, pédiatrie, oncologie pédiatrique
- MIMOZ Olivier, anesthésiologie – réanimation
- NEAU Jean-Philippe, neurologie
- ORIOT Denis, pédiatrie
- PACCALIN Marc, gériatrie
- PERAULT-POCHAT Marie-Christine, pharmacologie clinique
- PERDRISOT Rémy, biophysique et médecine nucléaire
- PIERRE Fabrice, gynécologie et obstétrique
- PRIES Pierre, chirurgie orthopédique et traumatologique
- RAMMAERT-PALTRIE Blandine, maladies infectieuses
- RICHER Jean-Pierre, anatomie
- RIGOARD Philippe, neurochirurgie
- ROBERT René, réanimation

- ROBLLOT France, maladies infectieuses, maladies tropicales
- ROBLLOT Pascal, médecine interne
- RODIER Marie-Hélène, parasitologie et mycologie (**retraite 01/03/2021**)
- SAULNIER Pierre-Jean, thérapeutique
- SCHNEIDER Fabrice, chirurgie vasculaire
- SILVAIN Christine, hépato-gastro-entérologie
- TASU Jean-Pierre, radiologie et imagerie médicale
- THIERRY Antoine, néphrologie
- THILLE Arnaud, réanimation
- TOUGERON David, gastro-entérologie
- WAGER Michel, neurochirurgie
- XAVIER Jean, pédopsychiatrie

Maîtres de Conférences des Universités-Praticiens Hospitaliers

- ALBOUY-LLATY Marion, santé publique (**en mission 2020/21**)
- ALLAIN Géraldine, chirurgie thoracique et cardio-vasculaire
- BEBY-DEFAUX Agnès, bactériologie – virologie (**en cours d'intégration PH**)
- BEN-BRIK Eric, médecine du travail (**en détachement**)
- BILAN Frédéric, génétique
- BOISSON Matthieu, anesthésiologie-réanimation et médecine péri-opératoire (**en mission 1 an à/c nov.2020**)
- CASTEL Olivier, bactériologie - virologie – hygiène
- CAYSSIALS Emilie, hématologie
- COUDROY Rémy, réanimation
- CREMNITER Julie, bactériologie – virologie
- DIAZ Véronique, physiologie
- FROUIN Eric, anatomie et cytologie pathologiques
- GACHON Bertrand, gynécologie-obstétrique
- GARCIA Magali, bactériologie-virologie
- JAVAUGUE Vincent, néphrologie
- KERFORNE Thomas, anesthésiologie-réanimation et médecine péri-opératoire
- LAFAY-CHEBASSIER Claire, pharmacologie clinique
- LIUU Evelyne, gériatrie
- MARTIN Mickaël, médecine interne
- PALAZZO Paola, neurologie (**en dispo 1 an**)
- PERRAUD Estelle, parasitologie et mycologie
- SAPANET Michel, médecine légale
- THUILLIER Raphaël, biochimie et biologie moléculaire

Professeur des universités

- PELLERIN Luc, biochimie et biologie moléculaire

Professeur des universités de médecine générale

- BINDER Philippe

Professeurs associés de médecine générale

- BIRAULT François
- FRECHE Bernard
- MIGNOT Stéphanie
- PARTHENAY Pascal
- VALETTE Thierry
- VICTOR-CHAPLET Valérie

Maîtres de Conférences associés de médecine générale

- AUDIER Pascal
- ARCHAMBAULT Pierrick
- BRABANT Yann
- JEDAT Vincent

Enseignants d'Anglais

- DEBAIL Didier, professeur certifié

Professeurs émérites

- CARRETIER Michel, chirurgie générale (08/2021)
- GIL Roger, neurologie (08/2023)
- GOMES DA CUNHA José, médecine générale (08/2021)
- GUILHOT-GAUDEFFROY François, hématologie et transfusion (08/2023)
- HERPIN Daniel, cardiologie (08/2023)
- KITZIS Alain, biologie cellulaire (16/02/2021)
- MARECHAUD Richard, médecine interne (24/11/2023)
- MAUCO Gérard, biochimie et biologie moléculaire (08/2021)
- RICCO Jean-Baptiste, chirurgie vasculaire (08/2022)
- SENON Jean-Louis, psychiatrie d'adultes (08/2023)
- TOUCHARD Guy, néphrologie (08/2021)

Professeurs et Maîtres de Conférences honoraires

- AGIUS Gérard, bactériologie-virologie
- ALCALAY Michel, rhumatologie
- ALLAL Joseph, thérapeutique (ex-émérite)
- ARIES Jacques, anesthésiologie-réanimation
- BABIN Michèle, anatomie et cytologie pathologiques
- BABIN Philippe, anatomie et cytologie pathologiques
- BARBIER Jacques, chirurgie générale (ex-émérite)
- BARRIERE Michel, biochimie et biologie moléculaire
- BECQ-GIRAUDON Bertrand, maladies infectieuses, maladies tropicales (ex-émérite)
- BEGON François, biophysique, médecine nucléaire
- BOINOT Catherine, hématologie – transfusion
- BONTOUX Daniel, rhumatologie (ex-émérite)
- BURIN Pierre, histologie
- CASTETS Monique, bactériologie -virologie – hygiène
- CAVELLIER Jean-François, biophysique et médecine nucléaire
- CHANSIGAUD Jean-Pierre, biologie du développement et de la reproduction
- CLARAC Jean-Pierre, chirurgie orthopédique
- DABAN Alain, oncologie radiothérapie (ex-émérite)
- DAGREGORIO Guy, chirurgie plastique et reconstructrice
- DESMAREST Marie-Cécile, hématologie
- DEMANGE Jean, cardiologie et maladies vasculaires
- DORE Bertrand, urologie (ex-émérite)
- EUGENE Michel, physiologie (ex-émérite)
- FAUCHERE Jean-Louis, bactériologie-virologie (ex-émérite)
- FONTANEL Jean-Pierre, Oto-Rhino Laryngologie (ex-émérite)
- GRIGNON Bernadette, bactériologie
- GUILLARD Olivier, biochimie et biologie moléculaire
- GUILLET Gérard, dermatologie
- JACQUEMIN Jean-Louis, parasitologie et mycologie médicale
- KAMINA Pierre, anatomie (ex-émérite)
- KLOSSEK Jean-Michel, Oto-Rhino-Laryngologie
- LAPIERRE Françoise, neurochirurgie (ex-émérite)
- LARSEN Christian-Jacques, biochimie et biologie moléculaire
- LEVILLAIN Pierre, anatomie et cytologie pathologiques
- MAIN de BOISSIERE Alain, pédiatrie
- MARCELLI Daniel, pédopsychiatrie (ex-émérite)
- MARILLAUD Albert, physiologie
- MENU Paul, chirurgie thoracique et cardio-vasculaire (ex-émérite)
- MORICHAU-BEAUCHANT Michel, hépato-gastro-entérologie
- MORIN Michel, radiologie, imagerie médicale
- PAQUEREAU Joël, physiologie
- POINTREAU Philippe, biochimie
- POURRAT Olivier, médecine interne (ex-émérite)
- REISS Daniel, biochimie
- RIDEAU Yves, anatomie
- SULTAN Yvette, hématologie et transfusion
- TALLINEAU Claude, biochimie et biologie moléculaire
- TANZER Joseph, hématologie et transfusion (ex-émérite)
- TOURANI Jean-Marc, oncologie
- VANDERMARCO Guy, radiologie et imagerie médicale

REMERCIEMENTS

A Monsieur le Professeur Rémy GUILLEVIN, Président du jury,

Vous me faites l'honneur d'avoir accepté de juger mon travail et de présider mon jury, je vous en remercie. Je vous remercie aussi pour votre encadrement bienveillant pendant mon internat.

A Monsieur le Docteur Benoît GIANELLI, Directeur de thèse,

Je te remercie d'avoir accepté de diriger ma thèse et de m'avoir confié ce sujet sur lequel j'ai vraiment apprécié travailler. Merci pour ta disponibilité, ton aide et ton encadrement de qualité tout au long de ce projet.

A Monsieur le Docteur Mathieu NAUDIN et Monsieur Pierre FAYOLLE,

Un grand merci pour votre aide précieuse à toutes les étapes de ce projet. J'ai énormément appris sur l'intelligence artificielle à vos côtés et je vous en remercie. Merci pour votre patience et votre disponibilité face à mes (très) nombreuses questions !

A Madame le Docteur Carole GUILLEVIN,

Je vous remercie pour vos conseils et remarques constructives sur mon travail.

A Monsieur de Professeur Jean-Philippe NEAU,

Vous me faites l'honneur d'être membre de mon jury, je vous en remercie.

A Monsieur le Docteur Nikolaos EFTHYMIADIS,

Merci d'avoir accepté d'être membre de mon jury. Je te remercie pour ta disponibilité.

Aux radiologues qui m'ont transmis leur savoir,

Je vous remercie pour tout ce que vous m'avez appris. Merci pour votre patience et votre disponibilité qui ont permis d'enrichir mes connaissances. Votre passion pour ce métier est devenue mienne. Un grand merci.

Aux manipulateurs radio,

Sans vous notre métier ne serait pas possible. Merci pour votre aide au quotidien.

A mes co-internes, à la « promo arc-en-ciel »,

Vous avez contribué à rendre mon internat plus agréable et plus humain. Merci pour ces riches années !

A mes amies Bourguignonnes et Poitevines,

Merci pour tous ces moments, pour votre présence, votre écoute et votre soutien.

Aux « Pictaviennes »,

Merci de m'avoir accueillie dans votre joli groupe comme vous l'avez fait. La vie à Poitiers est encore plus belle depuis que je vous ai rencontrées !

A ma belle-famille,

Merci de m'avoir accueillie dans votre famille avec autant de gentillesse et de générosité.

Soundous, quel plaisir de partager avec toi cette passion commune pour la médecine. Je te souhaite la plus belle des réussites.

A ma famille, à Maman Mouy, à mes tantes Souad et Fouzia,

Merci d'avoir toujours cru en moi et de m'avoir soutenue pendant ces études de médecine. Merci pour tous vos encouragements.

A Papa Haj,

J'ai pensé à toi tout au long de ce travail, tout au long de mes études. On en a rêvé quand j'étais enfant, tu m'appelais ton petit docteur. Voilà, je l'ai fait ! Tu n'es pas là pour le voir mais je sais que tu aurais été fier de moi...

A Salem, mon cher mari,

Merci de m'avoir accompagnée toutes ces années. Merci pour la force que tu m'apportes au quotidien. Merci pour ta présence, ton soutien et ton amour qui me sont essentiels.

A Zayd,

Tu es la plus belle chose qui soit arrivée dans ma vie, mon rayon de soleil. Tu m'émerveilles un peu plus chaque jour et te voir grandir fait de moi une maman comblée.

TABLE DES MATIERES

Remerciements

Table des matières

Index des Figures et Tableaux

Index des abréviations

I.	Introduction.....	11
II.	Matériels et méthodes	15
1-	Patients	15
2-	Acquisition des IRM.....	15
3-	Segmentation manuelle : vérité terrain	15
4-	Traitement des images	16
5-	Réseau de neurones	16
6-	Analyse statistique et métriques	18
III.	Résultats.....	19
1-	Caractéristiques des patients	19
2-	Segmentation	20
	Comparaison des segmentations automatique et manuelle	20
	Exemples de segmentations	21
	Variation du Dice score en fonction de la charge lésionnelle	23
IV.	Discussion	24
V.	Conclusion.....	27

Références bibliographiques

Annexes

Annexe A – Echelle EDSS

Annexe B – Critères NEDA

Annexe C – Score Rio

Annexe D – Protocole OFSEP

Résumé et mots clés

Serment

INDEX DES FIGURES ET TABLEAUX

Figure 1 : Évolution du nombre d'IRM en France pour 1000 habitants de 2006 à 2019 (Données publiques de l'OCDE concernant l'IRM)

Figure 2 : Place du réseau de neurones convolutifs dans la terminologie hiérarchique de l'intelligence artificielle

Figure 3 : (a) Représentation schématique d'un réseau neuronal artificiel par analogie avec (b) un réseau neuronal biologique

Figure 4 : Schéma du principe de fonctionnement d'un réseau de neurones convolutifs, issu de l'article de Fernandez-Maloigne et Guillevin de 2019

Figure 5 : architecture du réseau de neurones MPU-net ++

Figure 6 : Schéma représentant la formule du Dice score

Tableau 1 : Caractéristiques de la population totale

Figure 7 : Variation des valeurs normalisées de métriques en fonction du seuil pour un cas.

Tableau 2 : Moyennes des métriques d'évaluation dans la population de validation (écart-type)

Figure 8 : Exemple de bonne similarité des segmentations

Figure 9 : Exemple de discordance : segmentation plus large du réseau de neurones comparativement à la segmentation manuelle

Figure 10 : Exemple d'un cas avec une charge lésionnelle très faible favorisant un Dice score bas

Figure 11 : Dice score en fonction du volume de charge lésionnelle.

Figure 12 : Exemple de segmentation automatique de fins hypersignaux au contact des ventricules latéraux.

INDEX DES ABREVIATIONS

SEP : Sclérose En Plaques

IRM : Imagerie par Résonance Magnétique

EDSS : Expanded Disability Status Scale

OFSEP : Observatoire Français de la Sclérose en Plaques

OCDE : Organisation de Coopération et de Développement Economiques

CNN : Convolutional Neural Network

IA : Intelligence Artificielle

MPU-net : Minimally Parameterized U-net

I3M : Imagerie Métabolique Multi-noyaux Multi-organes

FLAIR : Fluid Attenuated Inversion Recovery

MICCAI BraTS : Medical Imaging Computing & Computer Assisted Intervention Brain Tumor Segmentation

Se : Sensibilité

Sp : Spécificité

VPP : Valeur Prédictive Positive

VPN : Valeur Prédictive Négative

PPVL : Valeur Prédictive Positive Lésionnelle

SensL : Sensibilité lésionnelle

LCS : Liquide céphalo-spinal

NEDA : No Evidence of Disease Activity

I. Introduction

La sclérose en plaque (SEP) est une maladie neurologique chronique démyélinisante atteignant le système nerveux central associant des phénomènes inflammatoires et neurodégénératifs. Avec une prévalence d'environ 83 cas pour 100 000, il s'agit de la pathologie du système nerveux central la plus fréquente et la plus invalidante du sujet jeune. La prévalence est plus élevée dans les pays de l'hémisphère Nord et chez les femmes (sexe ratio : environ 2/1) [1]. Il existe différentes formes évolutives de sclérose en plaques : la forme récurrente-rémittente, la forme secondairement progressive, la forme progressive-récurrente et la forme primaire-progressive [2].

Le diagnostic de sclérose en plaque repose sur les critères de McDonald révisés de 2017, se basant sur les notions de dissémination spatiale et temporelle [3].

Avec l'émergence et le développement de nouveaux traitements permettant de ralentir la progression de la maladie, un des grands challenges dans la prise en charge de la sclérose en plaque est d'utiliser des critères les plus sensibles et spécifiques possibles pour évaluer son évolution. Ainsi, en plus des critères d'évaluation clinique (Expanded Disability Status Scale - EDSS), l'IRM tient une place prépondérante dans le suivi de la sclérose en plaques [4].

Afin d'évaluer la stabilité ou la progression de la maladie, le suivi IRM recherche une majoration de la charge lésionnelle avec notamment l'apparition de nouvelles lésions ou l'augmentation en volume des lésions existantes. L'injection de produit de contraste gadolinés permet également de mettre en évidence d'éventuels rehaussements lésionnels témoignant de l'activité de la maladie. Cela permet d'adapter le traitement de fond si nécessaire, de changer de ligne de traitement ou de la poursuivre. Le score Rio est utilisé pour évaluer la réponse au traitement de fond en prenant en compte à 1 an : les rechutes, la progression du handicap (évalué par l'EDSS) et les lésions actives en IRM [5]. Schématiquement, les recommandations OFSEP (Observatoire Français de la Sclérose en Plaques) préconisent de réaliser une IRM au début du suivi, à l'initialisation du traitement de fond, 6 mois et 1 an après, puis tous les ans [6].

Dans la pratique courante en imagerie, l'évaluation IRM de l'évolution de la maladie par le radiologue implique une comparaison visuelle du nouvel examen au précédent, en recherchant l'apparition de nouvelles lésions ou la majoration en volume de celles existantes. Cette évaluation présente plusieurs limites. D'une part, il existe une part évidente de subjectivité, dépendante essentiellement du niveau d'expertise du radiologue. D'autre part, la tâche de comparaison, souvent considérée comme fastidieuse, prend un temps non négligeable (analyse comparative coupe par coupe). Un temps qu'il convient d'optimiser, en raison de l'augmentation de la demande d'imagerie malgré une démographie radiologique globalement constante (Figure 1).

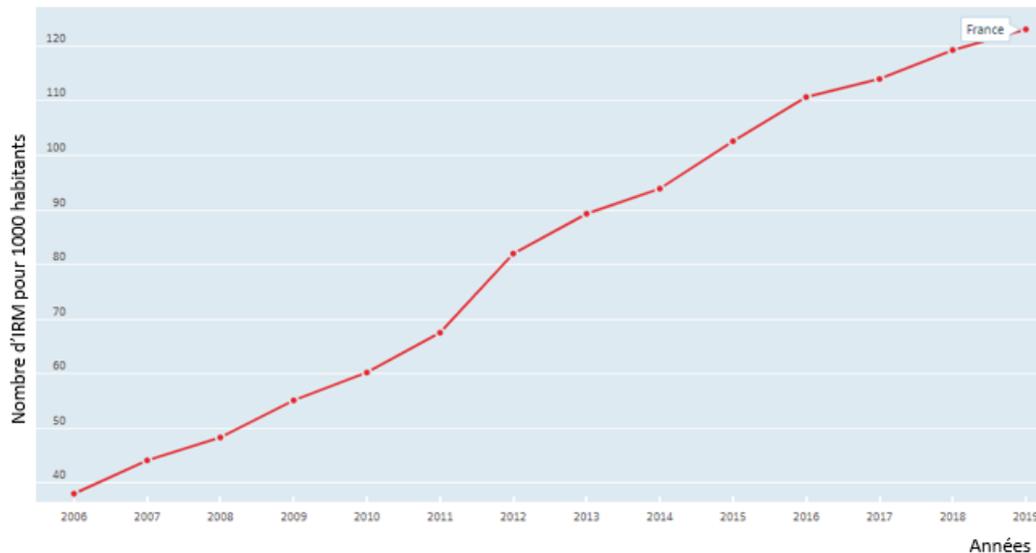


Figure 1 : Évolution du nombre d'IRM en France pour 1000 habitants de 2006 à 2019 (Données publiques de l'OCDE concernant l'IRM)

Pour diminuer les variabilités inter- et intra-observateurs et pour permettre une évaluation plus rapide et objective, en évitant des segmentations manuelles fastidieuses et difficiles, plusieurs méthodes de segmentation semi-automatiques ou entièrement automatiques ont été proposées dans la littérature [7].

Il existe trois principales catégories de segmentation, selon leur approche :

- Reconnaissance des lésions : lésions identifiées et segmentées en se basant sur les caractéristiques des signaux, le reste de l'image n'étant pas traité
- Classification hiérarchique : tous les points de l'image sont répartis en "tissus sains" et "autre", puis les points "autres" sont classés en "lésion" ou "non lésion"
- Classification directe : contrairement à l'approche précédente, la classe "lésion" est d'emblée introduite dans la première répartition.

Une quatrième catégorie consistant en la segmentation des lésions évolutives s'intéresse aux changements survenant sur une série temporelle d'images.

Le principe de segmentation automatisée utilise des réseaux de neurones convolutifs (CNN, *convolutional neural network*) [8]. Les réseaux de neurones convolutifs appartiennent à une sous-classe de la terminologie hiérarchique qui comprend l'intelligence artificielle (IA), l'apprentissage automatique (*machine learning*) et l'apprentissage profond (*deep learning*) (Figure 2).

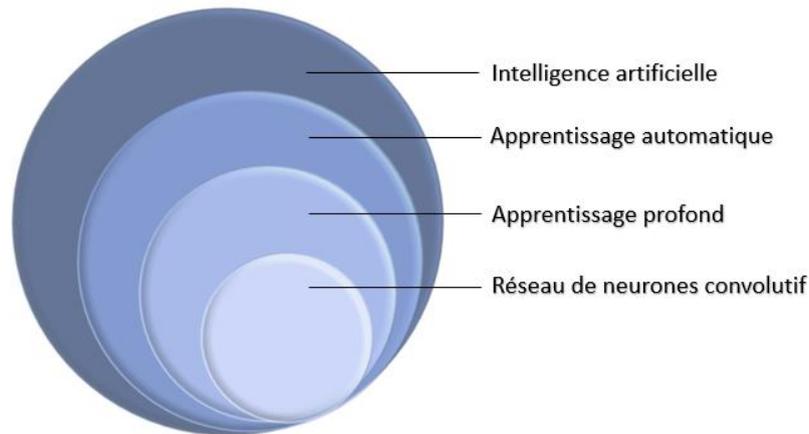


Figure 2 : Place du réseau de neurones convolutifs dans la terminologie hiérarchique de l'intelligence artificielle

L'intelligence artificielle est un terme générique qui définit des algorithmes capables de résoudre des problèmes nécessitant généralement l'intelligence humaine. L'apprentissage automatique est une sous-classe de l'intelligence artificielle consistant en la création d'algorithmes ayant la capacité d'apprendre sans être explicitement programmés. Ce qui différencie l'apprentissage profond de l'apprentissage automatique, c'est la capacité des algorithmes d'apprentissage profond d'apprendre par eux-mêmes quelles caractéristiques sont les meilleures pour la tâche de calcul attendue, sans que l'expert humain ne lui montre quelles caractéristiques choisir [9].

Par analogie avec le cortex cérébral, les réseaux de neurones artificiels sont modélisés comme des couches de neurones connectés entre eux, les sorties de certains neurones (*output*) pouvant devenir des entrées (*input*) pour d'autres (Figure 3).

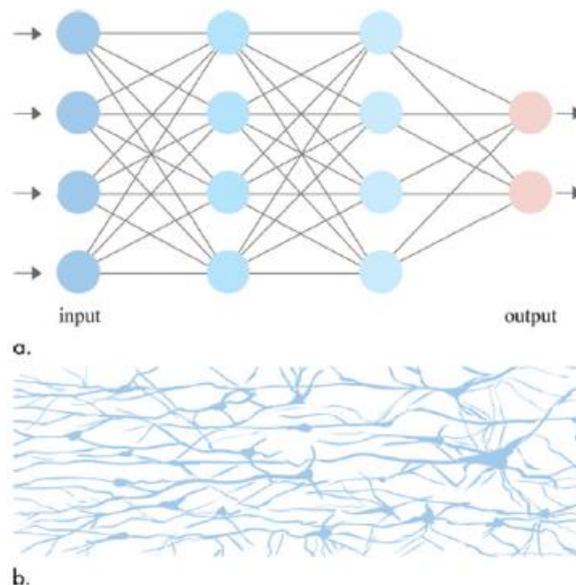


Figure 3 : (a) Représentation schématique d'un réseau neuronal artificiel par analogie avec (b) un réseau neuronal biologique [9]

Les réseaux de neurones convolutifs sont un sous-type de réseaux de neurones artificiels, inspirés du cortex visuel, avec un processus hiérarchique et non holistique, par le processus de convolution, multiplication point par point des petites matrices de poids (entrées) sur l'ensemble des images. Les réseaux de neurones convolutifs utilisent un outil appelé "*pooling*" permettant de réduire une image de grande taille tout en préservant les informations les plus importantes qu'elle contient. Les images sont ensuite comparées fragment par fragment, les fragments recherchés étant appelés les caractéristiques. Le réseau de neurones convolutifs cherche des caractéristiques approximatives qui se ressemblent dans deux images différentes, plutôt que de faire une comparaison entière image à image. Un schéma représentant le fonctionnement d'un réseau de neurones convolutif est représenté dans la Figure 4.

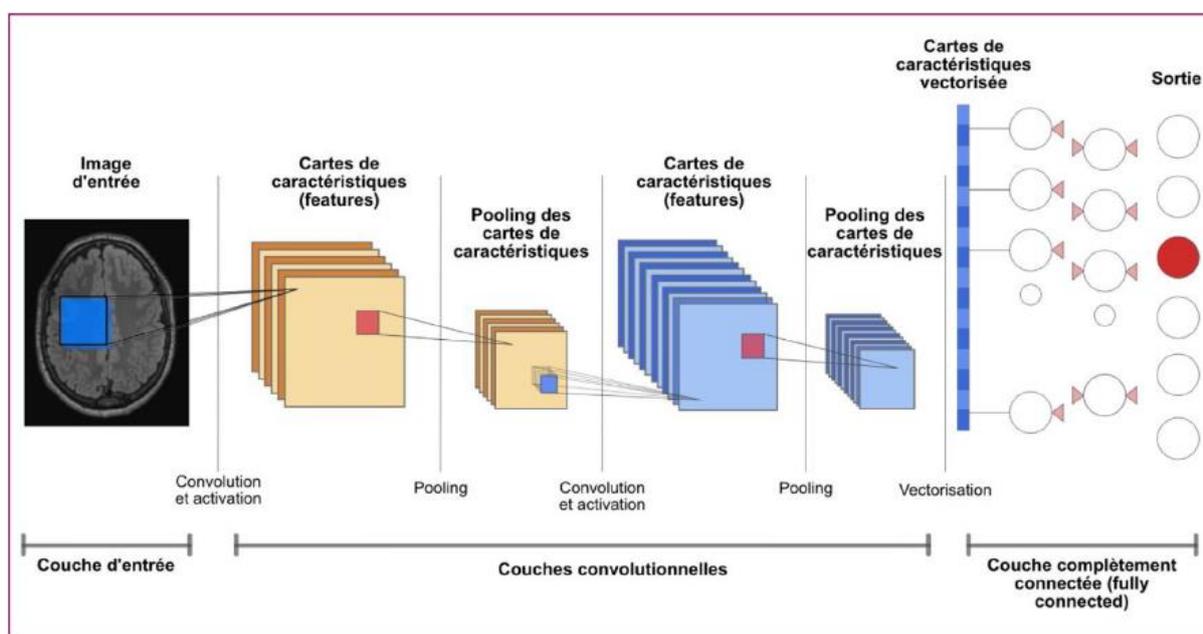


Figure 4 : Schéma du principe de fonctionnement d'un réseau de neurones convolutifs, issu de l'article de Fernandez-Maloigne et Guillevin de 2019 [10]

Nous proposons l'application d'un réseau de neurones convolutifs de type minimally parameterized U-net (MPU-net ++) permettant la segmentation automatisée des lésions de sclérose en plaque au sein de la substance blanche cérébrale. Ce réseau de neurones a été développé au sein du laboratoire I3M (XLIM - Site de Poitiers, Université de Limoges, France).

Le but de notre étude est de comparer les performances de ce réseau de neurones à la segmentation manuelle.

II. Matériels et méthodes

1- PATIENTS

Trente-neuf patients ont été inclus rétrospectivement à partir de la base de données du PACS (*Picture Archiving and Communication System*) du CHU de Poitiers. Il s'agissait de patients suivis au sein du service de neurologie de l'établissement, avec un diagnostic de sclérose en plaques établi (clinique, biologique et en imagerie). Tous les patients ont réalisé une IRM cérébrale de suivi de SEP au CHU de Poitiers. Différents niveaux de charge lésionnelle sont représentés dans la population de patients. Leurs caractéristiques sont résumées dans le tableau 1.

Les examens ont été répartis en 2 groupes distincts : 31 ont été utilisés pour la base d'entraînement et 8 pour la base de validation.

2- ACQUISITION DES IRM

Les examens utilisés dans cette étude ont tous été réalisés sur des IRM 3T (*Magnetom Verio* et *Skyra*, Siemens Healthineers, Erlangen, Allemagne), au sein du service d'Imagerie du CHU de Poitiers. Le protocole standardisé OFSEP (3D T1, axiale DWI, 3D FLAIR, +/- 3D T1 gadolinium) a été suivi pour le choix des séquences [6]. La séquence 3D FLAIR a été utilisée pour la segmentation.

3- SEGMENTATION MANUELLE : VERITE TERRAIN

Toutes les IRM ont été segmentées manuellement par un radiologue à l'aide du logiciel 3D Slicer [11]. La même procédure a été utilisée pour toutes les segmentations, en commençant par remettre l'examen dans son plan d'origine, afin de pouvoir segmenter dans le plan natif, évitant ainsi les effets de volumes partiels.

Il s'agissait de segmenter toutes les lésions de la substance blanche évocatrices de lésions de sclérose en plaques. Par exemple, les hypersignaux fins et linéaires au contact des ventricules n'ont pas été segmentés, ceux-ci étant aspécifiques et retrouvés très fréquemment chez les sujets sains.

Les segmentations ont été exportées au format Nifti (nii.gz) pour pouvoir être ainsi utilisées dans l'apprentissage du réseau de neurones.

4- TRAITEMENT DES IMAGES

Toutes les IRM ont été anonymisées et exportées au format DICOM avec post-traitement des données. Ces données ont été converties à l'aide de l'outil Dcm2niix au format NIFTI, nécessaire pour le traitement des images.

L'ensemble de ces traitements ont été effectués avec Python, un langage de programmation contenant de nombreuses bibliothèques permettant l'analyse et le traitement de l'image ainsi que des bibliothèques d'Intelligence Artificielle spécifiques à l'imagerie médicale.

Les cerveaux sont tout d'abord extraits des images par l'intermédiaire de la commande *Brain Surface Extractor* du logiciel *BrainSuite* [12]. Simultanément, les masques de ces cerveaux sont générés pour permettre de sélectionner avec précision quelle zone cérébrale à entraîner. Toutes les images, les masques et les segmentations sont rééchantillonnés avec des voxels isotropiques de 1mm de côté via la bibliothèque Pytorch [13].

Le logiciel *TorchIO medical imaging* a été utilisé afin d'augmenter artificiellement le nombre de données de la base d'apprentissage (data augmentation) [14]. Cette bibliothèque logicielle permet une approche experte sur les phénomènes qu'il est possible de trouver à l'IRM (biais de champs, déformations, translations, etc.). Comparée aux approches classiques, cela permet de prendre en compte l'ensemble des contraintes du moyen d'imagerie.

5- RESEAU DE NEURONES

Au-delà des méthodes classiques, la segmentation automatique nécessite l'utilisation de réseaux de neurones convolutifs. Le réseau de neurones U-Net est le plus connu et le plus utilisé en imagerie médicale, particulièrement bien applicable à la segmentation dans la sclérose en plaques.

Nous avons utilisé un réseau de neurones particulier dont l'architecture comportant un faible nombre de paramètres permet l'apprentissage à partir de peu de données et s'adapte donc très bien à notre étude.

Il s'agit du réseau de neurones de Fenneteau *et al.* [15], adapté à partir du U-Net de Isensee *et al.* [16]. Ce CNN d'Isensee *et al.* a obtenu la troisième place au *Medical Imaging Computing & Computer Assisted Intervention Brain Tumor Segmentation* (MICCAI BraTS). C'est une évolution du U-net : le U-net++. Comme la plupart des articles récents sur la segmentation de l'imagerie médicale, ce réseau de neurones utilise une architecture d'encodeur-décodeur (U-Net).

Le but du réseau de neurones est de reconnaître les voxels atteints par la maladie. Il a été initialement développé à partir d'IRM de sclérose en plaques déjà segmentées, issues de bases de données publiques. Celles-ci provenaient des challenges de segmentation des lésions de sclérose en plaques du MICCAI 2008, du challenge de segmentation de la sclérose en plaques 2015 du Symposium international d'imagerie biomédicale et de la base de données publiques de Lesjak [15]. Ces ensembles de données patients annotées rendus publics ont permis

d'explorer plus facilement la capacité des algorithmes d'apprentissage automatique tels que les réseaux de neurones convolutifs à traiter les données d'images médicales.

Il s'agit d'un réseau de neurones de type MPU-net++ (minimally parameterized U-net++), utilisant un faible nombre de paramètres d'apprentissage (37935 paramètres). Son architecture est présentée en Figure 5. L'architecture particulière du U-net est composée d'une voie descendante, dite d'encodage qui permet de contracter l'information au fur et à mesure de la descente (flèches bleues sur la Figure 5). Une fois au fond de l'architecture en U, la partie remontante, aussi appelée voie de décodage, va déconvoluer des informations préalablement contractées pour aboutir à une taille d'image similaire à celle d'entrée. Des informations issues de la phase de contraction sont ajoutées (*Copy-Concatenate*). Le modèle U-net ++ possède lui plusieurs voies de de décodage (flèches jaunes sur la Figure 5) qui permettent d'avoir une segmentation raffinée grâce à ces multiples connexions supplémentaires.

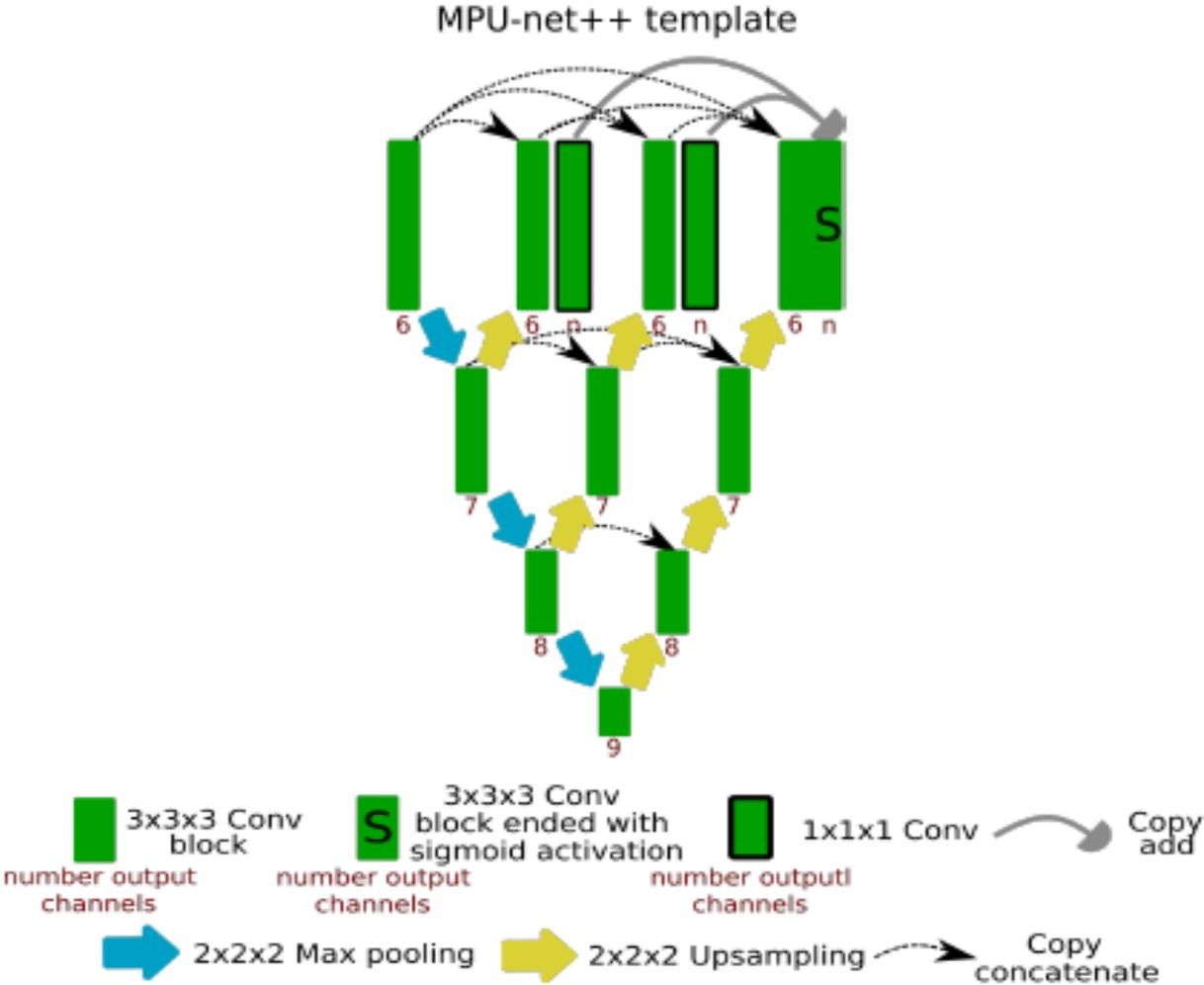


Figure 5 : architecture du réseau de neurones MPU-net ++

6- ANALYSE STATISTIQUE ET METRIQUES

Analyse statistique

Concernant l'analyse principale, nous avons comparé les segmentations manuelle et automatique, en utilisant principalement le Dice score ainsi que les autres métriques d'évaluation précédemment citées.

En analyse secondaire, nous avons étudié l'évolution du Dice score en fonction du volume, en utilisant le test de corrélation non paramétrique de Kendall.

Métriques

Pour la comparaison des segmentations, nous avons utilisé le coefficient de Dice ou Dice score [17]. Ce dernier permet d'évaluer la similarité de deux échantillons. Le Dice score a également été utilisé comme fonction de coût (Dice loss) au sein du CNN.

Le calcul du Dice score est un compromis entre la sensibilité (Se) et la valeur prédictive positive (VPP) (Figure 6).

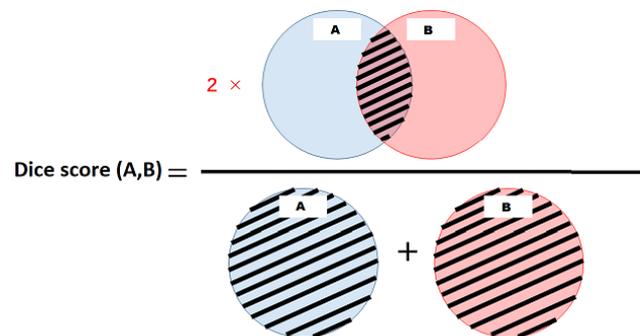
$$\begin{aligned} \text{Sensibilité} &= \frac{VP}{VP + FN} \\ \text{Valeur Prédictive Positive} &= \frac{VP}{VP + FP} \\ \text{Dice} &= \frac{2VP}{2VP + FP + FN} \\ \text{Dice} &= \frac{VPP - \text{Sensibilité}}{VPP + \text{Sensibilité}} \end{aligned}$$


Figure 6 : Schéma représentant la formule du Dice score

Nous avons également calculé les valeurs de sensibilité, spécificité (Sp) et de valeurs prédictives positive (VPP) et négative (VPN) à l'échelle du voxel dans la population de validation de 8 patients.

Concernant les autres métriques que nous avons calculées, les plus importantes sont :

- Mesure de surface :
 - Distance de Hausdorff
- Mesure des contours
 - Distance des contours
- Mesure de détection lésionnelle
 - Valeur Prédictive Positive Lésionnelle (*Lesion positive predictive value - PPVL*)
 - Sensitivité Lésionnelle (SensL)
- Mesure de la précision et du rappel
 - F1 Score

L'ensemble des métriques calculées sont issues d'un *challenge* MICCAI 2016 sur la SEP. Ce sont ces métriques qui ont permis de classer les meilleurs algorithmes [18].

III. Résultats

1- CARACTERISTIQUES DES PATIENTS

Trente-neuf patients ont été inclus rétrospectivement, dont 31 (79,5%) dans la population d'entraînement et 8 (20,5%) dans la population de validation. Les caractéristiques des patients inclus sont résumées dans le Tableau 1.

Tableau 1 : Caractéristiques de la population totale

Nombre de patients	39
Sexe	28F, 11M
Age, moyenne (limites)	53 (20-79)
Poids (kg), moyenne (limites)	65,5 (41-94)
Durée d'évolution de la maladie (années)	20,3 (3-37)
EDSS, moyenne (écart-type)	5,2 (2,08)

2- SEGMENTATION

Lors de la segmentation de chaque IRM par le réseau de neurones, plusieurs seuils sont testés ("*thresholds*"), allant de 0.1 à 0.9. Ceux-ci correspondent schématiquement à des niveaux de certitude. Plus le seuil est bas, plus le réseau de neurones segmentera de manière "large". A l'inverse, plus le seuil augmente, plus la segmentation sera restrictive et précise. Il en résulte donc une baisse de la sensibilité et une hausse de la spécificité lorsque le seuil augmente. Un exemple est donné en Figure 7.

Nous avons choisi d'utiliser le seuil 0.9 pour l'analyse des données qui semble être le meilleur compromis. Ce seuil permet en effet d'obtenir le meilleur Dice score avec un compromis idéal entre sensibilité et spécificité. L'objectif principal de cette étude étant de comparer les segmentations manuelles et automatiques, le choix du Dice score comme métrique principale semble être le plus approprié.

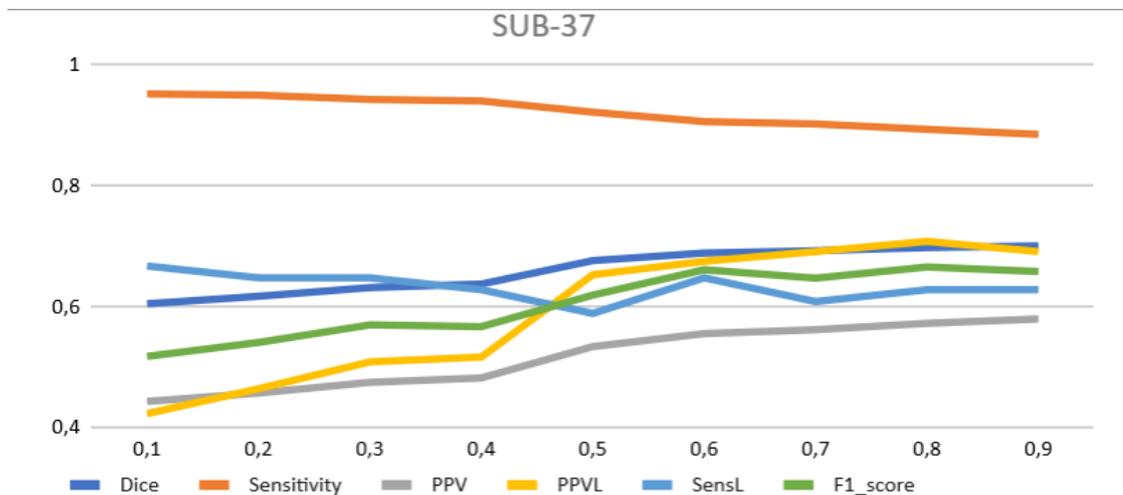


Figure 7 : Variation des valeurs normalisées de métriques en fonction du seuil pour un cas.

COMPARAISON DES SEGMENTATIONS AUTOMATIQUE ET MANUELLE

Les résultats des différentes métriques d'évaluation sont présentés dans le Tableau 2. La moyenne des Dice scores obtenus lors des comparaisons entre la segmentation manuelle et la segmentation automatique est de 0.64.

Tableau 2 : Moyennes des métriques d'évaluation dans la population de validation (écart-type)

Métrique	Moyenne (écart-type)
Dice score	0.637 (0.116)
Sensibilité	0.716 (0.151)
Spécificité	0.999 (0)
VPP	0.579 (0.113)
VPN	0.999 (0)
Volume d'erreur relatif	24.126 (20.545)
Distance Hausdorff	42.536 (13.660)
Distance moyenne contour	2.999 (1.986)
Distance Surface	0.439 (0.537)
PPVL	0.650 (0.129)
SensL	0.613 (0.083)
Score F1	0.626 (0.090)

EXEMPLES DE SEGMENTATIONS

Un exemple de segmentation satisfaisante est présenté dans la Figure 8, où nous pouvons voir une bonne concordance des segmentations avec quelques différences sur les pixels en bords de plaques.

Toutefois, quelques discordances de segmentation ont été notées sur certains cas. Dans l'exemple montré dans la Figure 9, le radiologue a sélectionné les plaques les mieux définies et les mieux circonscrites en franc hypersignal FLAIR alors que le réseau de neurones a également pris en compte dans sa segmentation les plages périlésionnelles en faible hypersignal FLAIR à contours flous mal limités.

Nous avons également obtenu un Dice score moins élevé dans l'exemple montré en Figure 10 qui comportait peu de plaques et de petit volume. Les segmentations se recoupent, mais le Dice score est fortement pénalisé par les voxels qui ne sont pas communs en bordure de plaque.

Dice score : 0.67
Radiologue = 4.138 cm³
IA = 4.291 cm³

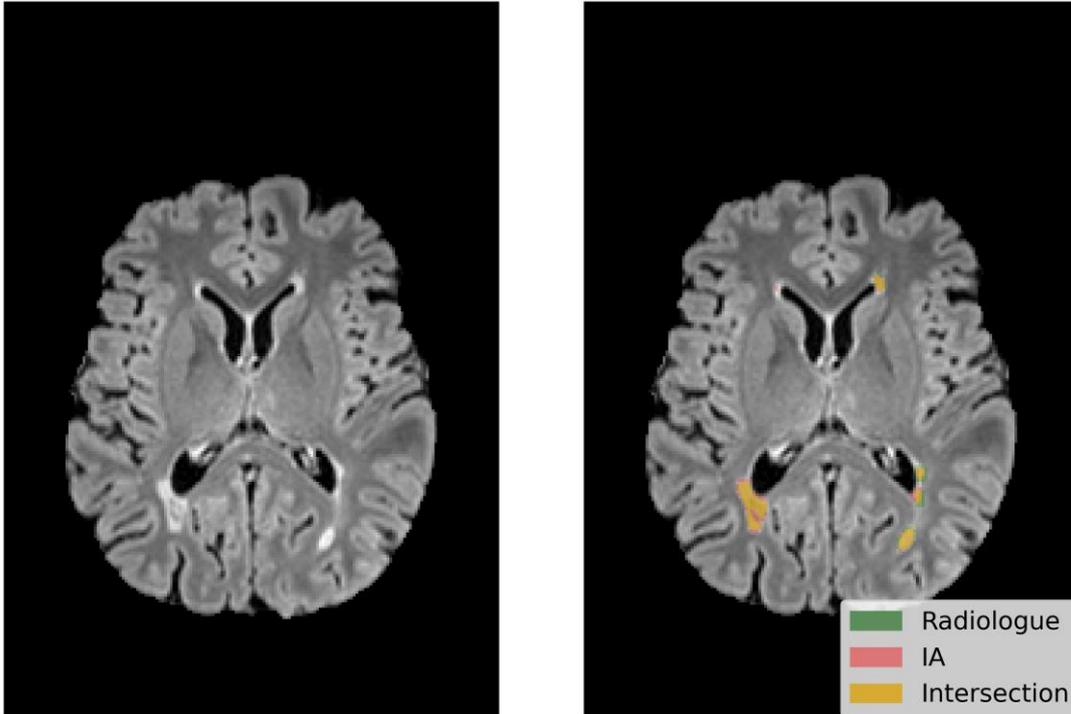


Figure 8 : Exemple de bonne similarité des segmentations.

Dice score : 0.65
Radiologue = 17.219 cm³
IA = 30.128 cm³

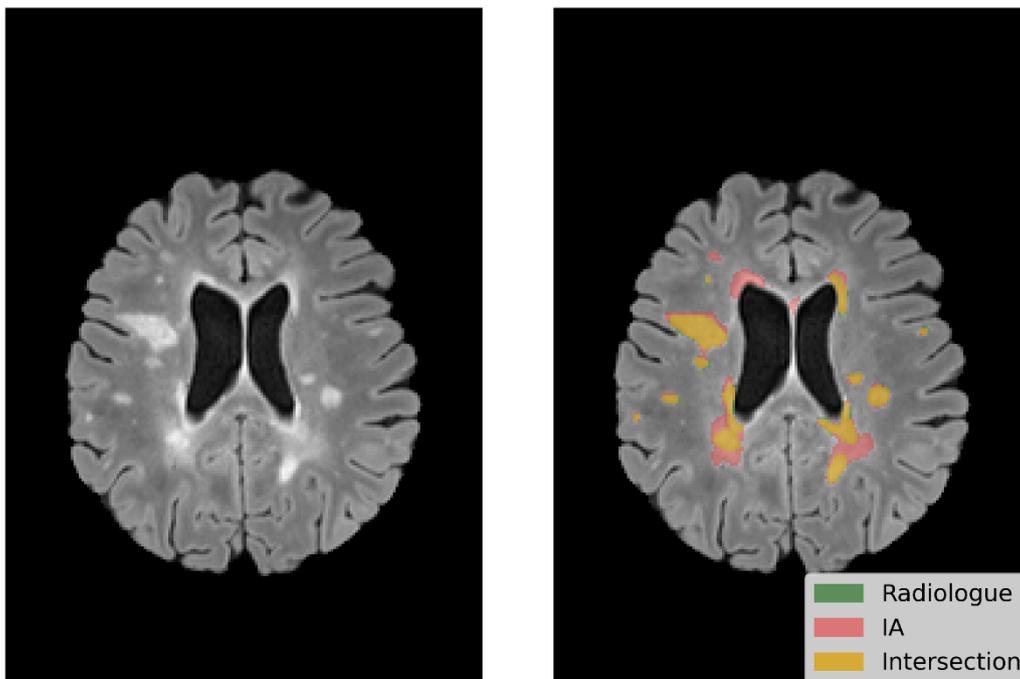


Figure 9 : Exemple de discordance : segmentation automatique plus large comparativement à la segmentation manuelle.

Dice score : 0.24
Radiologue = 0.507 cm³
IA = 0.567 cm³

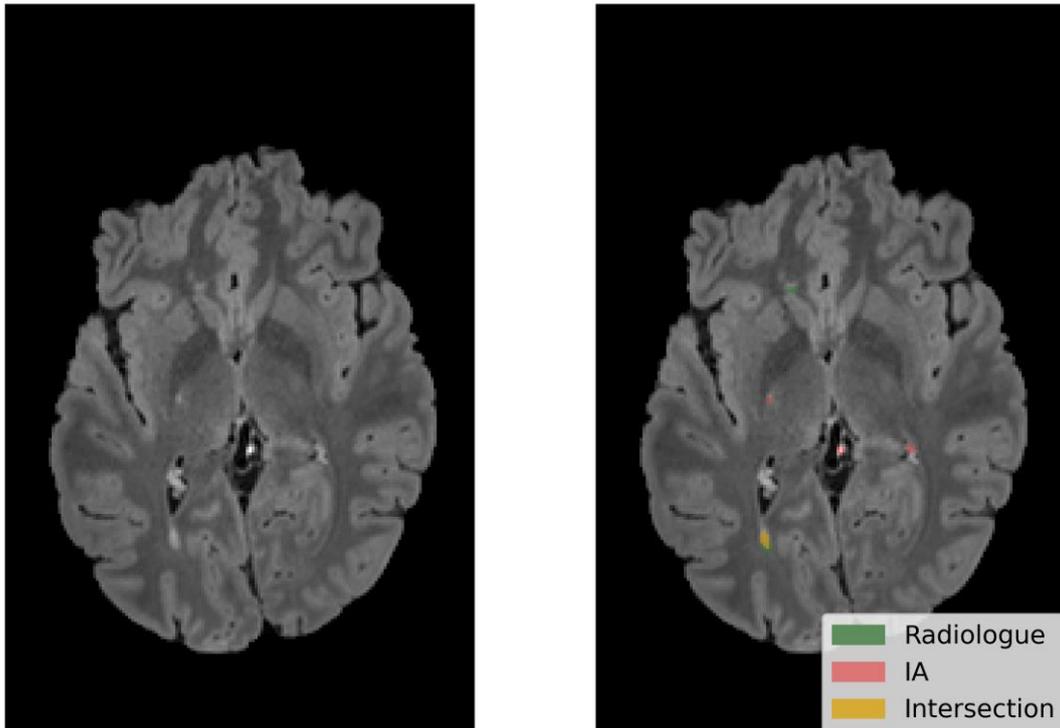


Figure 10 : Exemple d'un cas avec une charge lésionnelle très faible favorisant un Dice score bas.

VARIATION DU DICE SCORE EN FONCTION DE LA CHARGE LESIONNELLE

La Figure 11 représente l'évolution du Dice score en fonction du volume dans la population de validation.

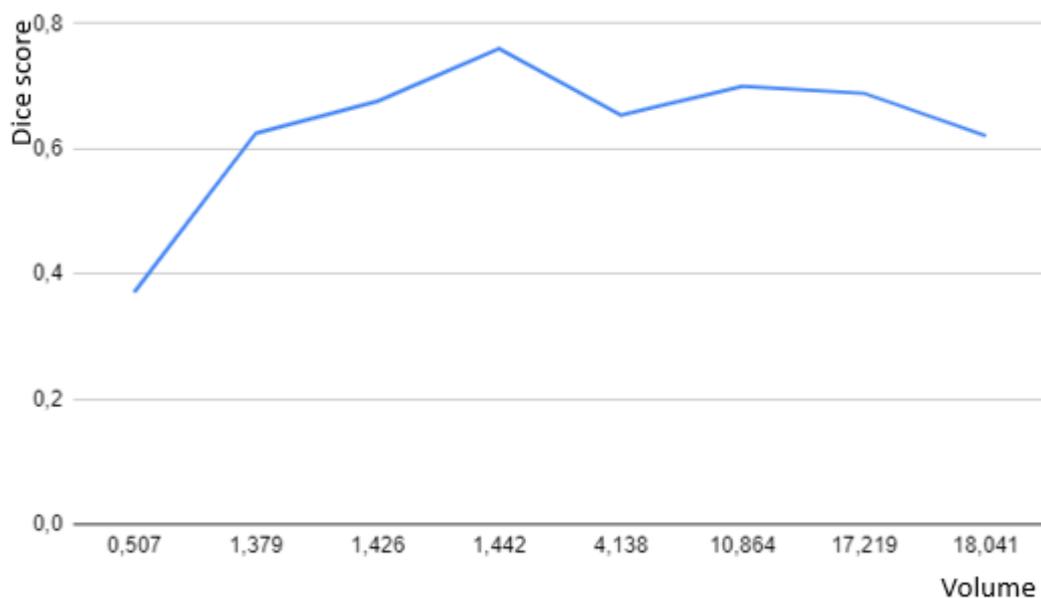


Figure 11 : Dice score en fonction du volume de charge lésionnelle.

Le test de corrélation de Kendall a été utilisé et montre un résultat non significatif, avec une estimation du coefficient de corrélation tau $\tau = 0,2143$ ($p = 0,5484$).

On remarque que le Dice score varie en fonction du volume (allant de 0.507 à 18.041 cm³), sans adopter une tendance claire dans notre population, avec néanmoins un Dice score particulièrement bas pour les volumes les plus faibles.

IV. Discussion

Nos résultats démontrent que le réseau de neurones étudié permet d'obtenir une segmentation comparable à celle réalisée par le radiologue, avec des performances très satisfaisantes comparativement aux publications récentes, sous réserve de populations différentes et d'une méthodologie différente limitant la comparaison. Dans l'étude de Valverde *et al.* [19], le Dice score obtenu était de 53,5%, de 56% et 57% dans l'étude de Valcarel *et al.* [20] et de 56,39% dans celle de Roy *et al.* [21].

Dans les exemples illustrés dans la partie résultats, nous avons montré que certaines segmentations étaient particulièrement satisfaisantes, avec une très bonne concordance des segmentations manuelle et automatique.

Quelques discordances ont toutefois été notées, notamment pour le cas montré en Figure 9, avec une segmentation automatique plus large que la segmentation manuelle, la première comprenant les plages mal limitées en faible hypersignal FLAIR entourant des plaques plus définies, auxquelles s'est limitée la segmentation manuelle. La définition des contours de la plaque étant relativement subjective dans ce cas, les deux options peuvent être considérées comme justes, mais cela impacte le Dice score et les autres métriques d'évaluation.

D'autre part, quelques erreurs ont été notées lors de la segmentation automatique. Le principal type est la segmentation inconstante du fin liseré en hypersignal FLAIR au contact des ventricules latéraux considéré comme non pathologique du fait de sa présence très fréquente chez les sujets sains (Figure 12). En effet, ces hypersignaux correspondent à une raréfaction de la substance blanche induite par une filtration accrue de LCS en périventriculaire : ils sont présents de façon physiologique à tout âge et ont tendance à s'accroître avec le vieillissement.

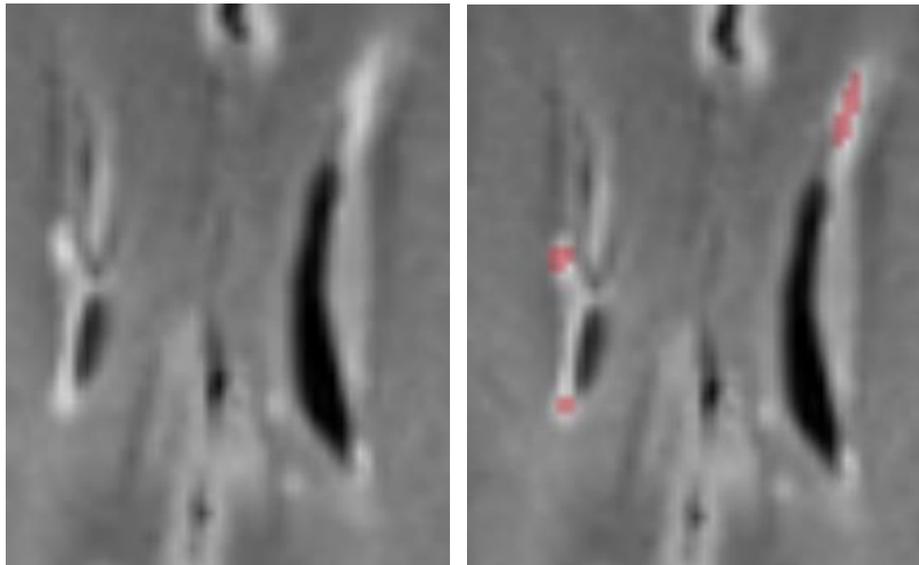


Figure 12 : Exemple de segmentation automatique de fins hypersignaux au contact des ventricules latéraux.

Le Dice score est la métrique d'évaluation de la similarité la plus utilisée en recherche dans le domaine de la segmentation automatisée par réseau de neurones. Néanmoins, celui-ci n'est pas parfait. En effet, il est nettement impacté lorsque la charge lésionnelle est faible, comme l'illustre l'exemple montré en Figure 10. Tel que démontré dans la littérature, il est habituel de rencontrer lors d'épreuves de segmentation un Dice score plus faible pour un volume lésionnel faible [22]. Nous n'avons pas retrouvé cette corrélation dans notre étude de façon significative, compte tenu d'un faible nombre de patients inclus, cependant nous observons une tendance à une corrélation positive entre Dice score et volume lésionnel, en particulier sur les faibles volumes (Figure 11).

Le faible nombre d'IRM segmentées et la segmentation réalisée par un unique radiologue sont les principales limites de notre étude. La faible quantité d'échantillons et la difficulté à réaliser un grand nombre d'annotations d'experts radiologues du fait du temps que la segmentation manuelle requiert, sont à ce jour les principales difficultés limitant l'entraînement et donc l'utilisation pratique des réseaux de neurones en imagerie médicale. La création de bases de données publiques de plus en plus fournies permet de pallier en partie à ce manque. Le développement de réseaux de neurones capables d'apprendre à partir de peu de données, comme c'est le cas pour celui que nous étudions, semble être une solution adaptée à cette problématique.

En effet, l'avantage de ce réseau de neurones est son architecture comportant un nombre de paramètres d'apprentissage limité. Ceci permet l'apprentissage à partir de peu de données, en diminuant le risque de sur-apprentissage (apprentissage "par cœur"). Il est admis que plus le nombre de paramètres est grand, plus le risque d'avoir un réseau de neurones capable d'apprendre par cœur

au lieu de reconnaître réellement les lésions est important. Dans une étude récente de Fenneteau *et al.*, il a été montré que, pour la tâche étudiée (segmentation de lésions de SEP), cette architecture simplifiée utilisant seulement 37 935 paramètres d'apprentissage apporte des résultats équivalents aux réseaux de neurones convolutifs les plus utilisés dont l'architecture complexe comporte plusieurs millions de paramètres [21]. Ceci constitue une réelle avancée dans la compréhension des réseaux de neurones et de leur architecture et permet d'envisager, par exemple, une application à des pathologies moins fréquentes. D'autre part, cela pourrait permettre une accessibilité plus aisée en pratique clinique, supportée par des serveurs plus légers, pouvant permettre une mise à disposition plus large du réseau de neurones, afin de mettre l'expertise des centres de références au profit des territoires.

En pratique clinique, le réseau de neurones intégré dans un logiciel d'aide au diagnostic pourrait être utilisé à chaque IRM d'un même patient, et générer une segmentation pour chaque examen du suivi afin de les comparer en étudiant les discordances ou *mismatch*, à la recherche de nouvelles lésions ou d'augmentation en taille des lésions existantes. En effet, si les paramètres d'apprentissage du réseau de neurones ne changent pas, la segmentation automatisée a l'avantage d'être parfaitement reproductible. Cela pourrait limiter les variabilités inter et intra-observateur et permettre une comparaison objective et reproductible des examens.

Il existe déjà sur le marché des logiciels d'aide au diagnostic utilisant une segmentation automatisée d'IRM cérébrales et en particulier de lésions de sclérose en plaques [23]. Mais l'avantage et la nouveauté qu'apporte le réseau de neurones de Fenneteau *et al.* est sa capacité à apprendre à partir de peu de données.

D'autre part, on peut se poser la question de la pertinence de l'utilisation de la charge lésionnelle comme variable d'évaluation du suivi. En effet, la charge lésionnelle n'est pas toujours corrélée à la sévérité clinique de la maladie. La grande majorité des études souligne l'absence de corrélation entre le score EDSS et les paramètres radiologiques mesurés au même instant [24]. En revanche, l'évolution radiologique précoce est prédictive du degré de handicap physique à long terme. Sur une cohorte de 140 patients ayant présenté un premier événement démyélinisant et suivis sur une période de 20 ans, l'évolution de la charge lésionnelle sur les séquences pondérées T2 au cours des 5 premières années était prédictive du passage en forme secondairement progressive à 20 ans [25].

La charge lésionnelle reste la variable de référence pour le suivi de la sclérose en plaque. L'adaptation du traitement de fond se fait selon les critères de NEDA (*No Evidence of Disease Activity*) et le score Rio [5]. Le NEDA-3 est basé sur des paramètres cliniques et radiologiques, prenant en compte la survenue ou non de poussée, l'aggravation ou non du handicap par le score EDSS et l'apparition ou l'augmentation en taille de lésions IRM en hypersignal T2 (ce qui équivaut à la charge lésionnelle), ainsi que le rehaussement de lésions après injection de produit

de contraste, témoignant de lésions actives. Un nouveau concept de NEDA-4 est actuellement à l'étude, ajoutant aux précédents critères du NEDA-3 l'évaluation de l'atrophie cérébrale (la perte de volume cérébral doit être inférieure à 0,4%/an), pouvant également s'effectuer à l'aide de l'intelligence artificielle.

Le réseau de neurone testé permet donc d'évaluer la charge lésionnelle de manière objective et reproductible. Mais, en se basant uniquement sur la séquence 3D FLAIR, il ne détecte pas les prises de contraste des lésions. Néanmoins, cette tâche étant peu sujette à variabilité inter et intra-observateur, elle semble moins nécessiter l'assistance par intelligence artificielle. De plus, les nouvelles recommandations de l'OFSEP ont significativement réduit les indications d'utilisation de produit de contraste gadolinés dans le cadre du suivi de sclérose en plaques (injection requise à l'initiation du traitement, 6 mois après et lors d'une nouvelle poussée) [6].

Conclusion

Le réseau de neurones étudié semble avoir des performances de segmentation très satisfaisantes, équivalentes aux réseaux de neurones habituels, avec l'avantage de nécessiter beaucoup moins de données d'apprentissage.

Nous avons montré une bonne concordance entre les mesures automatisées de l'IA et les segmentations manuelles, qui rendent compte de la charge lésionnelle. La mesure automatisée à l'aide de ce réseau de neurones pourrait donc être envisagée en pratique clinique courante dans l'évaluation du suivi en imagerie des scléroses en plaque avec pour avantages un gain de temps et également une estimation objective et reproductible de la charge lésionnelle.

Ces résultats préliminaires sont prometteurs, néanmoins le faible nombre de patients inclus est la principale limite de notre étude.

Nous envisageons la poursuite de cette étude d'une part en incluant davantage de patients et d'autre part en comparant des segmentations manuelles faites par plusieurs radiologues afin de confronter les résultats obtenus à l'aide de l'intelligence artificielle à la variabilité inter-observateurs humaine.

Références

1. Pugliatti M, Rosati G, Carton H, Riise T, Drulovic J, Vecsei L, et al. The epidemiology of multiple sclerosis in Europe. *Eur J Neurol*. juill 2006;13(7):700-22.
2. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. *Curr Opin Neurol*. déc 2018;31(6):752-9.
3. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. févr 2018;17(2):162-73.
4. Lublin FD, Reingold SC, National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis*. Defining the clinical course of multiple sclerosis: Results of an international survey. *Neurology*. 1 avr 1996;46(4):907-11.
5. Sormani M, Rio J, Tintorè M, Signori A, Li D, Cornelisse P, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler J*. avr 2013;19(5):605-12.
6. Brisset J-C, Kremer S, Hannoun S, Bonneville F, Durand-Dubief F, Tourdias T, et al. New OFSEP recommendations for MRI assessment of multiple sclerosis patients: Special consideration for gadolinium deposition and frequent acquisitions. *J Neuroradiol*. juin 2020;47(4):250-8.
7. Souplet J-C, Lebrun C, Chanalet S, Ayache N, Malandain G. Revue des approches de segmentation des lésions de sclérose en plaques dans les séquences conventionnelles IRM. *Rev Neurol (Paris)*. janv 2009;165(1):7-14.
8. Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJNL, Isgum I. Automatic Segmentation of MR Brain Images With a Convolutional Neural Network. *IEEE Trans Med Imaging*. mai 2016;35(5):1252-61.
9. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. *Radiology*. mars 2019;290(3):590-606.
10. Fernandez-Maloigne C, Guillevin R. L'intelligence artificielle au service de l'imagerie et de la santé des femmes. *Imag Femme*. déc 2019;29(4):179-86.
11. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. nov 2012;30(9):1323-41.
12. Shattuck DW, Leahy RM. BrainSuite: An automated cortical surface identification tool. *Med Image Anal*. 2002;14.
13. Chen KM, Cofer EM, Zhou J, Troyanskaya OG. Selene: a PyTorch-based deep learning library for sequence data. *Nat Methods*. avr 2019;16(4):315-8.
14. Pérez-García F, Sparks R, Ourselin S. TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput Methods Programs Biomed*. sept 2021;208:106236.
15. Fenneteau A, Bourdon P, Helbert D, Fernandez-Maloigne C, Habas C, Guillevin R. Apprentissage d'un CNN sur la segmentation des lésions de sclérose en plaques avec auto-supervision. :8.
16. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH. Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. In: Crimi A, Bakas S, Kuijff H, Menze B, Reyes M, éditeurs. *Brainlesion: Glioma, Multiple Sclerosis, Stroke*

and Traumatic Brain Injuries [Internet]. Cham: Springer International Publishing; 2018 [cité 5 juill 2021]. p. 287-97. (Lecture Notes in Computer Science; vol. 10670). Disponible sur: http://link.springer.com/10.1007/978-3-319-75238-9_25

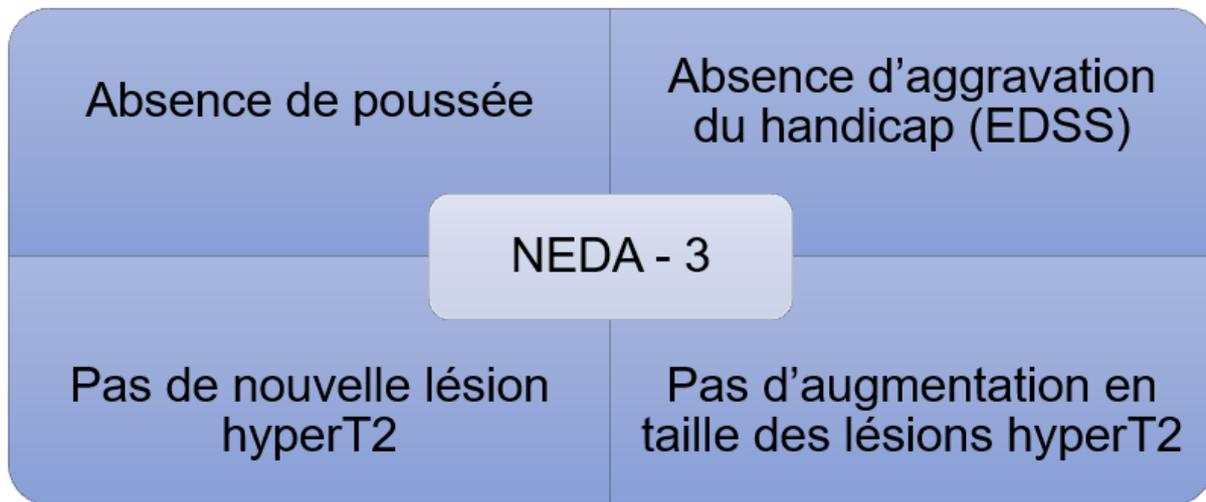
17. Eelbode T, Bertels J, Berman M, Vandermeulen D, Maes F, Bisschops R, et al. Optimization for Medical Image Segmentation: Theory and Practice When Evaluating With Dice Score or Jaccard Index. *IEEE Trans Med Imaging*. nov 2020;39(11):3679-90.
18. Commowick O, Istace A, Kain M, Laurent B, Leray F, Simon M, et al. Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure. *Sci Rep*. déc 2018;8(1):13650.
19. Valverde S, Cabezas M, Roura E, González-Villà S, Pareto D, Vilanova JC, et al. Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. *NeuroImage*. juill 2017;155:159-68.
20. Valcarcel AM, Linn KA, Vandekar SN, Satterthwaite TD, Muschelli J, Calabresi PA, et al. MIMoSA: An Automated Method for Intermodal Segmentation Analysis of Multiple Sclerosis Brain Lesions: Method For Inter-Modal Segmentation Analysis. *J Neuroimaging*. juill 2018;28(4):389-98.
21. Roy S, Butman JA, Reich DS, Calabresi PA, Pham DL. Multiple Sclerosis Lesion Segmentation from Brain MRI via Fully Convolutional Neural Networks. *ArXiv180309172 Cs* [Internet]. 24 mars 2018 [cité 28 sept 2021]; Disponible sur: <http://arxiv.org/abs/1803.09172>
22. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. déc 2015;15(1):29.
23. Schmidt P, Gaser C, Arsic M, Buck D, Förchler A, Berthele A, et al. An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *NeuroImage*. févr 2012;59(4):3774-83.
24. Bar-Zohar D, Agosta F, Goldstaub D, Filippi M. Magnetic resonance imaging metrics and their correlation with clinical outcomes in multiple sclerosis: a review of the literature and future perspectives. *Mult Scler J*. juill 2008;14(6):719-27.
25. Fisniku LK, Brex PA, Altmann DR, Miszkiel KA, Benton CE, Lanyon R, et al. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain*. 7 févr 2008;131(3):808-17.

ANNEXES

Annexe A : L'échelle EDSS (Expanded Disability Status Scale)

SCORE	CRITERES
0.0	Examen neurologique normal (tous systèmes fonctionnels (SF) à 0;; SF 1 mental acceptable).
1.0	Absence de handicap fonctionnel, signes minimes d'atteinte d'une des fonctions (SF 1, à l'exclusion du SF mental).
1.5	Absence de handicap fonctionnel, signes minimes dans plus d'un SF (plus d'un SF 1, à l'exclusion du SF mental).
2.0	Handicap minime d'un des SF (1 SF 2, les autres 0 ou 1).
2.5	Handicap minime dans 2 SF (2 SF 2, les autres 0 ou 1).
3.0	Handicap modéré dans un SF (1 SF score 3, les autres 0 ou 1) ;; ou atteinte minime de 3 ou 4 fonctions (3 ou 4 SF 2 ;; les autres 0 ou 1), mais malade totalement ambulatoire.
3.5	Totalement ambulatoire, mais atteinte modérée dans un SF (SF 3) et 1 ou 2 SF 2;; ou 2 SF 3 ;; ou 5 SF 2 (les autres 0 ou 1).
4.0	Malade totalement autonome pour la marche, vaquant à ses occupations 12h par jour malgré une gêne fonctionnelle relativement importante : 1 SF à 4 (les autres 0 ou 1), ou association de niveaux inférieurs dépassant les limites des degrés précédents. Capable de marcher 500 m environ sans aide ni repos.
4.5	Malade autonome pour la marche, vaquant à ses occupations la majeure partie de la journée, capable de travailler une journée entière, mais pouvant parfois être limité dans ses activités ou avoir besoin d'une aide minime, handicap relativement sévère : un SF 4 (les autres 0 ou 1), ou association de niveaux inférieurs dépassant les limites des degrés précédents. Capable de marcher 300m environ sans aide ni repos.
5.0	Capable de marcher environ 200 m sans aide ni repos, handicap suffisamment sévère pour entraver l'activité d'une journée normale. (En général un SF 5, les autres 0 ou 1, ou association de niveaux plus faibles dépassant ceux du grade 4.0).
5.5	Capable de marcher environ 100 m sans aide ni repos ;; handicap suffisamment sévère pour empêcher l'activité d'une journée normale. (En général un SF 5, les autres 0 ou 1, ou association de niveaux plus faibles dépassant ceux du grade 4.0).
6.0	Aide unilatérale (cane, canne anglaise, béquille) constante ou intermittente nécessaire pour parcourir environ 100 m avec ou sans repos intermédiaire. (En général association de SF comprenant plus de 2 SF 3+).
6.5	Aide permanente bilatérale (cannes, cannes anglaises, béquilles) nécessaire pour marcher 20 m sans s'arrêter. (En général association de SF comprenant plus de 2 SF 3+).
7.0	Incapable de marcher plus de 5 m même avec aide ;; essentiellement confiné au fauteuil roulant;; fait avancer lui-même son fauteuil et effectue le transfert;; est au fauteuil roulant au moins 12 h par jour. (En général association de SF comprenant plus d'un SF 4+;; très rarement, SF 5 pyramidal seulement).
7.5	Incapable de faire plus de quelques pas;; strictement confiné au fauteuil roulant;; a parfois besoin d'une aide pour le transfert;; peut faire avancer lui-même son fauteuil mais ne peut y rester toute la journée;; peut avoir besoin d'un fauteuil électrique. (En général association de SF comprenant plus d'un SF 4+).
8.0	Essentiellement confiné au lit ou au fauteuil, ou promené en fauteuil par une autre personne;; peut rester hors du lit la majeure partie de la journée;; conserve la plupart des fonctions élémentaires;; conserve en général l'usage effectif des bras. (En général SF 4+ dans plusieurs systèmes).
8.5	Confiné au lit la majeure partie de la journée ;; garde un usage partiel des bras ;; conserve quelques fonctions élémentaires. (En général SF 4+ dans plusieurs systèmes).
9.0	Patient grabataire ;; peut communiquer et manger. (En général SF 4+ dans plusieurs systèmes).
9.5	Patient totalement impotent, ne peut plus manger ou avaler, ni communiquer. (En général SF 4+ dans presque tous les systèmes).
10.0	Décès lié à la SEP.

Annexe B : Critères NEDA-3 (No Evidence of Disease Activity) [5]



Annexe C : Score Rio [5]

Critère IRM (0 ou 1)	= 0 si aucune nouvelle lésion T2 = 1 si le patient a > 5 nouvelles lésions T2
Critère clinique (0 à 2)	=0 si la patient n'a fait aucune poussée = 1 si le patient a fait 1 poussée = 2 si le patient a fait ≥ 2 poussées
Score modifié de Rio (0 à 3)	= somme des critères IRM et clinique

ANNEXE D : Protocole OFSEP [6]

 Protocole IRM OFSEP 2020		
	IRM cérébrale	IRM médullaire
Recommandées	Protocole OFSEP standard 3D T1 millimétrique DWI axiale + dADC 3D FLAIR millimétrique	Protocole OFSEP standard T2 sagittale
	Protocole OFSEP réduit (suivi trimestriel LEMP) DWI axiale + dADC 3D FLAIR millimétrique	
Optionnel	3D T1 gadolinium * millimétrique 2D TSE DP/T2 axiale ou 3D T2 DTI ≥ 15 directions 2D T2 EG (pour un 1er diagnostic) SWI 3D DIR	T1 sagittale gadolinium * T2 EG axiale T1 axiale gadolinium * STIR sagittale
* Quand injecter du gadolinium (0,1 mmol/kg) ? L'OFSEP recommande l'utilisation de gadolinium macrocyclique.		
<ul style="list-style-type: none"> - Au début du suivi (diagnostic, IRM antérieures non disponibles) - À l'initialisation d'un traitement de fond - Six mois après une initialisation de traitement de fond - [Recommandé] En cas de poussée - [Optionnel] Pour rechercher une activité en vue d'une indication thérapeutique 		

RESUME ET MOTS CLES

RESUME :

Introduction : Selon les recommandations de l'Observatoire Français de Sclérose En Plaques (OFSEP), les patients atteints de sclérose en plaques (SEP) nécessitent un suivi régulier par IRM. Afin de diminuer les variabilités d'analyse radiologique, intra- et inter-observateurs, et pour permettre une évaluation en imagerie plus rapide et objective de la charge lésionnelle de SEP en pratique clinique courante, nous proposons l'utilisation d'un réseau de neurones convolutifs de type MPU-net++. Dans le but d'évaluer les performances de ce réseau de neurones, la segmentation automatique a été comparée à la segmentation manuelle par un radiologue.

Matériels et méthodes : Trente-neuf patients ont été inclus rétrospectivement, dont 31 dans la population d'entraînement et 8 dans la population de validation. Nous avons comparé les segmentations automatique et manuelle dans la population de validation, en utilisant le Dice score et d'autres métriques d'évaluation. Nous avons également analysé l'évolution du Dice score en fonction du volume de la charge lésionnelle.

Résultats : La moyenne des Dice scores obtenus lors des comparaisons entre la segmentation manuelle et la segmentation automatique est de 0,64. Certaines segmentations sont très satisfaisantes, tandis que quelques discordances sont relevées. Par ailleurs, nous n'avons pas montré de corrélation significative entre Dice score et volume de la charge lésionnelle (coefficient de corrélation tau = 0,2143 ($p = 0,5484$)).

Conclusion : Le réseau de neurones étudié semble avoir des performances de segmentation très satisfaisantes, équivalentes aux réseaux de neurones habituels. Il présente l'avantage de nécessiter beaucoup moins de données d'apprentissage, permettant d'envisager une utilisation en pratique clinique courante. Par ailleurs, il permet un gain de temps d'analyse et une estimation objective et reproductible de la charge lésionnelle. Ces résultats préliminaires sont prometteurs, néanmoins le faible nombre de patients inclus est la principale limite de notre étude.

MOTS CLES :

Sclérose en plaques, IRM, Segmentation, Réseau de neurones convolutif, Intelligence artificielle

SERMENT



En présence des Maîtres de cette école, de mes chers condisciples et devant l'effigie d'Hippocrate, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la médecine. Je donnerai mes soins gratuits à l'indigent et n'exigerai jamais un salaire au-dessus de mon travail. Admis dans l'intérieur des maisons mes yeux ne verront pas ce qui s'y passe ; ma langue taira les secrets qui me seront confiés, et mon état ne servira pas à corrompre les mœurs ni à favoriser le crime. Respectueux et reconnaissant envers mes Maîtres, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses ! Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque !

