École des Mines de Paris

T H È S E

pour obtenir le titre de

Docteur en Sciences

de l'École des Mines de Paris

Spécialité : Informatique Temps Réel, Robotique et Automatique

présentée et soutenue par

Guillaume Dugas-Phocion

Segmentation d'IRM Cérébrales Multi-Séquences et Application à la Sclérose en Plaques

Thèse dirigée par Nicholas AYACHE

et préparée à l'INRIA Sophia-Antipolis, projet EPIDAURE

Soutenance prévue le 31 mars 2006

Jury

Nicholas	**AYACHE**	Directeur de thèse	INRIA Sophia
Christian	**BARILLOT**	Rapporteur	IRISA Rennes
Miguel Ángel	**GONZÁLEZ B.**	Examinateur	MEM-ISTB, Bern
Christine	**LEBRUN**	Examinateur	Hôpital Pasteur, Nice
Grégoire	**MALANDAIN**	Co-directeur	INRIA Sophia
Jean-François	**MANGIN**	Rapporteur	CEA, Orsay
Jean-Philippe	**THIRION**	Examinateur	Quantificare, Sophia

Remerciements

Je tiens tout d'abord à remercier les membres du jury : mon directeur de thèse, Nicholas Ayache, qui m'a suivi et soutenu dans les moments les plus difficiles de cette thèse ; Grégoire Malandain, co-directeur de thèse, vers lequel je me suis – très souvent – tourné lorsque trop de questions théoriques et pratiques envahissaient mon esprit. Je remercie Christian Barillot et Jean-Francois Mangin d'avoir consacré de leur temps à la lecture du manuscrit et d'avoir accepté d'être présent pendant la soutenance. Miguel Ángel González Ballester m'a apporté son soutien précieux pendant ma première année de thèse ; son éclairage sur le traitement de ce difficile problème qu'est l'effet de volume partiel et je le remercie chaleureusement. Je remercie l'entreprise QuantifiCare S.A. pour son parrainage de la thèse dans le cadre d'une bourse région, et Jean-Philippe Thirion d'avoir accepté de participer au jury.

Afin de rester en contact avec la réalité du monde médical, cette thèse a été réalisée en collaboration avec le docteur Christine Lebrun, du service de Neurologie du CHU Pasteur à Nice, que je remercie chaleureusement pour son expertise médicale, son aide précieuse lors de la validation et sa présence à la soutenance. Le docteur Caroline Bensa nous a fourni les informations sur les tests cliniques et a effectué un travail approfondi sur les corrélations entre les données issues de l'IRM et les tests cliniques, particulièrement les potentiels évoqués dans le cadre d'une étude sur le déficit cognitif léger. Je remercie également le docteur Stéphane Chanalet, du service de radiologie du CHU Pasteur Nice, qui nous a été d'un grand secours pour l'élaboration du protocole IRM et l'acquisition d'une base de données contenant à ce jour 43 patients et 82 instants.

Je remercie enfin tous mes collègues de bureau qui m'ont supporté pendant ces 3 ans et demi, Isabelle Strobant, colonne vertébrale du projet, qui a su avec beaucoup de tact apaiser les tensions. Mes pensées les plus amicales vont à Radu Stefanescu, qui a su me supporter comme co-bureau pendant plus de 2 ans et demi. J'ai une dette inestimable envers mes parents, sans qui je n'aurais jamais pu commencer ou terminer cette thèse. Enfin, mes pensées les plus amoureuses vont vers Jimena, sans qui j'aurais déjà tout abandonné et qui, en plus d'avoir été

un soutien psychologique incroyable, s’est investi sans compter son temps pour la finalisation de la soutenance et la rédaction de ce manuscrit.

Pour finir, et en guise de *post scriptum*, je remercie Plant, Fanch, Nono, Les Rabins Volants et Monmonc Serge.

Table des matières

Chapitre 1

Introduction

Dans ce manuscrit, nous nous sommes intéressés à l'analyse d'IRM cérébrales dans le cadre notamment du suivi de patients souffrant de sclérose en plaques (SEP). L'IRM est positionnée au niveau des examens complémentaires dans la démarche diagnostique de cette maladie ; elle joue également un rôle clé dans le suivi de l'état du patient et la quantification d'une réponse à une prise de médicaments. L'extraction automatique de quantificateurs pour la sclérose en plaques a donc de nombreuses applications potentielles, tant dans le domaine clinique que pour des tests pharmaceutiques. Par contre, la lecture de ces images est difficile de par la variabilité sur la taille, le contraste et la localisation des lésions : segmenter automatiquement les lésions de SEP en IRM est une tâche difficile.

Dans un premier temps, nous nous sommes consacrés à l'étude générale de la maladie, afin de situer le rôle de l'imagerie (chapitre 2). Alors que le chapitre 3 présente les différents prétraitements nécessaires à la robustesse du système, les chapitres 4 et 5 sont consacrés à la segmentation des tissus sains. Le chapitre 4 présente le modèle théorique dont est issu l'algorithme d'Espérance-Maximisation (EM) et ses applications en segmentation. Le chapitre 5 définit un modèle probabiliste des volumes partiels [40], ainsi qu'une nouvelle version de l'algorithme EM doté d'un traitement spécifique des points aberrants. L'utilisation des différentes séquences du protocole d'acquisition permet, comme le présente le chapitre 6, de raffiner le modèle et de spécialiser le processus pour la détection des lésions

de sclérose en plaques [42, 41]. Une évaluation quantitative des résultats est détaillée dans le chapitre 7 ainsi qu'une présentation des résultats de la chaîne de traitements [43]. Une discussion des résultats est enfin présentée dans le chapitre 8 avec des perspectives de recherche.

Synthèse des contributions

Les chapitres 4 et 5 ont donné lieu à un article accepté à l'oral dans la conférence *7th International Conference on Medical Image Computing and Computer Assisted Intervention* (*MICCAI04*) :

– Guillaume Dugas-Phocion, Miguel Ángel González Ballester, Grégoire Malandain, Christine Lebrun, and Nicholas Ayache. **Improved EM-Based Tissue Segmentation and Partial Volume Effect Quantification in Multi-Sequence Brain MRI**. In *Proc. of MICCAI04*, volume 1496 of Lecture Notes in Computer Science, pages 26 – 33, Saint-Malo, France, September 2004. Springer.

Ces informations sont ensuite utilisées pour la détection des lésions à l'aide de l'IRM T2 FLAIR : les paramètres de classes calculés à l'étape précédente permettent de calculer un seuillage automatique en IRM T2 FLAIR et d'obtenir ainsi une première estimation des lésions de sclérose en plaques. Le contourage peut ensuite être raffiné à l'aide des IRM T1 ou T2/DP : ce travail a donné lieu à une publication dans la conférence ISBI 2004, ainsi qu'à deux publications acceptées en tant que poster dans des conférences spécialisées dans la sclérose en plaques :

– Guillaume Dugas-Phocion, Miguel Ángel González Ballester, Christine Lebrun, Stéphane Chanalet, Caroline Bensa, Grégoire Malandain, and Nicholas Ayache. **Hierarchical Segmentation of Multiple Sclerosis Lesions in Multi-Sequence MRI**. In *International Symposium on Biomedical Imaging : From Nano to Macro (ISBI'04)*, Arlington, VA, USA, April 2004.

– Guillaume Dugas-Phocion, Miguel Ángel González Ballester, Christine Lebrun, Stéphane Chanalet, Caroline Bensa, and Grégoire Malandain. **Segmentation automatique des hypersignaux de la substance blanche (HSB)**

sur des IRM T2 FLAIR de patients atteints de forme rémittente de sclérose en plaques. In *Journées de Neurologie de Langue Française*, Strasbourg, France, April 2004.

– Guillaume Dugas-Phocion, Miguel Ángel González Ballester, Christine Lebrun, Stéphane Chanalet, Caroline Bensa, Marcel Chatel, Nicholas Ayache, and Grégoire Malandain. **Automatic segmentation of white matter lesions in T2 FLAIR MRI of relapsing-remitting multiple sclerosis patients**. In 20th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), Vienna, Austria, October 2004.

Il reste ensuite à utiliser correctement l'information multi-séquences pour obtenir un bon contourage des lésions de SEP, et pas uniquement des hypersignaux de la substance blanche en IRM T2 FLAIR. Une région d'intérêt est calculée à partir de l'IRM T1 : elle permet d'éliminer les artefacts de flux dans les IRM dérivées du T2. Les labélisations, fournies par l'algorithme de segmentation présenté au chapitre 5, affinent le contourage des lésions. Une publication dans l'ECTRIMS 2005 aborde ce processus :

– G. Dugas-Phocion, C. Lebrun, S. Chanalet, M. Chatel, N. Ayache, and G. Malandain. **Automatic segmentation of white matter lesions in multi-sequence MRI of relapsing-remitting multiple sclerosis patients**. In *21th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS)*, Thessaloniki, Greece, September 2005.

Chapitre 2

IRM et sclérose en plaques

2.1 Définition, caractérisation

2.1.1 Introduction

La sclérose en plaques, après l'épilepsie, est la maladie neurologique touchant l'adulte jeune la plus fréquente, dont la cause exacte est encore inconnue actuellement. Historiquement, Charcot fournissait une description des lésions liées à cette pathologie dès 1868 [27], et les progrès fulgurants de la médecine au cours du XXe siècle n'ont apporté que peu d'éléments supplémentaires à la compréhension de la maladie. Ainsi, non seulement la cause de la maladie n'est pas déterminée mais le diagnostic peut être délicat et différé par rapport au début de la maladie.

Les deux sexes sont atteints, avec une proportion nettement plus élevée de femmes (environ deux fois plus de femmes que d'hommes, soit un *sex ratio* de 2/1). Les premiers symptômes apparaissent généralement entre vingt et quarante ans, de manière extrêmement exceptionnelle avant dix ans ou après cinquante ans. Toutes les populations ne sont pas atteintes dans les mêmes proportions, notamment selon leur répartition géographique. Actuellement, la sclérose en plaques affecte environ 2 millions de personnes dans le monde, dont environ 50 000 en France [88, 143].

2.1.2 Études épidémiologiques de la sclérose en plaques : difficultés et résultats

2.1.2.1 Méthodologie

Il est difficile de mener correctement une étude épidémiologique sur la sclérose en plaques. Les deux principales raisons sont liées à la nature de la maladie : il n'y a pas de marqueur spécifique de la SEP, et son diagnostic est retardé par rapport au début de la maladie. Les méthodes de recensement utilisées lors d'études épidémiologiques doivent donc tenir compte des difficultés diagnostiques propres à cette maladie.

L'utilisation des certificats de mortalité n'a que peu d'intérêt car la SEP est rarement la cause directe de la mort. L'incidence mesure le nombre de nouveaux cas apparaissant au sein d'une population pendant une durée donnée. Elle ne peut pas être considérée comme un reflet correct du nombre de cas de SEP en raison du délai de diagnostic de la maladie. L'estimation de la prévalence, c'est-à-dire du nombre de cas connus à un moment donné au sein d'une population, semble plus raisonnable ; mais elle aussi est sujette à critique. La découverte d'autres maladies virales aux symptômes similaires à la SEP, et l'absence de marqueurs spécifiques ont souvent conduit à la reconsidération du diagnostic. En outre, il faut tenir compte des flux de population, et du biais statistique induit par une éventuelle variation du délai de diagnostic de la maladie, en fonction des évolutions des techniques de dépistage.

Enfin, l'outil de recensement des SEP ne constitue pas la seule difficulté de comparaison des études épidémiologiques entre elles. Il faut également prêter attention à la source des informations : qualité de la couverture sanitaire, développement de l'informatisation des dossiers hospitaliers, éventuelle création de services spécialisés dans la prise en charge de la SEP. Tout cela modifie les résultats des analyses statistiques et doit être pris en compte.

2.1.2.2 Facteurs environnementaux et génétiques

En 1938, Steiner fut l'un des premiers à noter l'existence d'un gradient de distribution nord-sud de la SEP. Certaines zones comme la Scandinavie, l'Écosse, l'Europe du Nord, le Canada et le nord des États-Unis souffrent d'une forte prévalence, d'environ 100 pour 100 000 habitants, alors que la prévalence demeure relativement basse, inférieure à 20 pour 100 000 habitants, autour de la Méditerranée et au Mexique. La maladie est exceptionnelle en Afrique noire et en Asie. Kurtzke proposa même de diviser l'hémisphère nord en trois zones, de prévalence variable selon la latitude. En revanche, au sein d'un même pays et donc pour une même latitude, on observe de fortes variations de prévalence, par exemple au sein de la Hongrie, de la Suisse, de l'Italie et de l'Afrique du Sud. Il est donc nécessaire de faire intervenir deux facteurs plus spécifiques : un facteur d'environnement (climat, mode de vie, exposition à des virus), et une susceptibilité génétique.

L'hypothèse d'un facteur d'environnement se vérifie par exemple grâce à l'étude des migrations de populations entre des zones de prévalences inégales. Très schématiquement, ceux qui migrent après l'âge de 15 ans ont le risque de la région d'origine, ceux qui migrent avant l'âge de 15 ans ont le risque de la région d'arrivée. Pour prendre un exemple précis, des Caucasiens d'origine californienne et des Orientaux originaires du Japon se sont installés à Hawaï. La comparaison des prévalences montre que les Orientaux acquièrent un risque plus important de développer une SEP s'ils émigrent du Japon à Hawaï tandis que le risque de développer la maladie diminue chez les Caucasiens qui quittent la Californie pour Hawaï.

L'inégalité géographique pourrait être due aux différences génétiques des populations. Par exemple, les Japonais sont une population à faible prévalence. Cette hypothèse se vérifie dans d'autres cas où la zone géographique est à moyenne ou forte prévalence : on relève la rareté chez les noirs américains au nord comme au sud des USA. D'autre part, pour vérifier l'importance du facteur génétique, des études sont menées pour mesurer la prévalence au sein d'une famille, ou mieux encore, entre jumeaux monozygotes. Le nombre de familles multi-cas ne paraît pas être dû au hasard : on estime le risque de 2% pour la famille proche d'un

patient (parent : 2,75% ; enfant 2,5% ; fratrie 4% ; oncle ou tante 2% ; neveu ou nièce 1.5% ; cousin germain 1,75%), soit un facteur de 50 par rapport aux sujets non-apparentés. Des études menées sur des jumeaux dont un est porteur de la maladie montrent pour les dizygotes (jumeaux dont le code génétique est différent) une concordance de 2% et pour les monozygotes une concordance de 20% [88, page 21]. Ceci fait apparaître à la fois l'importance du facteur génétique dans le déterminisme de la maladie et le fait qu'il ne s'agit pas d'une maladie héréditaire, auquel cas il y aurait une concordance de 100% pour les jumeaux monozygotes. Il existe donc clairement une susceptibilité d'origine génétique, mais qui ne suffit pas pour que la maladie se déclare.

L'épidémiologie de la sclérose en plaques est un problème complexe dont la présentation ci-dessus n'est qu'un résumé. Pour une information plus détaillée, on peut consulter [3, 34, 126].

2.1.3 Examen macroscopique et microscopique de la maladie

Les caractères neuropathologiques classiques de la sclérose en plaques, pour l'essentiel, sont connus depuis les descriptions de Charcot [26]. Les plaques, qui ont donné leur nom à la maladie, sont des lésions focales du système nerveux central dont les aspects microscopiques sont variables, mais qui comportent très habituellement une démyélinisation et souvent une réaction gliale intense. Rappelons que le système nerveux central et périphérique est un réseau fonctionnel, composé de dizaines de milliards de cellules, les neurones, qui communiquent entre elles par des signaux électriques – le long des axones – et chimiques – via le système synaptique. L'axone, prolongement du neurone pouvant atteindre un mètre de long, permet d'acheminer le signal électrique généré par le corps du neurone. La myéline est un constituant lipidique qui permet d'accélérer le transfert de l'information électrique le long de l'axone. Cette gaine est composée par d'autres cellules du cerveau : les oligodendrocytes – 10 fois plus nombreux que les neurones – qui composent la substance blanche. Le rôle essentiel de cette substance blanche est d'augmenter la vitesse de transmission de l'influx nerveux : sans elle, l'information circulerait 100 fois plus lentement. On comprend donc ai-

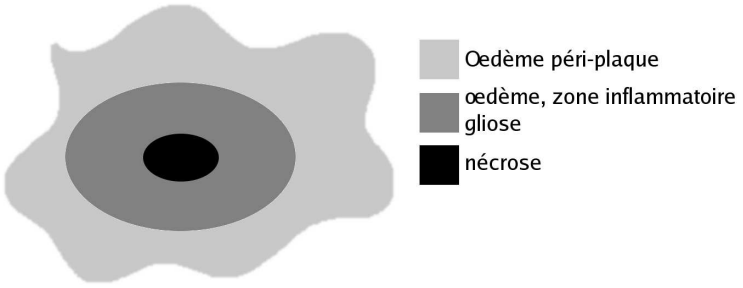


FIG. 2.1 – *Une zone lésée se présente sous la forme d'une plaque sclérosée. Une plaque est généralement hétérogène et peut se composer de nécrose (le liquide céphalo-rachidien vient remplir cette cavité nécrotique), d'œdème, de gliose (zone démyélinisée totalement) et d´une zone inflammatoire ; tout autour de la plaque, une couronne œdémateuse peut apparaître.*

sément que les plaques, signe de démyélinisation autour des axones, perturbent le fonctionnement du cerveau, puisque les réseaux de fibres passant par ces plaques sont temporairement hors d'usage. Par ailleurs, la démyélinisation s'accompagne d'une inflammation qui n'est pas sans conséquence sur le fonctionnement du cerveau : ceci se traduit, le plus souvent, par une "poussée clinique".

Une étude histologique *post mortem* d'un cas suspect de SEP nécessite l'examen des lésions présentes dans le parenchyme cérébral, mais aussi dans des secteurs clés du système nerveux central tels que la moelle épinière ou le nerf optique. Les plaques sont la plupart du temps visibles à l'œil nu : elles apparaissent comme des foyers grisâtres dont la taille varie de celle d'une tête d'épingle à de larges plages occupant la quasi-totalité d'un hémisphère. L'examen microscopique montre que les fibres nerveuses n'ont pas disparu, mais elles ont été dénudées de leur manchon de myéline. Une plaque est généralement hétérogène et peut se composer de nécrose (le liquide céphalo-rachidien (LCR) vient remplir cette cavité nécrotique), d'œdème, de gliose (zone totalement démyélinisée) et d'une zone inflammatoire ; tout autour de la plaque, une couronne œdémateuse peut apparaître (figure 2.1).

Ces plaques se situent essentiellement dans la substance blanche et sont répar-

ties de manière très irrégulière d'un cas à l'autre. Certaines plaques sous-corticales peuvent cependant atteindre la subtance grise ; d'autre touchent directement la matière grise, comme le cortex ou le tronc cérébral. Les lésions les plus nombreuses et les plus volumineuses du cerveau se trouvent dans le voisinage des ventricules cérébraux, essentiellement dans les régions postérieures : ce sont les plaques périventriculaires, les plus caractéristiques de la maladie. Les nerfs optiques et la moelle épinière sont aussi des zones privilégiées d'évolution des plaques.

Les autres altérations du système nerveux central appréciables macroscopiquement sont plus rares et peu spécifiques : dilatation ventriculaire, épendymite granuleuse [117]. Une atrophie cérébrale forte est également visible dans environ la moitié des cas de SEP anciennes [73].

2.2 Démarche diagnostique pour la sclérose en plaques

Les plaques apparaissent et régressent (parfois disparaissent) au cours de l'évolution de la maladie. Pendant certaines périodes appelées poussées, des symptômes peuvent apparaître brutalement, puis s'atténuer et même disparaître lors de la phase de rémission. C'est à la fréquence de ces poussées, mais surtout à la qualité des rémissions que tient la gravité de la maladie. Il existe plusieurs formes différentes de sclérose en plaques dont la gravité est variable. Certaines sont particulièrement aiguës : elles se caractérisent par une symptomatologie brève, où les lésions sont le siège de phénomènes inflammatoires intenses. Mais de telles formes sont plutôt rares. Les formes les plus courantes sont les suivantes :

- formes rémittentes pures avec des poussées plus ou moins nombreuses : l'invalidité résiduelle est variable, ainsi que l'intervalle entre les poussées (de quelques mois à plus de 10 ans). Elles représentent plus de 80% de l'ensemble des patients.
- formes progressives, soit d'emblée dans 10% des cas, soit après une période plus ou moins longue de forme rémittente. La maladie et l'invalidité s'aggravent sur une période supérieure à un an.
- formes rémittentes-progressives : progression chronique de l'invalidité avec

épisodes d'aggravation suivis d'une amélioration progressive.

Tous les intermédiaires existent entre les formes bénignes (10% des cas) correspondant à un handicap non significatif après 10 ans d'évolution et les formes graves diminuant rapidement l'autonomie. Globalement, la diminution de l'espérance de vie est faible, de l'ordre de 10 à 15 % par rapport aux sujets sains.

2.2.1 Symptomatologie

La symptomatologie révélatrice de la SEP est très variée et dépend de l'âge du patient. Notons que la première poussée est isolée dans 45 à 65% des cas. Les symptômes les plus fréquents et les plus faciles à mettre en évidence sont les manifestations physiques : signes oculaires (baisse de la vision d'un œil, douleurs orbitaires, vision double), troubles moteurs (troubles de l'équilibre, paralysies etc.), troubles sensoriels (fourmillements, névralgies, hyperpathies). La gravité des symptômes d'un patient atteint de sclérose en plaques peut être quantifiée grâce à des échelles de référence. Actuellement, c'est l'échelle EDSS (*Expanded Disability Status Scale*), mise au point par le Dr Kurtzke [79], qui est utilisée par les neurologues.

2.2.2 Schéma diagnostique

En l'absence de signe biologique ou clinique précis de la maladie, le diagnostic de SEP ressemble plus à un diagnostic d'élimination qu'à un diagnostic positif. D'autres maladies peuvent répondre à ces critères, comme par exemple la maladie de Lyme ou certaines tumeurs cérébrales. Plusieurs paramètres entrent en ligne de compte.

Le nombre de poussées précédentes donne généralement une bonne indication sur le diagnostic ; un syndrome cliniquement isolé – une seule poussée – ne peut pas être directement diagnostiqué comme une SEP. L'examen clinique joue un rôle déterminant : son but est d'évaluer l'état de santé du patient grâce à son interrogatoire et son examen objectif. L'interrogatoire, bien que subjectif, est indispensable car il donne des informations sur les symptômes à durée courte, tels

que le syndrome de Lhermitte[1]. Par contre, dans le cas où la description faite par le patient est peu évocatrice, seules les données de l'examen clinique pratiqué par un médecin au moment de la poussée permettent de conclure à l'existence ou non de celle-ci. Un examen neurologique complet permet de vérifier les fonctions motrices, les réflexes, la sensibilité, la coordination du mouvement, la statique, les fonctions génito-sphinctériennes etc ; un examen général complémentaire peut être utile. Le cumul des signes cliniques correspondant à des zones neurologiques éloignées fait généralement penser à la sclérose en plaques.

Enfin, pour la plupart des maladies neurologiques, et notamment en cas de doute sur les données issues de l'examen clinique, il peut être extrêmement utile de recourir à d'autres examens complémentaires. Pour la sclérose en plaques, les examens biologiques, radiologiques et électrophysiologiques ont un triple intérêt. Au début, ils permettent d'écarter d'autres diagnostics possibles ; ultérieurement, ils contribuent au diagnostic de SEP qui peut être, selon les cas, définie ou probable. Enfin, dans le cadre de protocoles de recherche intéressant des groupes de malades, ils peuvent participer à évaluer l'efficacité des traitements. Essentiellement répartis entre l'examen du liquide céphalo-rachidien, les potentiels évoqués et l'imagerie par résonance magnétique, les examens complémentaires apportent un faisceau d'arguments pour confirmer le diagnostic.

2.3 L'IRM incluse dans la démarche diagnostique

2.3.1 Utilisation de l'IRM

Pour la sclérose en plaques, l'IRM est très certainement l'examen complémentaire le plus utilisé. Elle donne une indication sur la nature des tissus, et fait notamment apparaître les zones démyélinisées qui sont évocatrices de la sclérose en plaques [105]. Par contre, l'IRM n'est pas spécifique pour la sclérose en plaques. Elle possède une valeur prédictive positive, mais il existe plusieurs cas où les images contiennent des zones lésionnelles d'apparence identique aux lésions

[1]Sensation de piqûres d'aiguille venant de la colonne sur les bras et les jambes qui apparaît lorsque l'on baisse la tête de façon à ce que le menton touche la poitrine.

de sclérose en plaques [14], mais de nature complètement différente en réalité. L'IRM se situe bien donc dans la catégorie des examens complémentaires.

L'IRM contribue au diagnostic de la maladie [160, 170], au suivi des patients au cours du temps [52], et à l'évaluation des traitements [98]. Son intérêt pour le pronostic dans le cadre de la sclérose en plaques semble réel, mais aucun travail n'en apporte une preuve formelle. L'utilisation de l'IRM pour la SEP est un sujet d'étude à part entière : beaucoup de travaux sont publiés sur ce sujet, et l'utilisation de l'IRM évolue nettement au cours des années [150, 161, 47], notamment grâce à l'apparition de nouvelles séquences [51] et à l'augmentation de la résolution des images [44]. L'intégralité du parenchyme cérébral – par exemple la matière blanche d'apparence normale – est également un sujet d'études récentes [100].

Différents protocoles sont utilisés en IRM pour voir les lésions de sclérose en plaques. De nombreux articles montrent la corrélation qui existe entre les lésions et leur aspect IRM [170, 169, 132, 57], et il est impossible d'être exhaustif à ce sujet. Au cours de la maladie, les zones démyélinisées apparaissent comme des zones d'hypersignal sur des images pondérées en T2 qui reste encore aujourd'hui une modalité très importante pour la sclérose en plaques : les critères de Barkhof, présentés ci-après, font référence à cette séquence. Certaines modalités sont aujourd'hui couramment utilisées par les médecins : IRM pondérée en T1, avec ou sans injection de produit de contraste, densité de protons (IRM DP). L'ensemble constitue l'IRM dite conventionnelle. D'autres séquences ont été très étudiées ces dernières années mais, bien que la littérature associée soit abondante, ne sont pas utilisées en clinique.

2.3.2 Les critères IRM pour la sclérose en plaques

À la première poussée, le diagnostic est assez difficile à poser de manière sûre et définitive. Même si le clinicien peut avoir des présomptions avec les différents éléments cliniques et para-cliniques, il est important de ne pas inquiéter le patient inutilement. À l'inverse, il est aujourd'hui prouvé que la mise en place d'un traitement de manière précoce permet de ralentir significativement l'évolution de la

maladie. C'est pourquoi il est devenu indispensable de définir des critères de diagnostic précis qui permettent dans la majeure partie des cas de ne pas se tromper. Il y a d'une part les critères spécifiques sur les IRM, d'autre part des critères plus complets fondés sur l'ensemble des examens disponibles.

Les critères IRM utilisés aujourd'hui sont ceux de Barkhof [11] qui semblent plus précis que les critères plus anciens [115, 49]. L'IRM du patient est dite suggestive si au moins 3 des critères suivants sont rencontrés :

- une lésion rehaussée par le gadolinium en IRM T1 ou 9 lésions hyperintenses T2,
- au moins une lésion sous-tentorielle, c'est-à-dire dans la zone située en dessous de la tente du cervelet,
- au moins une lésion juxtacorticale – à proximité du cortex,
- au moins 3 lésions périventriculaires – à proximité des ventricules.

Les critères de Poser [125] ne tenaient pas suffisamment compte de l'information importante issue des examens complémentaires, et notamment de l'IRM. Ils ont été remplacés par les critères de McDonald [97] : à partir de l'ensemble des éléments récoltés, on cherche à poser le diagnostic suivant l'idée que la sclérose en plaques présente à la fois une dissémination spatiale – les lésions peuvent apparaître à plusieurs endroits du cerveau, dans la mœlle épinière ou dans les nerfs optiques – et une dissémination temporelle - les lésions sont évolutives et d'âge différent. Ces critères illustrent ce mode de pensée : un tel diagnostic sera posé si on se trouve dans l'un des deux cas suivants :

- le patient a fait deux poussées, avec deux symptômes différents (ou plus) ;
- le patient a fait deux poussées, mais avec un seul symptôme identique lors des deux poussées. La dissémination spatiale n'est pas prouvée, et c'est l'IRM qui est utilisée dans ce cas :
  - l'IRM est suggestive selon les critères de Barkhof ou ;
  - l'IRM est douteuse selon les critères de Barkhof, et le liquide céphalorachidien contient des anticorps spécifiques à la sclérose en plaques.
- le patient a fait une seule poussée, mais avec deux symptômes simultanés. La dissémination temporelle n'étant pas prouvée, deux critères sont néces-

saires pour fixer le diagnostic :
  - une deuxième poussée survient ;
  - l'IRM est évolutive.
- si une seule poussée survient avec un seul symptôme, ni la dissémination spatiale ni la dissémination temporelle ne sont prouvées. Il faut alors recourir aux deux précédents critères conjointement pour aboutir au diagnostic positif de sclérose en plaques.

2.3.3 IRM conventionnelle, IRM multi-séquences

L'importance de l'IRM dans le cadre de l'étude de la sclérose en plaques s'est fortement accrue ces dernières années, tant pour l'aide au diagnostic que pour la mesure d'efficacité des traitements et d'évolution de la maladie dans le cadre d'essais cliniques. Quelques techniques d'IRM dites conventionnelles, ont démontré leur utilité : IRM pondérée en T2, T1 avec prise de produit de contraste et leurs dérivés. Cependant, les nombreuses études statistiques effectuées récemment ont montré une faible corrélation entre les métriques calculées à partir de segmentation de lésions en IRM conventionnelle, et les manifestations cliniques de la maladie [50]. De nouvelles techniques d'acquisition ont fait leur apparition, et les métriques correspondantes semblent être plus adaptées à l'observation de l'atteinte tissulaire spécifique à la sclérose en plaques [52].

2.3.3.1 Les limites de l'IRM conventionnelle

Une mesure classique consiste à extraire les quantificateurs globaux issus des IRM T2/DP. Sur ces séquences, il est établi que les hypersignaux permettent d'évaluer l'atteinte non-spécifique des tissus : œdème, gliose, démyélinisation réversible, atteinte axonale irréversible. Il a été également montré que les quantificateurs issus de l'étude des lésions cérébrales en IRM T2 lors d'un syndrome cliniquement isolé ont une forte valeur prédictive pour l'apparition d'une sclérose cliniquement définie et l'évolution de la maladie sur 10 ans : 30% des patients présentant une IRM anormale développent une SEP dans l'année qui suit [19],

50% après 5 ans et 80% après 10 ans [110]. À titre de comparaison, 10 à 15% seulement des patients présentant une IRM normale développeront une SEP dans les 10 années suivantes. De plus, l'apparition de nouvelles lésions visibles en T2 3 mois plus tard augmente fortement le risque d'apparition d'une SEP cliniquement définie [19]. La corrélation entre le nombre de lésions visibles en T2 ou la charge lésionnelle T2, avec le temps séparant la première de la deuxième poussée est forte. Dans ce cas, l'IRM conventionnelle est donc un biomarqueur pour le développement de SEP cliniquement définie [11].

En revanche, la plupart des études croisées concernant les formes rémittentes de SEP – les plus courantes – et utilisant les quantificateurs issus de l'IRM T2 ont présenté des résultats significatifs du point de vue statistique, mais avec des corrélations relativement faibles. La valeur prédictive de la charge lésionnelle T2, notamment, s'est avérée relativement faible en ce qui concerne l'anticipation de l'atteinte permanente du patient [75]. Cette différence entre les syndromes cliniquement isolés, et les formes rémittentes de SEP demeure difficile à expliquer. Il est possible que la faible valeur prédictive ait pour cause une période de rémission après la poussée : afin d'étudier la conséquence de chaque poussée, il conviendrait d'introduire un délai important entre la poussée et l'acquisition IRM.

Les lésions hypointenses en IRM T1 correspondent aux zones de démyélinisation. Bien qu'une forte corrélation existe entre ces lésions et les échelles de type EDSS [170], un tel quantificateur est difficile à exploiter dans le cadre de la sclérose en plaques. Le problème principal est que l'atteinte tissulaire est extrêmement variable d'une lésion à l'autre [84] : une lésion hypointense en T1 sur deux disparaît après la poussée, ce qui montre que la destruction des tissus n'était qu'apparente. Enfin, les lésions sont très difficiles à détecter dans les zones critiques, comme le tronc cérébral, la mœlle épinière ou le nerf optique [55].

Il apparaît donc nécessaire de définir de nouveaux quantificateurs, issus de nouvelles modalités d'acquisition, et qui seraient plus sensibles, plus spécifiques et surtout plus adaptés à la maladie. Actuellement, aucune des métriques issues de l'IRM conventionnelle ne respecte les prérequis nécessaires pour être un substitut d'un marqueur clinique de la maladie (voir chapitre 9.1 pour plus de détails).

Plusieurs nouvelles modalités ont pourtant un très fort potentiel : le transfert de magnétisation et la spectroscopie.

2.3.3.2 Le potentiel des nouvelles modalités : l'IRM-MTR

Parmi les nouvelles modalités d'acquisition, **l'IRM de transfert de magnétisation** (IRM-MTR) est probablement la plus prometteuse pour la sclérose en plaques. Cette technique est basée sur les interactions entre les protons qui bougent librement et ceux dont le mouvement est réduit. Dans le liquide céphalo-rachidien, ces deux états correspondent respectivement aux protons présents dans l'eau et dans les macro-molécules de myéline, pour simplifier. Une impulsion d'une fréquence légèrement différente de la fréquence de résonance est appliquée, ce qui sature la magnétisation des protons les moins mobiles ; cet excès de magnétisation est transféré aux protons les plus mobiles, donc réduit l'intensité de la magnétisation observable. La perte de signal dépend de la densité des macro-molécules dans un tissu. Donc, un rapport de transfert de magnétisation bas (MTR, ou *Magnetization Transfer Ratio*) reflète une capacité réduite pour les macro-molécules dans le LCR d'échanger de la magnétisation avec les molécules d'eau aux alentours, ce qui montre dans le cas de la SEP une perte de myéline et une réduction de la masse axonale.

Cette modalité est intéressante pour plusieurs raisons. Tout d'abord, elle fournit une mesure quantitative et continue, en relation directe avec l'atteinte tissulaire : bien que cette atteinte ne soit pas *a priori* spécifique à la SEP, des études *post-mortem* sur l'humain et le petit animal montrent une relation stricte entre la perte de signal et l'atteinte axonale et le degré de démyélinisation [168]. En outre, l'étude peut alors être menée sur la totalité du cerveau et permet d'estimer une atteinte générale de la matière blanche d'apparence normale [28]. Cependant, l'IRM MTR présente des inconvénients. Le souci majeur est la nécessité de couplage à l'acquisition avec une modalité conventionnelle, ce qui pose des problèmes de reproductibilité. En l'absence de toute étude rétrospective au long cours, ceci explique probablement la faible utilisation clinique actuelle.

L'IRM spectroscopique est également très intéressante car elle fournit des

IRM T2 FSE IRM DP IRM T1 3D IRM T2 FLAIR

FIG. 2.2 – Exemple des 4 séquences utilisées dans le protocole d'acquisition : T2 FSE, Densité de Protons (DP), T1 3D et T2 FLAIR. Les lésions de sclérose en plaques sont visibles dans toutes les séquences, mais particulièrement en T2 FLAIR où les hypersignaux sont très visibles.

informations quantitatives et continues sur les deux aspects majeurs de la pathologie, à savoir l'inflammation associée à l'attaque de la myéline et l'atteinte axonale [116]. Cette dernière peut notamment être estimée via l'examen des changements des mesures du N-acétyl-aspartate, essentiellement localisé dans les neurones et leurs ramifications (dendrites et axones). Une réduction de la résonance de cette molécule révèle directement un changement du métabolisme des neurones. Par contre le rapport signal sur bruit est assez faible, et les durées d'acquisition sont beaucoup trop importantes pour autoriser une utilisation en milieu clinique.

2.3.3.3 De la nécessité d'une analyse multi-séquences

Au final, même si ces technologies semblent beaucoup plus adaptées à la maladie que l'IRM conventionnelle, elles demeurent trop récentes pour être utilisées en milieu clinique. Alors que les moyens informatiques se généralisent de plus en plus en milieu hospitalier, il est donc nécessaire d'exploiter au maximum, si possible de manière automatique, l'IRM conventionnelle. Ceci impose donc un traitement multi-séquences des images, pour plusieurs raisons.

Un système de traitement d'images médicales spécialisé dans le traitement d'une pathologie doit être robuste à une altération de la qualité des images ou à un changement de l'appareil d'acquisition, pour ne prendre que deux exemples. Une segmentation multi-séquences des tissus sains permet au système de prendre

connaissance de l'état des images (qualité, résolution) et d'obtenir des informations chiffrées et fiables sur la caractérisation des tissus dans toutes les séquences disponibles. Ensuite, la détection puis le contourage des différentes lésions doit exploiter séparément chacune des différentes séquences : pour l'étude de la sclérose en plaques, le couple T2/DP donne l'atteinte tissulaire globale, œdème et gliose compris ; le T2 FLAIR renforce l'œdème, le T1 montre les zones potentiellement nécrotiques et l'atteinte axonale.

Le système idéal est donc une chaîne de traitements globale. Un prétraitement sur les images permet en premier lieu de supprimer les différents artefacts en effectuant une normalisation spatiale et en intensité, pour que le traitement lui-même s'effectue dans le même repère : le chapitre 3 donne les détails de ces prétraitements. Le chapitre 4 explique ensuite la méthode générale, basée sur l'algorithme EM, pour obtenir un masque grossier du parenchyme cérébral pour ensuite segmenter la matière blanche, la matière grise et le liquide céphalo-rachidien. Le chapitre 5 précise la méthode présentée au chapitre 4 pour mieux tenir compte d'un artefact d'acquisition, les volumes partiels, qui gêne tout particulièrement l'étude des lésions de SEP. Une fois que les tissus potentiellement pathologiques sont obtenus, un pipeline de post-traitements, présenté dans le chapitre 6, permet d'augmenter la spécificité du système en séparant les véritables lésions des différents artefacts. Enfin une évaluation quantitative montre l'efficacité du système dans le chapitre 7, et de nombreuses perspectives de recherche sont présentées dans le chapitre 8.

Chapitre 3

Prétraitements

Un neurologue qui travaille en milieu hospitalier, sans dispositif informatique, doit aujourd'hui regarder les IRM sur des planches. Elles ont l'aspect de grandes radiographies sur lesquelles sont disposées régulièrement plusieurs coupes. En raison du coût du passage à l'IRM, on préfère souvent un ensemble de coupes fines mais non jointives (voir paragraphe 3.2), ce qui réduit la portée de l'analyse des images. De nombreux facteurs peuvent fausser le diagnostic du praticien : ce sont ces artefacts que les prétraitements vont chercher à corriger.

3.1 Introduction rapide à la résonance magnétique nucléaire

Pour montrer l'origine des différents artefacts d'acquisition, résumons en quelques phrases le processus. Précisons que les paragraphes suivants ne sont qu'un simple résumé d'un système fort complexe, mais suffisent à montrer l'origine des artefacts d'acquisition.

La résonance magnétique nucléaire est une technique en développement depuis une cinquantaine d'années. Le phénomène physique a été conceptualisé en 1946 par Bloch et Purcell, parallèlement. Ils ont obtenu le prix Nobel en 1952. Cette technique a été largement utilisée par les chimistes, puis les biologistes. Les premiers développements en imagerie datent des années 1973. Les premières

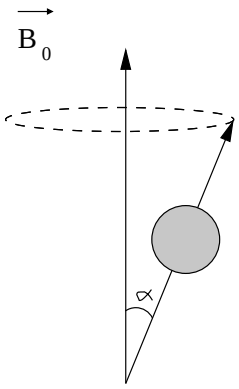


FIG. 3.1 – Le spin représenté ici a un mouvement de précession autour du champ magnétique $\overrightarrow{B_0}$ à la fréquence ω_0 dite fréquence de Larmor.

images chez l'homme ont été réalisées en 1979. Aujourd'hui, après 15 ans d'évolution, l'IRM est devenue une technique majeure de l'imagerie médicale moderne.

L'IRM est basée sur l'interaction des protons, notamment ceux présents dans l'eau, sous l'effet d'un champ magnétique extérieur. Le comportement du proton est assimilable à celui d'un aimant de moment magnétique $\overrightarrow{\mu}$. Au repos, la résultante (somme des moments magnétiques) $\overrightarrow{M} = \sum_v \overrightarrow{\mu}$ est nulle (figure 3.2, a). Lorsque les protons sont plongés dans un champ magnétique $\overrightarrow{B_0}$, leurs moments magnétiques de spin $\overrightarrow{\mu}$ s'alignent localement sur la direction de $\overrightarrow{B_0}$. En pratique, en appliquant des lois de la mécanique quantique qui seront passées sous silence dans ce manuscrit, seulement deux positions sont possibles : chaque spin fait un angle α ou $\pi - \alpha$ avec le champ $\overrightarrow{B_0}$. Ces spins sont alors animés d'un mouvement de précession autour de $\overrightarrow{B_0}$ à une fréquence précise dépendant directement de $\overrightarrow{B_0}$, la fréquence de Larmor (figure 3.1). L'orientation parallèle est la plus probable, car le niveau d'énergie est plus bas qu'en position antiparallèle. Ceci génère une résultante $\overrightarrow{M}$ orientée dans la direction du champ $\overrightarrow{B_0}$ (figure 3.2, b).

Le principe de la résonance magnétique nucléaire consiste en l'application d'une onde électromagnétique $\overrightarrow{B_1}$ à la fréquence de résonance, c'est-à-dire à la fréquence de Larmor, pour que l'apport énergétique soit possible : c'est l'impulsion RF, ou impulsion radio-fréquence. Le moment magnétique macroscopique $\overrightarrow{M}$ évolue au cours du temps suivant un mouvement que l'on peut décrire comme

une spirale assortie d'un basculement [18]. L'angle de basculement de $\overrightarrow{M}$ varie en fonction de l'énergie émise, c'est-à-dire de la durée et de l'amplitude de cette excitation. On peut déterminer un angle de 90 degrés (figure 3.2, c), ou 180 degrés en adaptant ces deux paramètres (figure 3.2, d).

Lors de l'arrêt de l'excitation par l'impulsion RF, la dynamique de retour des spins à l'état d'équilibre crée une onde électro-magnétique qui peut être mesurée par une antenne. La mesure IRM est en fait la mesure du temps de relaxation de ce signal : elle dépend de l'intensité du champ magnétique constant $\overrightarrow{B_0}$ mais également de la nature des tissus. Plus précisément, il se trouve que des phénomènes physiques différents caractérisent la relaxation de l'aimantation longitudinale m_z et de l'aimantation transverse m_{xy}. Ces phénomènes sont caractérisés par deux échelles temporelles que l'on a coutume de paramétrer par deux constantes de temps T1 – relaxation de m_z – et T2 – relaxation de m_{xy}. D'autres variations sont possibles : par exemple, l'imagerie T2 FLAIR est un T2 avec annulation du signal des liquides par une première impulsion à 180 degrés.

Tout le processus présenté ci-dessus est un résultat obtenu sur un volume global. Un codage spatial peut s'effectuer dans l'espace des fréquence ou des phases : des détails supplémentaires sur ces différentes techniques de codage sont disponibles dans le livre *The Basics of MRI* [1].

Plusieurs facteurs peuvent entraîner une dégradation de la qualité de l'image. Ainsi, un champ $\overrightarrow{B_0}$ non homogène va provoquer un biais spatial de reconstruction [152]. Certaines distortions géométriques sont également parfois constatées, surtout lorsque le champ magnétique $\overrightarrow{B_0}$ est élevé. L'interaction entre le champ oscillant $\overrightarrow{B_1}$ et les différents tissus ou des objets ferromagnétiques peut introduire des variations de l'intensité (inhomogénéités RF). Des artefacts dits de flux peuvent apparaître : il suffit qu'un volume en mouvement qui traverse une coupe subisse une pulsation radio-fréquence via $\overrightarrow{B_1}$, mais soit sorti de la coupe quand le signal est enregistré. Cela peut induire des variations du signal qui nous gêneront particulièrement dans l'analyse des images T2 FLAIR par la suite, comme il sera

[1] *The Basics of MRI*. Joseph P. Hornak, Ph.D. Copyright © 1996-2006 J.P. Hornak. All Rights Reserved.

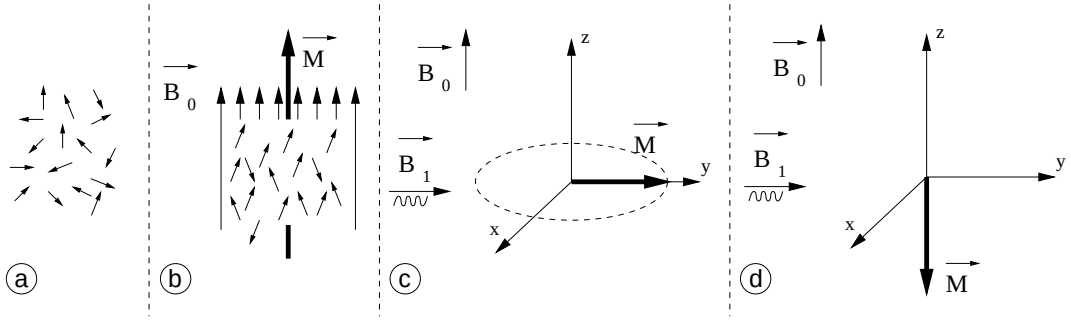


FIG. 3.2 – Sans champ magnétique, la résultante des moments est nulle car l'orientation des spins est aléatoire (a). Sous l'effet d'un champ magnétique $\overrightarrow{B_0}$, il y a un excédent des spins en position parallèle (angle α) par rapport aux autres (angle $\pi - \alpha$), ce qui crée une résultante des moments $\overrightarrow{M}$, colinéaire à $\overrightarrow{B_0}$ (b). Il est possible d'obtenir, en fonction de la durée et de l'amplitude de l'excitation $\overrightarrow{B_1}$, d'obtenir un angle de basculement de 90 degrés (c) ou de 180 degrés (d) de la résultante $\overrightarrow{M}$.

montré dans le chapitre 6.

D'autres facteurs peuvent altérer la qualité des IRM : des mouvements intempestifs, par exemple des mouvements oculaires, peuvent induire des échos dans les images. Outre ces problèmes à l'acquisition, la qualité, la résolution, le contraste des images n'est pas fixé à l'avance, ce qui complique une analyse automatique des images. Il est donc indispensable de générer une chaîne de prétraitements qui permettra au système de segmentation de travailler sur des bases solides : acquisition, positionnement du repère spatial, normalisation des images. C'est le but de ce chapitre.

3.2 Acquisition et lecture des images

3.2.1 Lecture des images

En pratique, les IRM sont fournies sous la forme d'un ensemble d'images 2D qui, mises dans la bonne géométrie, vont former une image tridimensionnelle. Pour assembler ces différentes images, un ensemble de paramètres est disponible

pour chaque coupe suivant une norme, la norme DICOM. Dans notre cas, ces paramètres nous sont donnés dans un ensemble de fichiers, les en-têtes DICOM. L'interprétation de ces paramètres n'est pas aisée : l'information y est redondante et pas toujours cohérente. Certains champs sont véritablement importants pour la reconstruction de l'image, c'est-à-dire le placement des coupes dans la bonne géométrie ; d'autres permettent d'obtenir des informations cliniques sur le patient ou le protocole d'acquisition lui-même. L'examen de l'annexe C de la partie 3 de la documentation générale sur DICOM permet d'en savoir un peu plus sur la définition des différents champs[2].

Pour reconstruire l'image tri-dimensionnelle à partir des différentes coupes 2D, il nous faut insérer le repère 2D de chaque coupe dans le repère 3D de l'imageur. L'orientation du repère 2D, la position de son origine, ainsi que la taille des voxels et l'épaisseur des coupes sont donc nécessaires. Ces 4 paramètres sont donnés par 4 champs DICOM : *Pixel Spacing*, *Image Orientation (Patient)*, *Image Position (Patient)*, et *Slice Thickness*, et permettent de calculer la taille du voxel lors de l'acquisition (figure 3.3). Ces voxels représentent les volumes unitaires dans lesquels le signal IRM est calculé.

- *Image Position (Patient)* donne la position, dans le repère IRM, de l'origine du repère de la coupe 2D.
- *Image Orientation (Patient)* donne, dans le repère IRM, les 2 vecteurs du repère de la coupe 2D.
- *Pixel Spacing* donne, dans le repère IRM, la taille des deux vecteurs du repère de la coupe 2D.
- *Slice Thickness* donne l'épaisseur de la coupe, ce qui permet de voir si les coupes sont jointives ou non.

3.2.2 Reconstruction des images

Avant toute reconstruction, il faut tout d'abord s'assurer que les coupes sont bien jointives. Pour des raisons économiques, les IRM sont parfois fournies sous

[2]La documentation complète est téléchargeable à l´adresse suivante : *http://medical.nema.org/dicom/2004.html*

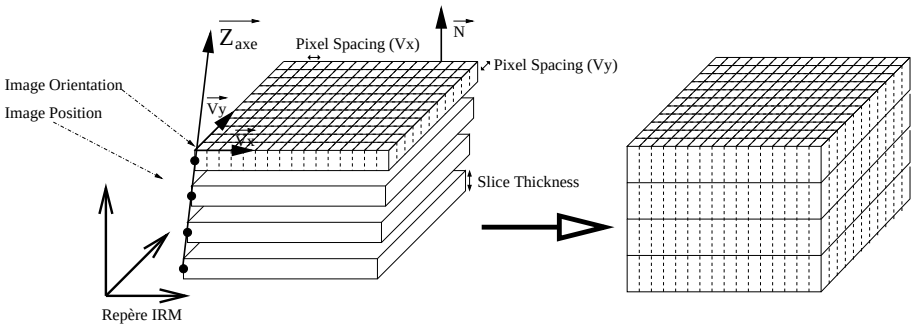


FIG. 3.3 – La reconstruction des IRM doit se faire avec précaution. Les coupes ne sont pas obligatoirement jointives, et rien de garantit que l'axe de progression de l'origine du repère 2D – $\overrightarrow{Z_{axe}}$ – soit perpendiculaire au plan de coupe. Dans tous les cas, les voxels issus de l'acquisition coupe à coupe (à gauche) ne seront pas automatiquement les mêmes que ceux de l'image finale qui sera traitée par le système (à droite).

forme de coupes fines (épaisseur de l'ordre du millimètre) mais non jointives, ce qui assure des images de bonne qualité, mais plus difficilement compatibles avec une exploitation tri-dimensionnelle, puisque des données sont manquantes pour reconstituer le volume. Il faut donc faire attention que l'épaisseur des coupes soit compatible avec la distance inter-coupe.

Pour insérer proprement une coupe dans le repère IRM, il suffit de placer le repère de la coupe 2D dans un repère fixe à l'aide des coordonnées ci-dessus, et de construire un volume de coupe d'épaisseur donnée par *Slice Thickness* (figure 3.3). Parfois, l'axe de progression de l'origine du repère 2D ne suit pas la direction donnée par la normale au plan de chaque coupe : il faut alors rééchantillonner les images 2D pour construire un volume 3D. Dans notre protocole d'acquisition, les séquences dérivées du T2 – T2, densité de protons, T2 FLAIR – sont particulièrement touchées par ce phénomène, l'angle $(\overrightarrow{N}, \overrightarrow{Z_{AXE}})$ peut atteindre 3 degrés. Ce rééchantillonnage est pourtant dommageable, car il crée des effets de volume partiel supplémentaires.

La solution choisie est de faire comme si ce rééchantillonnage n'était pas nécessaire, tout en conservant une matrice de transformation applicable à l'image en cas de besoin (figure 3.4). Le détail du calcul de cette matrice est donné en

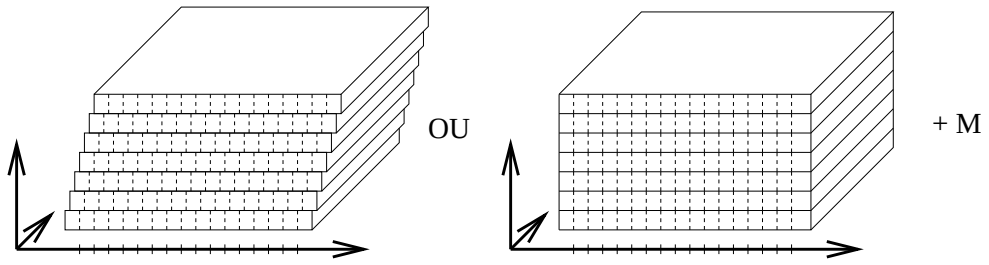


FIG. 3.4 – Un rééchantillonnage des coupes (figure de gauche) va rajouter des problèmes de volumes partiels. Une solution consiste à faire comme si le repère image était orthogonal et à garder une matrice de transformation affine (figure de droite) que l'on utilise, par exemple, pour recaler les différentes séquences entre elles.

annexe de ce manuscrit (chapitre 9). Ainsi, pour tous les traitements – comme l'algorithme EM présenté ci-après – qui ne portent que sur une seule séquence, les images avant rééchantillonnage sont préférées ; par contre, pour le recalage des différentes séquences, les images reconstruites – après rééchantillonnage – sont utilisées.

3.2.3 Protocole d'acquisition

La base de données d'images utilisée dans le cadre de cette étude a été réalisée au CHU Pasteur (Nice). Elle porte sur 34 patients et 9 témoins, pour un total de 83 acquisitions. Chacune de ces dernières suit le protocole d'acquisition suivant :

- T1 3D (TE=1.7, TR=8, FA=20, taille des voxels : 1.016*1.016*0.8 mm^3). Sa résolution étant la meilleure de toutes les séquences, il sera utilisé pour une étude du LCR, très sujet aux volumes partiels.
- T2/DP FSE (TE=8/100, TR=5000, ETL=16, taille des voxels 0.937, 0.937, 2 mm^3). Cette séquence fournit deux images recalées de manière intrinsèque. Fondamentale pour l'étude de la SEP, elle sera utilisée pour la segmentation en tissus ainsi que pour raffiner le contourage des lésions.
- T2 FLAIR (TE=150, TR=10000, TI=2200, taille des voxels 0.937*0.937*4 mm^3). Malgré sa faible résolution en z et des artefacts de flux particulière-

ment visibles, le T2 FLAIR est indispensable à la détection des lésions de SEP, dont le contraste est excellent sur cette séquence. Il fournit cependant une sur-segmentation des lésions et ne peut être utilisé seul pour la sclérose en plaques.

3.3 Normalisation spatiale

D'une manière très générale, on peut dire que le recalage permet d'établir une relation géométrique entre des objets représentés dans des images. Les usages du recalage sont multiples en traitement de l'image [156, 157, 65, 134]. Le cadre de notre étude pratique est cependant assez restreint. Les données de départ sont un ensemble d'acquisitions multi-séquences pour plusieurs patients en des instants différents.

Dans notre étude, l'enjeu le plus important est d'aligner deux images du même patient, acquises au même instant mais avec des paramètres d'acquisition différents, et ce afin de pouvoir les utiliser conjointement pour produire une segmentation robuste et fiable. Les méthodes principales utilisées dans le chapitre 4 se placent dans un espace multi-séquences, où chaque voxel de coordonnées fixées représente le même volume dans les deux images. Il est donc indispensable d'avoir une méthode la plus précise possible, puisqu'un écart de plus d'un voxel dans le recalage va terriblement gêner l'analyse des images et la segmentation.

Pour les besoins de l'algorithme de segmentation, nous aurons également besoin d'une labélisation *a priori* des images, c'est-à-dire de la probabilité *a priori* d'appartenir à certaines classes (matière blanche, matière grise, LCR). De telles informations sont contenues dans un atlas statique, qu'il faut recaler sur les images à traiter.

3.3.1 Présentation de la méthode utilisée

Du point de vue algorithmique, il existe une grande diversité de méthodes et de mises en œuvre : il est difficile d'établir un état de l'art exhaustif dans ce domaine

tant il est étudié au niveau de la communauté internationale depuis plusieurs années. On peut citer [20, 90, 66, 101, 123] qui font un résumé de la plupart des méthodes développées. Deux catégories de mise en œuvre conceptuellement très différentes apparaissent : les méthodes géométriques et les méthodes iconiques.

Les méthodes géométriques consistent à rechercher une transformation entre deux sous-ensembles de primitives – appelés aussi amers – extraits dans chaque image : points [158, 157], courbes [154], surfaces [102], repères orientés, etc. Il est possible d'utiliser un cadre stéréotaxique ou des marqueurs qui apparaissent dans chaque image comme des amers géométriques évidents. L'extraction d'amers intrinsèques (appartenant à l'objet) peut être réalisée manuellement par un expert à l'aide du logiciel approprié. On peut enfin chercher à automatiser cette extraction d'amers intrinsèques, essentiellement pour des raisons de reproductibilité, mais le problème majeur consiste alors à rendre cette segmentation suffisamment robuste.

Les méthodes iconiques, qui utilisent les images dans leur totalité, consistent à optimiser un critère de ressemblance, appelé aussi mesure de similarité, fondé sur des comparaisons locales de l'intensité des deux images. Si les algorithmes de recalage géométrique cherchent à mettre en correspondance les objets représentés dans les images, les algorithmes de recalage iconique ont pour but de trouver une transformation entre les images des deux objets. Parmi les méthodes iconiques, on peut encore distinguer deux classes d'approche : les méthodes iconiques pures qui ne réalisent un recalage que par comparaison entre les intensités des points de mêmes coordonnées [31, 156], et les méthodes iconiques mixtes qui cherchent en chaque point quel est le point le plus proche dans l'autre image, au sens d'un critère sur l'intensité, dans un voisinage donné [112]. Il existe également des techniques intermédiaires entre les recalages iconiques et les recalages géométriques [69, 22].

Pour des raisons de robustesse, et parce qu'aucun marqueur ou cadre stéréotaxique n'était présent dans nos images, les méthodes de recalage iconique ont été préférées. La méthode dite de recalage par bloc (*block-matching*) présentée dans [112] a ensuite été choisie, notamment sur un critère de rapidité – quelques minutes par recalage – ainsi que pour sa capacité à gérer le multi-séquences et

les points aberrants, présents via les lésions de sclérose en plaques. L'utilisation exhaustive sur une base de 43 patients n'a posé aucun problème, et le recalage a toujours été suffisant pour pouvoir effectuer le processus de segmentation présenté dans la chaîne de traitements.

3.3.2 Recalage et IRM multi-séquences

L'étude des IRM multi-séquences telle que la présentent les chapitres 4 et 5 de cette thèse est basée sur l'analyse des intensités dans l'espace multi-séquences, où chaque voxel de coordonnées identiques représente le même volume dans toutes les images. Dans une situation pareille, on comprend bien que la qualité du recalage entre les séquences est très importante, et peut avoir une grande influence sur la qualité de la segmentation.

L'histogramme conjoint est un outil très pratique pour visualiser un espace multi-séquences : il associe à chaque couple d'intensités possibles une valeur une proportion de voxels qui peut être interprétée comme une densité de probabilité discrète. Il sera utilisé dans toute la suite de ce manuscrit pour montrer les résultats de l'algorithme EM. Si les images sont bien recalées, dans le cas de deux séquences différentes acquises au même instant sur le même patient, l'histogramme conjoint pourra être interprété en fonction des intensités des différents tissus dans les deux images représentées. Par contre, si les images ne sont pas recalées ou incorrectement recalées, des artefacts vont survenir et rendre l'histogramme conjoint totalement inutilisable (figure 3.5).

Les séquences utilisées sont les IRM T2 / densité de protons (DP), T1 et T2 FLAIR. Les IRM T2 et DP sont intrinsèquement recalées, puisque ce sont deux échos différents de la même séquence. Par contre, l'image T1 n'a même pas la même résolution : dans le protocole d'acquisition utilisé (section 3.2.3), l'épaisseur des coupes est deux fois plus grande en T2/DP qu'en T1. Il faudra donc regarder avec attention l'influence des problèmes de recalage sur le résultat, et particulièrement sur la segmentation en tissus (section 8.1.1).

En outre, tout ceci suppose qu'il existe une transformation rigide entre les différentes modalités, ce qui en pratique n'est pas toujours le cas. Un artefact

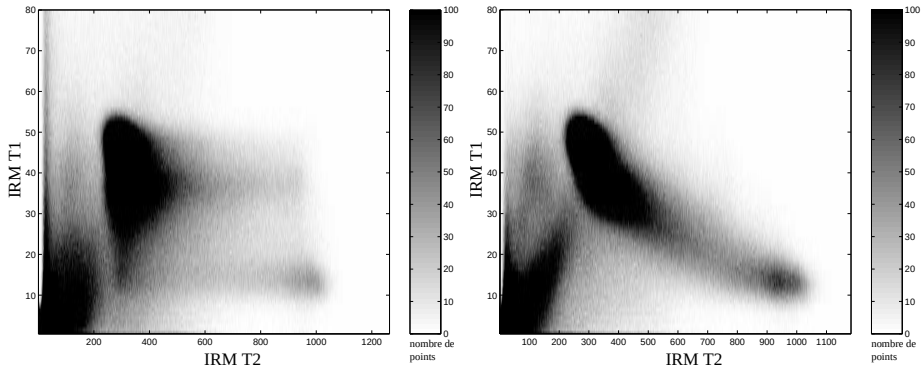


FIG. 3.5 – Histogrammes conjoints de deux séquences IRM (T2 FSE et T1) avant (à gauche) et après (à droite) recalage. Un non-alignement est visible dans l'histogramme conjoint via des bandes horizontales rajoutées, ainsi qu'une perte de contraste de la zone représentant le liquide céphalo-rachidien (LCR) en bas à droite de l'histogramme.

de reconstruction, un mouvement oculaire ou une inclinaison de l'axe Z (voir figure 3.3) dû à une mauvaise lecture des images crée un décalage entre certaines modalités qui nuit beaucoup à tout système de segmentation basé sur un modèle voxellique. C'est pour éviter ces problèmes que la segmentation en tissus sera effectuée sur les deux images T2/DP, qui sont déjà recalées (chapitres 4 et 5). Une extension à d'autres séquence est toujours possible, mais il convient alors de mesurer précisément l'influence des artefacts de recalage sur chacune des étapes du pipeline, ainsi que sur le résultat final. Les images T1 et T2 FLAIR doivent bien sûr recalées sur les images T2/DP, car ces séquences seront très utiles pour la segmentation des lésions (chapitre 6).

3.3.3 Recalage et atlas statistique

Comme nous allons le voir par la suite, le processus de segmentation est basé sur une analyse statistique de l'intensité de chaque voxel des images dans l'espace multi-séquences. Un atlas *a priori* est utilisé : il donne la probabilité *a priori* pour chaque voxel d'appartenir à l'une des classes que l'on cherche à segmenter. L'atlas, mis à disposition par le *Montreal Neurological Institute* [46, 48] et largement

utilisé via le package SPM[3], permet une segmentation *a priori* du parenchyme cérébral en trois classes : matière blanche, matière grise et liquide céphalo-rachidien. D'un point de vue pratique, il a été construit à partir de 305 IRM T1 [45], à partir desquelles une segmentation semi-automatique a été effectuée, en trois classes. Plusieurs images sont fournies : une image du patient moyen[4] dans les modalités T2 et T1, une image des *a priori* pour chaque classe, ainsi que pour le masque du parenchyme cérébral. Pour utiliser correctement les informations fournies par l'atlas statistique, il est nécessaire de placer dans le même repère les images de l'atlas et du patient. Il est préférable de recaler l'image de l'atlas sur le patient, afin de préserver la qualité des IRM originales.

Ce recalage présente des difficultés, par rapport à un recalage classique. D'abord, il doit être de même type que celui qui a servi à la construction de l'atlas. Idéalement, en prenant en compte le fait que tous les algorithmes de recalage ne sont pas équivalents, il doit être exactement le même. Cependant, pour l'atlas statistique du MNI, un recalage affine a été utilisé, ce qui limite le nombre de degrés de liberté et réduit du même coup les différences entre les différents algorithmes possibles. Pour cette raison, nous avons cherché à utiliser la méthode la plus robuste à notre disposition.

Les recalages géométriques sont à exclure dans ce cas précis, car l'atlas statistique n'offre aucun amer géométrique, et il est impossible d'en détecter dans l'image du patient moyen ; le recalage par blocs est une solution viable. Le contour du parenchyme cérébral doit être bien contrasté, pour maximiser les chances de succès du recalage : il est donc préférable de choisir le T2 pour ce recalage, de par le contraste entre le LCR cortical – forte intensité en T2 – et la matière grise. De plus, en T1, la présence d'une couche de gras entre le crâne et les méninges peut altérer grandement le résultat du recalage. La figure 3.6 présente une des erreurs qui a motivé le choix du T2 pour le recalage de l'atlas.

[3] SPM est un package matlab, accessible via l'URL *http ://www.fil.ion.ucl.ac.uk/spm/software/*

[4] La moyenne est ici la moyenne arithmétique des IRM. La construction des *a priori* suit le même processus : c'est la moyenne arithmétique des labélisations binaires.

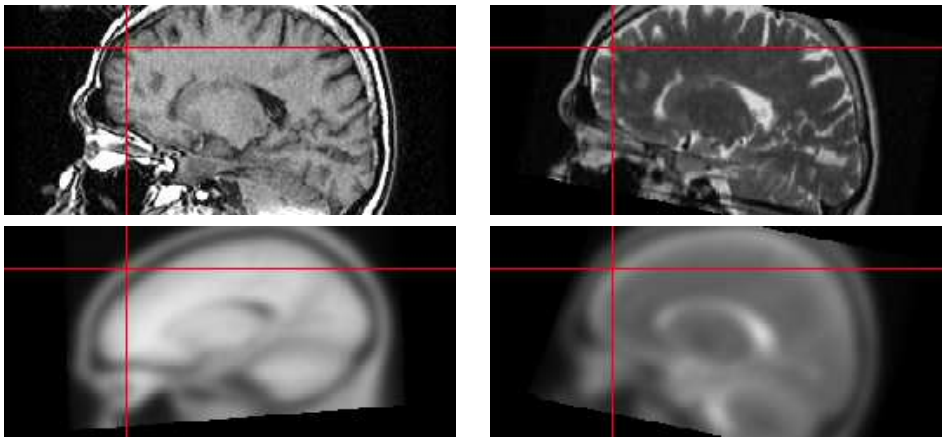


FIG. 3.6 – L'utilisation du T2 est préférable lors du recalage de l'atlas (en bas) sur le patient (en haut) par une méthode basée sur un recalage par blocs. En utilisant le T1 (à gauche), le recalage donne de mauvais résultats environ une fois sur 10, alors qu'en utilisant le T2 (à droite), nous n'avons noté aucun problème sur une base de 83 instants.

3.4 Normalisation en intensité

Dans la section 3.1, les principaux défauts des images IRM ont été présentés. Outre les distortions géométriques que nous ne traiterons pas ici, une conséquence directe de ces problèmes est que deux voxels de composition biologique identique n'ont pas nécessairement la même intensité. Des méthodes sont classiquement utilisées pour limiter ces variations lors de l'acquisition [104]. Elles sont en général capables de corriger une grande partie des défauts propres à l'aimant, mais pas le biais dû au sujet lui-même : l'hypothèse de composition biologique identique au sein d'une même classe est souvent audacieuse, et un biais tissulaire est généralement présent.

Ainsi, de nombreux algorithmes cherchent à corriger *a posteriori* le biais dans les images 3D obtenues par résonance magnétique. Ces algorithmes reposent généralement sur une modélisation de l'image en fonction de l'intensité "réelle" du voxel qui traduit les caractéristiques physiques du tissu. Mathématiquement parlant, étant donné un voxel i de coordonnées v_i dans l'IRM, son intensité y_i est considérée comme étant reliée à l'intensité réelle x_i suivant :

$$y_i = b_i x_i + \epsilon_i^{mes}$$

Le biais b_i est généralement considéré comme une fonction des coordonnées v_i, variant lentement dans le volume, et supposé multiplicatif, en accord avec la nature intrinsèque des phénomènes physiques sous-jacents [152]. Plusieurs méthodes sont disponibles pour éliminer ce biais.

- Le champ de biais peut dans un premier temps être éliminé seul, avec un critère de type entropique [92]. Il est également possible de supprimer le biais dans plusieurs images issues du même appareil, ce qui augmente la robustesse de l'algorithme [81].
- La correction du champ de biais peut être également couplée au processus de segmentation. De nombreuses méthodes choisissent cette voie, car la segmentation est souvent le but final à atteindre : la technique consiste à uniformiser l'intensité IRM observée dans un même tissu, et à alterner segmentation et correction du biais. Le système présenté par Wells, basé sur l'algorithme d'Espérance-Maximisation (EM) [175] a connu un franc succès et a inspiré de nombreux successeurs [165, 1, 127, 128].

Puisque les méthodes proposées dans les chapitres suivants sont basées sur l'algorithme EM, il paraissait naturel de choisir cette dernière technique de correction de biais. VipBiasCorrection, une implémentation de l'algorithme présenté dans [92], est attractif, mais fonctionne mal sur les IRM densité de protons et T2 FLAIR. Enfin, comme nous le verrons, l'introduction du modèle de volume partiel nécessite comme hypothèse de base l'uniformité de l'intensité au sein des classes ne contenant pas de volumes partiels : les méthodes de correction de biais couplées à une segmentation tendent vers cette uniformité, et leur choix parait d'autant plus adapté. Dans la pratique, la méthode que nous utilisons est celle de [129].

Extraction du biais en IRM T2 FLAIR : l'étape de segmentation en tissus de la chaîne de traitements (chapitre 5) fournit une segmentation plus fine que [175]. La faible résolution de l'image et la faiblesse du contraste matière blanche / matière grise pousse à utiliser cette segmentation pour calculer un biais spatial

spécifique en IRM T2 FLAIR, *ie* un biais coupe à coupe pour pallier aux problèmes posés par l'artefact osseux. Comme la segmentation en tissus est fixée, ce biais peut se calculer en une seule étape : il suffit de minimiser les variations basse-fréquence de l'intensité au sein de chacune des 3 classes – matière blanche, matière grise, LCR. Nous verrons dans le chapitre 6 que cette correction du biais a une très grande importance sur la qualité de la détection des lésions de SEP.

3.5 Présentation de la chaîne de prétraitements

Les différents prétraitements présentés ci-dessus permettent de placer les images (patient et atlas) dans un unique espace, où le positionnement des tissus et leur signal IRM sont uniformisés :

- **Décodage des images** : la lecture attentive des en-têtes DICOM permet de reconstruire les images dans un repère orthogonal direct et de fixer leur orientation (section 3.2). Pour ce faire, chaque coupe doit être correctement placée dans le repère IRM en utilisant l'orientation du plan de coupe (marqueurs DICOM *Image Position* et *Image Orientation)*.
- **Recalage multi-séquences intra-patient** : à coordonnées égales, chaque voxel des différentes séquences doit représenter le même volume dans le repère IRM (Section 3.3.2). Tout algorithme de recalage rigide fonctionne ; une attention particulière doit être portée à la limitation du nombre de rééchantillonnages appliqués, qui altèrent la qualité des images. Dans notre cas, les images T2/DP du patient ne sont pas rééchantillonnées, et les IRM T2 FLAIR et T1 ne sont rééchantillonnées qu'une fois.
- **Recalage atlas statistique / patient** : il permet d'obtenir un *a priori* sur la localisation des différents tissus et est un élément clé de la robustesse de la chaîne de traitements. Le choix de l'algorithme de recalage rigide / affine est difficile, car la qualité des images à mettre en correspondance varie énormément. Les méthodes basées sur des appariements par bloc donnent de bons résultats.
- **Extraction des inhomogénéités spatiales** : la correction de biais utilisée

maximise l'uniformité du signal au sein d'une même classe : matière blanche, matière grise, LCR.

Les lésions de sclérose en plaques ont un signal très variable : le type de maladie (rémittente ou secondaire progressive), la localisation de la lésion (périventriculaire, juxtacorticale, corticale) ou le type de lésion (œdémateuse, nécrotique, trou noir) peuvent influencer grandement le processus de détection de la lésion. Vu les données du problème, il paraît indispensable d'avoir un bon *a priori* sur l'environnement des lésions, à savoir les tissus sains. De plus, toute caractérisation des tissus sains permettra d'obtenir, par simple exclusion, une pré-segmentation des lésions. C'est le but des chapitres 4 et 5.

Chapitre 4

Segmentation en tissus

Suite aux différents prétrairements présentés dans le chapitre précédent, l'ensemble des images a donc été placé dans un repère spatial unique, à savoir le repère de l'atlas statistique pour permettre l'utilisation d'un *a priori* sur les labélisation. Le problème majeur est donc maintenant de segmenter les images, et ce afin d'obtenir une cartographie des lésions éventuelles de matière blanche. Il ne faut pas perdre cet objectif de vue : même si la segmentation en tissus que ce chapitre va développer est une étape cruciale dans le processus de détection des lésions, elle n'est pas un but en soi. Dans le choix de la technique, il faudra donc un modèle particulièrement robuste à la qualité des images et à la présence de points aberrants – les lésions de SEP – mais aussi un modèle qui fournit des informations directement exploitables dans la détection et le contourage des lésions.

4.1 Choix de la méthode

Le but principal du processus de segmentation est le partitionnement d'une image en régions – également appelées classes – homogènes du point de vue d'un certain nombre de caractéristiques ou critères. En imagerie médicale, la segmentation est très importante, que ce soit pour l'extraction de paramètres ou de mesures sur les images, ainsi que pour la représentation et la visualisation. Beaucoup d'applications nécessitent une classification des images en différentes régions anato-

miques : graisse, os, muscles, vaisseaux sanguins, alors que d'autres recherchent une détection des zones pathologiques, comme les tumeurs, ou des lésions cérébrales de matière blanche. De nombreux articles et livres [155, 10, 120] font le résumé des différentes techniques de segmentation disponibles pour l'imagerie médicale ; ce paragraphe ne se veut pas la synthèse de toute cette littérature.

Les techniques de segmentation peuvent être divisées en classes de multiples manières : la plus commune est de séparer les méthodes de segmentation de régions qui recherchent un critère d'homogénéité au sein d'une zone, et les méthodes de détection de contours, qui se concentrent sur l'interface entre ces régions. L'arrivée d'outils mathématiques et la diversité des problèmes posés a cependant souvent brisé cette dichotomie : nous parlerons d'abord des méthodes classiques qui ont initié la segmentation, et poursuivrons par les différents développements plus récents : les approches statistiques, la logique floue, les modèles déformables.

4.1.1 Méthodes classiques

Parmi les méthodes classiques, le seuillage est souvent la plus simple, et rarement la plus efficace. De nombreuses techniques de seuillage ont été développées [72, 153] : certaines sont basées sur une analyse d'histogramme, d'autres sur des propriétés locales, comme la moyenne ou l'écart type. Si le seuillage permet de binariser un résultat final, ce qui est toujours nécessaire, il n'est bien souvent pas suffisant pour obtenir une segmentation acceptable, bien que de récentes applications utilisent un seuillage adaptatif local, notamment sur des images aux structures floues et peu visibles, telle des mammographies [87]. Le seuillage local est également utilisé comme étape finale, car il permet un traitement adaptatif du résultat d'une suite de filtres [70].

Toujours dans les méthodes classiques, la croissance de région part d'un pixel de l'image, appelé graine, qui appartient à la structure à segmenter, et recherche à partir de là des pixels d'intensité similaires, en suivant un critère d'homogénéité au sein de la structure à segmenter. Ces techniques ont trouvé leur utilité pour des images cardiaques [151], ou angiographiques [71] ; leur robustesse laisse cepen-

dant à désirer, et on leur préfère un usage dans le cadre des Interfaces Homme / Machine et des segmentations semi-automatiques [109]. La méthode dite des lignes de partage des eaux (*watershed*) [15] est également une méthode de segmentation de région très intéressante, car elle permet d'introduire une forme de détection de contour dans la segmentation de région en utilisant le gradient de l'image à segmenter comme représentation topographique [137]. Très dépendante de l'initialisation, elle convient mal pour les images fortement bruitées : le passage au gradient augmente encore ce bruit et rend le résultat difficilement exploitable. Les lignes de partage des eaux sont cependant très utilisées dans le cadre des segmentations interactives, sur des structures bien contrastées [83] ou dans un système semi-automatique [25].

4.1.2 Logique floue, ensembles statistiques

La théorie des ensembles flous n'est que la généralisation de la théorie des ensembles présentée par Zadeh en 1965 comme une expression mathématique de l'incertitude [178]. Sans entrer dans les détails, elle consiste à définir une mesure floue par une mesure de probabilité Π appliquée sur une partie Y d'un ensemble X ($\Pi(\emptyset) = 0$, $\Pi(X) = 1$) , représentable par une distribution de possibilité sur l'ensemble X. Ceci permet de travailler sur un espace de caractéristiques, comme des coefficients d'ondelettes [12], et de faire plus facilement de la fusion de données, puisque toutes les informations sont représentées par des distributions de possibilité. Pour obtenir un résultat final, il faut appliquer aux distributions un algorithme de partitionnement. La solution la plus simple est un algorithme dérivé du *fuzzy c-means*, qui fournit un partitionnement non-supervisé, puisque la formation des différentes classes n'est pas guidée par un expert. Cet algorithme, célèbre par sa simplicité, est souvent utilisé comme première étape à un processus de segmentation [7], comme base à un algorithme plus complexe [119, 118], ou à titre de comparaison [177].

La littérature est riche quand à l'utilisation des techniques dérivées des ensembles flous appliquées à la segmentation [16, 67, 163]. Une forme de recherche de graphe associée à un système basé sur la logique floue conduit à un système de

segmentation perfectionné [163] qui a mené à une étude sur une très grande base de données appliquée à la sclérose en plaques [162] dans le cadre d'un test clinique, ce qui souligne la fonctionnalité de la méthode. Des critères flous peuvent être facilement introduits dans un processus de segmentation différent, à base de modèles déformables, voir par exemple [33]. Enfin, on ne compte plus le nombre de technique d'extraction des inhomogénéités dans les IRM à base de *fuzzy c-means* [1, 86, 144].

D'un point de vue statistique, l'algorithme de *fuzzy c-means* peut être vu comme un partitionnement flou basé sur l'écart à la moyenne. Dans [68], Hathaway présente l'algorithme d'Espérance Maximisation (EM) comme une minimisation alternée du même critère que celui utilisé dans *fuzzy c-means* auquel on ajoute une pénalité sous un terme entropique. Désormais, l'estimation des paramètres du modèle (moyenne, matrice de covariance, probabilité *a priori*) fait partie intégrante de l'algorithme, et les extensions sont multiples. Même si d'autres études l'ont précédé [58, 85], Wells a été le premier à intégrer une correction des inhomogénéités des IRM à partir des bases mathématiques posées par Dempster [39]. De nombreux travaux ont alors suivi, notamment sur l'utilisation des champs de Markov pour ajouter des contraintes spatiales [141, 164], intégrer des modèles de volume partiel [60, 40] ou segmenter des lésions de tout type [103, 43, 167]. De récentes études ont également montré qu'il était possible de coupler un système de segmentation basé sur l'EM avec un algorithme de recalage non-rigide, pour spécifier plus précisément un *a priori* et intégrer une labélisation plus précise de structures anatomiques [124]. Facile à implémenter et mathématiquement solide, l'EM est une base intéressante pour la segmentation, et en particulier la segmentation de tissus cérébraux.

4.1.3 Modèles déformables

Les modèles déformables introduits par [77], encore appelés *snakes*, contours actifs ou ballons, ont connu de nombreuses applications en segmentation, mais aussi en animation, suivi de contour ou modélisation géométrique. Ils sont en effet généralement plus robustes au bruit et aux éventuelles discontinuités dans les

contours de l'image. Ils permettent en outre une interaction relativement aisée avec l'utilisateur ainsi que l'introduction de connaissances *a priori* concernant la forme de l'objet recherché. Enfin, on peut obtenir en théorie une segmentation sub-voxellique de l'image. Deux grandes familles de modèles déformables existent : les modèles paramétriques [77, 56] et les les modèles géométriques [91, 24].

Les premiers sont les plus anciens et nécessitent une représentation paramétrique ou discrète. Les seconds, fondés sur la théorie d'évolution des courbes et la méthode des ensembles de niveau (*level sets*) introduite par Osher et Sethian [111, 145, 146], utilisent une représentation implicite du modèle et permettent des changements de topologie. Pitiot nous donne une très bonne description des paradigmes associés aux modèles déformables [121]. Cependant, si de récentes études montrent que les modèles déformables et les ensembles de niveau sont parfaitement utilisables dans le cadre de la segmentation de tissus cérébraux [9, 114, 138, 33], leur mise en œuvre reste difficile et les temps de calcul sont souvent très longs. De plus, la variabilité des tissus cérébraux, notamment des replis du cortex, est très difficile à intégrer dans un modèle de ce type, et est un sujet d'étude à part entière. Quand à la segmentation des lésions de SEP, qui reste notre but final, elle est difficile par ce biais, car il n'existe pas d'*a priori* suffisamment précis sur la forme des différentes lésions. Il est cependant possible d'introduire ce genre de contrainte dans un système semi-automatique, ou en intégrant la forme des structures environnantes [122], mais le processus est très lourd à mettre en œuvre.

4.1.4 Discussion

Rappelons brièvement les données du problème final, la segmentation des lésions de sclérose en plaques en IRM multi-séquences. Ces lésions sont présentes essentiellement dans la matière blanche, et sont couramment définies dans la littérature médicale comme des hypersignaux de la substance blanche. Mais comme nous le verrons plus en détail dans le chapitre 6, les lésions de SEP n'ont pas une signature fixe : selon leur localisation (à proximité du cortex, près des ventricules,

sous la tente du cervelet), selon le type de maladie observé (rémittente, secondaire progressive), le signal des lésions est difficile à prévoir.

Il est toujours possible d'utiliser une grande base de données de patients et de témoins dans le cadre d'un système expert ou basé sur des réseaux de neurones. De telles applications ont déjà vu le jour, et leur fonctionnalité a été montrée [179, 181]. De telles méthodes, bien que disponibles, sont cependant restreintes à l'utilisation de grandes bases de données pour les besoins d'apprentissage du système, et requièrent une certaine uniformité dans la qualité des images initiales. Nous avons préféré choisir une solution moins restrictive par rapport aux données utilisées.

Des techniques basées sur des modèles déformables existent, mais leur paramétrisation est délicate, les temps de calcul sont très longs et la qualité des images dont nous disposons nous laisse penser que des solutions basées sur l'intensité et travaillant à l'échelle du voxel sont envisageables. L'algorithme d'Espérance Maximisation est souvent pris en exemple pour sa robustesse quand il est contraint par un atlas statistique comme probabilité *a priori*. La preuve de convergence, la possibilité d'y inclure une correction des inhomogénéités de l'image ainsi que des contraintes spatiales en font un algorithme de choix pour la segmentation des tissus cérébraux.

Une faiblesse de l'EM est pourtant la nécessité que l'atlas soit représentatif des patients à segmenter, ce qui n'est pas acquis d'avance. Lors de la segmentation par un EM, par exemple, d'images néonatales comme celles étudiées dans [127], un atlas statistique constitué à partir de patients adultes ne convient pas. Lorsque la variabilité entre les patients est trop importante, d'autres méthodes comme les classificateurs *kNN* supervisés donnent de meilleurs résultats [35]. Cette méthode consiste à comparer chaque voxel à segmenter à l'ensemble d'une base d'apprentissage, et à sélectionner les k éléments de cette base les plus proches pour obtenir une labélisation finale. Ceci permet d'utiliser facilement une base d'apprentissage labélisée, mais sans phase d'apprentissage, comme pour un réseau de neurones par exemple. Cependant, la sclérose en plaques étant une maladie touchant à large majorité les adultes de 20 à 40 ans, l'espace des images à segmenter s'en trouve

fortement restreint, et l'atlas fourni par le MNI convient à notre étude. Dans la suite de ce chapitre, une présentation de l'algorithme EM sera effectuée, ainsi que son application à la segmentation des IRM multi-séquences en tissus cérébraux. Une discussion sera menée pour montrer les faiblesses de l'algorithme qu'il faudra pallier.

4.2 L'algorithme EM

4.2.1 Présentation de l'algorithme

L'algorithme d'Espérance-Maximisation est un algorithme proposé par Dempster dans [39]. Il se situe dans un cadre beaucoup plus général que la segmentation. L'algorithme que nous présentons dans ce paragraphe a été présenté par Wells dans [175] puis développé par Van Leemput dans [165, 166], et concerne plus particulièrement la segmentation d'IRM cérébrales. Les images à segmenter sont un ensemble d'IRM multi-séquences ; le résultat désiré est une labélisation en un nombre K de classes. Dans la version la plus simple de l'algorithme qui est présentée ici, $K = 3$: matière blanche, matière grise, LCR.

La distribution de l'intensité de chaque classe de tissus est approximée par une gaussienne G_{μ_k,Σ_k} de moyenne μ_k et de matrice de covariance Σ_k. La probabilité *a priori* π_i^k pour chaque voxel i d'appartenir à une classe k est disponible sous la forme d'un atlas statistique. Les probabilités *a posteriori* γ_i^k, calculées pour chaque voxel, sont les labélisations recherchées. Les données du problème sont les suivantes :

- pour chaque voxel i, l'intensité des images x_i ;
- pour chaque voxel i, la probabilité *a priori* π_i^k d'appartenir à la classe k.

Les résultats à calculer sont les suivants :

- paramètres des gaussiennes associées à chaque classes (μ_k, Σ_k) ;
- pour chaque voxel i, probabilité *a posteriori* γ_i^k d'appartenir à la classe k.

Si les paramètres de classes sont connus, estimer la labélisation revient à calculer, pour chaque voxel, la probabilité *a posteriori* d'appartenir à chaque classe : ce

calcul est possible par une loi de Bayes (équations 4.1). Si les labélisations sont connues, les paramètres de classes sont, comme nous allons le voir dans la preuve de convergence, directement calculable à partir des labélisations (équation 4.2). L'algorithme alterne en fait entre le calcul des paramètres de classes – étape de Maximisation – et la mise à jour des labélisations – étape d'Espérance – (figure 4.1).

Algorithme :

- Étape préliminaire : initialiser les γ_i^k à la valeur donnée par les π_i^k ;
- (1) calcul des paramètres de classes μ_k, Σ_k :

$$\left.\begin{array}{rcl} \mu_k & = & \dfrac{\sum_{i=1}^{N} \gamma_i^k x_i}{\sum_{i=1}^{N} \gamma_i^k} \\ \Sigma_k & = & \dfrac{\sum_{i=1}^{N} \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^{N} \gamma_i^k} \end{array}\right\} \tag{4.1}$$

- (2) mise à jour des probabilités *a posteriori* γ_i^k

$$\gamma_i^k = \frac{\pi_i^k G_{\mu_k,\Sigma_k}(x_i)}{\sum_{l=1}^{K} \pi_i^l G_{\mu_l,\Sigma_l}(x_i)} \tag{4.2}$$

- Itérer (1) et (2) jusqu'à convergence.

4.2.2 Preuve de convergence

L'algorithme d'Espérance Maximisation (EM) peut être présenté de multiples manières. Hathaway [68] présente ainsi cet algorithme comme une minimisation alternée d'un critère interprété comme la vraisemblance complète de l'image par rapport au modèle de mixture, pénalisée par un terme entropique mesurant l'interaction entre les mixtures. Nous allons reprendre une démonstration analogue disponible dans [54] qui conduit à la présentation de l'EM comme une minimisation alternée de la vraisemblance, à l'aide de l'introduction d'une variable cachée Z.

Soit x_i les données observables du problème, réalisations de la variable aléa-

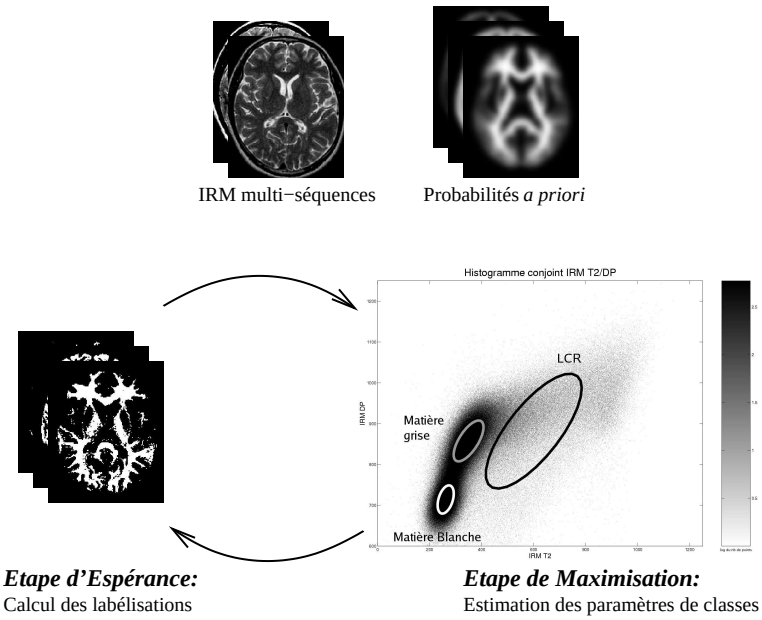


FIG. 4.1 – Schéma simplifié de l'algorithme EM : alternance entre l'étape d'espérance – calcul des probabilités *a posteriori* – et l'étape de maximisation – estimation des paramètres de classes.

toire X. Soit Z un ensemble de variables cachées associées à X. On note Θ l'ensemble des paramètres du modèle utilisé pour l'estimation de la distribution de probabilité de X. D'un point de vue pratique et pour faire le lien entre les applications futures et la théorie présente, les x_i seront les intensités des voxels dans l'espace multi-séquences des images, alors que Θ sera le couple moyenne / matrice de covariance (μ, Σ) qui définit chaque gaussienne associée à la distribution de l'intensité de chaque classe de tissus. Z contiendra la labélisation *a posteriori* des voxels.

Les paramètres peuvent être estimés par un critère basé sur la maximisation de la vraisemblance des paramètres pour un jeu de données. La log-vraisemblance s'écrit comme suit :

$$L(\Theta) = \log p(X|\Theta) \tag{4.3}$$

L'estimation des paramètres Θ par maximisation de la log-vraisemblance est

donc :

$$\hat{\Theta} = \arg\max_{\Theta} L(\Theta) = \arg\max_{\Theta} \log p(X|\Theta) \tag{4.4}$$

Cette formulation de la vraisemblance (équation 4.3) est difficile à maximiser dans le cas général. Toute l'astuce de l'EM consiste extraire une nouvelle formulation de la vraisemblance, plus facile à maximiser à l'aide d'une minimisation alternée. Introduisons donc la variable Z dans l'équation 4.4 en utilisant la loi du calcul des probabilités conditionnelles (équation 4.5) :

$$p(X|\Theta) = \frac{p\big((X,Z)|\Theta\big)}{p\big(Z|(X,\Theta)\big)} \tag{4.5}$$

Soient $\tilde{P}(Z)$, la distribution de probabilité de la variable Z, et $P_\Theta(Z) = p\big(Z|(X,\Theta)\big)$:

$$\begin{aligned} L(\Theta) &= \log \frac{p\big((X,Z)|\Theta\big)}{P_\Theta(Z)} \frac{\tilde{P}(Z)}{\tilde{P}(Z)} \\ &= \log \frac{p\big((X,Z)|\Theta\big)}{\tilde{P}(Z)} + \log \frac{\tilde{P}(Z)}{P_\Theta(Z)} \end{aligned} \tag{4.6}$$

L'équation 4.6 étant valable quel que soit Z, nous prenons l'espérance de l'équation précédente sur la variable Z :

$$\begin{aligned} L(\Theta) &= E_Z\Big(\log \frac{p\big((X,Z)|\Theta\big)}{\tilde{P}(Z)}\Big) + E_Z\Big(\log \frac{\tilde{P}(Z)}{P_\Theta(Z)}\Big) \\ \mathcal{L}(Z,\Theta) &= E_Z\Big(\log \frac{p\big((X,Z)|\Theta\big)}{p(Z)}\Big) \\ L(\Theta) &= \mathcal{L}(Z,\Theta) + KL\big(\tilde{P}(Z)||P_\Theta(Z)\big) \end{aligned} \tag{4.7}$$

Le deuxième terme de l'équation 4.7 est la divergence de Kullback-Leibler entre les distributions de probabilité $\tilde{P}(Z)$ et $P_\Theta(Z)$. Une propriété majeure de

cette divergence est d'être toujours positive, et de s'annuler lorsque les deux distributions sont égales. Le premier terme de l'équation 4.7 $\mathcal{L}(Z,\Theta)$ est donc une borne inférieure de la log-vraisemblance quelle que soit la distribution de Z, et on a $L(\Theta) = \mathcal{L}(Z,\Theta)$ quand $\tilde{P}(Z) = P_\Theta(Z)$. Maximiser $L(\Theta)$ par rapport à Θ est donc équivalent à maximiser $\mathcal{L}(Z,\Theta)$ quand $\tilde{P}(Z) = P_\Theta(Z)$: c'est le principe de l'algorithme EM [39, 74, 17], qui est en fait une minimisation alternée :

1. **Initialisation** des paramètres $\Theta^{(0)}$ à l'itération 0.
2. **Étape d'Espérance**, calcul de Z : $\tilde{P}(Z^{(t+1)}) = P_{\Theta^{(t)}}(Z)$, en fonction des paramètres $\Theta^{(t)}$ calculés à l'itération précédente.
3. **Étape de Maximisation**, calcul de Θ : $\Theta^{(t+1)} = \arg\max_\Theta \mathcal{L}(Z^{(t+1)},\Theta)$, en utilisant la variable $Z^{(t+1)}$ calculée à l'étape précédente.
4. **Itérer 2 et 3** jusqu'à convergence.

On itère ainsi les étapes 2 et 3 jusqu'à ce que le critère de convergence soit atteint : généralement, ce critère est un critère de variation sur la log-vraisemblance : $|L(X|\Theta^{(t+1)}) - L(X|\Theta^{(t)})| \leq \epsilon$.

Dans l'étape d'Espérance, à l'itération $t+1$, les propriétés de la divergence de Kullback-Leibler permettent le calcul de la distribution de probabilité de Z.

$$\tilde{P}(Z^{(t+1)}) = p\big(Z|(X,\Theta^{(t)})\big)$$

Dans l'étape de Maximisation, $\mathcal{L}(Z^{(t+1)},\Theta)$ est maximisé par rapport à Θ (cf équation 4.7).

$$\begin{aligned}\mathcal{L}(Z^{(t+1)},\Theta) &= \sum_Z \tilde{P}(Z^{(t+1)}) \log \frac{p\big((X,Z^{(t+1)})|\Theta\big)}{\tilde{P}(Z^{(t+1)})} \\ &= \sum_Z p\big(Z^{(t+1)}|(X,\Theta^{(t)})\big) \log p\big((X,Z^{(t+1)})|\Theta\big) \\ &\quad + H(Z^{(t+1)}) \end{aligned} \tag{4.8}$$

Dans l'équation précédente 4.8, $H(Z)$ est l'entropie de la variable Z, et ne dépend pas de Θ. La maximisation de $\mathcal{L}(Z,\Theta)$ par rapport à Θ peut être réécrite

comme la maximisation d'une fonctionnelle appelée Q-fonction :

$$\begin{aligned} Q^{(t+1)}(\Theta) &= \sum_{Z} p(Z^{(t+1)}|(X,\Theta^{(t)}))\log p\big((X,Z^{(t+1)})|\Theta\big) \qquad (4.9)\\ Q^{(t+1)}(\Theta) &= E_{Z|X}\Big[\log p\big((X,Z^{(t+1)})|\Theta\big)\Big] \end{aligned}$$

Q est ici la log-vraisemblance complète moyenne, c'est-à-dire la vraisemblance des données X si la variable Z était observée et fixée par rapport à X. De plus, Q peut être plus facilement maximisée par rapport à Θ que la fonctionnelle initiale de la log-vraisemblance, et maximiser $Q(\Theta)$ conduit à la maximisation du critère initial $L(\Theta)$ pour Z bien choisi. La maximisation de $Q(\Theta)$ n'est pas toujours facile à mettre en œuvre. Fort heureusement, il n'est pas strictement nécessaire de maximiser la borne inférieure par rapport à Θ : tout accroissement sur Q suffit pour la convergence. Cet algorithme s'appelle l'EM Généralisé [39].

4.2.3 Application aux modèles de mixture de gaussiennes et à la segmentation des IRM cérébrales

Dans un modèle de mixture [99], la distribution de probabilité de $X|\Theta$ est définie comme une somme pondérée de K fonctions paramétriques, ayant pour ensemble de paramètres Θ :

$$\begin{aligned} p(X|\Theta) &= \sum_{k=1}^{K} p\big((X,Z=k)|\Theta\big)\\ &= \sum_{k=1}^{K} p(Z=k)\cdot p\big(X|(Z=k,\Theta)\big)\\ &= \sum_{k=1}^{K} \pi_k p\big(X|(Z=k,\Theta)\big) \end{aligned}$$

Les $\{\pi_k\}$ sont les proportions de chaque classe : $\sum_{k=1}^{K} \pi_k = 1$. Ces paramètres sont inconnus et K, le nombre de composants, est un hyperparamètre à ajuster.

Dans le cas particulier des mixtures de gaussiennes, $p\big(X|(Z=k,\Theta)\big)$ est une combinaison de K gaussiennes de paramètres $\{\mu_k, \Sigma_k\}$:

$$\begin{aligned} p\big(X=x|(Z=k,\Theta)\big) &= G_{\mu_k,\Sigma_k}(x) \\ &= \frac{1}{(2\pi)^{d/2}|\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)} \end{aligned}$$

d est ici la dimension des données X. Avec un tel modèle, les paramètres sont :

$$\Theta = \{\pi_1, \ldots, \pi_{K-1}, \mu_1, \ldots, \mu_K, \Sigma_1, \ldots, \Sigma_K\}$$

L'algorithme EM fournit ainsi une solution très élégante à l'estimation des paramètres Θ en fonction des données x. En effet, dans le cas d'un modèle de mixture de gaussiennes, la variable cachée Z devient alors l'appartenance de chaque échantillon à l'une des K classes. Si cette information était connue, c'est-à-dire si les données étaient labélisées, l'estimation serait facile.

Soit $\gamma_i^k = p(Z_i = k|X = x_i, \theta_k)$ la probabilité *a posteriori* d'obtenir une labélisation $Z_i = k$ en fonction des paramètres θ_k de la classe k, et de l'échantillon x_i. Les γ_i^k sont donc la labélisation *a posteriori* des données. L'étape d'espérance s'obtient facilement en utilisant une loi de Bayes :

$$\begin{aligned} \gamma_i^k &= \frac{p(Z_i=k)p(X=x_i|Z_i=k;\theta_k)}{p(X=x_i;\theta_K)} \\ &= \frac{\pi_k G_{\mu_k,\Sigma_k}(x_i)}{\sum_{l=1}^{K}\pi_l G_{\mu_l,\Sigma_l}(x_i)} \end{aligned} \tag{4.10}$$

Le calcul de l'étape de Maximisation consiste en la dérivation de la fonctionnelle Q définie par l'équation 4.9. Nous passons sous silence les calculs disponibles dans [54] et ne donnons que les résultats, pour une estimation complète de

la matrice de covariance :

$$\left.\begin{array}{rcl} \pi_k & = & \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\gamma_i^k \\ \mu_k & = & \dfrac{\sum_{i=1}^{N}\gamma_i^k x_i}{\sum_{i=1}^{N}\gamma_i^k} \\ \Sigma_k & = & \dfrac{\sum_{i=1}^{N}\gamma_i^k (x_i-\mu_k)(x_i-\mu_k)^T}{\sum_{i=1}^{N}\gamma_i^k} \end{array}\right\} \quad (4.11)$$

En rajoutant une contrainte sur l'EM, à savoir $\forall k,\ \Sigma_k = \sigma^2 I_d$, l'algorithme EM peut être comparé à l'algorithme de partitionnement *fuzzy c-means* [16]. Par contre, les labélisations γ_i^k ne sont pas des labélisations floues ; les proportions des γ_i^k des différentes classes pures n'ont *a priori* rien à voir avec la proportion en termes de volume de chacun des tissus purs dans le voxel en question. Ceci nous a poussé à introduire un modèle de volume partiel spécifique. Des détails seront fournis au chapitre 5.

Avec les équations 4.11, il devient très facile de construire un algorithme basique de segmentation d'IRM. En effet, le bruit IRM suit une distribution de probabilité Ricienne [78], et il est possible de pré-filtrer les IRM en suivant un tel modèle [108], mais le modèle de bruit Gaussien demeure une bonne approximation du modèle Ricien quand le rapport signal sur bruit est suffisant, typiquement RSB > 3 [149]. Il suffit donc de choisir un certain nombre K de tissus à segmenter dans l'image et de prendre les intensités des voxels x_i comme données[1]. Après une initiation des γ_i^k qui reste à définir, il suffit d'alterner les étapes de Maximisation (équations 4.11) et d'Espérance (équation 4.10).

- Étape préliminaire : initialiser les labélisations des images (aléatoirement, ou avec un atlas statistique)
- (1) Étape de Maximisation : calcul des paramètres de classes μ_k, Σ_k et π_k
- (2) Étape d'Espérance : mise à jour des labélisations γ_i^k
- Itérer (1) et (2) jusqu'à convergence.

[1] Il est possible d'intégrer les coordonnées des voxels dans les données, comme décrit dans [142], ce qui permet une régularisation spatiale ; comme nous nous sommes concentrés sur l'obtention des paramètres de classes, garder uniquement l'intensité des voxels comme donnée semblait plus raisonnable.

Si l'atlas statistique est utilisé comme probabilité *a priori*, l'algorithme change peu : π_k devient π_i^k puisque la probabilité *a priori* dépend de la position du voxel, et sa réestimation n'est plus nécessaire, car sa valeur est fixée. L'algorithme est alors identique à celui présenté dans la sous-section 4.2.1 (équations 4.1 et 4.2). La labélisation des classes devient prévisible, puisqu'elle correspond au label fourni par l'atlas, et la convergence est grandement facilitée ; par contre, le résultat dépend assez fortement de l'atlas. Il faut donc faire attention que ce dernier soit représentatif des images à traiter.

Deux applications majeures sont utilisées dans le cadre de cette thèse : la segmentation du masque du cerveau, et la segmentation du parenchyme cérébral en tissus : matière grise, matière blanche, liquide céphalo-rachidien (LCR).

4.3 Segmentation du masque du cerveau en IRM T2/DP

La segmentation du masque du parenchyme cérébral par un algorithme EM sur des IRM n'est pas une nouveauté : Kapur a proposé une méthode similaire en 1996 [76], et Wells avant elle [175]. D'autres méthodes basées sur l'analyse d'histogrammes sont également disponibles, telle la méthode de J. F. Mangin [94, 93] utilisable via le logiciel Brainvisa [36, 136, 135, 95, 93]. Une utilisation des modèles déformables permet encore d'améliorer la qualité des résultats [8]. Dans le cadre de notre étude, il faut cependant préciser que la segmentation du masque du cerveau est une des premières étapes du processus de segmentation : il est donc indispensable d'avoir une assurance sur la qualité du résultat, même si d'autres méthodes permettent d'obtenir une segmentation plus fine.

La méthode que nous présentons dans ce paragraphe est très similaire à [76]. Elle est en fait un algorithme EM à 4 classes - matière blanche, matière grise, LCR, autre - appliquée sur l'intégralité de l'image double-écho T2/DP. De par la présence d'un atlas statistique pour les 4 classes, la convergence est rapide, et démontrée puisque dans ce cas, nous ne sortons pas du cadre original de l'EM. L'atlas du MNI, disponible via le module SPM de Matlab, fournit les probabilités

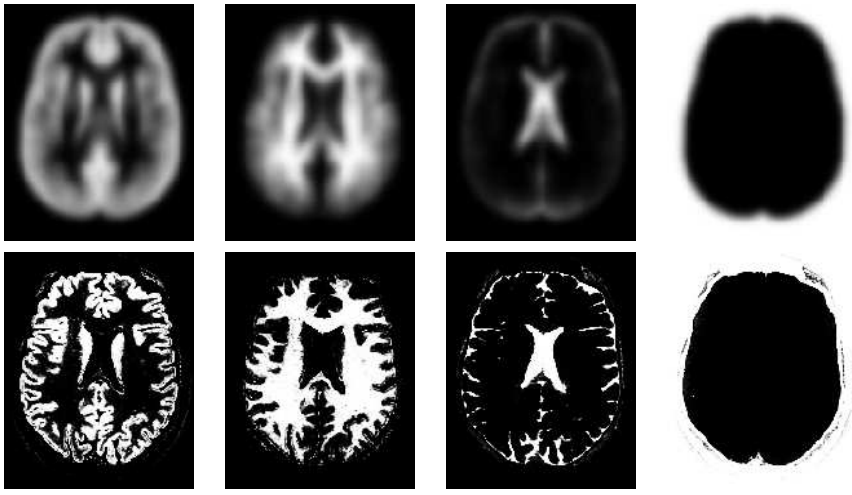


FIG. 4.2 – Segmentation du masque du parenchyme cérébral à partir des IRM T2 / DP. L'atlas du MNI fournit les *a priori*. De gauche à droite, les labélisations *a priori* (haut) et *a posteriori* (bas) sont représentées : matière grise, matière blanche, LCR, hors cerveau. La présence d'artefacts hors du cerveau lui-même, et une sur-segmentation du LCR qui sera constatée plus loin, empêche ces segmentations d'être utilisées directement pour segmenter les lésions de SEP.

a priori d'appartenance à 4 classes, en plus d'une image T2 d'un patient moyen. L'opération de segmentation consiste donc simplement à passer d'une probabilité *a priori* à une probabilité *a posteriori* (figure 4.2). Comme on peut le voir en comparant les images originales (IRM T2 et DP) et le masque obtenu, des opérations de morphologie mathématique simples sont suffisantes pour obtenir un masque du parenchyme cérébral (figure 4.3).

Ces opérations doivent cependant être menées avec précaution. Deux problèmes sont constatés : une sur-segmentation du masque dans les zones grasses entre le parenchyme et le crâne, et une sous-segmentation du masque. Le premier effet vient simplement du fait que la signature en intensité du parenchyme cérébral ne lui est pas spécifique. La présence d'un atlas statistique nous assure une certaine qualité du résultat, mais rien ne garantit que la segmentation du cerveau sera propre après un simple EM.

La sous-segmentation a, quant à elle, une origine plus subtile. L'atlas disponible via le logiciel, fourni par le MNI, a en effet été obtenu à partir de segmen-

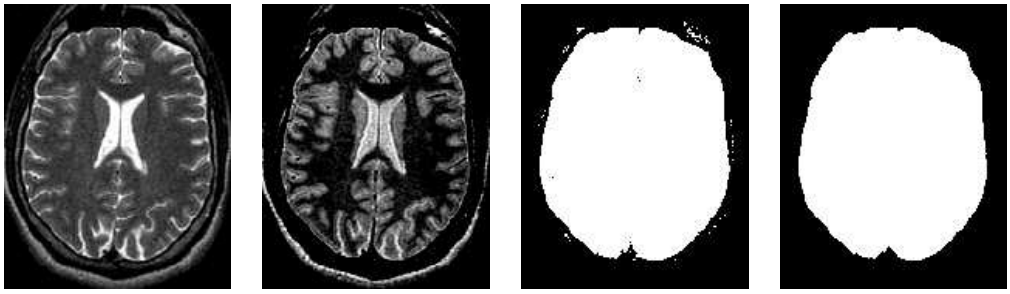


FIG. 4.3 – De gauche à droite : IRM T2 et DP, masque obtenu après utilisation de l'EM, avant et après morphologie mathématique.

tation d'un certain nombre d'IRM en matière blanche, matière grise et LCR : d'autres artefacts - lésions, vaisseaux notamment - n'ont pas été labélisés. Ceci a pour conséquence qu'au sein du parenchyme cérébral, la somme des probabilités a *priori* pour la matière blanche, la matière grise et le LCR n'est pas égale à 1 au sein du parenchyme cérébral, ce qui explique la sous-segmentation du masque. Corriger l'atlas n'est pas possible, car forcer la somme des probabilité à 1 sur le bord de l'atlas n'a pas de sens. Les opérations à mener sont donc les suivantes (en connexité 26) :

- érosion, extraction de la plus grande composante connexe, dilatation,
- dilatation, suppression des trous, érosion.

La conjonction de ces deux étapes donne de bons résultats pour l'ensemble des images de la base de travail. Le masque est précis, même sur les coupes du cervelet. Il peut présenter certaines irrégularités, auxquelles il faudra faire attention dans le modèle de segmentation en tissus. Un tel masque nous permet cependant de poser ces irrégularités comme des points aberrants, ce qui facilite leur traitement dans un processus statistique comme celui de l'EM.

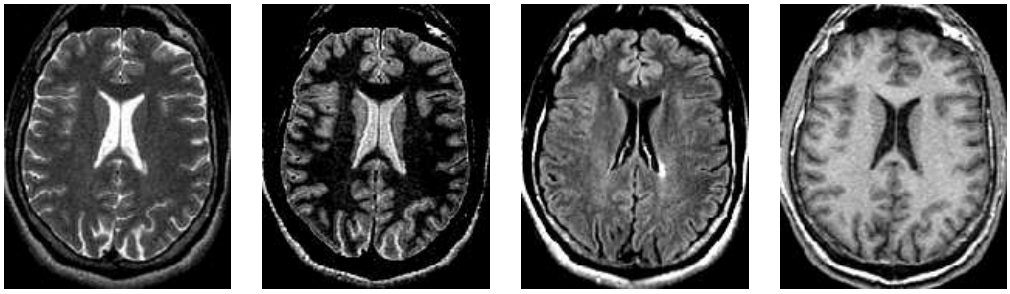


FIG. 4.4 – Présentation des différentes séquences disponibles pour la chaîne de traitements (de gauche à droite) : vue axiale des IRM T2, DP, T2 FLAIR, T1. La modalité préférée pour la segmentation des lésions de SEP est le couple T2/DP : c'est donc sur elle que portera l'analyse lors de la segmentation en tissus et de la présentation du modèle de volume partiel. Les autres modalités seront utilisées *a posteriori.*

4.4 Segmentation en tissus.

4.4.1 Algorithme

Le but de cette segmentation en tissus sains est leur caractérisation, afin de pouvoir construire un premier processus de détection des lésions de SEP. Même si les 4 séquences (T2 FSE / DP, T1, T2 FLAIR) sont disponibles, la segmentation ne sera menée dans un premier temps que sur le couple T2/DP, alors que les autres séquences – T1, T2 FLAIR – montreront tout leur intérêt pour la détection des lésions et leur spécification : ceci permet de n'effectuer la segmentation que sur des images vierges de tout rééchantillonnage. En outre, nous ferons délibérément l'impasse sur certains processus décrits dans la littérature pour améliorer la qualité visuelle de la segmentation – les contraintes sur le voisinage, par exemple – puisque notre but est d'obtenir un critère sur l'intensité qui sera raffiné ensuite par contraintes spatiales.

Dans sa formulation la plus simple, le processus de segmentation prend un ensemble multi-séquences à segmenter – les deux images T2 / densité de protons, auxquelles le masque binaire du cerveau a été appliqué – et fournit en sortie 3 labélisations : matière blanche, matière grise, LCR. Comme indiqué dans les

équations 4.10 et 4.11, l'algorithme EM donne deux grands résultats :

- la labélisation des segmentations via les γ_i^k
- l'estimation des paramètres du modèle μ_k et Σ_k.

4.4.2 Présentation des résultats

Pour éliminer dans un premier temps l'influence des lésions de SEP, un témoin a été choisi pour cet exemple, et les résultats sont présentés dans la figure 4.5. Pour la segmentation du masque du cerveau, seule la segmentation était intéressante. Une mauvaise estimation des paramètres de classes, ou un mauvais modèle était sans importance tant que la segmentation était valide. Par contre, dans le cas de la segmentation en tissus, avoir une bonne estimation des paramètres est primordial, car c'est à partir de là que les lésions de SEP vont être segmentées. Pour observer la qualité de la segmentation dans la figure 4.5, il faut donc en permanence regarder deux espaces reliés.

- L'espace des intensités : visualisé via l'histogramme conjoint entre les deux modalités, il est représenté par les paramètres de classes μ_k et Σ_k. Une méthode pour visualiser ces paramètres est de tracer l'estimateur de confiance donné par $p(Z = z) = cte$. Dans un espace 2D, cet estimateur est en fait l'ellipse de Mahalanobis donné par $\frac{1}{2}(X - \mu_k)^T \Sigma_k^{-1}(X - \mu_k) = \lambda$. Dans toutes les figures similaires à la figure 4.5, les ellipses vérifieront $\lambda = 1$.
  Pour avoir une meilleur visibilité de l'histogramme, celui-ci est en fait un log-histogramme : la couleur de chaque pixel de l'image correspond au log du nombre d'occurences du couple (I_{T2}, I_{DP}) correspondant dans le couple d'image IRM T2 / IRM DP. Ceci permet essentiellement de visualiser sur le même histogramme le LCR avec le reste des tissus, bien que le nombre de voxels correspondant au LCR soit très faible.
- L'espace des images, dans lequel les segmentations sont effectivement visualisées. Dans la figure 4.6, les images représentant des segmentations sont en fait les labélisations γ_i^k dont les valeurs sont comprises entre 0 (noir) et 1 (blanc).

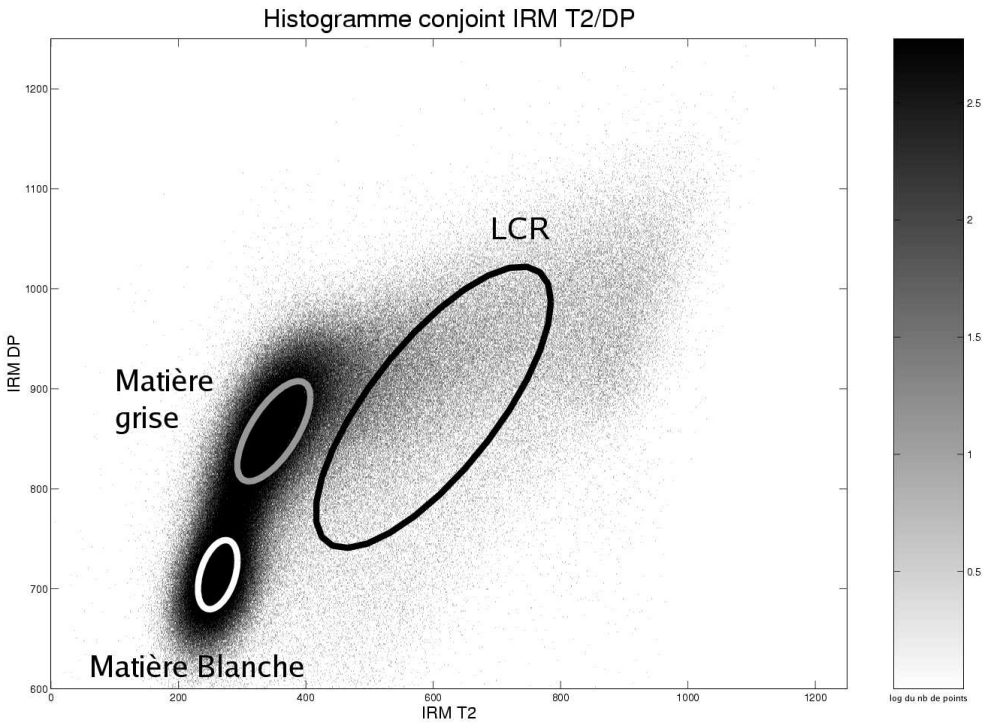


FIG. 4.5 – **Résultat de la segmentation par un algorithme EM à 3 classes** (matière blanche, matière grise, LCR) du couple IRM T2/DP, sur lequel on a appliqué un masque du parenchyme cérébral : visualisation de l'espace des intensités. La variance de la classe LCR est très grande par rapport à celle de la classe matière blanche : le modèle semble inadapté, car ces variances devraient être de taille comparable.

Si les images suivent le modèle de bruit Ricien – qui, rappelons le, est approximé dans cette étude par un modèle gaussien (cf partie 4.2.3) – les différentes classes devraient présenter les mêmes variances, sous réserve d'uniformité du signal au sein de la classe. Or, si cela est globalement vrai pour la matière blanche et la matière grise, on observe une sur-segmentation du LCR, essentiellement au niveau des sillons corticaux, et une surestimation flagrante de la variance de la classe correspondante.

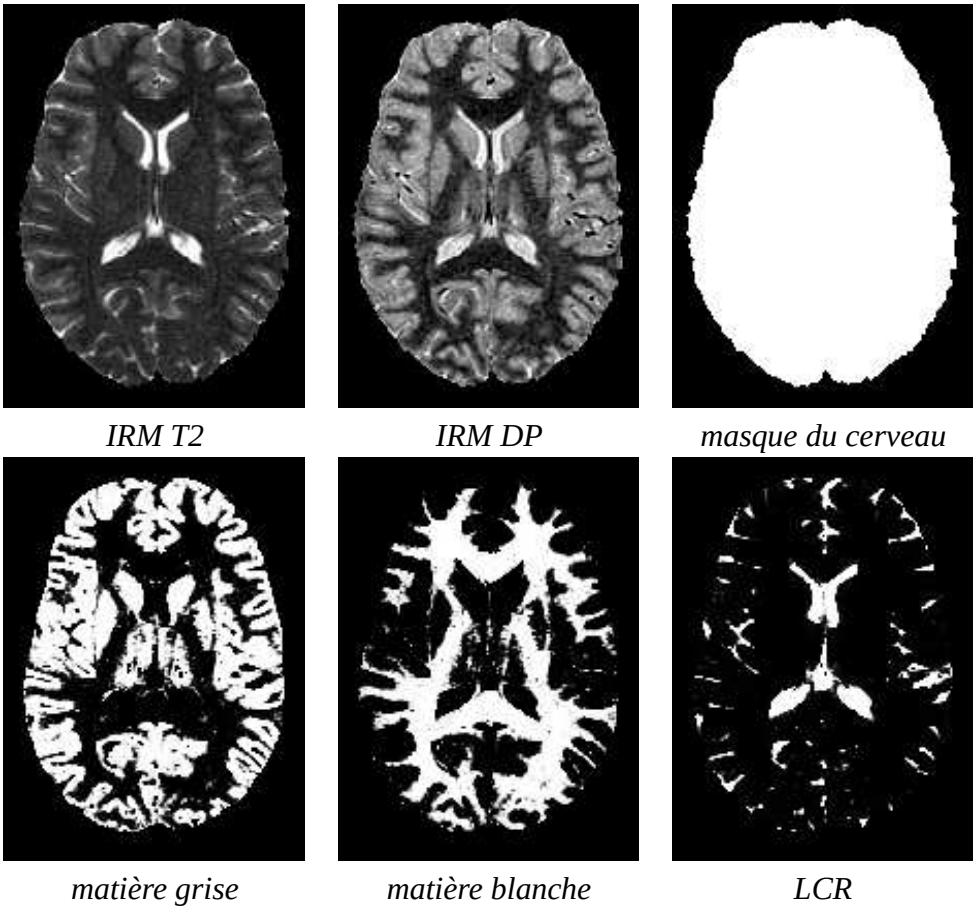


FIG. 4.6 – **Résultat de la segmentation par un algorithme EM à 3 classes** (matière blanche, matière grise, LCR) du couple IRM T2/DP, sur lequel on a appliqué un masque du parenchyme cérébral : visualisation de l'espace des images.
Sur cet exemple, une surestimation du LCR cortical est constaté : le LCR est très touché par le phénomène de volume partiel, ce qui conduit à une sur-segmentation des sillons corticaux.

4.4.3 Discussion

Le modèle courant donne bien une segmentation des tissus du parenchyme cérébral. Dans la littérature, de nombreuse méthodes permettent de résoudre ce problème, en introduisant des améliorations à la base mathématique fourni par l'EM : introduction de contraintes locales via des champs de Markov, séparation des points aberrants, utilisation d'un atlas labélisé pour identifier différentes structures, utilisation d'un atlas statistique plus précis, etc. A la différence de ces études, notre but est d'obtenir une bonne estimation des différents paramètres de classes, et pas uniquement d'obtenir une bonne labélisation, car la première segmentation des lésions se fera à partir de ces paramètres. Il convient donc de modifier le modèle lui-même en se posant la question suivante : pourquoi le système courant conduit-il à une sur-évaluation du volume du LCR, et particulièrement, comme nous le verrons plus tard, du LCR cortical ? Plusieurs solutions sont alors possibles.

- L'hypothèse d'uniformité du signal au sein de la classe est-elle vraie ? Il est établi que, par exemple, le signal des noyaux gris centraux n'est pas exactement le même que le signal du cortex. En quoi cela gêne-t-il le modèle ?
- La résolution des images n'est pas infinie : l'échantillonnage de l'image couplé à des replis importants dans les frontières inter-tissus, notamment au niveau du cortex, fausse l'estimation du nombre de classe et invalide le modèle de bruit gaussien.
- Enfin, résumer une segmentation du parenchyme cérébral en 3 tissus est peut-être grossier. Faut-il rajouter un certain nombre de classes ? Comment intégrer une nouvelle classe tout en gérant l'atlas statistique ?

Le but du chapitre suivant est de répondre à ces questions. Nous y verrons notamment comment l'introduction d'un modèle de volume partiel couplé à une segmentation des vaisseaux permet de valider le modèle initial.

Chapitre 5

Modèle de volume partiel

Les artefacts d'acquisition spécifiques à l'IRM sont nombreux. Biais spatial, reconstruction imparfaite, mouvements oculaires sont souvent la cause de l'échec des différentes techniques de segmentation automatique en imagerie médicale. Lorsque la surface de la structure à segmenter est importante par rapport à son volume, la résolution finie des images conduit à des effets de volume partiel qui gênent la segmentation : les voxels qui englobent la frontière vont contenir deux types de tissus. Selon la méthode, l'effet peut être plus ou moins gênant, mais si les volumes partiels ne sont pas pris en compte, une incertitude apparaît sur le placement de cette frontière, et les volumes mesurés peuvent être affectés grandement, avec une erreur de l'ordre de 20 à 60 % [59, 106]. Une bonne estimation des volumes partiels est donc indispensable à une bonne segmentation, et comme souvent, c'est la conjugaison de la correction des différents artefacts qui rend un algorithme de segmentation fonctionnel.

Dans notre cas, cet effet de volume partiel (*Partial Volume Effect*, ou PVE en anglais) nous gêne particulièrement pour la segmentation de la matière grise et du LCR : il fausse grandement l'estimation des paramètres de ces classes. Les replis du cortex induisent pour la matière grise une importante surface de contact avec les autres tissus – matière blanche, LCR. Le LCR cortical sera particulièrement touché par cet effet : l'épaisseur des sillons corticaux est souvent inférieure à la résolution des images, de l'ordre du millimètre. Enfin, les lésions n'échapperont

pas à cet effet, avec la difficulté supplémentaire de non-uniformité tissulaire au sein des lésions. Pour celles-ci, un traitement particulier sera nécessaire, en fonction de la sensibilité demandée au système de segmentation. Ce traitement sera présenté plus en détail dans le chapitre 8.

5.1 Signature des volumes partiels : image synthétique

Afin de voir l'influence des volumes partiels sur la segmentation dans le cadre d'un modèle bayésien, il est intéressant de regarder les résultats de l'algorithme EM à deux classes appliqué à une image simulée. Soit une image 2D, et une labélisation sur cette image en 2 classes. L'interface entre ces deux classes prend la forme d'une sinusoïde, et ce afin d'accentuer la longueur de l'interface par rapport à la surface occupée par chacune des classes (figure 5.1). D'un point de vue pratique, cette image est construite en haute résolution (3000*3000 pixels), et la labélisation est binaire. L'image est alors rééchantillonnée à une résolution plus faible : 300*300 et 60*60. Du bruit est rajouté dans les trois images, et les histogrammes correspondants sont calculés. Pour que l'histogramme soit suffisamment lisse, chaque image est générée plusieurs fois avec un bruit ajouté différent, et seule la moyenne des histogrammes est représentée dans la figure 5.1, ce qui permet de comparer plus facilement les images haute et basse résolution, sans avoir à lisser l'histogramme.

Plusieurs remarques sont à faire sur cette expérience. Tout d'abord, le volume de la classe noire est non-négligeable : 20% de l'image totale. Par contre, la frontière entre les deux classes est très grande : dans l'image basse résolution, la quasi-totalité des voxels appartenant à cette classe sont désormais des volumes partiels. Un algorithme EM classique, qui cherche à plaquer une mixture de deux gaussiennes, échoue à modéliser la distribution de probabilité générée par ce cas très simple : la variance de la classe noire est bien trop grande, et l'estimation de la moyenne en pâtit. L'EM est célèbre en statistique, car toute distribution de probabilité peut être approximée par une mixture d'un nombre suffisant de gaussiennes.

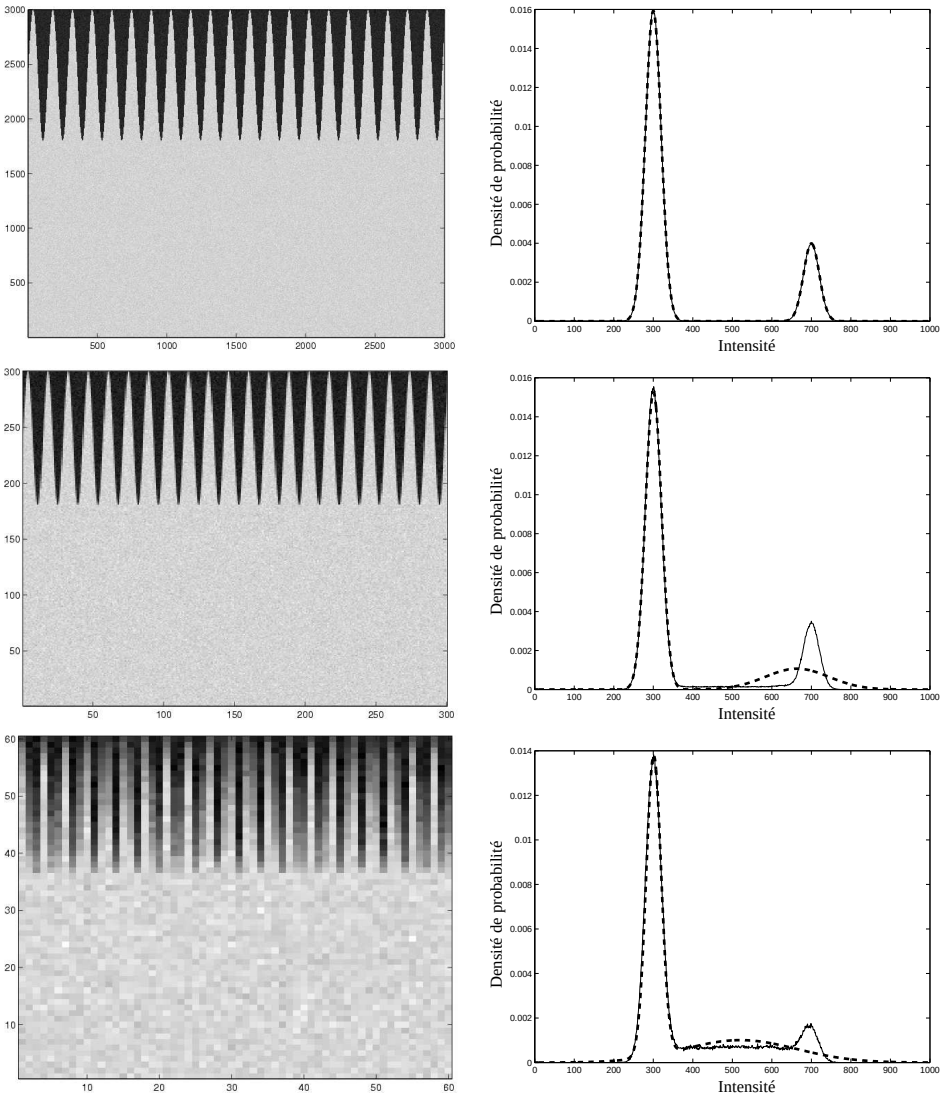


FIG. 5.1 – Images synthétiques à diverses résolutions (à gauche) à partir desquelles l'effet de volume partiel est simulé. L'histogramme de chaque image (à droite) est visualisé (trait continu noir). Un EM bi-classe simple est alors appliqué : la distribution de probabilité issue de cette segmentation est visualisée dans l'histogramme en pointillés épais. Dans l'image haute résolution, seules deux gaussiennes sont visibles dans l'histogramme. Dans l'image basse résolution, les volumes partiels sont visualisés comme un plateau entre les deux gaussiennes, et l'estimation de la distribution de probabilité est faussée, particulièrement pour la classe noire, fortement affectée par les volumes partiels. Pour situer cette image par rapport à une IRM cérébrale, la classe blanche représente la matière grise, alors que la classe noire représente le LCR qui présente une grande interface avec la matière grise par rapport à son volume.

Par contre, cela ne résoudrait pas notre problème de profiter de ce résultat, car la différenciation entre les deux classes et les volumes partiels n'est pas pour autant facilitée. Il nous faut donc un modèle plus contraint, qui nous permette de récupérer les paramètres des classes où le nombre de voxels purs – ne contenant qu'un type de tissus – est faible par rapport aux voxels partiels qui contiennent plusieurs types de tissus.

5.2 Choix de la méthode

Les volumes partiels sont considérés avec intérêt depuis longtemps dans tous les types d'imagerie, et particulièrement en imagerie médicale. Dans [176], Wu propose déjà un modèle Bayésien associé à des contraintes locales via une énergie de Gibbs pour segmenter des images de microscopie RMN, une variante haute résolution de l'IRM. En ce qui concerne l'IRM cérébrale, les volumes partiels sont encore plus flagrants et gênants, car la résolution des images est souvent de l'ordre des structures à observer : les sillons corticaux ont souvent une épaisseur de l'ordre du millimètre, et le volume visualisé en IRM sous forme d'un voxel est un mélange de LCR cortical et des méninges environnantes.

Dès 91, Choi propose de modéliser chaque voxel comme une mixture de tissus en IRM [30]. L'utilisation d'un modèle Bayésien ou de ses dérivés – estimateurs MAP, EM, *fuzzy c-means* – nécessite souvent l'introduction d'un modèle de volume partiel, car une segmentation voxellique est plus sujette à cet effet qu'une segmentation se basant sur des formes et des contours. En plus d'une estimation du biais spatial, Shattuck [148] ajoute une segmentation des volumes partiels à l'aide d'un *a priori* spatial de type Markov dans un modèle Bayésien. Van Leemput [167] intègre ce paradigme dans un EM avec un *a priori* spatial similaire, mais ne fournit des résultats qu'en 2D et les paramètres sont difficiles à ajuster. Dans [107], Noe propose un modèle plus simple où les classes partielles à l'interface entre deux tissus sont clairement identifiées : la composition de chaque classe partielle en termes de pourcentage est fixée, ainsi que le nombre de classes partielles. Nous nous inspirons fortement de cette méthode pour intégrer un tel modèle dans

un algorithme de type EM, comme le montre la suite de ce chapitre. D'autres techniques plus subtiles peuvent être utilisées pour ce problème. Laidlaw [80] utilise des histogrammes locaux – à l'échelle du voxel – associés à des contraintes locales. Gonzalez [60, 61] propose une étude plus complète de l'estimation des pourcentages de chaque tissu au sein d'un voxel, à l'aide de la méthode dite des intervalles de confiance. Ceci permet d'avoir une bonne idée de l'erreur sur le résultat, ce qui est très important en termes de calcul de volumes, par exemple. Un telle ligne a été suivie dans la littérature, comme le montrent les études de Chiverton [29].

De nombreuses techniques sont donc à notre disposition pour résoudre ce problème. Dans le choix de la méthode, il faut une fois encore tenir compte du but final, à savoir l'amélioration de l'estimation des paramètres de classes dans le cadre de la segmentation des lésions. Le modèle en question doit être également facile à intégrer à l'algorithme EM, et doit tolérer la présence des points aberrants que sont les lésions de SEP. Cela nous permettra par la suite de situer les lésions dans l'histogramme conjoint et donc d'obtenir une première segmentation. Pour ces multiples raisons, la méthode proposée par Noe est donc choisie comme base de travail. Sa qualité principale est de nous assurer d'optimiser les paramètres de classes des tissus qui nous concernent : la matière blanche, la matière grise et le LCR. Cependant, comme nous allons le voir, la simple introduction d'un nombre fixe de classes partielles n'est pas suffisant pour résoudre le problème.

5.3 Modèle de bruit et volumes partiels

Pour avoir une bonne idée du problème à résoudre, il est indispensable de reconstruire toutes les étapes d'acquisition, depuis les caractéristiques physiologiques des tissus jusqu'à l'image. Ensuite, dans un cadre EM, il sera possible d'énumérer toutes les approximations à faire pour obtenir un compromis entre un modèle simple et un résultat satisfaisant.

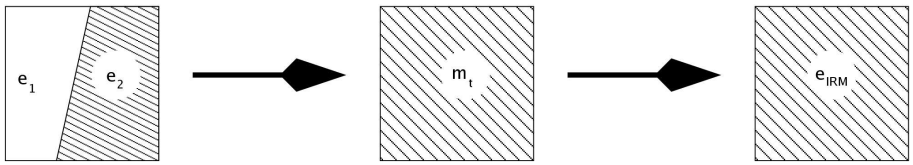


FIG. 5.2 – Bruit tissulaire, bruit IRM et volumes partiels. Le bruit tissulaire modélise la variabilité du signal au sein d'une même classe, issue d'une variabilité physiologique des tissus. Le processus d'acquisition se compose d'un rééchantillonnage qui donne un signal moyen m_t et y ajoute un bruit e_{IRM}, qui est admis comme gaussien quand le rapport signal / bruit est assez élevé.

Présentation du problème : la suite d'opérations présentée dans ce paragraphe se situe dans le cadre d'une segmentation d'un volume en deux classes, d'intensités théoriques I_1 et I_2. Ces classes présentent tout d'abord une certaine variabilité, prise en compte sous la forme d'un bruit tissulaire propre à chaque classe, e_1 et e_2. Dans une segmentation IRM, ce bruit tissulaire prendra compte des différences physiologiques entre des zones d'une même classe, comme le cortex et les noyaux gris centraux, ou la matière blanche normale et le corps calleux.

Il faut ensuite ajouter un rééchantillonnage des images qui va donner un signal moyen, dans chaque voxel, auquel se superpose un bruit d'acquisition e_{IRM} (figure 5.2). Soit $I_{1,mes}$ et $I_{2,mes}$ les intensités des classes pures ; I_{mes} l'intensité du voxel contenant un pourcentage α du tissu 1 et un pourcentage $1-\alpha$ du tissu 2.

$$I_{1,mes} = I_1 + e_1 + e_{IRM}$$

$$I_{2,mes} = I_2 + e_2 + e_{IRM}$$

$$I_{mes} = \alpha(I_1 + e_1) + (1-\alpha)(I_2 + e_2) + e_{IRM} \tag{5.1}$$

e_{IRM} suit une densité approximée par une gaussienne. Les bruits biologiques e_1 et e_2 sont spatialement corrélés, puisque dus aux propriétés intrinsèques de l'objet imagé. Dans la littérature, pour des raisons de simplicité, ils sont cependant souvent considérés comme blancs, dépendant de la structure considérée, stationnaires et gaussiens et de faible variance. α est la proportion de la classe 1 dans

tous les voxels, purs ou partiels. α, dans le cas général, est inconnu et difficile à estimer comme le montre Gonzalez [60].

Approximations choisies :

- $e_{IRM} = 0$. Cette approximation consiste à concentrer le bruit IRM dans les valeurs e_1 et e_2 : elle permet d'accélérer les calculs dans l'EM, car les volumes partiels, à α fixé, suivent une distribution de probabilité gaussienne fixe et connue, comme nous le verrons dans le paragraphe suivant.
- Pour les voxels contenant les deux classes de tissus α est uniformément distribué. Cette assertion est une approximation : si $\alpha \in]0, 1[$, il n'y a aucune raison que α soit uniformément distribué pour les voxels partiels. Le cas présenté sur la figure 5.3 est d'ailleurs assez représentatif de l'interface matière grise / LCR, puisque l'épaisseur des sillons corticaux est inférieur à un millimètre, alors que la résolution de l'image est de 1 millimètre. Cependant forcer la distribution de probabilité de α comme uniforme est dans notre cas un gage de stabilité algorithmique. Plus de détails sont donnés au paragraphe 5.6.

5.4 Intégration des volumes partiels dans l'EM

5.4.1 Présentation du modèle utilisé

Hors volumes partiels, le modèle de bruit gaussien s'applique donc à chaque classe : la distribution de probabilité de l'intensité de chaque classe reste une gaussienne. La variance de chaque classe inclut le bruit tissulaire et le bruit IRM. Il nous reste donc à modéliser la distribution de probabilité des volumes partiels, en gardant les classes pures originales. Pour simplifier l'étude, on considère uniquement les voxels contenant deux types de tissus. Cette approximation est courante et suffit largement dans le cadre de notre étude : étudier la frontière matière blanche / matière grise ou matière blanche / LCR est intéressante, mais le nombre de voxels contenant 3 types de tissus est négligeable par rapport au reste des voxels partiels.

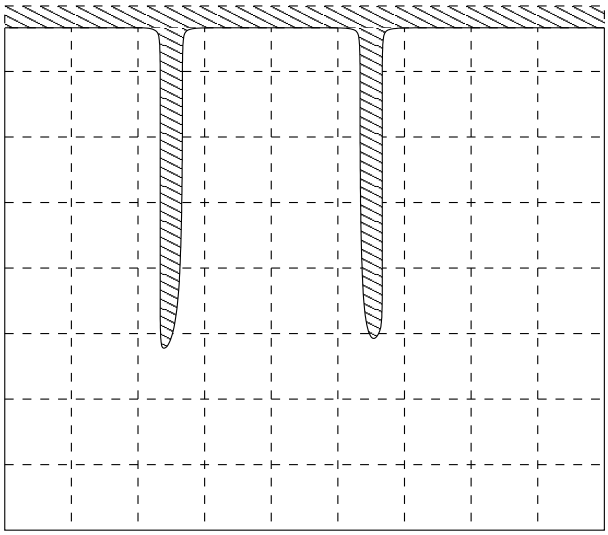


FIG. 5.3 – Exemple de volumes partiels où la distribution de probabilité de α n'est pas uniforme pour les voxels partiels. Dans ce cas, chaque volume partiel contient environ un tiers de la classe rayée, contre deux tiers de la classe blanche. Ce genre de problème survient quand les volumes partiels contiennent un repliement de la frontière entre les deux classes.

Dans un cadre EM, la solution la plus simple reste de modéliser les classes partielles comme une mixture de gaussiennes. La solution la plus facile à intégrer à l'EM a été proposée par Noe [107] ; nous nous servons de ses travaux comme base de travail. Les approximations formulées au paragraphe 5.3 conduisent à l'expression suivante :

$$I_{PVE} = \alpha I_1 + (1 - \alpha) I_2$$

I_1 et I_2 ont alors des distributions de probabilité gaussiennes centrées sur les moyennes μ_1 et μ_2, mais avec leurs variances propres Σ_1 et Σ_2. Cette approximation présente l'avantage de pouvoir facilement s'insérer dans l'algorithme présenté dans le chapitre 4, puisque les moyennes et variances sont simplement les variances des classes pures dont est issue la classe partielle en question. L'étape consiste ensuite à prendre des valeurs fixées de α : par exemple, $\alpha \in [0.2, 0.4, 0.6, 0.8]$. Pour chaque valeur fixée de α, I_{PVE} suit alors une distribution de probabilité gaus-

sienne aux paramètres suivants :

$$\begin{aligned} \forall\alpha &\in [0.2, 0.4, 0.6, 0.8] \\ \mu_\alpha &= \alpha\mu_1 + (1-\alpha)\mu_2 \quad (5.2) \\ \Sigma_\alpha &= \alpha^2\Sigma_1 + (1-\alpha)^2\Sigma_2 \quad (5.3) \end{aligned}$$

Une petite comparaison entre l'équation 5.3 et l'expérience synthétique menée dans le paragraphe 5.1 est intéressante. En effet, si on était sûr que le modèle de bruit gaussien était respecté sur toute l'image, et hors biais tissulaire, la variance du bruit serait indépendante de la classe associée. Or, si $\Sigma_1 = \Sigma_2 = \Sigma$, $\Sigma_{\alpha=0.5} = 0.5\Sigma$. Nous sommes conscients des limites de calcul de ce modèle de volume partiel : il nous permet d'avoir, comme on le verra par la suite, une bonne estimation des paramètres des classes pures. L'objectif est de prendre en compte le biais tissulaire en prenant des variances différentes pour chaque classe : si, au final, les variances de toutes les classes sont comparables, rien n'interdit de refaire une estimation finale de la variance globale à toute l'image pour mieux s'adapter au modèle original de bruit IRM.

5.4.2 Intégration dans l'algorithme de segmentation

Pour intégrer un tel modèle dans un algorithme de type EM, il reste encore à fixer la probabilité *a priori* pour les classes partielles. Il est toujours possible de supposer qu'il n'existe pas d'*a priori* spatial pour les volumes partiels, et de poser un *a priori* indépendant de la position i du voxel. Dans la littérature, nombreux sont ceux qui choisissent des contraintes locales à l'aide d'une énergie de Gibbs pour intégrer les volumes partiels. Cependant, nous avons évincé délibérément cette dernière méthode, car elle privilégie davantage les segmentations que les paramètres de classes. Dans le cadre d'une étude des lésions de SEP à partir des intensités, les paramètres de classes seront utilisés pour obtenir une première segmentation des lésions : il est préférable de repousser les contraintes de voisinage le plus tard possible dans la chaîne algorithmique pour être sûr que le modèle utilisé pour la segmentation soit valide. La solution choisie est donc de fixer les *a*

priori des classes partielles comme des fractions de l'atlas statistique.

Soient π_i^{MG}, π_i^{LCR} les *a priori* pour les classes pures, $\pi_i^{MG/LCR,k}$ les *a priori* pour les classes partielles ; a_i^{MG}, a_i^{MB}, a_i^{LCR} les trois atlas statistiques fournis pour les classes matière blanche, matière grise, LCR. Les paramètres β sont à estimer.

$$\begin{aligned} \pi_i^{MG} &= \beta^{MG} \cdot a_i^{MG} \\ \forall j \in [1,3],\ \pi_i^{MG/LCR,j} &= \beta_j^{MG/LCR} \cdot a_i^{MG} \\ \forall j \in [4,6],\ \pi_i^{MG/LCR,j} &= \beta_j^{MG/LCR} \cdot a_i^{LCR} \\ \pi_i^{LCR} &= \beta^{LCR} \cdot a_i^{LCR} \end{aligned}$$

Arbitrairement, on choisit pour cet exemple 6 classes partielles pour l'interface matière grise / LCR. Trois classes partielles auront comme *a priori* une fraction de l'atlas du LCR, les trois autres, une fraction de l'atlas de la matière grise. Le choix d'une combinaison linéaire des deux atlas est également possible, mais le calcul des β est alors beaucoup plus difficile, car le système n'est plus linéaire.

Pour prendre en compte ces contraintes de façon correcte, il faudrait en théorie intégrer les équations ci-dessus dans la formulation de l'étape de Maximisation, c'est-à-dire la dérivée de la Q fonction présentée dans l'équation 4.9. Comme $\gamma_i^k = p\big(Z|(X,\Theta)\big)$, $Q(\Theta)$ se calcule comme suit :

$$\begin{aligned} Q(\Theta) &= \sum_N \sum_Z p\big(Z|(X,\Theta)\big)\ \log p\big((X,Z)|\Theta\big) \\ &= \sum_{i=1}^{N} \sum_{k=1}^{K} \gamma_i^k \log p(Z_i = k, X = x_i|\theta_k) \end{aligned}$$

Une application directe de la loi de Bayes nous donne :

$$p(Z_i = k, X = x_i|\theta_k) = p(X = x_i|Z_i = k, \theta_k)p(Z_i = k|\theta_k)$$

$p(X = x_i|Z_i = k, \theta_k)$ est une gaussienne de paramètres (μ_k, Σ_k), ce qui nous donne une expression de la Q fonction suivante :

$$\begin{aligned} Q(\Theta) &= \sum_{i=1}^{N}\sum_{k=1}^{K}\gamma_i^k \log p(X = x_i|Z_i = k, \theta_k) + \sum_{i=1}^{N}\sum_{k=1}^{K}\gamma_i^k \log p(Z_i = k|\theta_k) \\ &= \sum_{i=1}^{N}\sum_{k=1}^{K}\gamma_i^k\big(-\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x_i-\mu_k)^T\Sigma_k^{-1}(x_i-\mu_k)\big) \\ &\quad + \sum_{i=1}^{N}\sum_{k=1}^{K}\gamma_i^k \log \pi_i^k \end{aligned}$$

Afin que la somme des *a priori* fasse 1 quelle que soit la position du voxel i dans le parenchyme cérébral, deux contraintes sont à ajouter sur les paramètres β :

$$\beta^{MG} + \sum_{j=1}^{3}\beta_j^{MG/LCR} = 1 \tag{5.4}$$

$$\beta^{LCR} + \sum_{j=4}^{6}\beta_j^{MG/LCR} = 1 \tag{5.5}$$

Ces contraintes sont appliquées à l'aide de deux multiplicateurs de Lagrange :

$$\begin{aligned} Q'(\Theta) &= Q(\Theta) - \lambda\left(\beta^{MG} + \sum_{j=1}^{3}\beta_j^{MG/LCR} - 1\right) \\ &\quad -\eta\left(\beta^{LCR} + \sum_{j=4}^{6}\beta_j^{MG/LCR} - 1\right) \end{aligned}$$

$$\begin{aligned}
\frac{\partial Q'(\Theta)}{\partial \beta^{MG}} &= \sum_{i=1}^{N} \frac{\gamma_i^{MG}}{\beta^{MG}} - \lambda = 0 \\
\forall j \in [1,3],\ \frac{\partial Q'(\Theta)}{\partial \beta_j^{MG/LCR}} &= \sum_{i=1}^{N} \frac{\gamma_i^{MG/LCR,j}}{\beta_j^{MG/LCR}} - \lambda = 0 \\
\forall j \in [4,6],\ \frac{\partial Q'(\Theta)}{\partial \beta_j^{MG/LCR}} &= \sum_{i=1}^{N} \frac{\gamma_i^{MG/LR,j}}{\beta_j^{MG/LCR}} - \eta = 0 \\
\frac{\partial Q'(\Theta)}{\partial \beta^{LCR}} &= \sum_{i=1}^{N} \frac{\gamma_i^{LCR}}{\beta^{LCR}} - \eta = 0
\end{aligned}$$

Les paramètres β sont faciles à obtenir :

$$\lambda = \sum_{i=1}^{N} \gamma_i^{MG} + \sum_{j=1}^{3} \sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \tag{5.6}$$

$$\eta = \sum_{i=1}^{N} \gamma_i^{LCR} + \sum_{j=4}^{6} \sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \tag{5.7}$$

$$\beta^{MG} = \left(\sum_{i=1}^{N} \gamma_i^{MG} \right) / \lambda \tag{5.8}$$

$$\forall j \in [1,3],\ \beta_j^{MG/LCR} = \left(\sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \right) / \lambda \tag{5.9}$$

$$\forall j \in [4,6],\ \beta_j^{MG/LCR} = \left(\sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \right) / \eta \tag{5.10}$$

$$\beta^{LCR} = \left(\sum_{i=1}^{N} \gamma_i^{LCR} \right) / \eta \tag{5.11}$$

Les autres paramètres du modèle sont alors les moyennes et variances des classes pures et partielles. Cependant, les calculs consécutifs sont alors très lourds, notamment pour l'estimation des matrices de covariances. Il est beaucoup plus

simple d'ajouter une étape de calcul des paramètres des classes partielles à partir des paramètres des classes pures, qui sont estimés de la même manière que dans un EM classique. On obtient donc un algorithme à 3 étapes au lieu de deux :

– Étape d'Espérance : labélisation des images, calcul des γ_i^k pour toutes les classes, y compris les classes partielles, suivant l'équation :

$$\gamma_i^k = \frac{\pi_i^k G_{\mu_k,\Sigma_k}(x_i)}{\sum_{l=1}^{K} \pi_l^k G_{\mu_l,\Sigma_l}(x_i)}$$

– Étape de Maximisation : estimation des moyennes et covariances pour les classes pures uniquement :

$$\begin{aligned} \mu_k^{pure} &= \frac{\sum_{i=1}^{N} \gamma_i^k x_i}{\sum_{i=1}^{N} \gamma_i^k} \\ \Sigma_k^{pure} &= \frac{\sum_{i=1}^{N} \gamma_i^k (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^{N} \gamma_i^k} \end{aligned}$$

– Étape des classes partielles : calcul arithmétique des paramètres des classes partielles (équations 5.2 et 5.3) ;
– Estimation de l'*a priori* pour chacune des classes à l'aide des équations 5.6 à 5.11.

Les calculs restent très faciles à intégrer dans une implémentation dont la caractéristique principale reste la rapidité. Pour visualiser les effets de cette modification, comme dans la partie 4.4, les images d'entrée sont le couple T2/DP, images originales intrinsèquement recalées. La seule difficulté reste le choix du nombre de classes partielles. Dans la littérature, il existe un certain nombre de critères qui permettent de choisir le nombre de classes dans un algorithme de type EM, mais les données du problème sont différentes car, cette fois-ci, les classes ne sont pas équivalentes. Dans un premier temps, le nombre de classes partielles est arbitrairement fixé à 6.

Il est très difficile d'évaluer les résultats. Dans un modèle basé sur l'intensité, l'effet de volume partiel n'est pas le seul phénomène qui gêne l'estimation des paramètres de classes, dont on ne connaît pas de valeur standard. Aucune

segmentation manuelle ne peut valider un tel modèle : la plupart des segmentations manuelles disponibles ne fournissent qu'une séparation entre les tissus et ne tiennent pas compte des volumes partiels. Notre seul moyen de validation est donc l'estimation des paramètres de classe et particulièrement de la matrice de covariance : c'est sur ce point que portera l'estimation de la qualité de la segmentation. À la vue des résultats présentés dans la figure 5.4, l'inclusion du modèle de volumes partiels n'apporte pas les résultats escomptés. Certes, l'essentiel du LCR cortical et ventriculaire est labélisé comme des volumes partiels dans la segmentation, mais la variance du LCR est toujours disproportionnée. Il manque une forme de rejet des points aberrants, qui est présentée dans le paragraphe suivant.

5.5 Ajout d'une classe supplémentaire pour les points aberrants

La méthode classique pour intégrer un rejet des points aberrants consiste souvent à intégrer dans l'algorithme une classe à la distribution de probabilité uniforme. Ceci s'apparente généralement au calcul d'une distance de Mahalanobis par rapport à chaque classe. Cette méthode, appliquée dans le logiciel EMS, résultat de la thèse de Van Leemput, est intéressante quand aucun *a priori* n'existe sur l'intensité de ces points aberrants. Or, dans le cas des IRM T2 de patients atteints de sclérose en plaques, deux points gênent particulièrement la segmentation en tissus : les lésions de SEP et les vaisseaux, qui sont des hyposignaux situés à proximité du LCR et segmentés comme tel dans le système présenté au paragraphe précédent (figure 5.5). Dans un premier temps, et ce afin d'observer les réactions du système, une seule classe *Points Aberrants* sera ajoutée pour traiter le problème des vaisseaux. Dans un deuxième temps, le problème des lésions sera abordé.

Comme le signal des vaisseaux est relativement constant en IRM T2/DP, imposer une distribution de probabilité uniforme pour la classe ajoutée ne se justifie plus : une distribution de probabilité gaussienne est plus adaptée. Au sujet de l'*a priori* spatial, aucun atlas statistique n'est disponible. Cependant, les vaisseaux

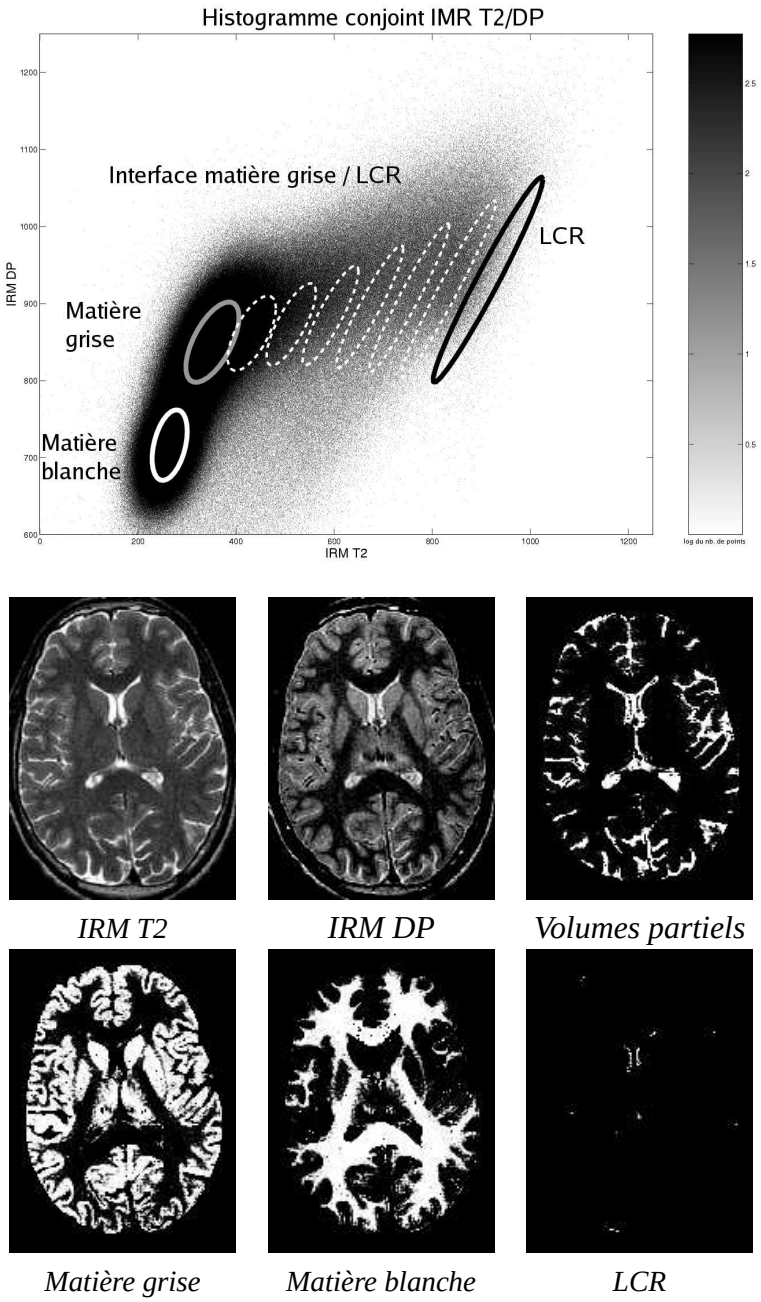


FIG. 5.4 – Segmentation en tissus sains après application du masque du parenchyme cérébral. Cette fois-ci, en plus des 3 classes matière blanche, matière grise et LCR, l'interface matière grise / LCR a été modélisée par 6 classes partielles (en pointillé sur l'histogramme). Sur cette image, la quasi totalité du LCR est en fait un ensemble de volumes partiels. Par contre, l'estimation de la moyenne et de la variance de la classe LCR est manifestement faussée : la variance est anormalement élevée, particulièrement sur l'axe IRM DP.

se situent grossièrement à proximité du LCR : l'*a priori* pour la nouvelle classe est donc une fraction de l'atlas du LCR. D'un point de vue algorithmique, rien n'est changé. Un nouvel *a priori* $\pi_i^{LCR,2}$, calculé à partir d'un coefficient $\beta^{LCR,2}$ est associé à la nouvelle classe : $\pi_i^{LCR,2} = \beta^{LCR,2} \cdot a_i^{LCR}$. La contrainte imposée par l'équation 5.5 est donc très simplement modifiée en l'équation 5.17, et l'algorithme ne change absolument pas. Tout se passe en fait comme si la classe LCR avait été séparée en deux : LCR et *Points Aberrants*.

$$\pi_i^{MG} = \beta^{MG} \cdot a_i^{MG} \tag{5.12}$$

$$\forall j \in [1,3],\ \pi_i^{MG/LCR,j} = \beta_j^{MG/LCR} \cdot a_i^{MG} \tag{5.13}$$

$$\forall j \in [4,6],\ \pi_i^{MG/LCR,j} = \beta_j^{MG/LCR} \cdot a_i^{LCR} \tag{5.14}$$

$$\pi_i^{LCR} = \beta^{LCR} \cdot a_i^{LCR} \tag{5.15}$$

$$\pi_i^{LCR,2} = \beta^{LCR,2} \cdot a_i^{LCR} \tag{5.16}$$

$$\beta^{LCR} + \beta^{LCR,2} + \sum_{j=4}^{6} \beta_j^{MG/LCR} - 1 = 0 \tag{5.17}$$

$$\beta^{MG} + \sum_{j=1}^{3} \beta_j^{MG/LCR} - 1 = 0 \tag{5.18}$$

Ce qui donne les formules suivantes pour les β :

$$\lambda = \sum_{i=1}^{N} \gamma_i^{MG} + \sum_{j=1}^{3} \sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \tag{5.19}$$

$$\eta = \sum_{i=1}^{N} \gamma_i^{LCR,1} + \sum_{i=1}^{N} \gamma_i^{LCR,2} + \sum_{j=4}^{6} \sum_{i=1}^{N} \gamma_i^{MG/LCR,j} \tag{5.20}$$

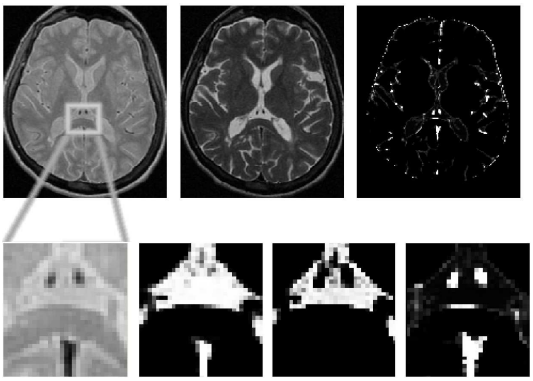


FIG. 5.5 – Visualisation de la segmentation des vaisseaux en IRM. De gauche à droite et de haut en bas : IRM DP et T2, segmentation finale des hyposignaux. Zoom : IRM DP, segmentation du LCR avant et après introduction de la classe supplémentaire. Ces hyposignaux qui, labélisés comme du LCR, sont la cause principale, avec les volumes partiels, de la mauvaise estimation des paramètres de classes.

$$
\begin{aligned}
\beta^{MG} &= \left(\sum_{i=1}^{N} \gamma_i^{MG}\right)/\lambda \\
\forall j \in [1,3],\ \beta_j^{MG/LCR} &= \left(\sum_{i=1}^{N} \gamma_i^{MG/LCR,j}\right)/\lambda \\
\forall j \in [4,6],\ \beta_j^{MG/LCR} &= \left(\sum_{i=1}^{N} \gamma_i^{MG/LCR,j}\right)/\eta \\
\beta^{LCR,1} &= \left(\sum_{i=1}^{N} \gamma_i^{LCR,1}\right)/\eta \\
\beta^{LCR,2} &= \left(\sum_{i=1}^{N} \gamma_i^{LCR,2}\right)/\eta
\end{aligned}
$$

La visualisation des résultats sur l'histogramme (figures 5.6 et 5.7) nous permet de dire que les objectifs ont été atteints, en ce qui concerne le calcul des paramètres de classes. Le signal dans les ventricules est extrêmement variable, et les volumes partiels ne sont que l'une des causes de cette variabilité. Ce que

nous donne le système est le signal du LCR à l'intérieur des ventricules, c'est-à-dire le LCR pur. Les artefacts de flux à l'interface liquide / solide, la présence des méninges, les *plexus choroïdes* dans les ventricules affectent le signal du LCR et le rendent particulièrement difficile à prévoir ; ces artefacts sont inclus par le système dans l'ensemble des volumes partiels. On obtient donc une modélisation finalement acceptable.

5.6 Détails d'implémentation

D'un point de vue pratique, plusieurs points sont à préciser pour une implémentation efficace du système. L'EM est parfois difficile à adapter à un problème de segmentation à cause de l'initialisation difficile à ajuster, et d'une convergence hasardeuse. Les variables à estimer sont les suivantes :

- les paramètres β_i^k qui gouvernent les *a priori* π_i^k selon les équations 5.12 à 5.16 ;
- les paramètres de classes : moyenne μ_i^k et matrice de covariance Σ_i^k pour chaque classe, pure ou partielle ;
- les labélisations *a posteriori* de chaque classe γ_i^k.

Initialisation : l'EM étant en deux étapes, il suffit d'initialiser soit les paramètres de classes, soit la labélisation *a posteriori*. Comme les paramètres de classes sont finalement ce que l'on cherche, il est plus simple de fournir l'initialisation de la labélisation en prenant les valeurs fournies par l'atlas. Il faut néanmoins faire attention à ce que pour la première itération, rien ne différencie la classe LCR de la classe *Points aberrants* : ajouter une valeur aléatoire dépendant de la position du voxel aux labélisations de toutes les classes permet de ne pas tomber dans un minimum local. L'algorithme se présente donc comme suit.

1. Estimation des moyennes et matrices de covariance des classes pures (matière blanche, matière grise, LCR, Points aberrants) : elle se fait simplement avec les équations 4.1.
2. Choix de la classe LCR : les classes LCR et *Points aberrants* sont différenciées par leur moyenne dans l'image T2 : la classe à la moyenne la plus

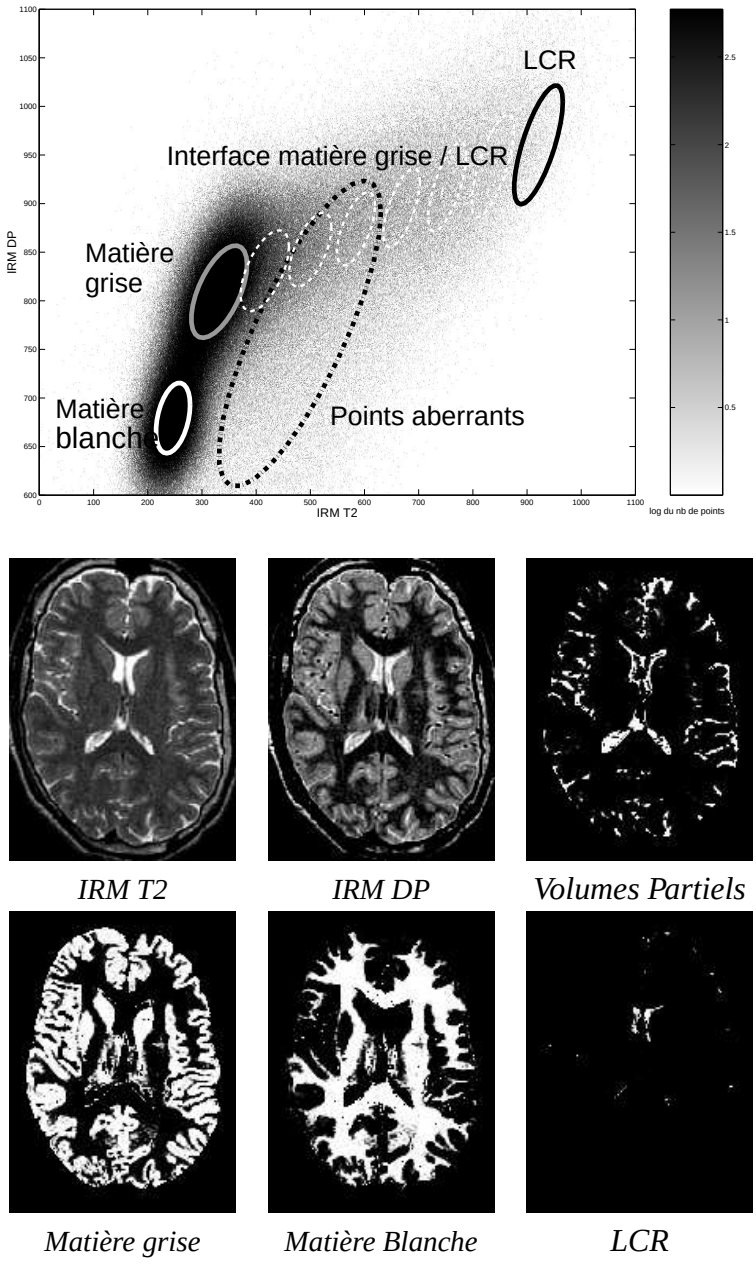


FIG. 5.6 – Segmentation en tissus sains après application du masque du parenchyme cérébral, en utilisant un modèle de volume partiel (pointillés blancs) ainsi qu'une classe contenant les points aberrants (pointillés noirs dans l'histogramme). Malgré la variabilité du signal du LCR, le système en a extrait une classe à la variance comparable à celle de la matière grise et de la matière blanche, et la moyenne correspond à celle du LCR ventriculaire, ou les volumes partiels sont inexistants.

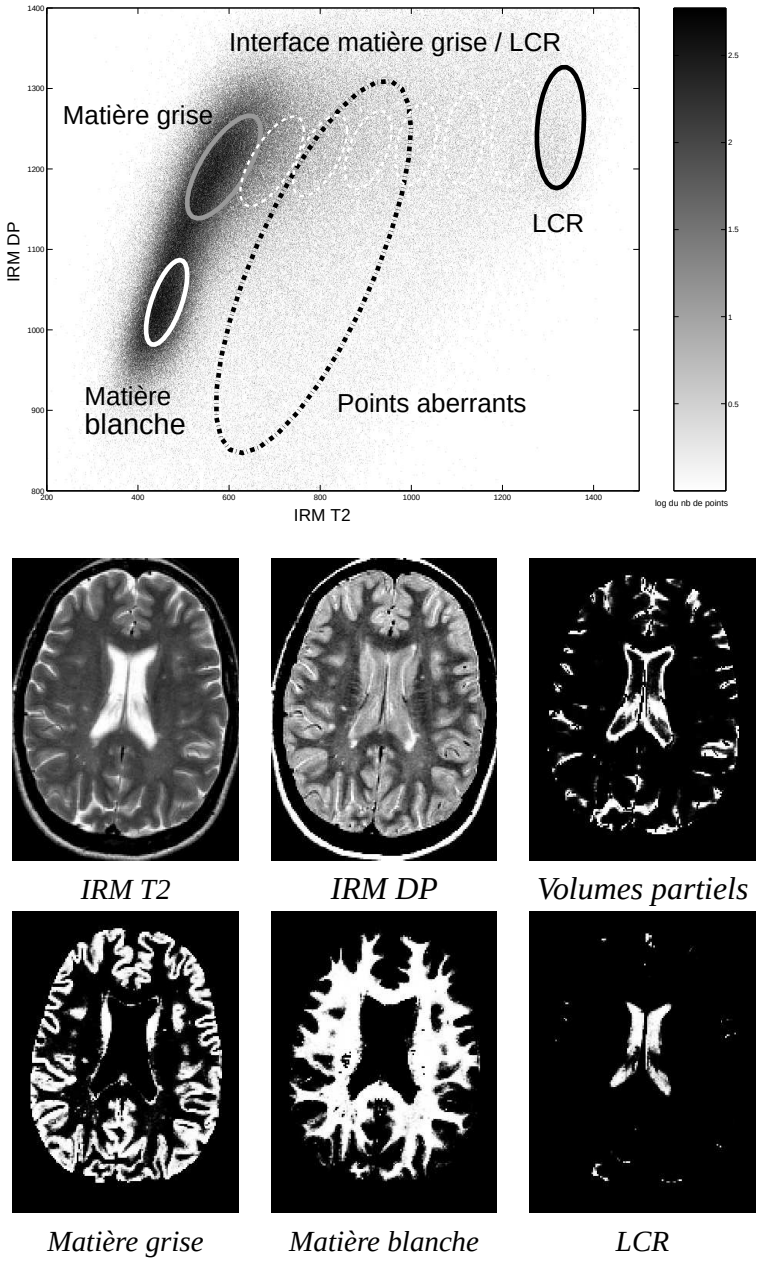


FIG. 5.7 – La segmentation en tissu a ici été appliquée sur des images obtenues avec un protocole d'acquisition identique, mais dans un centre différent et donc avec une machine différente. On observe des différences sensibles sur la qualité des images : les volumes partiels semblent plus importants, et le signal du LCR est pratiquement le même que celui de la matière grise en IRM DP. Le résultat est raisonnable, tant sur la qualité de la segmentation que sur la variance des classes pures.

grande est la classe LCR.

3. Calcul des moyennes et matrices de covariance des classes partielles avec les équations 5.2 et 5.3.
4. Remise à jour des coefficients β, à partir des paramètres intermédiaires λ et η.
5. Calcul de la log-vraisemblance globale de l'image, test du critère d'arrêt.

Les résultats sur les calculs des paramètres de classes sont présentés dans le tableau 5.1. Avec la correction de l'effet de volume partiel et l'intégration de la classe Points aberrants, les résultats sont convenables pour le calcul des variances. Plus précisément, la variance du LCR est maintenant comparable à la variance des autres classes pures. Même si la segmentation n'est peut-être pas parfaite – aucun contrôle précis n'a été effectué sur la segmentation – les paramètres de classes ont maintenant un sens, et peuvent être utilisés par la suite.

Pour encore améliorer la stabilité de l'algorithme, il est intéressant de noter que la classe Points aberrants a souvent une variance très importante par rapport aux variances des autres classes : le résultat est alors que la classe se superpose avec certaines classes partielles, qui ne sont alors plus sur un pied d'égalité : certaines classes partielles se retrouvent avec leur *a priori* voisin de zéro. Une grande partie des itérations de l'algorithme est en outre consacrée à estimer les différents β, alors que seuls les paramètres des classes sont intéressants pour une future segmentation des lésions. La solution à ces deux problèmes est de placer artificiellement les classes partielles sur un pied d'égalité en égalisant leurs *a priori* :

$$\begin{aligned}
&\forall j_1 \neq j_2 \in \{PVE\}, \quad && \textstyle\sum_i \pi_i^{MG/LCR,j_1} && = \sum_i \pi_i^{MG/LCR,j_2} \\
&\forall j_1 \neq j_2 \in \{PVE\}, && \beta^{MG/LCR,j_1} \textstyle\sum_i a_i^{(j_1)} && = \beta^{MG/LCR,j_2} \sum_i a_i^{(j_2)}
\end{aligned}$$

Cette contrainte est néanmoins difficile à implémenter dans l'algorithme, car elle reviendrait à rajouter de nombreux multiplicateurs de Lagrange et rendrait le système final non linéaire. Il est possible de reproduire une contrainte similaire par

Écart type (T2)	Matière grise	Matière blanche	LCR
Sans modèle de volume partiel	68.5	32	160.2
Avec PVE, mais sans points aberrants	49.7	33.9	167.9
Avec PVE et avec points aberrants	46.7	34.1	58.4

Écart type (DP)	Matière grise	Matière blanche	LCR
Sans modèle de volume partiel	53.3	41.8	145.2
Avec PVE, mais sans points aberrants	49.9	41.2	159.6
Avec PVE et avec points aberrants	49.7	40.4	60.5

TAB. 5.1 – Estimation de l'écart type des 3 classes pures en IRM T2 (en haut) et en IRM DP (en bas), calculée sur l'ensemble de la base. L'écart type de la classe LCR a été divisée par 3 avec l'ajout de la classe 'Points aberrants', ce qui montre son importance pour la validité du modèle de volume partiel.

une modification légère de l'algorithme, en remplaçant les $\sum_i \gamma_i^{MG/LCR,j}$ par des valeurs moyennées, égales pour toutes les classes partielles, dans le calcul des β, λ et η. Cet artifice d'implémentation permet d'accélérer fortement la convergence de l'algorithme et de stabiliser les différentes classes partielles les unes par rapport aux autres.

Chapitre 6

Segmentation des lésions de SEP

A ce moment de l'étude, plusieurs points importants ont été réglés dans la résolution du problème initial et vont nous assurer une certaine stabilité dans la chaîne de traitements. En effet, comme il sera montré dans la suite, la segmentation des lésions de sclérose en plaques est un problème difficile et mal posé, souffrant d'une forte variabilité intra- et inter-experts et très sensible au protocole d'acquisition utilisé. Il semble donc indispensable de placer l'ensemble des images multi-séquences de tous les patients dans le même espace de travail. Les étapes de prétraitement ont permis une uniformisation spatiale des images, alors que la segmentation en tissus et le modèle de volume partiel présentés dans les 2 chapitres précédents a permis une uniformisation en intensité : chaque tissu possède maintenant, dans le cadre d'un modèle de mixtures de gaussiennes, une caractérisation dans l'espace des intensités en IRM T2/DP et une segmentation.

Le gros avantage d'avoir choisi une telle direction est de factoriser la variabilité et la difficulté, pour se concentrer sur la partie la plus difficile du problème, la segmentation des lésions. Comme nous avons pu le voir dans le chapitre précédent, le modèle utilisé est effectivement vérifié sur l'ensemble de la base de données, ce qui est rassurant pour la suite. Il est par contre indispensable d'être capable de situer l'influence des différents prétraitements sur le résultat, afin de quantifier les différentes sources d'erreur possibles. Une étude de certaines étape de la chaîne est prévue dans le chapitre 8 de ce manuscrit. Maintenant que le cadre

de travail est fixé, le problème doit être posé et clairement défini. Une méthode de segmentation spécifique au problème initial doit également être présentée. C'est le but de ce chapitre.

6.1 Petit état de l'art

Comme le résume très bien R. Sharma dans le chapitre 5 de [155], les causes de la difficulté de la segmentation des lésions de SEP en IRM sont relativement connues. Bien que l'IRM fournisse un excellent contraste pour les différents tissus du parenchyme cérébral (matière blanche, matière grise, LCR), les lésions ne sont pas toujours bien contrastées et leur segmentation est rendue plus difficile par l'effet de volume partiel avec les tissus environnants, particulièrement le LCR. Une des particularités de l'IRM est la possibilité de modifier le contraste des différents tissus en manipulant les différents paramètres d'acquisition. Le résultat dépendra donc essentiellement de la capacité du protocole d'acquisition utilisé pour isoler les différentes lésions.

Tout d'abord, nous avons choisi de travailler sur un ensemble de séquences à un instant fixé ; il est possible d'utiliser plusieurs instants pour ainsi détecter les lésions évolutives. Les travaux de Rey [133, 134] et de Thirion [157, 23] détaillent avec beaucoup de pédagogie les différentes voies possibles à partir d'une série temporelle d'IRM T2. Un traitement multi-séquence de séries temporelles est également possible et a fait l'objet de plusieurs études dans la littérature [147, 2, 128, 130]. Le traitement des lésions évolutives est très important dans un cadre d'aide au diagnostic ou de suivi de l'état du patient [82] ; ce n'est néanmoins pas le sujet actuellement traité dans ce manuscrit.

Les méthodes de segmentation de lésions de SEP sont proches des méthodes de segmentation de tissus présentées dans le chapitre 4. Udupa construit un système basé sur la logique floue [163, 162] spécialisé dans la détection des lésions de SEP. Zijdenbos présente dans le système INSECT une chaîne de traitements basée sur un réseau de neurones, particulièrement adapté au problème [181, 180] : l'étape d'apprentissage se base sur un ensemble de segmentations manuelles four-

nies par plusieurs experts. Ceci assure un résultat fiable si le protocole permet une identification des lésions à partir de leur signal, sans nécessairement avoir besoin de contraintes locales. EMS, présenté par van Leemput [164] permet une segmentation multi-séquences basée sur l'intensité intégrant des contraintes locales à l'aide de champs de Markov. Pachai propose une approche plus pyramidale du traitement des lésions de SEP [113]. Anbeek propose une analyse statistique à l'aide d'un classificateur *kNN* de coupler intensité et contraintes locales pour segmenter les lésions de matière blanche [4, 6, 5].

Mais plus que la méthode, le choix de la séquence est le plus important. Le T2 FLAIR permet une sensibilité accrue hors de la fosse postérieure par rapport au T2 FSE [139], mais toute technique utilisant cette modalité est sujette à une sur-segmentation des lésions dans toutes les régions présentant une forte vascularité ou un flux de LCR. L'imagerie MTR est également très importante pour la SEP car le signal fournit des informations quantitatives en plus d'un masque des lésions [140], et permet également de limiter les artefacts de flux en FLAIR [53]. La charge lésionnelle extraite du duo T2 FSE / densité de proton reste la mesure de base de l'atteinte générique de la maladie [52], mais la signature des lésions est souvent proche, dans cette modalité, de la signature de la matière grise [171]. Au final, la méthode de segmentation utilisée devra analyser et traiter chaque séquence séparément pour espérer obtenir un résultat exploitable. Une caractérisation de chaque séquence du point de vue de la pathologie doit donc être précisée, pour s'assurer que la chaîne de traitements correspond à une réalité dans les images.

6.2 Présentation des lésions et des séquences

Notons tout d'abord que le processus de segmentation des lésions se sépare en deux phases, dont la difficulté n'est pas la même pour un processus manuel – purement humain – ou automatique : la détection et le contourage [162].

Détection : La première phase consiste à dire si la lésion existe ou non dans une zone précise de l'image. Les critères IRM actuellement utilisés en milieu clinique demandent une réponse binaire sur l'existence ou non d'un certain type de lésion : "y a-t-il une lésion périventriculaire en IRM T2 ?" ou "une lésion juxtacorticale est-elle apparue ?". Ce processus de détection nécessite une grande expérience de la part du praticien, car elle implique un grand nombre de critères, plus ou moins objectifs et parfois difficiles à quantifier. Par exemple, dans un contexte de diagnostic, le praticien aura plutôt tendance à éliminer la lésion en cas de doute, pour éviter de déclencher inutilement un traitement coûteux aux effets secondaires handicapants. Cet assertion est moins vraie dans le cadre d'un test pharmaceutique où les résultats sont construits à partir d'une grande base d'image : une surestimation du nombre de lésions sur certains patients n'est pas préjudiciable à l'étude.

Contourage : La seconde phase consiste à la détermination du volume de la lésion, une fois que l'on est sûr que celle-ci existe à cet endroit. La difficulté est différente de celle de la phase de détection : si la lésion est ancienne, son contour sera généralement bien contrasté, la zone œdémateuse sera quasi-inexistante et le contourage sera alors facile. Par contre, pour toute lésion jeune, il faut définir précisément le contour, même si l'intensité au sein de la lésion varie progressivement. Dans les deux cas, un contourage manuel est difficile à reproduire et consomme énormément de temps au praticien.

La difficulté de la détection et le caractère fastidieux du contourage font le succès des systèmes semi-automatiques d'extraction de quantificateurs pour la sclérose en plaques. La conception d'un tel système n'est pas simple, et différentes interfaces homme / machine peuvent conduire à des résultats différents [109]. Les travaux de cette thèse ont pour but de résoudre les deux problèmes, et donc d'obtenir un système de segmentation automatique. Il est cependant intéressant de savoir comment transformer un tel système en y rajoutant une interaction avec le praticien.

Différentes lésions pour différentes maladies Plusieurs variantes de sclérose en plaques existent, chacune ayant une signature différente en IRM. Pour

garder une signification clinique à cette étude, il faut donc être sûr du cadre dans lequel elle se trouve. Ainsi, de nombreux articles ont été publiés sur les syndromes cliniquement isolés et le caractère prédictif de l'IRM. Il a en effet été montré que les quantificateurs issus de l'étude des lésions cérébrales en IRM T2 lors d'un syndrome cliniquement isolé ont une forte valeur prédictive pour l'apparition d'une sclérose cliniquement définie [19, 110, 11].

Pour les formes aiguës de la SEP, le diagnostic est encore différent. Par exemple, pour les scléroses en plaques à révélation tardive (touchant des patients de 60 ans et plus), la coexistence de pathologies fréquentes (hypertension artérielle, cervicarthrose) associées à la SEP rend l'interprétation de l'IRM et des potentiels évoqués délicate et le diagnostic n'en est que plus difficile. Les SEP progressives, qui se caractérisent par une aggravation relativement continue de l'état général du patient, doivent être traitées différemment du point de vue de l'IRM : le signal de la matière blanche n'est plus uniforme, et les lésions sont plus difficiles à identifier (figure 6.1).

Les patients de la base de données utilisée pour cette étude sont pratiquement tous atteints de forme rémittente de SEP. Dans cette pathologie, la plus courante, les lésions sont bien individualisées par rapport au reste des tissus environnants, même si leur signal est variable. Elles peuvent être plus ou moins contrastées, avec une zone nécrotique au centre et un œdème en périphérie.

- Les lésions périventriculaires, caractéristiques de la maladie, sont le plus souvent assez grosses (de 5 à 10 mm de diamètre), de forme variable, avec un contraste important dans la zone centrale. Ce sont les lésions qui respectent le plus la structure en couronne présentée au chapitre 2 dans la figure 2.1. Souvent faciles à détecter par le contraste de la partie centrale, elles sont difficiles à contourer car leur forme peut être complexe.
- Les lésions juxtacorticales sont généralement petites et peu contrastées et de forme approximativement sphérique ou ellipsoïdale. Leur détection n'est pas toujours facile, car si la lésion est plus petite que l'épaisseur de coupe de la séquence que l'on regarde, il en résulte un contraste médiocre : ce problème survient tout particulièrement en IRM T2 FLAIR. Une bonne détec-

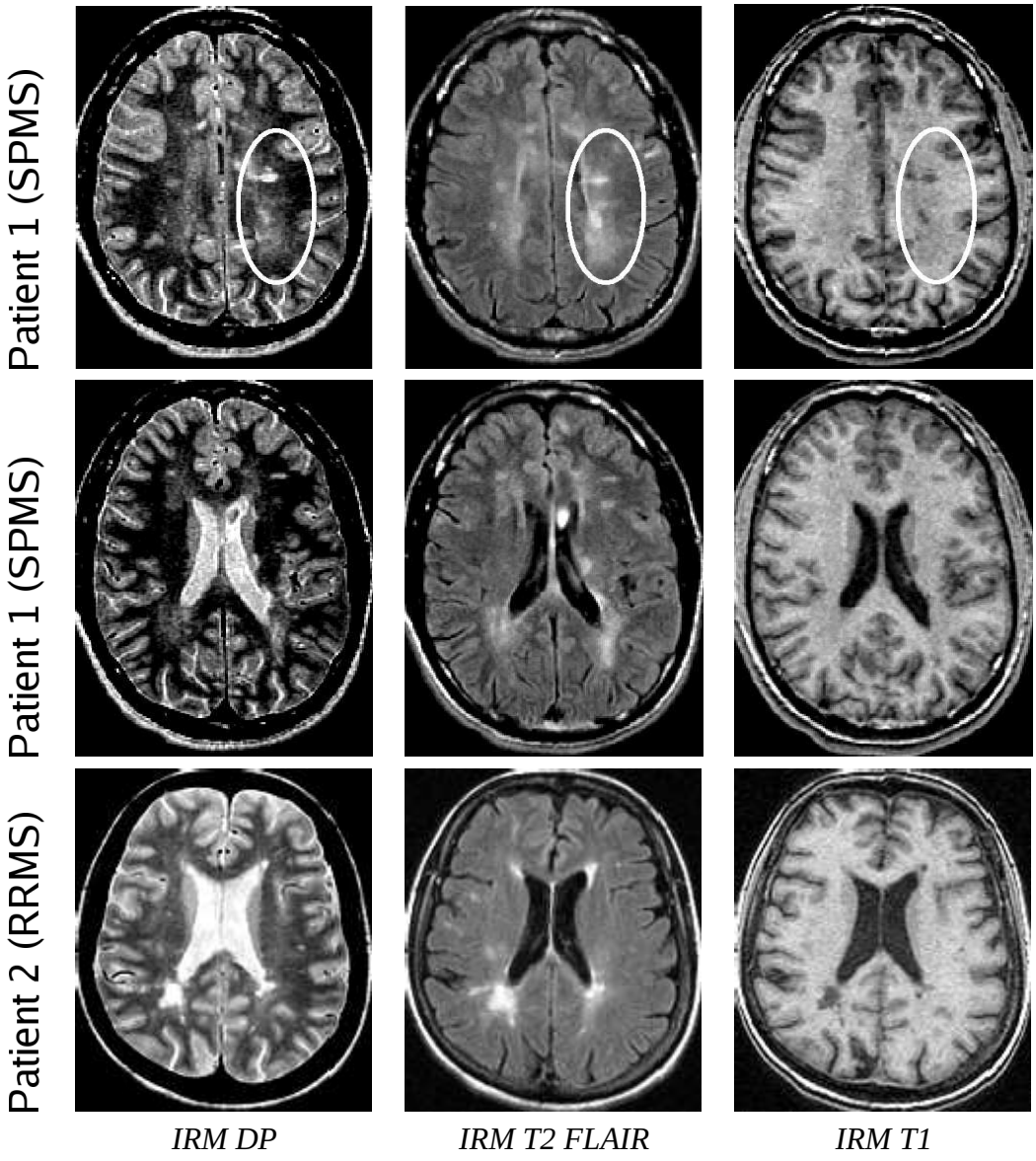


FIG. 6.1 – Les deux formes de maladie sont présentes chez ces deux patients : forme secondaire progressive (SPMS, deux premières lignes), et rémittente (RRMS, en bas), avec les modalités correspondantes (de gauche à droite) : IRM DP, T2 FLAIR et T1. Dans le premier cas, les lésions ont un signal extrêmement variable, mais dont l'intégralité doit être signalé comme lésions, y compris le léger hypersignal visible en T2 FLAIR et en IRM DP (lésion entourée sur la première ligne). Par contre, dans la forme rémittente, le contraste des images est bien établi, et seule la partie centrale des lésions est à prendre en compte en IRM T2/DP.

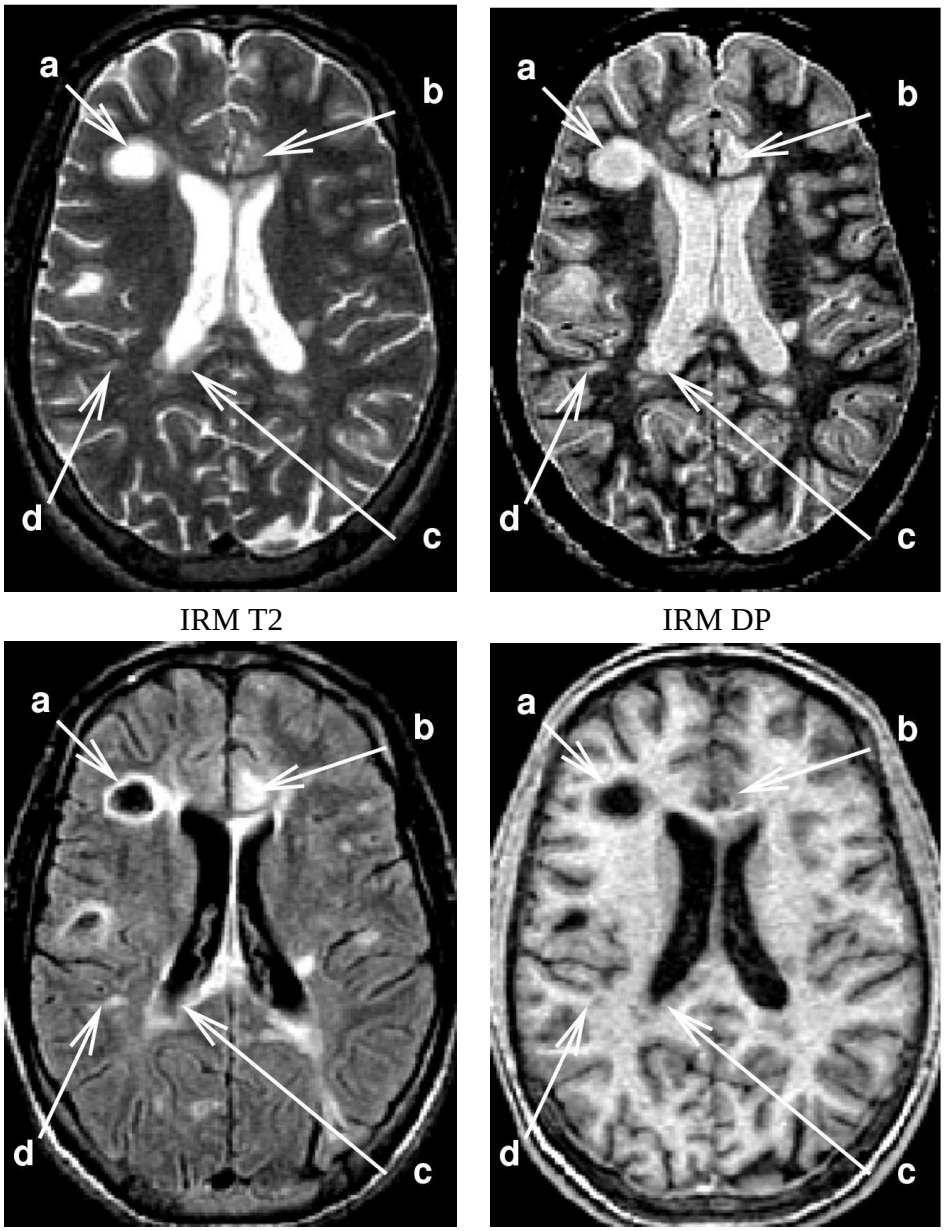


FIG. 6.2 – Visualisation des différents types de lésions de SEP en IRM multi-séquences : (a) nécrotique, (b) corticale, (c) périventriculaire, (d) juxta-corticale. Chacune de ces lésions va avoir un signal différent, le traitement doit donc être adapté.

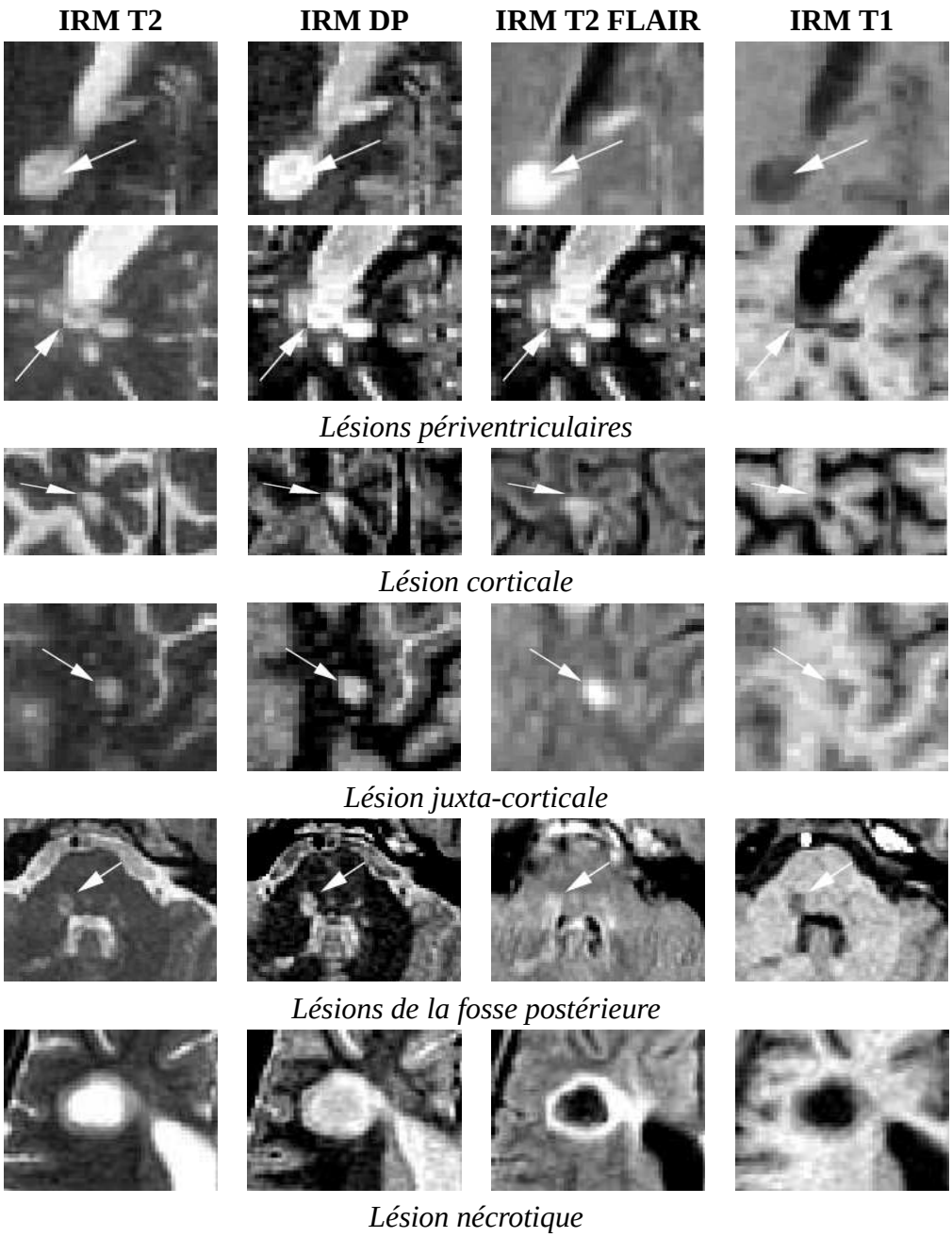


FIG. 6.3 – Différentes lésions au sein d'une forme rémittente de SEP. Les lésions périventriculaires (deux premières lignes) sont les plus caractéristiques de la maladie. Le centre des lésions nécrotiques possède un signal identique à celui du LCR. Le contraste des lésions de la fosse postérieure (dans le cervelet) est différent de celui des autres lésions, notamment en T2 FLAIR, et ces lésions doivent être traitées séparément.

tion de ces lésions est très importante, car l'apparition d'une lésion juxtacorticale est un critère important pour l'établissement d'une SEP cliniquement définie.

– Les lésions corticales sont des lésions juxtacorticales qui se sont étendues à la matière grise environnante, ou des lésions nées dans le cortex. Le contraste de ces lésions est variable en IRM conventionnelle, mais elles sont très bien détectées en IRM de diffusion. Ce ne sont pas des lésions de matière blanche, mais elles sont prises en compte pour la SEP dans le calcul des charges lésionnelles.
– Les lésions nécrotiques sont des lésions anciennes dont la partie centrale s'est nécrosée au point de devenir liquide. Ceci a pour conséquence principale que la partie nécrosée de la lésion possède alors un signal analogue à celui du LCR, ce qui rend plus difficile la détection dans notre système, basé sur l'information fournie par l'intensité en premier lieu. Quand la lésion est périventriculaire, la partie nécrosée n'est pas matériellement distinguable des ventricules, et est généralement ignorée.

Cette description nous permet de caractériser les différentes lésions dans l'espace des images. De cette information, on peut dire que tout système de segmentation automatique aura une grande difficulté à récupérer toutes ces lésions avec une qualité égale sur la segmentation et le contourage.

6.3 Intérêt du multi-séquences pour la segmentation des lésions de SEP

Maintenant que la caractérisation par l'image des lésions est faite, il est temps de s'intéresser à la manière dont cette information va être intégrée au processus de segmentation. Ce problème est plus compliqué qu'il ne le paraît, car certaines de ces informations sont plus qualitatives que quantitatives. En outre, il faut absolument que les mêmes informations soient intégrées à la segmentation automatique et au processus de validation. Pour l'instant, nous nous contentons de résumer les différents critères quantitatifs à appliquer directement, et à leur intégration dans la

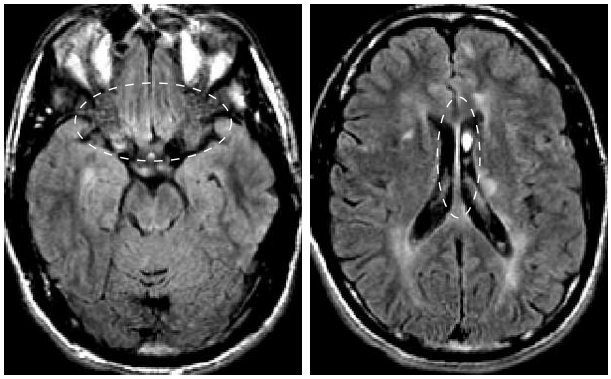


FIG. 6.4 – L'IRM T2 FLAIR est une séquence très pratique pour l'analyse de la sclérose en plaques, mais dont il faut se méfier pour une extraction correcte des quantificateurs IRM. Certains artefacts gênent la détection des hypersignaux de la substance blanche – potentielles lésions de SEP – notamment l'artefact osseux (à gauche) et les artefacts de flux (à droite)

chaîne de traitements ; la validation sera explicitée dans le chapitre suivant.

Propriétés des différentes séquences

IRM T2 FLAIR Cette séquence fournit un bon contraste pour la plupart des lésions hors fosse postérieure. A part le cas très particulier des lésions à composante nécrotique, toutes les lésions sont visibles comme un hypersignal fort par rapport à toutes les autres classes du parenchyme cérébral. Par contre, le volume des hypersignaux de la substance blanche en T2 FLAIR est une surestimation du volume lésionnel en IRM T2/DP. La détection va donc pouvoir se baser sur cette modalité, mais ne peut s'en contenter, malgré un bon contraste lésions / tissus sains. En outre, divers artefacts vont gêner la détection des lésions et induire un certain nombre de faux positifs. Parmi eux, deux sont à prendre en compte plus particulièrement (figure 6.4).

– *Les artefacts de flux* : ils conduisent à un hypersignal à tous les endroits présentant un flux de liquide – LCR, sang ou lymphe. Schématiquement, lors du calcul du signal IRM sur un volume donné, cet artefact apparaît quand le volume de liquide mesuré à l'excitation ne se situe plus dans le voxel

en question à la relaxation des spins. D'un point de vue pratique, cet effet apparaît entre les deux ventricules, à proximité des cornes ventriculaires et plus généralement à l'interface parenchyme / LCR.

– *L'artefact osseux* est un hypersignal qui survient dans toutes les séquences dérivées du T2 dans la zone temporo-basale. Il se caractérise par un hypersignal relativement localisé : il peut survenir d'une coupe à l'autre. Comme on peut le voir sur la figure 6.4, le signal IRM à cet endroit s'apparente au signal d'une lésion de SEP. Il faudra donc tenir compte de ce phénomène.

IRM T2 / DP Cette double séquence est à l'heure actuelle la séquence indispensable à toute étude sur la SEP. Bien que le contraste des lésions y soit relativement aléatoire, elle présente l'avantage de signaler l'atteinte non-spécifique des tissus, ce qui est intéressant pour quantifier l'état général du patient. L'extraction de la charge lésionnelle dans cette modalité est donc un de nos buts à atteindre. D'un point de vue signal, l'intensité des lésions se situe entre la matière grise pour les lésions les moins contrastées et les volumes partiels matière grise / LCR pour les plus contrastées, avec un possible hypersignal en IRM DP par rapport au LCR. Cet hypersignal, bien que certains articles le posent comme base de travail, n'est pas toujours valide selon le protocole d'acquisition (figure 6.5). La détection n'est donc pas facile dans cette modalité, et nécessite l'ajout de critères supplémentaires – spatiaux, morphologiques – difficiles à intégrer vu la complexité des structures cérébrales. Il est certes toujours possible d'utiliser le T1, d'une résolution élevée, pour y inclure ces critères topologiques mais comme nous le verrons dans le paragraphe suivant, le problème est loin d'être facile. Il sera évoqué en termes de perspectives.

IRM T1 Cette modalité ne pose pas de contraintes particulières. Dans le protocole utilisé, la résolution est grande – taille du voxel : 0.8*0.8*1 mm – ce qui nous permet d'obtenir une meilleure segmentation du LCR en termes de volumes partiels. Nous verrons dans la section suivant que cette information est utilisée pour réduire les artefacts de flux en FLAIR. Du point de vue des lésions, les hyposignaux en IRM T1 donnent l'atteinte axonale. Ces hyposignaux sont un sous-

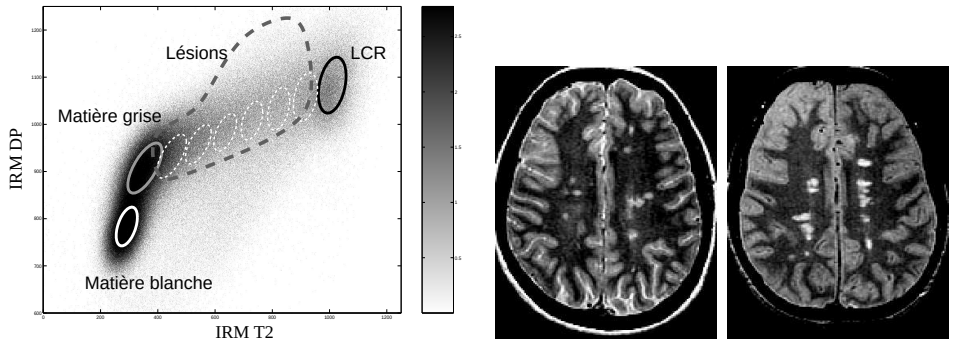


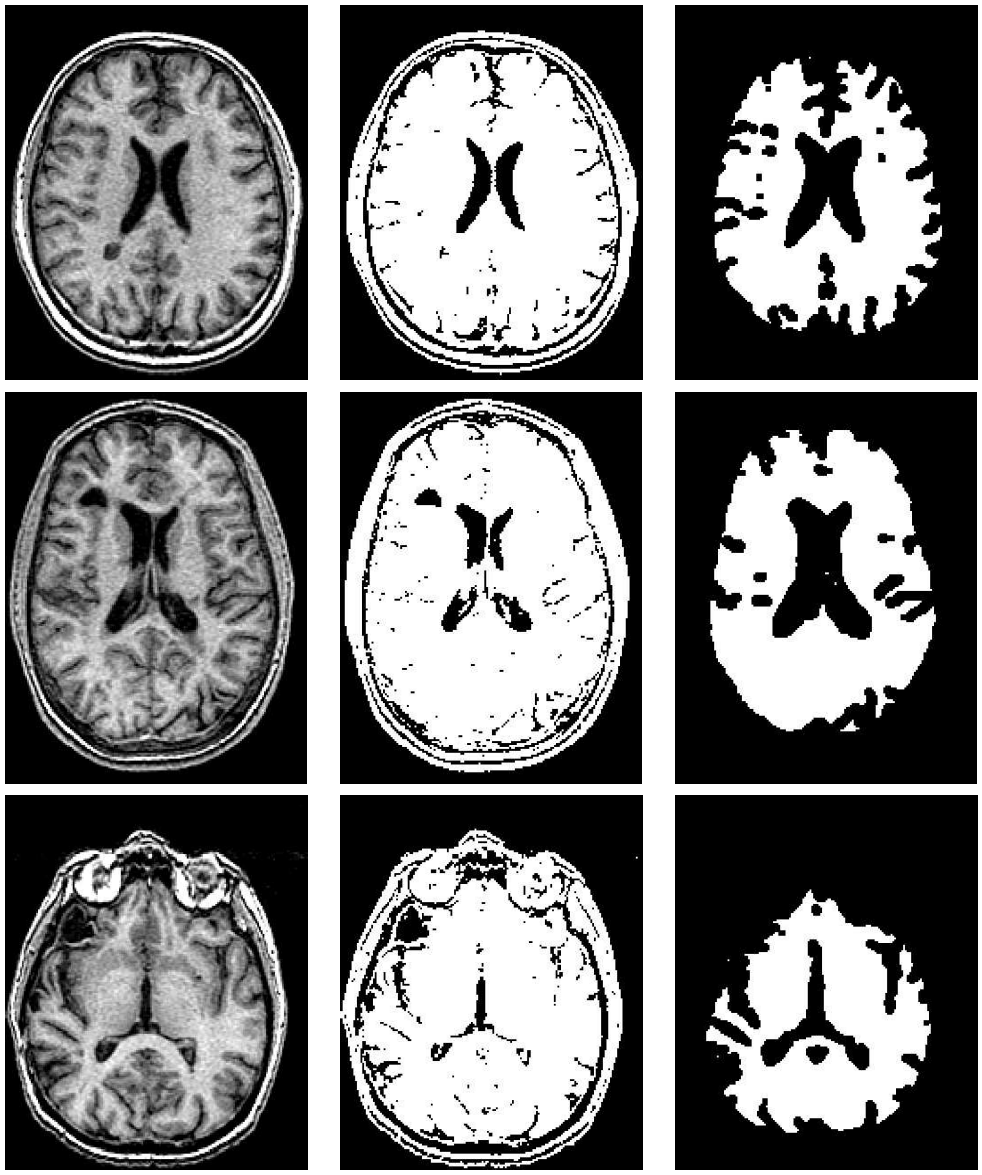
FIG. 6.5 – En IRM T2 / DP, le contraste des lésions de SEP varie. En règle générale, les lésions se situent sur la zone délimitée en gris foncé en larges pointillés dans l'histogramme conjoint (à gauche). Comme on peut le voir, cette zone recoupe fortement les volumes partiels matière grise / LCR. La détection est donc très difficile à réaliser sur cette modalité.

ensemble des lésions : la séquence sera donc traitée *a posteriori*, comme le couple T2/DP, pour le contourage des lésions.

6.4 Finalisation algorithmique

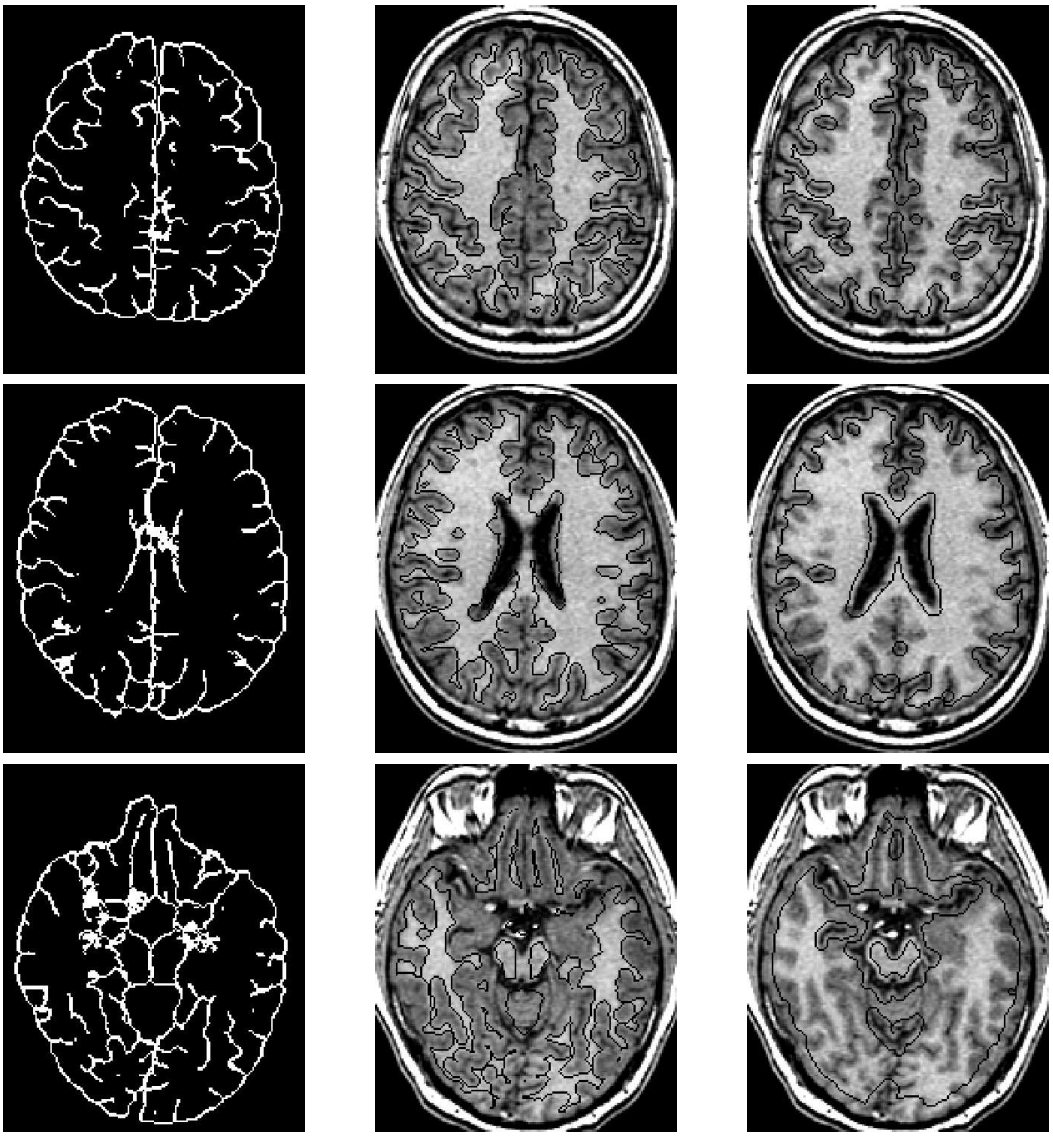
Les deux premières parties de ce chapitre présentent les caractéristiques des différentes lésions dans les séquences disponibles. Ces informations sont issues de la littérature médicale d'une part (cf. chapitre 2) et de l'expérience sur les différents protocoles d'acquisition auxquels nous avons été confrontés, ainsi que de la communication avec le Dr Lebrun au CHU Pasteur à Nice. Avant de présenter la version finale de la chaîne de traitements, il est important de présenter la version "algorithmique" de ces points : comment employer ces informations pour un protocole de segmentation – manuelle ou automatique ?

Détection des lésions en IRM T2 FLAIR Tout d'abord, le FLAIR possède une forte capacité de détection des lésions supra-tentorielles. Il sera donc utilisé comme première étape de détection. Les segmentations issues du modèle présenté dans le chapitre 5 permettent d'obtenir une caractérisation des différents tissus sains (matière blanche, matière grise, LCR) : moyenne et écart type dans l'image



IRM T1 *Seuillage* $\mu_{LCR} + \sigma_{LCR}$ Masque final

FIG. 6.6 – Calcul de la région d'intérêt pour la segmentation des lésions de SEP à partir de l'IRM T1. Les segmentations en tissus obtenues au chapitre précédent permettent de calculer un seuillage sensible, qui nous assure de ne pas éliminer de lésions, hormis les *trous noirs* (première ligne). Des opérations de morphologie mathématique (élimination des trous, plus grande composante connexe, érosion finale) permettent d'ajouter une marge de sécurité à toutes les interfaces liquides / solides, susceptibles de faux positifs dans toutes les modalités issues du T2. Les *trous noirs* non connexes au LCR ne sont pas pour autant éliminés (deuxième ligne). Par contre, si une lésion est directement connexe au LCR, une connaissance médicale supplémentaire est nécessaire et la lésion n'est pas prise en compte (troisième ligne).



Sillons (Brainvisa) | *Masque issu des sillons (seuillage sur la carte de distance)* | *Masque issu du seuillage de l'IRM T1*

Patient : MCI01

FIG. 6.7 – Visualisation des différentes possibilités pour la région d'intérêt issue du T1 : il s'agit ici d'éliminer les artefacts de flux à la frontière parenchyme/LCR. Un seuillage sur l'IRM T1 donne une région d'intérêt peu précise (à droite). Le système de détection des sillons de *Brainvisa* [36, 136, 135, 95, 93] permet d'obtenir un masque plus précis (colonne centrale) en seuillant la carte de distance aux sillons obtenus (à gauche). Cependant, la segmentation des sillons étant obtenue à partir de la frontière matière blanche/matière grise, certaines lésions sont alors éliminées de la région d'intérêt (image centrale). Le seuillage sur le T1 est donc choisi.

FLAIR. À partir de ces moyennes, un seuillage $\lambda = \mu_{MG}+3\sigma_{MG}$ permet d'obtenir une pré-segmentation des lésions [42]. Avec cette méthodologie, le FLAIR devient la pierre angulaire de la détection des lésions. Il est alors important de corriger les artefacts dont il souffre. L'artefact osseux sera corrigé par une simple correction de biais coupe à coupe, spécifique au FLAIR. Pour ce faire, la même technique que celle évoquée dans le paragraphe 3.4 du chapitre 3 est utilisée, en prenant comme segmentation les labélisations issues des chapitres 4 et 5. Cette méthode, simple, s'avère fonctionnelle sans pour autant réduire de façon significative le contraste des lésions très étendues.

Au sujet des artefacts de flux, dans toutes les modalités issues du T2 – IRM T2/DP, IRM T2 FLAIR – l'hypersignal des lésions périventriculaires au niveau des cornes des ventricules est souvent interprété comme un effet des volumes partiels ou un artefact de flux. En conséquence, une marge de sécurité de 1 à 2 millimètres autour des ventricules, dans laquelle aucune lésion ne peut se situer, permet de s'affranchir de ces problèmes. Cette marge de sécurité est calculée à l'aide de l'IRM T1.

Spécification de la région d'intérêt en IRM T1 Ceci est fait par une segmentation du parenchyme cérébral par un seuillage sur l'image T1 auquel on a appliqué le même masque du cerveau obtenu au paragraphe 4.3. Ce seuil est calculé de la même manière que le seuil en FLAIR : le couple de paramètres moyenne / matrice de covariance est estimé pour chacun des trois tissus sains à partir des labélisations précédemment obtenues en IRM T2/DP : matière blanche, matière grise, LCR. Pour éviter que les lésions visibles en T1 soient exclues du parenchyme cérébral, un seuil très sensible est utilisé, de l'ordre de $\mu_{LCR} + \sigma_{LCR}$. Ce seuil est appliqué sur l'IRM T1 originale, dont la résolution est élevée. Après une régularisation habituelle du masque à base de morphologie mathématique (connexité 26 : plus grande composante connexe, élimination des trous), une érosion du parenchyme (calculée avec une carte de distance à 2 mm) nous donne une région d'intérêt dans laquelle les lésions de SEP seront recherchées (figure 6.6). Il est possible d'utiliser des critères plus complexes pour extraire cette région d'intérêt.

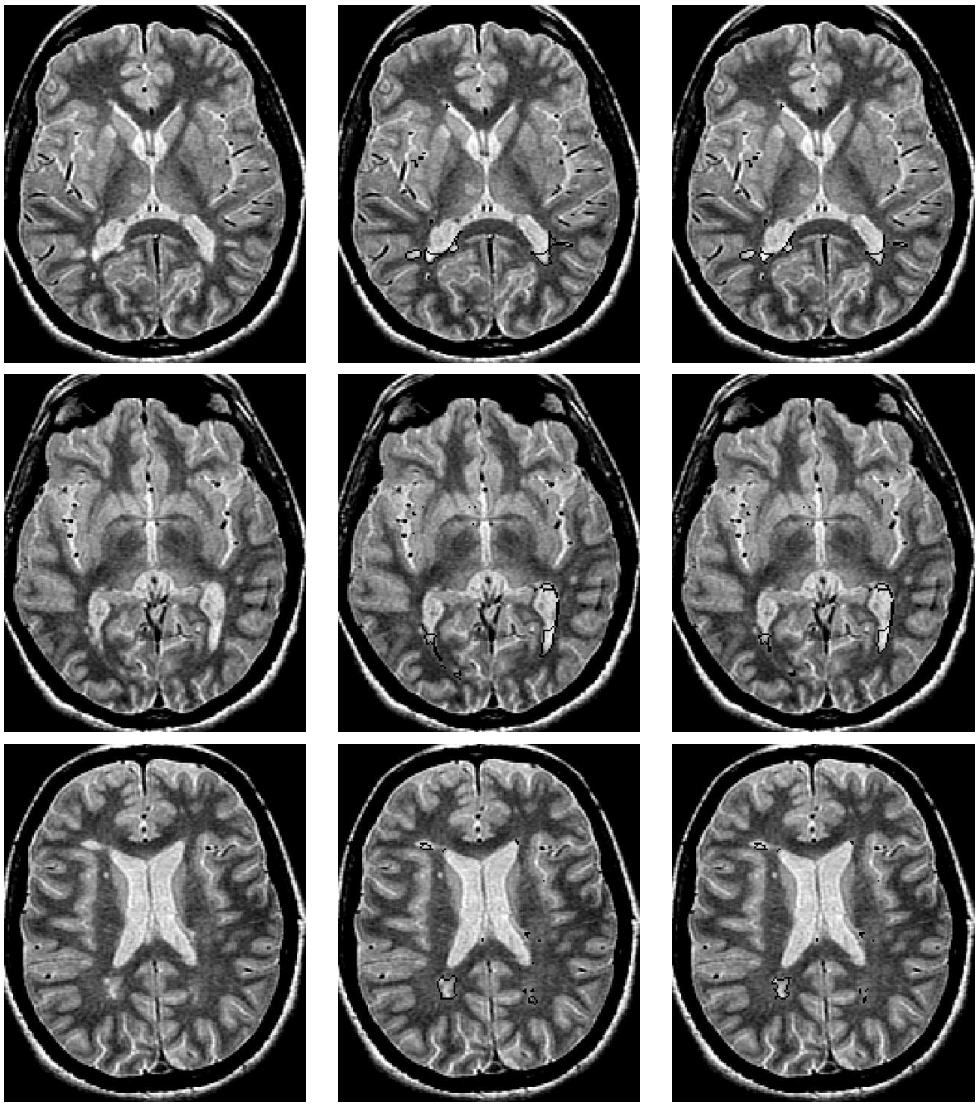
Ainsi, le système Brainvisa [36, 136, 135, 95, 93] permet une segmentation des sillons corticaux à partir de l'interface matière blanche / matière grise, à l'aide de calculs sur les maillages. Les résultats sont impressionnants sur les témoins. Ce système est cependant difficile à utiliser sur les images des patients atteints de SEP, car les lésions ont un signal de matière grise en T1 : l'interface matière blanche / lésion est alors, d'un point de vue signal, la même que l'interface matière blanche / matière grise, et la segmentation des sillons corticaux, visualisé par une surface holomorphe à une sphère pour chaque hémisphère cérébral, empiète sur certaines lésions (figure 6.7).

Contourage des lésions en IRM T2/DP Il ne reste plus qu'à appliquer le masque binaire sur la segmentation obtenue en IRM T2/DP pour extraire une segmentation finale des lésions. Comme on peut le voir dans l'histogramme de la figure 6.5, seuls deux tissus sont à exclure totalement, d'un point de vue signal : tout voxel labélisé à partir de son intensité comme de la matière blanche ou du LCR ne peut pas être une lésion. L'intersection entre le masque binaire des hypersignaux en IRM T2 FLAIR et la réunion de la matière grise, des volumes partiels et des points aberrants hyperintenses en IRM DP donne donc les lésions de SEP.

6.5 Présentation de la chaîne de traitements

La chaîne de traitements comprend donc les étapes suivantes.

1. Acquisition des images : reconstruction 3D de toutes les séquences à partir des coupes (chapitre 3, paragraphe 3.2).
2. Recalage rigide intra-patient des différentes séquences : T1 sur T2 et T2 FLAIR sur T2 (chapitre 3, paragraphe 3.3.2).
3. Recalage rigide puis affine de l'atlas statistique – image T2 moyenne – sur l'IRM T2 du patient. D'un point de vue pratique, cette étape permet également de réduire la taille des images au strict nécessaire, puisque l'atlas statistique nous donne la localisation du cerveau dans l'image (paragraphe 3.3.3).



IRM DP *Avant application* *Après application*

FIG. 6.8 – Extraction de lésions de SEP en IRM T2/DP à partir de la segmentation en tissus fournie au chapitre 5. La segmentation des lésions de SEP à partir du T2 FLAIR est présentée sur la coupe axiale de l'IRM DP, avant (milieu) et après (droite) application de la segmentation en tissus. Le T2 FLAIR fournit en général une sur-segmentation des lésions. La suppression des voxels labélisés comme matière blanche et LCR permet de coller au couple T2/DP, tout en utilisant les qualités de détection du T2 FLAIR.

4. Segmentation du masque du parenchyme cérébral (paragraphe 4.3).
5. Segmentation en tissus : labélisation des images, extraction des paramètres de classes pour la matière blanche, matière grise, LCR et interface matière grise / LCR en IRM T2/DP (chapitre 5).
6. Estimation du biais spatial à partir des segmentations obtenues à l'étape 5. Réestimation des paramètres de classes en IRM T2/DP.
7. Extraction d'un biais coupe à coupe en IRM T2 FLAIR. Calcul d'un prémasque des lésions par un seuillage automatique en IRM T2 FLAIR (paragraphe 6.4).
8. Calcul des paramètres de classes en IRM T1 à partir des labélisations obtenues précédemment. Calcul de la région d'intérêt pour la recherche des plaques (paragraphe 6.4).
9. Élimination des voxels labélisés comme matière blanche et LCR à partir de la segmentation en tissus obtenue à l'étape 6. Élimination des lésions dont la taille est trop petite.

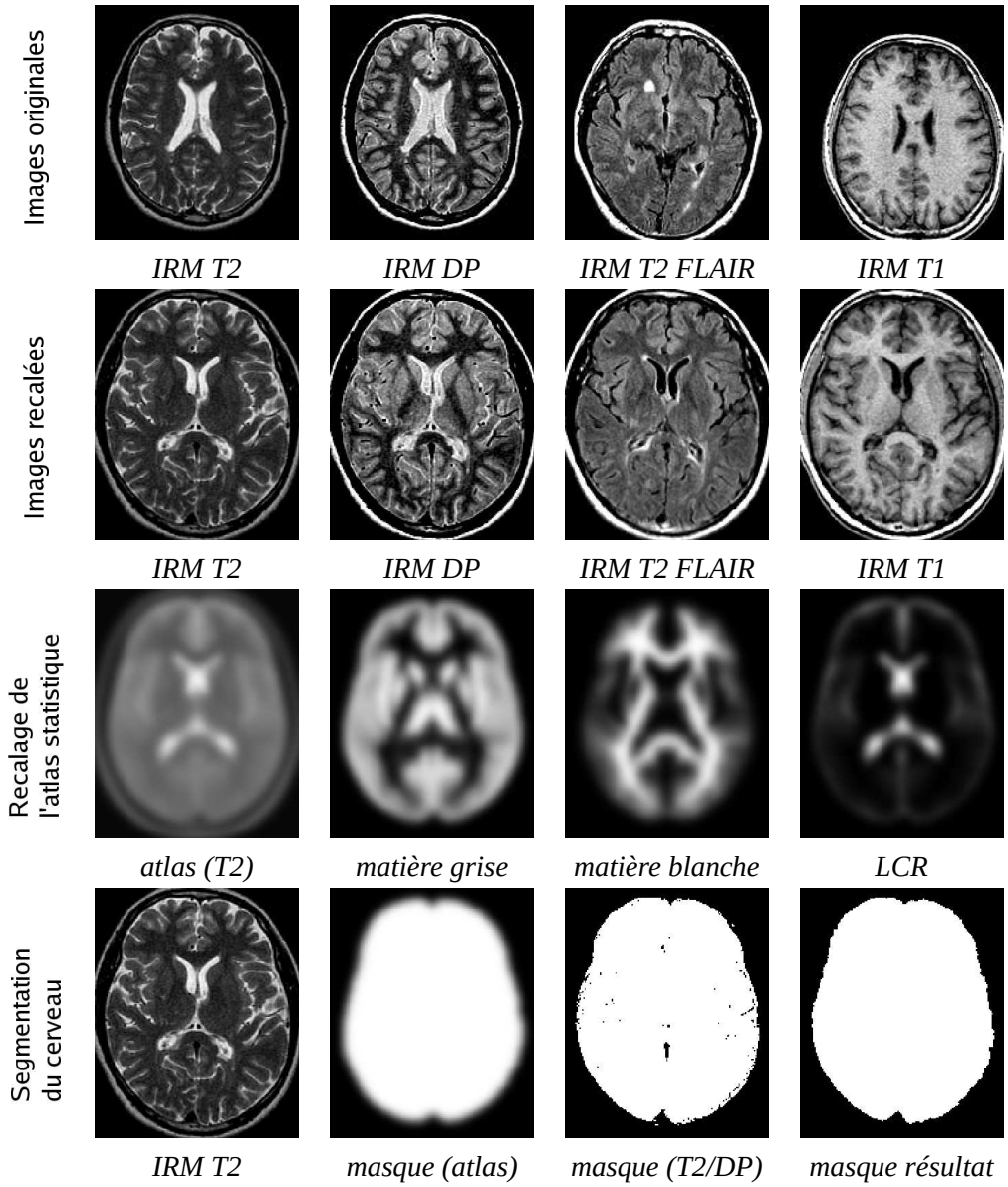


FIG. 6.9 – Présentation de l'ensemble des prétraitements. Le recalage de l'atlas statistique est un recalage rigide puis affine entre les IRM T2 de l'atlas (image moyenne) et du patient. La segmentation du masque est obtenue par un algorithme EM à 4 classes et une régularisation du masque à l'aide d'opérations de morphologie mathématique.

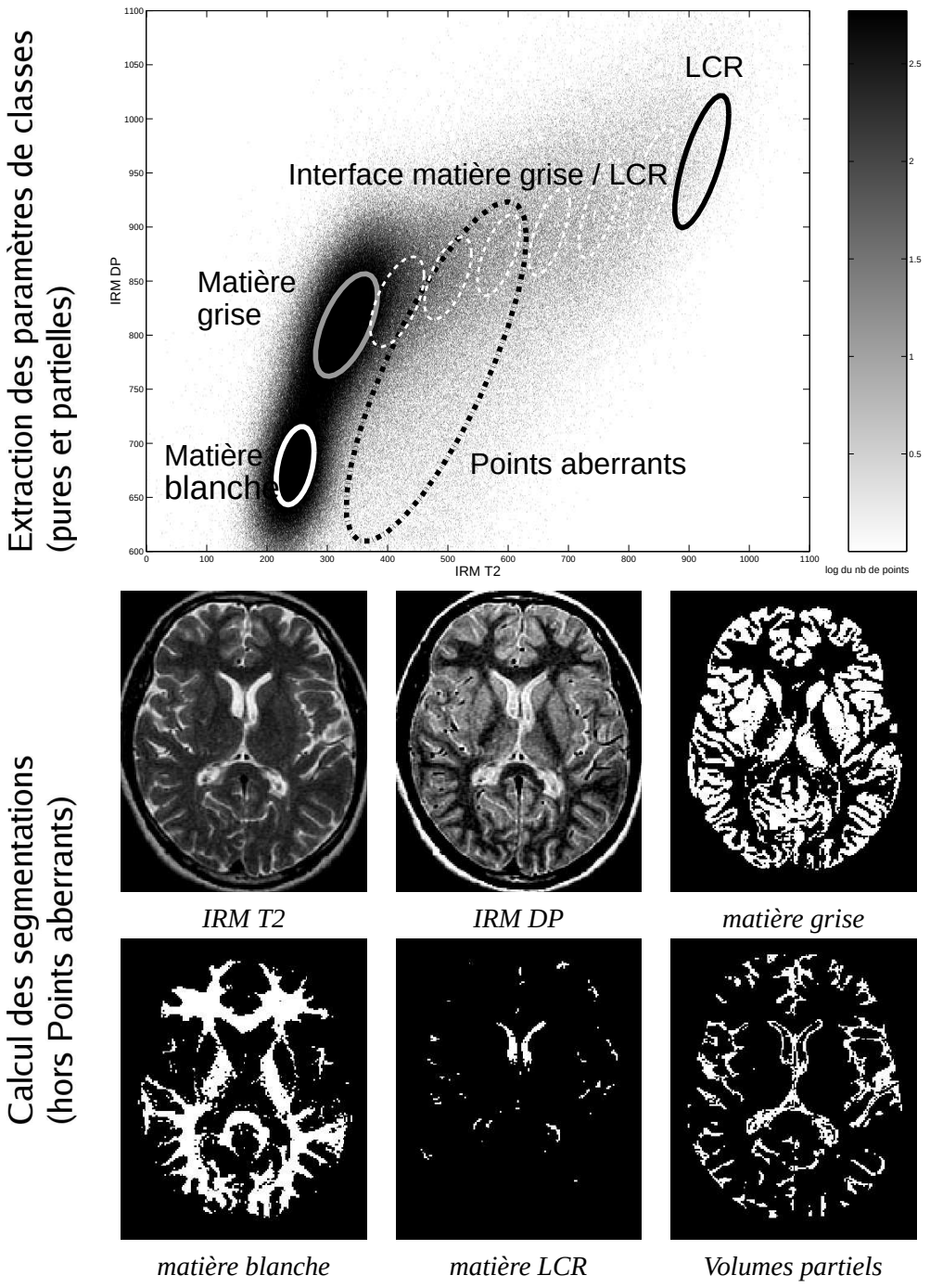


FIG. 6.10 – Segmentation en tissus sur le couple IRM T2/DP. L'extraction des paramètres de classes est effectuée à partir de l'algorithme EM modifié suivant les indications données dans le chapitre 5. Par contre, les segmentations sont obtenues *a posteriori*, sans tenir compte de la classe *Point aberrants*, et avec un seuil d'appartenance à chaque classe de 3σ.

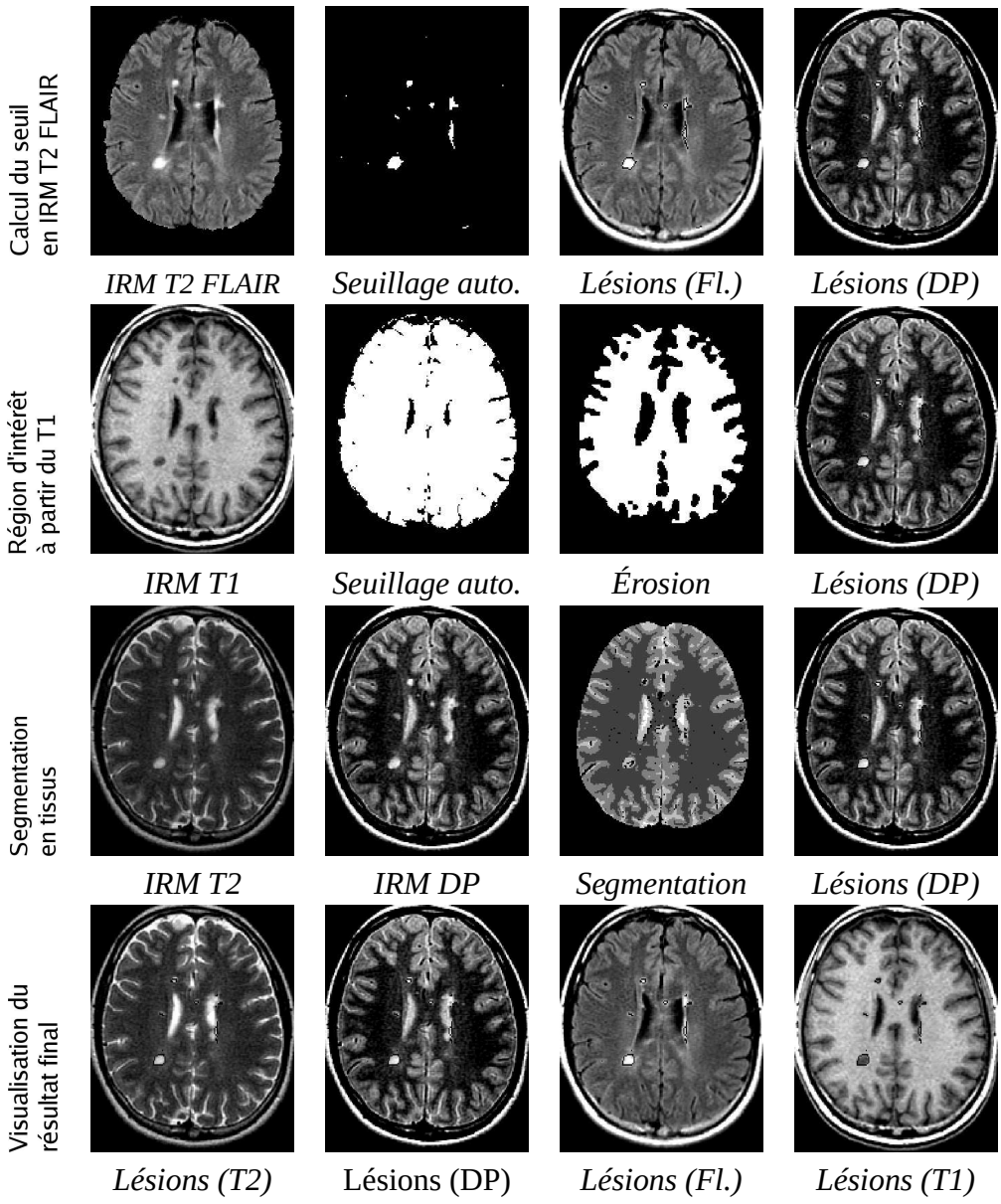


FIG. 6.11 – Visualisation de la chaîne de traitement après la segmentation en tissus. Les segmentations issues du couple T2/DP sont utilisées pour calculer un seuillage en T2 FLAIR (détection des lésions), et en T1 (segmentation du parenchyme cérébral). Les voxels labélisés comme de la matière blanche ou du LCR sont alors supprimés du masque des lésions dans la version finale du résultat.

Chapitre 7

Présentation des résultats et évaluation

En règle générale, effectuer une évaluation d'un système de traitement d'image médicale, quel qu'il soit, n'est pas une chose aisée. La résolution, le contraste, la finesse des images varient en fonction du protocole d'acquisition, voire de la machine utilisée. Pour un problème de segmentation, les structures à détecter sont plus ou moins visibles en fonction de leur taille et de leur contraste : une variabilité intra- et inter-expert est donc présente, et ne facilite pas l'établissement d'un résultat optimal dont l'algorithme de segmentation automatique doit se rapprocher, même si de récentes publications montrent que l'établissement d'un tel étalon or est possible [172, 173].

Dans le cas de la segmentation des lésions de SEP, le problème est encore plus compliqué. Les buts d'un système de segmentation automatique d'IRM cérébrales pour la SEP sont en effet multiples : aide au diagnostic, suivi de l'évolution de l'état du patient, test clinique, etc. Pour comparer le masque binaire obtenu par segmentation automatique avec la segmentation manuelle, une simple comparaison voxel à voxel ne suffit donc pas à mesurer la qualité du travail produit. D'un autre côté, la reproductibilité d'un quantificateur est difficile à juger, d'autant plus que nous n'avions à notre disposition qu'un seul expert pour l'évaluation, ce qui nous a interdit tout étude de la variabilité inter-experts. Avec plusieurs experts, un

algorithme de type STAPLE [172] aurait pu grandement nous aider dans l'établissement d'un étalon or.

La chaîne de traitements a été présentée au cours des chapitres précédents. L'ensemble des prétraitements, explicités au cours du chapitre 3, place les images dans un espace spatial fixé ; la segmentation en tissus, obtenue au cours des chapitres 4 et 5, permet d'automatiser les différentes étapes du chapitre 6 et de raffiner le traitement des inhomogénéités en fournissant une segmentation en tissus plus précise. Il reste donc à choisir les quantificateurs qui vont permettre une évaluation objective et à expliciter les résultats sur la segmentation des lésions.

7.1 Choix du protocole d'évaluation : établissement d'un étalon or

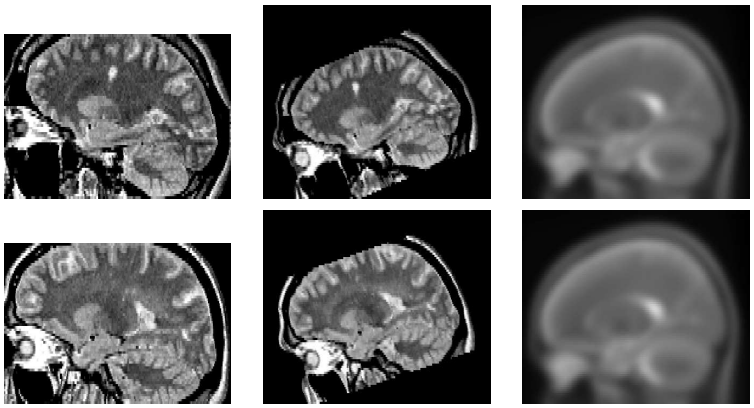
Comme le paragraphe 6.2 du chapitre 6 le montre, la segmentation se décompose en deux problèmes : la détection des lésions et leur contourage. Avant de présenter le choix de la méthode d'évaluation, il est important de savoir ce que l'on cherche. Les critères de Barkhof nécessitent essentiellement des informations sur la présence ou non de différentes lésions : c'est donc la détection qui est ici mise en œuvre. Par contre, pour le calcul de la charge lésionnelle T2 (volume des lésions en IRM T2/DP), le contourage a une grande importance. Nous avons donc essayé d'évaluer les deux points séparément.

Plusieurs directions sont possibles pour une évaluation. Une solution simple consiste à soumettre les résultats de l'algorithme de segmentation automatique au médecin, et de lui faire consigner ses modifications. Cette solution présente l'avantage d'être facile à mettre en œuvre, mais n'est valable que pour un seul système : toute modification de l'algorithme nécessite une nouvelle évaluation. La possibilité la plus pérenne est de faire construire une segmentation des lésions au médecin lui-même, et de prendre le résultat comme étalon or. Si plusieurs experts sont disponibles pour l'évaluation, l'étalon or peut être calculé à l'aide d'un algorithme de type STAPLE [172, 89, 173], mais ce n'est pas notre cas. Le processus de segmentation manuelle étant très long (de 2 à 5 heures par patient), il

est impossible pour des raisons pratiques de demander au médecin de réitérer plusieurs fois le contourage des lésions. Le processus peut être accéléré à l'aide d'une segmentation semi-automatique : le temps de segmentation par patient passe alors à 45 minutes en moyenne [63]. Mais ceci peut influencer le résultat, car il est difficile de construire une bonne interface homme / machine pour un résultat optimal [109].

Comme nous avons préféré garder la solution de l'étalon or à comparer avec les résultats obtenus par segmentation automatique, nous avons opté pour la solution suivante. Un opérateur aux connaissances suffisantes sur les IRM de SEP, mais sans étiquette d'expert effectue les segmentations manuelles. Le logiciel utilisé, *manualSegmentation*, fournit des régions d'intérêt. Les segmentations sont effectuées sur l'IRM DP, mais toutes les autres séquences – T2 FSE, T2 FLAIR, T1 – sont visibles lors du contourage des lésions pour assurer le caractère multi-séquences de l'évaluation. Pour se rapprocher le plus possible des conditions cliniques, les segmentations manuelles sont fournies à l'expert sous la forme de livrets, où toutes les coupes sont visibles. Chaque page contient les 4 séquences où les contours des lésions sont visibles en surimpression. Un deuxième livret est fourni sans les lésions pour que l'intégralité de l'image soit visible. L'expert médical peut alors annoter les livrets et effectuer une série de corrections dont la trace est gardée. L'opérateur peut alors corriger ses segmentations, et soumettre ses corrections au médecin. Au bout d'un certain nombre d'itérations, le processus converge et la segmentation manuelle est alors étiquetée comme étalon or. Ce processus présente l'avantage de ne pas gaspiller le temps passé par l'expert, tout en diminuant la variabilité intra-expert puisque plusieurs versions de la segmentation lui sont fournies.

10 patients de la base de données ont été choisis et leurs IRM ont été ainsi segmentées. Les segmentations ont été effectuées sur la coupe axiale de l'IRM DP, et ont pris environ 5 heures par patient, notamment pour assurer une cohérence en z (axe normal aux coupes axiales). Avant d'être contourées, les images ont été recalées sur le patient moyen fourni par l'atlas du MNI. Ceci permet d'effectuer le contourage sur un nombre fixe de coupes : le temps de travail sur chaque patient



Images originales *Images recalées* *Patient moyen*

FIG. 7.1 – Une coupe coronale de deux patients différents est ici visualisée, avant et après recalage sur le patient moyen. Avant d'effectuer le contourage manuel des lésions, les images sont recalées sur un patient moyen fixé. S'assurer que chaque patient a le même nombre de coupes à segmenter permet d'uniformiser la sensibilité du contourage manuel et ainsi de faciliter le travail de l'opérateur.

est alors grossièrement le même (figure 7.1). Chacune des lésions a également été classée dans 4 catégories :

- lésions juxtacorticales,
- lésions corticales,
- lésions périventriculaires,
- lésions de la fosse postérieure.

Nous avons également labélisé les lésions nécrotiques, dont le signal en IRM T2 FLAIR est différent de celui des autres lésions. Ces catégories ne sont pas exclusives ; elles permettent par contre de préciser les résultats pour différentes lésions dont les caractéristiques – taille, contraste, localisation – varient fortement.

7.2 Introduction à l'évaluation quantitative

Prenons le cas d'un test pharmaceutique sensé détecter si un sujet est sain (test négatif) ou malade (test positif). Pour valider l'efficacité du test, ce dernier est appliqué sur une base de sujets, sains ou malades. 4 populations vont se distinguer.

	Malade	Sain	Quantificateur
Test positif	VP	FP	$Vpp = \frac{VP}{VP+FP}$
Test négatif	FN	VN	$Vpn = \frac{VN}{VN+FN}$
Quantificateur	$Se = \frac{VP}{VP+FN}$	$Sp = \frac{VN}{VN+FP}$	

TAB. 7.1 – Évaluation de la qualité d'un test pharmaceutique. Sur des sujets sains ou malades (colonne), le test peut être positif ou négatif (ligne). Une bonne sensibilité est le signe qu'une large majorité de malades sont détectés comme malades – bonne détection – alors qu'une bonne spécificité indique que peu de sujets sains sont détectés comme malades – peu de fausses alertes.

- **Vrais positifs** (VP) : nombre de sujets malades détectés comme malades.
- **Faux négatifs** (FN) : nombre de sujets malades détectés comme sains.
- **Faux positifs** (FP) : nombre de sujets sains détectés comme malades.
- **Vrai négatifs** (VN) : nombre de sujets sains détectés comme sains.

Si le test est idéal, il n'y aura ni faux positif, ni faux négatif. Pour évaluer la qualité du test, quatre quantificateurs sont possibles (tableau 7.1).

- **Sensibilité** (Se) : si le sujet est malade, probabilité qu'il soit détecté comme malade.
- **Spécificité** (Sp) : si le sujet est sain, probabilité qu'il soit détecté comme sain.
- **Valeur prédictive positive** (Vpp) : si le test est positif, probabilité que le sujet soit effectivement malade.
- **Valeur prédictive négative** (Vpn) : si le test est négatif, probabilité que le sujet soit effectivement sain.

En pratique, quand un médecin reçoit le résultat d'un examen complémentaire, positif ou négatif, il ne sait pas si le patient souffre de l'affection qu'il cherche à diagnostiquer ou non, et les probabilités qui l'intéressent s'expriment de la manière suivante : quelle est la probabilité de présence de la maladie M chez ce patient, sachant que l'examen a donné un résultat positif (ou négatif) ?

Ce sont donc les valeurs prédictives qui correspondent aux préoccupations des médecins, et elles pourraient sembler les "meilleurs" paramètres d'évaluation. Pourtant, en réalité, c'est la sensibilité et la spécificité qui sont le plus souvent uti-

lisées pour évaluer les examens complémentaires. La raison en est la suivante : la sensibilité d'un examen pour une affection repose sur la définition de la population des "malades", et est donc caractéristique de la maladie et du signe. En particulier, elle n'est pas susceptible de varier d'un centre à l'autre (d'un service hospitalier spécialisé à une consultation de médecin généraliste, par exemple). Le même raisonnement peut s'appliquer à la spécificité, si on considère qu'elle repose aussi sur la définition de la maladie.

Les valeurs prédictives, au contraire, sont fonctions des proportions respectives de malades et de non-malades dans la population (de la prévalence de la maladie). Or ces proportions sont dépendantes des centres considérés ; les valeurs prédictives des examens varient donc d'un centre à l'autre pour une même maladie, ce qui explique qu'elles sont moins utilisées comme paramètre d'évaluation, même si elles sont intéressantes à connaître pour un centre donné.

Courbe ROC Lorsqu'un examen fournit des résultats de type continu, il faut déterminer le meilleur seuil entre les valeurs pathologiques et les valeurs normales. L'idéal serait d'obtenir une sensibilité et une spécificité égales à 1. Ce n'est généralement pas possible, et il faut tenter d'obtenir les plus fortes valeurs pour ces deux paramètres, sachant qu'ils varient en sens inverse. On s'aide pour ce choix d'un outil graphique, la courbe ROC (*Receiver Operating Characteristics*, figure 7.2) qui est le tracé des valeurs de la sensibilité Se en fonction de $1 - Sp$. Il suffit alors de chercher le point de la courbe qui se rapproche le plus du point de coordonnées $(Se = 1, 1 - Sp = 0)$.

7.3 Évaluation quantitative de la segmentation des lésions de SEP

Remarques : Les lésions de la fosse postérieure sont peu visibles en IRM T2 FLAIR : comme cette séquence a été choisie pour la détection des lésions, la plupart des lésions de la fosse postérieure ne sont pas détectées. Parfois, la correction du biais sur cette séquence permet de faire légèrement ressortir ces lésions. Lors

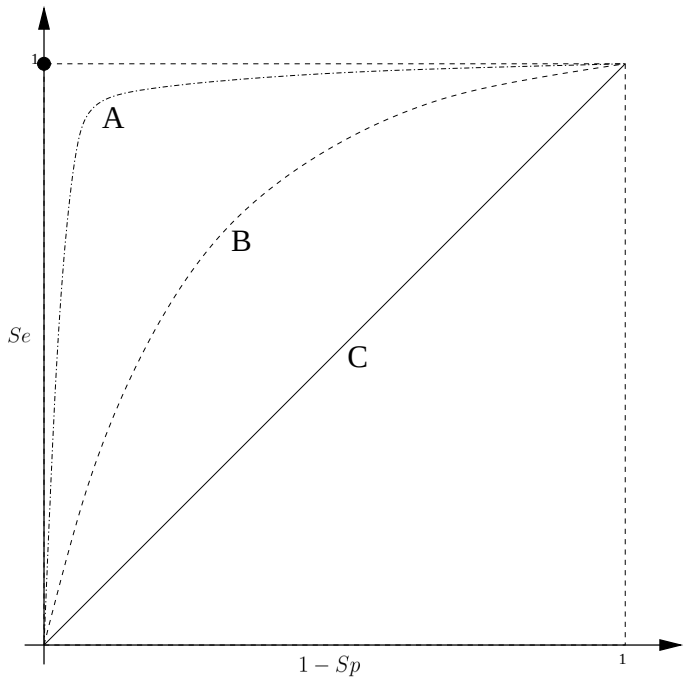


FIG. 7.2 – Courbe ROC (*Receiver Operating Characteristics*) pour 3 examens fictifs. Lorsqu'un examen fournit des résultats de type continu, il faut tenter d'ajuster le meilleur seuil pour obtenir les plus fortes valeurs de sensibilité et de spécificité, sachant qu'elles varient en sens inverse. Dans cette figure, la courbe A correspond à un bon critère diagnostic pour lequel on peut obtenir simultanément des valeurs élevées de sensibilité et de spécificité.

des statistiques effectuées dans ce chapitre, nous avons donc délibérément éliminé les lésions de la fosse postérieure pour mieux coller au domaine d'utilisation de l'IRM T2 FLAIR. Le traitement de ces lésions sera abordé dans le chapitre 8, sous formes de perspectives. Nous avons évalué séparément la détection et le contourage : les quantificateurs sont différents car les valeurs habituelles de sensibilité et de spécificité ne sont pas directement exploitables.

Les contourages manuels, comme le montre la figure 7.1, ont été réalisés sur les IRM des patients recalées sur le patient moyen. Par contre, les résultats chiffrés qui vont suivre sont calculés à partir des labélisations binaires reportées sur les images originales : le recalage sur le patient moyen étant un recalage affine, ceci permet d'éviter de fausser le calcul des volumes.

7.3.1 Détection

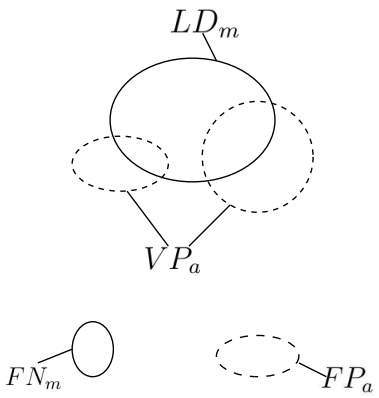
Nous avons choisi la lésion comme unité. Sur les contourages manuels comme sur les segmentations automatiques des lésions de SEP, une labélisation binaire des voxels est obtenue, et une extraction en composantes connexes est effectuée. Chacune des composantes connexes dans les segmentations automatiques et dans les contourages manuels est appelée respectivement *lésion automatique* et *lésion manuelle*.

Dans la liste suivante, on définit les termes essentiels, l'acronyme entre parenthèses étant le quantificateur associé : ainsi, VP_a est le nombre de vrais positifs parmi l'ensemble des lésions automatiques et $vol(VP_a)$ est le volume représenté, en voxels, dans la labélisation binaire automatique. Les IRM suivent toutes le même protocole, ce qui explique le choix du voxel comme unité : ses dimensions sont 1*1*2 mm^3. Les valeurs suivantes sont présentées de manière graphique dans la figure 7.3, et de manière analogue à la section 7.2 dans le tableau 7.2.

- **Vrais positifs** (VP_a) : nombre de lésions automatiques ayant une intersection non-nulle avec une des lésions manuelles.
- **Faux positifs** (FP_a) : nombre de lésions automatiques n'ayant aucune intersection avec aucune des lésions manuelles.
- **Lésions automatiques** (L_a) : nombre de composantes connexes dans l'image binaire issue de la segmentation automatique. La formule simple $VP_a + FP_a = L_a$ découle des trois définitions ci-dessus.
- **Lésions détectées** (LD_m) : nombre de lésions manuelles ayant une intersection non-nulle avec une des lésions automatiques.
- **Faux négatifs** (FN_m) : nombre de lésions manuelles n'ayant aucune intersection avec aucune des lésions automatiques.
- **Lésions manuelles** (L_m) : nombre de composantes connexes dans l'image binaire issue du contourage manuel. La formule simple $LD_m + FN_m = L_m$ découle des trois définitions ci-dessus.

Des données précédentes, on peut extraire les quantificateurs suivants.

- **Sensibilité** (Se) : rapport du nombre de lésions détectées sur le nombre total



$$Se = \frac{LD_m}{LD_m + FN_m} = \frac{LD_m}{L_m}$$

$$Vpp = \frac{VP_a}{VP_a + FP_a} = \frac{VP_a}{L_a}$$

FIG. 7.3 – Calcul des quantificateurs pour évaluer la détection : comparaison entre les lésions manuelles (trait continu) et les lésions automatiques (en pointillé). Par rapport au tableau classique de validation d'un test pharmaceutique (tableau 7.1), deux difficultés se créent : la notion de faux négatif n'est pas accessible, et les vrais positifs sont définis comme des **vrais positifs** automatiques VP_a ou des **lésions détectées** LD_m selon que l'on regarde les lésions automatiques ou les lésions manuelles. La **sensibilité** Se reste d'actualité, mais la spécificité n'est pas calculable faute de vrais négatifs : le poids des faux positifs est regardé via la **valeur prédictive positive** Vpp.

de lésions manuelles :

$$Se = \frac{LD_m}{LD_m + FN_m} = \frac{LD_m}{L_m}$$

– **Sensibilité volumique** (SeV) : rapport du volume de lésions détectées sur le volume total de lésions manuelles :

$$SeV = \frac{vol(LD_m)}{vol(LD_m) + vol(FN_m)} = \frac{vol(LD_m)}{vol(L_m)}$$

– **Valeur prédictive positive** (Vpp) : rapport du nombre de vrai positifs sur le nombre total de lésions automatiques :

$$Vpp = \frac{VP_a}{VP_a + FP_a} = \frac{VP_a}{L_a}$$

Détection : calcul des quantificateurs

	Lésion manuelle	Volume labélisé comme sain	Quantificateur
Lésion automatique	VP_a / LD_m	FP_a	$Vpp = \frac{VP_a}{VP_a+FP_a}$
Volume détecté comme sain	FN_m		
Quantificateur	$Se = \frac{LD_m}{LD_m+FN_m}$		

TAB. 7.2 – Pour la détection, les unités sont le nombre de lésions manuelles et le nombre de lésions automatiques. Selon que l'on regarde l'un ou l'autre, vrais positifs du tableau 7.1 deviennent alors les **vrais positifs** VP_a (lésions automatiques) et les **lésions détectées** LD_m (lésions manuelles).

– **Valeur prédictive positive volumique** ($VppV$) : rapport du volume de ces vrais positifs sur le volume total des lésions automatiques :

$$VppV = \frac{vol(VP_a)}{vol(VP_a) + vol(FP_a)} = \frac{vol(VP_a)}{vol(L_a)}$$

Bien que le nombre de vrais négatifs (voir tableau 7.1) ne soit pas calculable quand l'unité est la lésion, on peut considérer que l'ensemble des vrais négatifs est le volume matière blanche $vol(MB)$[1] ni labélisé ni détecté comme lésion. On peut alors définir la **spécificité volumique** SpV par analogie avec la spécificité comme suit :

$$1 - SpV = \frac{vol(FP_a)}{vol(FP_a) + vol(MB)}$$

Les quantificateurs SeV et $VppV$ permettent de mieux se rendre compte du poids des faux positifs et des faux négatifs en termes de volume : ainsi, si les faux positifs sont nombreux en termes de composantes connexes, mais de faible volume, la valeur prédictive positive sera médiocre, mais la valeur prédictive po-

[1] D'un point de vue pratique, ce volume est calculé à l'aide de la segmentation de la matière blanche fournie au chapitre 5.

Détection : nombre de lésions

Patient	VP_a	FP_a	L_a	LD_m	FN_m	L_m
MCI01	25	9	34	28	9	37
MCI02	8	7	15	11	9	20
MCI04	5	13	18	5	1	6
MCI06	27	10	37	31	14	45
MCI07	47	10	57	56	11	67
Valeur moyenne : par patient	22.4	9.8	32.2	26.2	8.8	35

TAB. 7.3 – Visualisation de la qualité de la détection, en termes de nombre de lésions. Il y a en moyenne $VP_a = 22.4$ vrais positifs et $FP_a = 9.8$ faux positifs parmi les $L_a = 32.2$ lésions automatiques, et $LD_m = 26.2$ lésions détectées et $FN_m = 8.8$ faux négatifs parmi les $L_m = 35$ lésions manuelles.

sitive volumique sera très proche de 1, comme c'est le cas sur nos images.

Les résultats sur l'ensemble de la base de données sont présentés dans les tableaux 7.3, 7.4 et 7.5 pour 5 patients représentatifs des différents problèmes rencontrés et dont l'intégralité des coupes est disponible en annexe de ce manuscrit.

Au regard des chiffres présentés dans le tableau 7.5, le système de segmentation peut être qualifié de sensible : $Se = 74.9$ % des lésions sont détectées, et ces lésions représentent $SeV = 95.4$ % de la masse lésionnelle globale. $1 - Vpp = 29.5$ % des lésions automatiques sont des faux positifs, mais ils représentent $1 - VppV = 5.7$ % de la masse lésionnelle globale détectée. Il est intéressant de noter que l'ensemble des faux positifs représente un volume relativement stable, compris entre 150 et 350 voxels, qui représentent $1 - SpV = 0.115$ % du volume de la matière blanche.

Pour la segmentation des lésions de sclérose en plaques, une bonne sensibilité est très importante. Dans le cadre d'un système d'aide au diagnostic, par exemple, un faux négatif – lésion manuelle non détectée par le système – doit être cherché par le médecin dans l'ensemble de l'image, ce qui coûte énormément de temps. Par contre, un faux positif ou un mauvais contourage peut être corrigé *a posteriori* : une correction manuelle effectuée par un expert peut être ajoutée, couplée à un système de segmentation semi-automatique, par exemple.

Détection : volumes lésionnels (voxels)

Patient	$v(FP_a)$	$v(L_a)$	$v(FN_m)$	$v(L_m)$	$v(MB)$
MCI01	308	3120	317	4723	242885
MCI02	200	4734	396	7873	249848
MCI04	359	1881	3	1456	227522
MCI06	304	5665	601	7945	252798
MCI07	185	8448	481	19000	199928
Volume moyen par patient	271.2	4769.6	359.6	8199.4	234596
Volume moyen par lésion	27.67	148.1	40.9	234.3	

TAB. 7.4 – Visualisation de la qualité de la détection en termes de volume lésionnel. Les $FP_a = 9.8$ faux positifs du tableau 7.3 ne représentent que $v(FP_a) = 271.2$ voxels parmi les $v(L_a) = 4769.6$ voxels détectés comme lésion ; les $FN_m = 8.8$ faux négatifs du tableau 7.3 ne représentent que $v(FN_m) = 359.6$ voxels parmi les $v(L_m) = 8199.4$ voxels labélisés comme lésion lors des contourages manuels.

Détection : quantificateurs globaux (pourcentage)

Patient	Se	SeV	$1 - Vpp$	$1 - VppV$	$1 - SpV$
MCI01	75.7	93.3	26.5	9.9	0.127
MCI02	55	94.6	46.7	4.2	0.080
MCI04	83.3	99.8	72.2	19.1	0.158
MCI06	68.9	92.4	54.1	5.4	0.120
MCI07	83.6	97.5	17.5	2.2	0.092
Ensemble des patients	**74.9**	**95.6**	**30.5**	**5.7**	**0.115**

TAB. 7.5 – Visualisation de la qualité de la détection en termes de quantificateurs. $Se = \mathbf{74.9}$ % des lésions sont détectées, et ces lésions représentent $SeV = \mathbf{95.6}$ % du volume lésionnel segmenté manuellement ; $1 - Vpp = \mathbf{30.5}$ % des lésions automatiques sont des faux positifs, mais ces faux positifs ne représentent que $1 - VppV = \mathbf{5.7}$ % du volume lésionnel détecté et $1 - SpV = \mathbf{0.115}$ % du volume de la matière blanche.

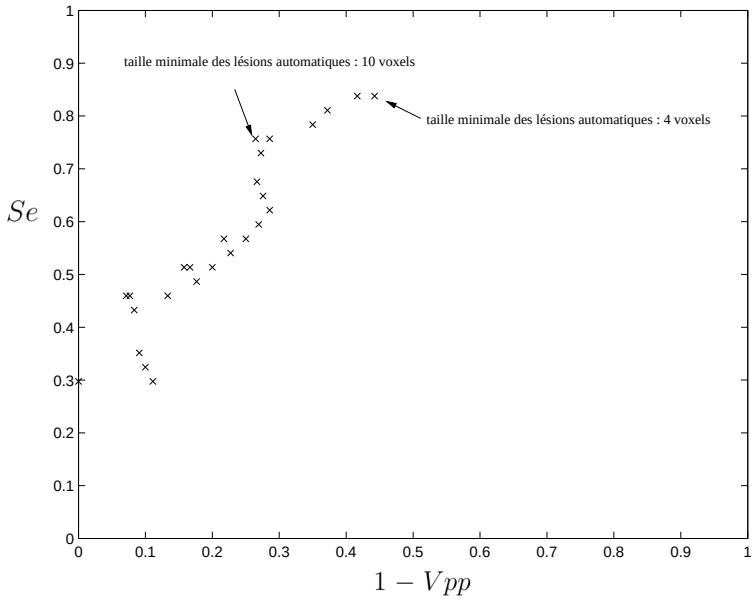


FIG. 7.4 – Évolution du couple sensibilité / valeur prédictive positive en fonction de la taille minimale des lésions automatiques pour le patient MCI01. Pour une taille minimale des lésions automatiques fixée à 4 voxels, la sensibilité atteint $Se = 83.8$ % pour un taux de faux positifs de $1 - Vpp = 44.2$ % ; les lésions détectées représentent alors $SeV = 97.5$ % du volume des lésions manuelles et les faux positifs obtenus, $1 - VppV = 13.6$ % du volume total des lésions automatiques.
Les résultats présentés dans ce chapitre ont été calculés pour une taille minimale des lésions automatiques de 10 voxels pour équilibrer sensibilité et valeur prédictive positive. Pour ce patient, les quantificateurs sont alors les suivants : $Se = 75.7$ %, $1 - Vpp = 26.5$ % ; $SeV = 93.3$ %, $1 - VppV = 9.9$ %.

Qualité de la détection en fonction du type de la lésion : nous avons vu dans le paragraphe 7.1 que les lésions manuelles étaient différenciées en 5 catégories : juxtacorticales, corticales, périventriculaires, fosse postérieure et nécrotiques. Les lésions de la fosse postérieure, invisibles en IRM T2 FLAIR, ont été écartées (voir paragraphe 7.3). Voici les sensibilités Se pour chacune de ces catégories :

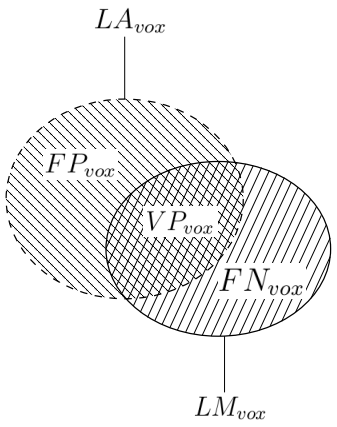
- lésions juxtacorticales : 61.4 % ;
- lésions corticales : 81.6 % ;
- lésions périventriculaires : 94.6 % ;
- lésions nécrotiques : 100 %.

Les lésions juxtacorticales sont les plus touchées par les faux négatifs : 38.6% des lésions juxtacorticales ne sont pas détectées. La résolution de la séquence IRM T2 FLAIR – 1*1*4 mm^3 – et le faible volume des lésions manuelles concernées – une dizaine de voxels pour les plus petites – sont ici à mettre en cause : les effets de volume partiel réduisent fortement le contraste des lésions, et donc la qualité de la détection. Par contre, les lésions périventriculaires sont les mieux détectées et représentent 69.5% de la charge lésionnelle. Malgré leur faible contraste en IRM T2 FLAIR, les lésions corticales sont bien détectées. Bien que seule leur périphérie soit segmentée correctement, toutes les lésions nécrotiques sont détectées.

7.3.2 Contourage

Après avoir étudié la qualité de la détection, il faut examiner de plus près la qualité du contourage. Pour ce faire, reprenons les définitions analogues à celles proposées dans la section 7.2. L'unité de travail est le voxel. Pour chaque définition, l'acronyme est le quantificateur associé : VP_{vox} est le volume de vrais positifs, en voxels. Les valeurs suivantes sont présentées de manière graphique dans la figure 7.5, et de manière analogue à la section 7.2 dans le tableau 7.6.

- **Voxels automatiques** (A_+) **:** nombre de voxels détectés comme lésion.
- **Voxels manuels** (M_+) **:** nombre de voxels labélisés comme lésion dans les contourages manuels.
- **Vrais positifs** (VP_{vox}) **:** nombre de voxels détectés comme lésion, et labélisés comme tel dans les contourages manuels.



$$Se_{vox} = \frac{VP_{vox}}{VP_{vox} + FN_{vox}} = \frac{VP_{vox}}{LM_{vox}}$$

$$Vpp_{vox} = \frac{VP_{vox}}{VP_{vox} + FP_{vox}} = \frac{VP_{vox}}{LA_{vox}}$$

FIG. 7.5 – Calcul des quantificateurs pour évaluer le contourage : comparaison entre les lésions manuelles (trait continu) et les lésions automatiques (en pointillé). Pour n'évaluer que le contourage, il ne faut s'intéresser qu'aux lésions manuelles et aux lésions automatiques dont l'intersection est non-nulle. L'unité est ici le voxel : les vrais positifs sont représentés par un nombre unique VP_{vox}. Comme pour la détection, la sensibilité voxellique Se_{vox} et la valeur prédictive voxellique Vpp_{vox} permettent d'observer la qualité du contourage.

- **Faux positifs** (FP_{vox}) **:** nombre de voxels détectés comme lésion, mais labélisés comme tissu sain dans les contourages manuels.
- **Faux négatifs** (FN_{vox}) **:** nombre de voxels labélisés comme lésion dans les contourages manuels, mais détectés comme tissu sain.

A partir de ces données, on peut estimer la qualité du contourage (figure 7.5) :

- **Sensibilité voxellique** (Se_{vox}) **:** calculée sur l'ensemble des lésions manuelles détectées, elle permet d'estimer les sous-segmentations du système par rapport aux contourages manuels :

$$Se_{vox} = \frac{VP_{vox}}{VP_{vox} + FN_{vox}} = \frac{VP_{vox}}{LM_{vox}}$$

- **Valeur prédictive positive voxellique** (Vpp_{vox}) **:** calculée sur l'ensemble des lésions automatiques ayant une intersection avec une des lésions manuelles, elle permet d'estimer les sur-segmentations du système par rapport

Contourage : calcul des quantificateurs

Voxels	Labélisés comme lésion	Labélisés comme sains	Quantificateur
Détectés comme lésion	VP_{vox}	FP_{vox}	$Vpp_{vox} = \frac{VP_{vox}}{VP_{vox}+FP_{vox}}$
Détectés comme sains	FN_{vox}		
Quantificateur	$Se_{vox} = \frac{VP_{vox}}{VP_{vox}+FN_{vox}}$		

TAB. 7.6 – Pour le contourage, l'unité est le voxel : les vrais positifs sont donc représentés cette fois-ci par un unique nombre VP_{vox}, qui est le nombre de voxels labélisés et détectés comme lésion. Seule la qualité du contourage est évaluée : on ne s'intéresse qu'aux lésions manuelles et automatiques ayant une intersection non-nulle. Parler de vrais négatifs n'a donc pas de sens ici : les faux positifs FP_{vox} sont donc évalués via la valeur prédictive positive voxellique Vpp_{vox}.

aux contourages manuels.

$$Vpp_{vox} = \frac{VP_{vox}}{VP_{vox} + FP_{vox}} = \frac{VP_{vox}}{LA_{vox}}$$

A titre de comparaison, un quantificateur similaire, appelé *similarity index* a été utilisé à la place du couple (Se_{vox}, Vpp_{vox}) dans [4] :

$$Si = \frac{2 * VP_{vox}}{2 * VP_{vox} + FP_{vox} + FN_{vox}}$$

Nous avons cependant préféré séparer l'examen des deux valeurs FN_{vox} et FP_{vox}, et c'est pourquoi deux quantificateurs sont fournis au lieu d'un.

Les résultats sont visualisés dans les tableaux 7.7 et 7.8. La valeur prédictive positive voxellique est de 70.4 % : 29.6 % de la charge lésionnelle calculée automatiquement correspond à des faux positifs, en termes de voxels, parmi les lésions automatiques ayant une intersection avec une lésion manuelle. La sensibilité voxellique est de 40.4 % en moyenne. Si une érosion 2D est effectuée sur la labélisation manuelle, cette sensibilité voxellique passe à 49.2 % ; pour les lésions périventriculaires, ce chiffre passe de 50.0 à 68.0 %, ce qui montre que le système tend à sous-segmenter la plupart des lésions. Les résultats sont donc perfectibles

Contourage : volumes lésionnels (voxels)

Patient	VP_{vox}	FN_{vox}	FP_{vox}
MCI01	1865	2541	947
MCI02	2674	4803	1860
MCI04	915	538	607
MCI06	4106	3238	1255
MCI07	6278	12241	1985
Valeur moyenne	3167.6	4672.2	1330.8

TAB. 7.7 – Visualisation de la qualité du contourage : on regarde ici les lésions manuelles détectées et les lésions automatiques qui sont des vrais positifs.

Contourage : quantificateurs globaux (pourcentage)

Patient	Se_{vox}	$1 - Vpp_{vox}$
MCI01	42.3	33.7
MCI02	34.5	41.0
MCI04	62.3	39.9
MCI06	55.9	23.4
MCI07	33.9	24.4
Ensemble des patients	**40.4**	**29.6**

TAB. 7.8 – Visualisation de la qualité du contourage : on regarde ici les lésions manuelles détectées et les lésions automatiques qui sont des vrais positifs. Parmi les voxels labélisés comme lésion dans les contourages manuels, $Se_{vox} = \mathbf{40.4}$ % sont détectés comme lésion ; parmi les voxels détectés comme lésion, $1 - Vpp_{vox} = \mathbf{29.6}$ % sont des faux positifs.

Contourage : Sensibilité voxellique par type de lésion (pourcentage)

Type de lésion	Se_{vox}	Se_{vox} (érosion des lésions manuelles)
Lésions péri-ventriculaires	52.0	68.0
Lésions juxta-corticales	41.9	68.1
Lésions corticales	40.8	51.0
Lésions nécrotiques	19.2	20.4
Ensemble des lésions	**40.4**	**49.2**

TAB. 7.9 – Tableau de résultats, hors lésions de la fosse postérieure. La sensibilité voxellique est calculée au sein de chaque type de lésion – périventriculaire, juxta-corticale, corticale ou nécrotique. Si on érode les segmentations manuelles (colonne de droite), la sensibilité voxellique augmente sensiblement, ce qui montre que le système peine à contourer la périphérie des lésions.

en termes de contourage. Nous tenterons de trouver quelques explications à ces résultats et présenter des perspectives de recherche dans le chapitre suivant.

7.4 Visualisation d'un patient : MCI01

Parmi les IRM de la base de données, certaines sont plus faciles à traiter que d'autres. Une autre maladie peut se superposer à la sclérose en plaques, ce qui peut prêter à confusion. Même si toutes les lésions sont des lésions de SEP, des problèmes à l'acquisition peuvent diminuer la qualité des images. Certaines lésions, plus difficiles à segmenter, peuvent altérer le résultat. Le patient dont les coupes axiales sont affichées dans ce paragraphe représente un patient standard, sans problème particulier : les lésions y sont correctement contrastées, les images sont de bonne qualité. Certes, la variabilité sur la structure d'une lésion fait parfois échouer l'algorithme, dans la détection ou le contourage ; mais on ne trouve pas de lésions périventriculaires que l'on pourrait confondre avec les artefacts de flux des cornes des ventricules, ou de lésions nécrotiques démesurées connexes au LCR cortical : globalement, les résultats sont satisfaisants. D'autres images plus délicates à segmenter seront présentées dans le chapitre suivant, ce qui nous permettra de conclure tout en proposant des perspectives de recherche intéressantes.

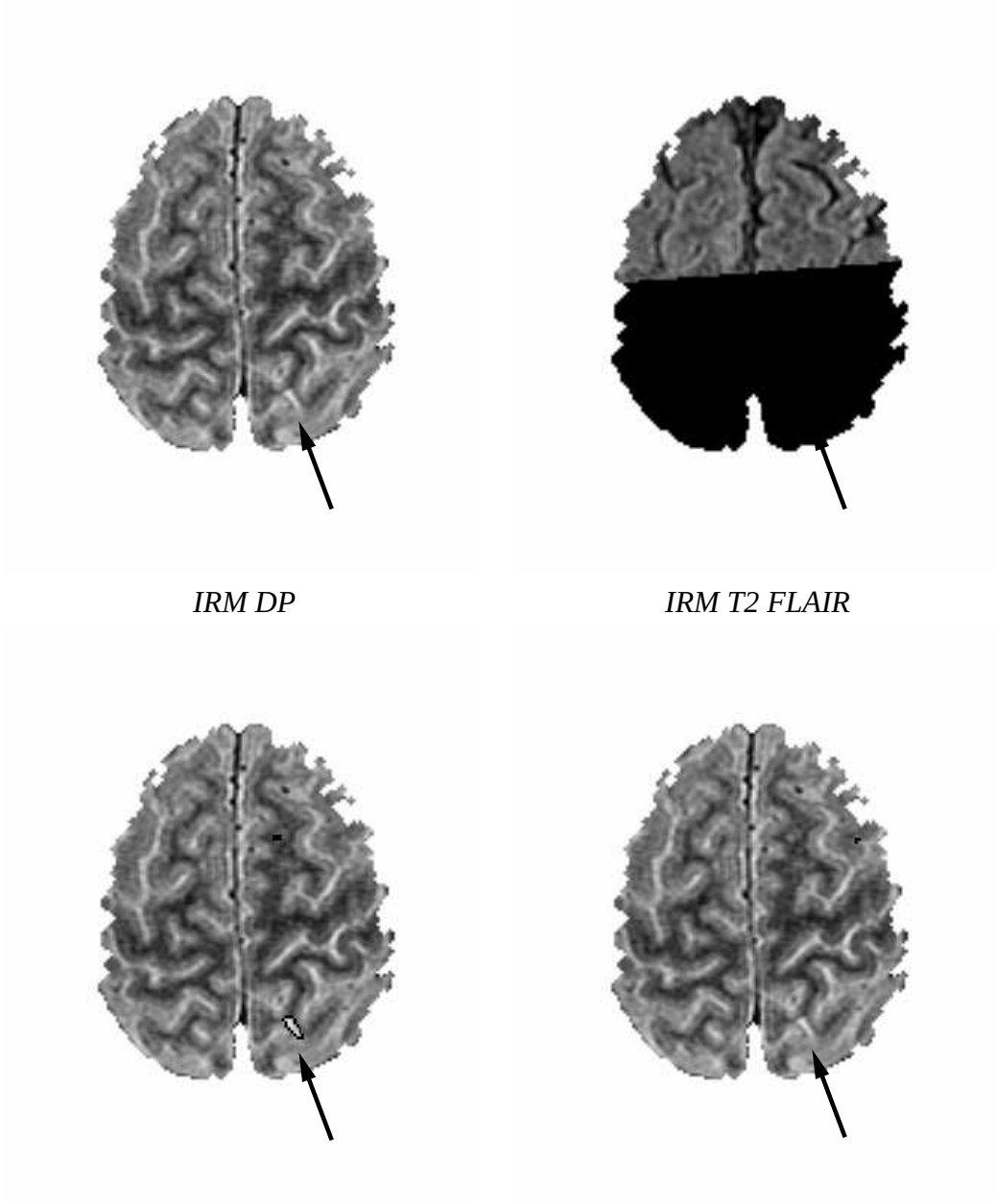


FIG. 7.6 – Parfois, l'IRM T2 FLAIR est tronquée et l'information multi-séquences est alors incomplète. La détection étant basée sur cette séquence, elle échoue pour la lésion corticale droite (flèche).

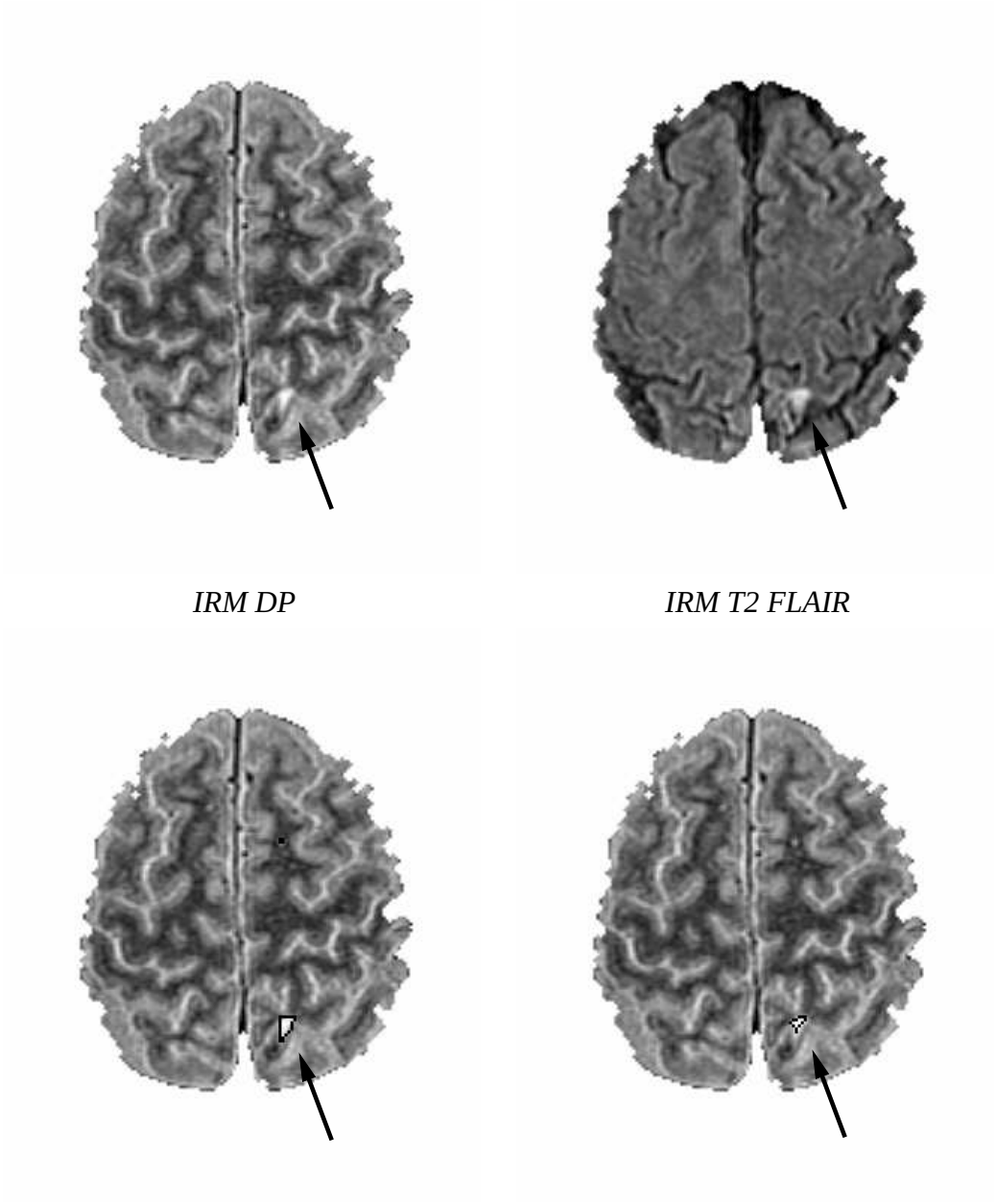


FIG. 7.7 – Lorsque la donnée multi-séquences est complète, la détection est correctement effectuée, alors que ce n'était pas le cas dans la figure 7.6. La chaîne de traitement produit dans notre cas une légère sous-estimation du volume de la lésion.

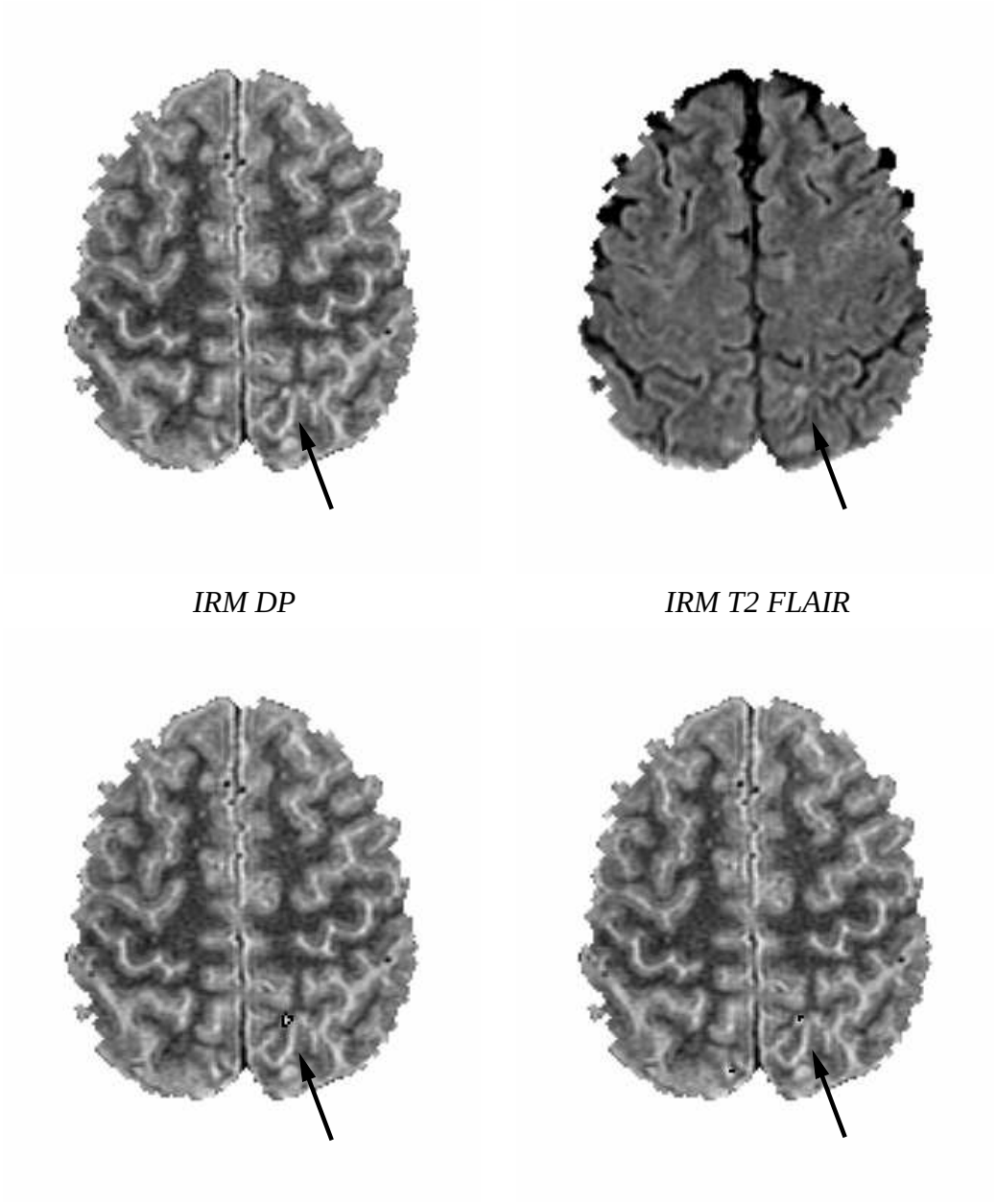


FIG. 7.8 – La lésion corticale (flèche) est bien détectée, mais comme dans la figure 7.7, il y a ici sous-estimation du volume lésionnel.

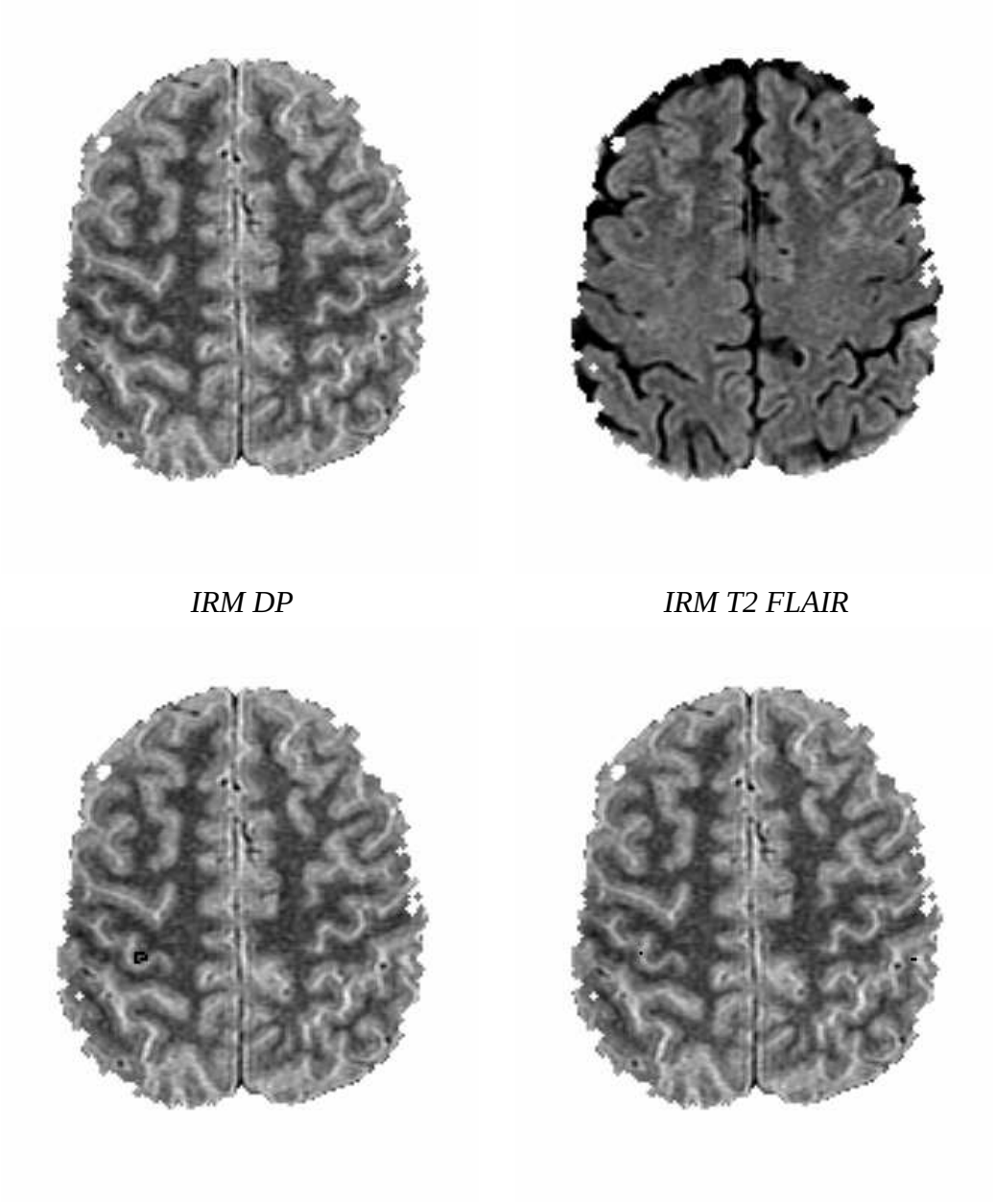


FIG. 7.9 – Pas de lésions dans cette image.

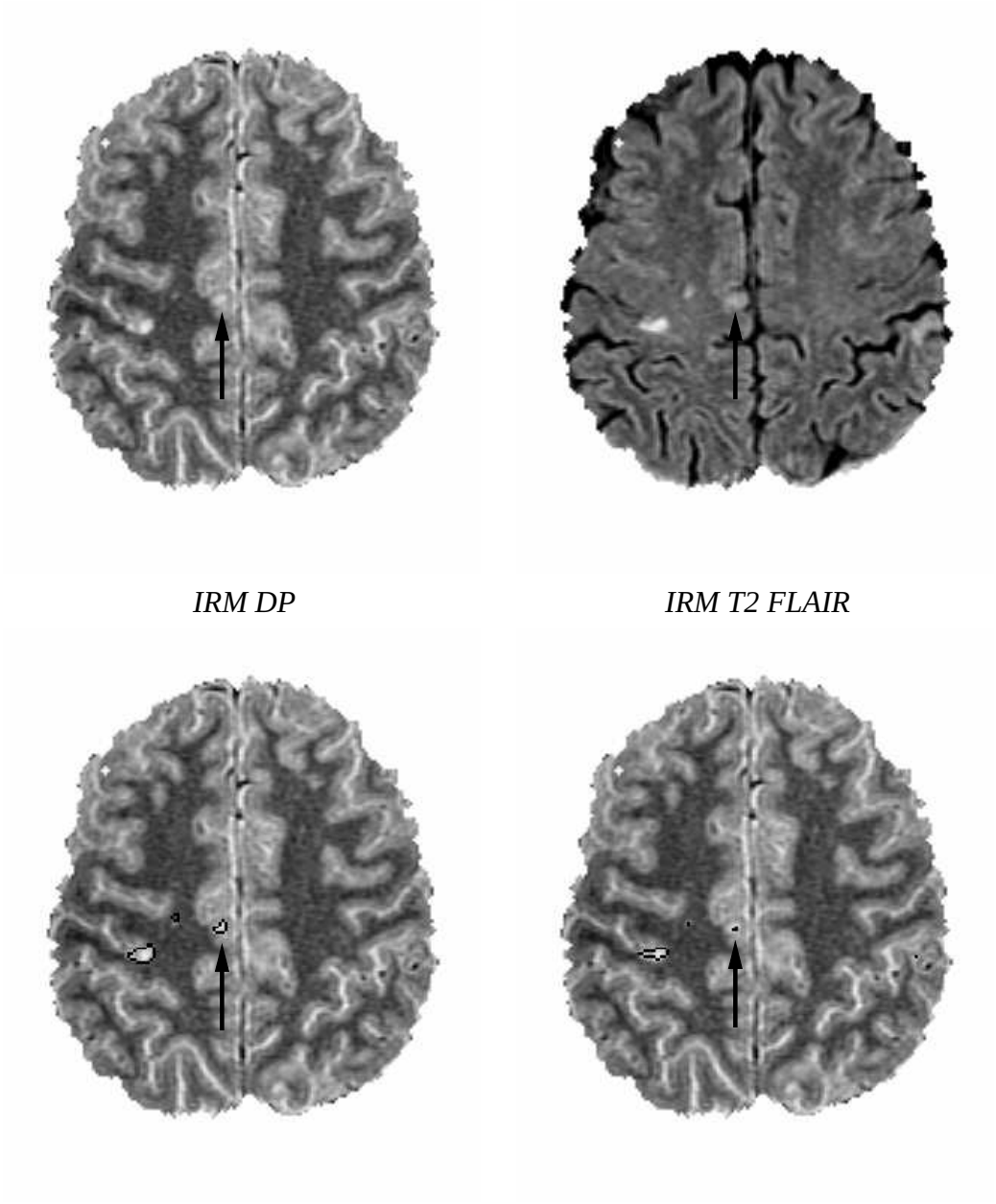


FIG. 7.10 – Le FLAIR conduit souvent à la sous-estimation des lésions corticales, comme on le voit ici sur cette coupe (flèche). La lésion juxta-corticale, à côté, est difficile à détecter par sa taille, mais elle est suffisamment contrastée en FLAIR, comme on le verra dans la figure 7.11.

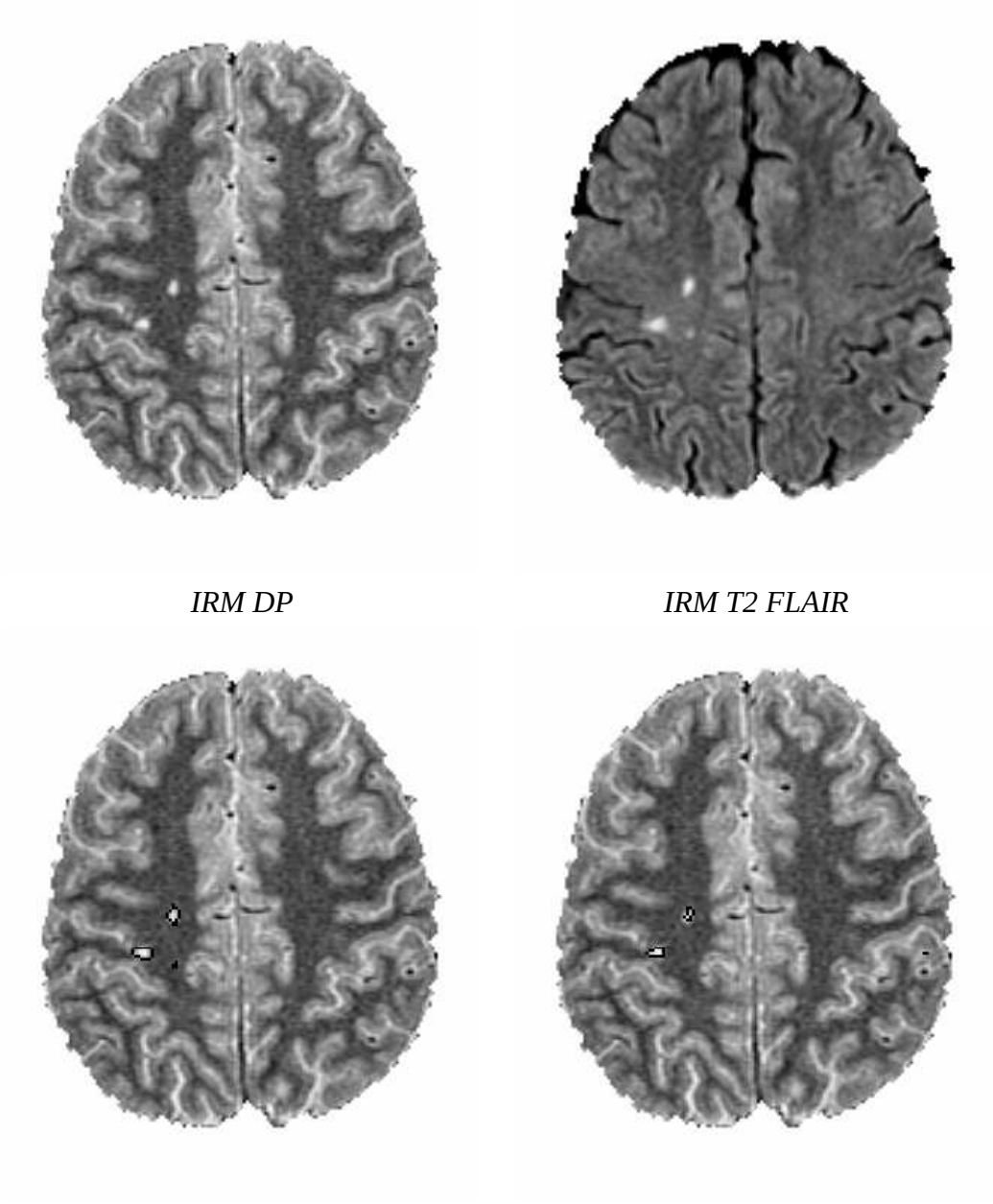


FIG. 7.11 – Sur ces 3 lésions, la segmentation automatique semble plus spécifique que la segmentation manuelle.

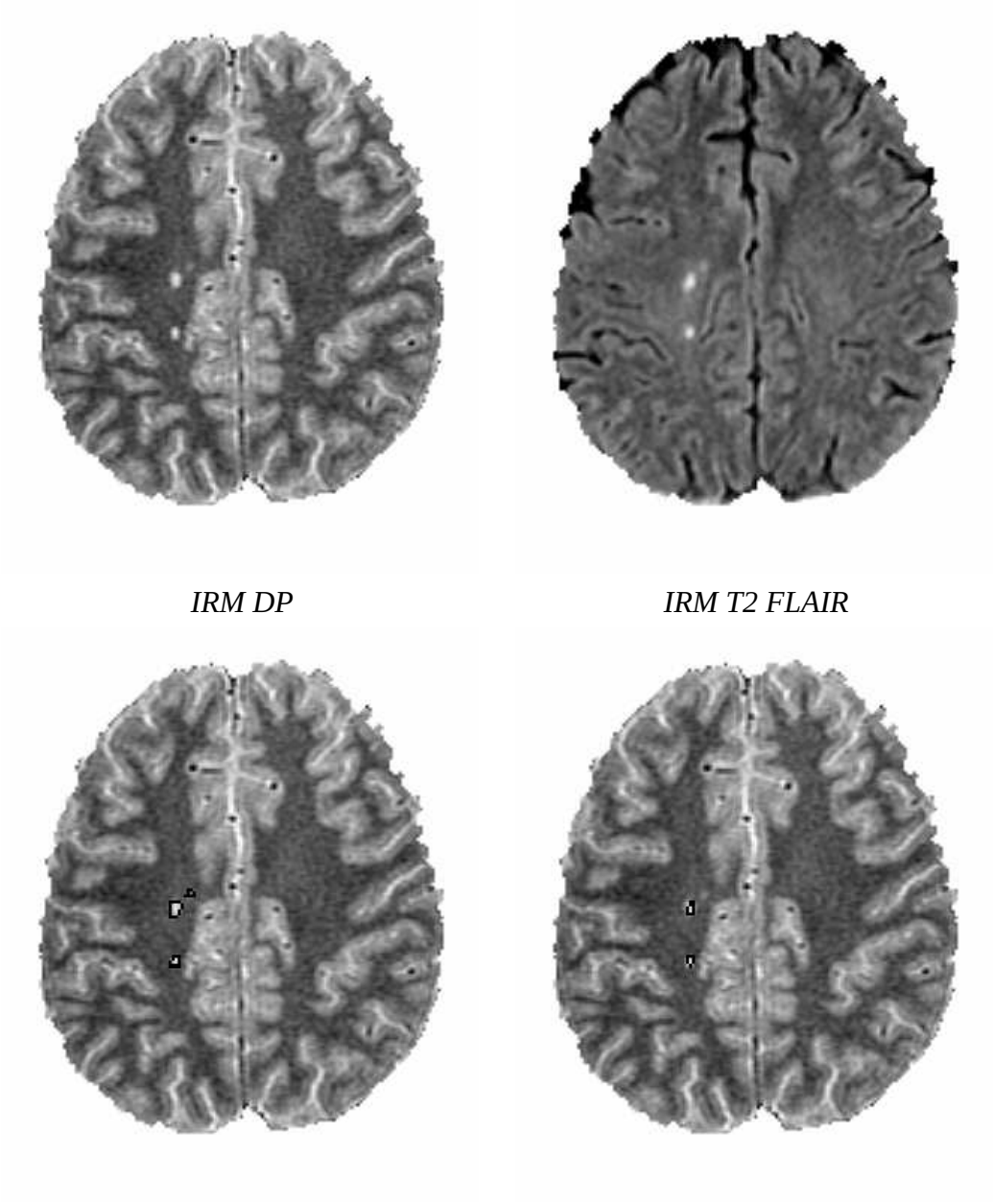


FIG. 7.12 – Une lésion juxta-corticale de faible volume – quelques voxels – est ici omise par la chaîne de traitements. Les deux autres sont détectées correctement.

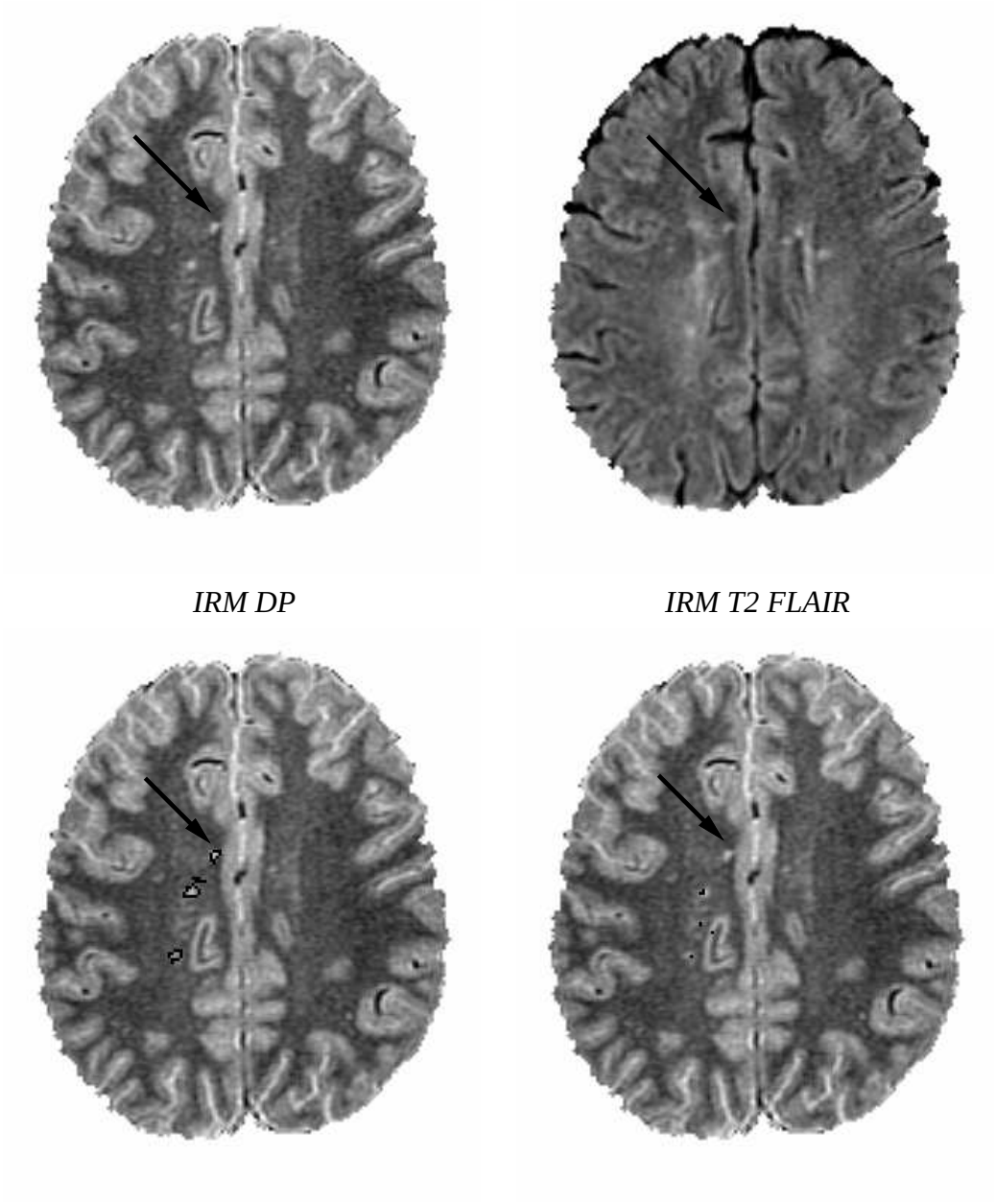


FIG. 7.13 – L'épaisseur de coupe est ici la source du faux négatif sur cette coupe (flèche). Il est à noter que les nombreuses techniques de seuillage adaptatif ne résoudraient pas le problème, à cause des artefacts de flux tout autour du masque du parenchyme cérébral.

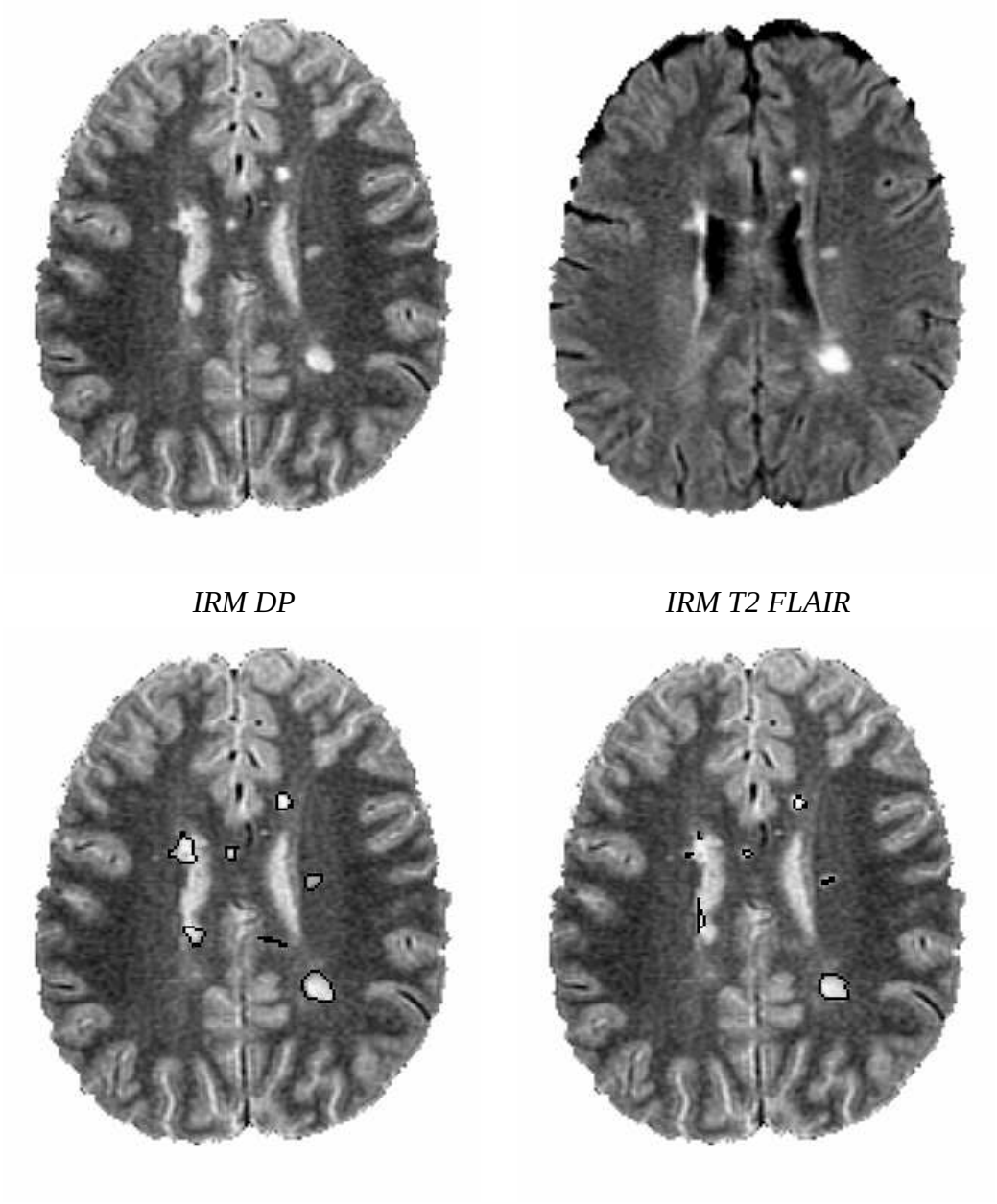


FIG. 7.14 – Les effets de volume partiel combinés aux artefacts de flux sur les séquences dérivées du T2 (en l'occurrence DP et T2 FLAIR) nous ont conduit à établir une zone de 1 à 2 millimètres autour de l'interface parenchyme / LCR, calculée à partir du T1, zone dans laquelle ne peuvent se situer les lésions. Ceci explique dans cette coupe le mauvais contourage de certaines lésions péri-ventriculaires par la chaîne de traitement.

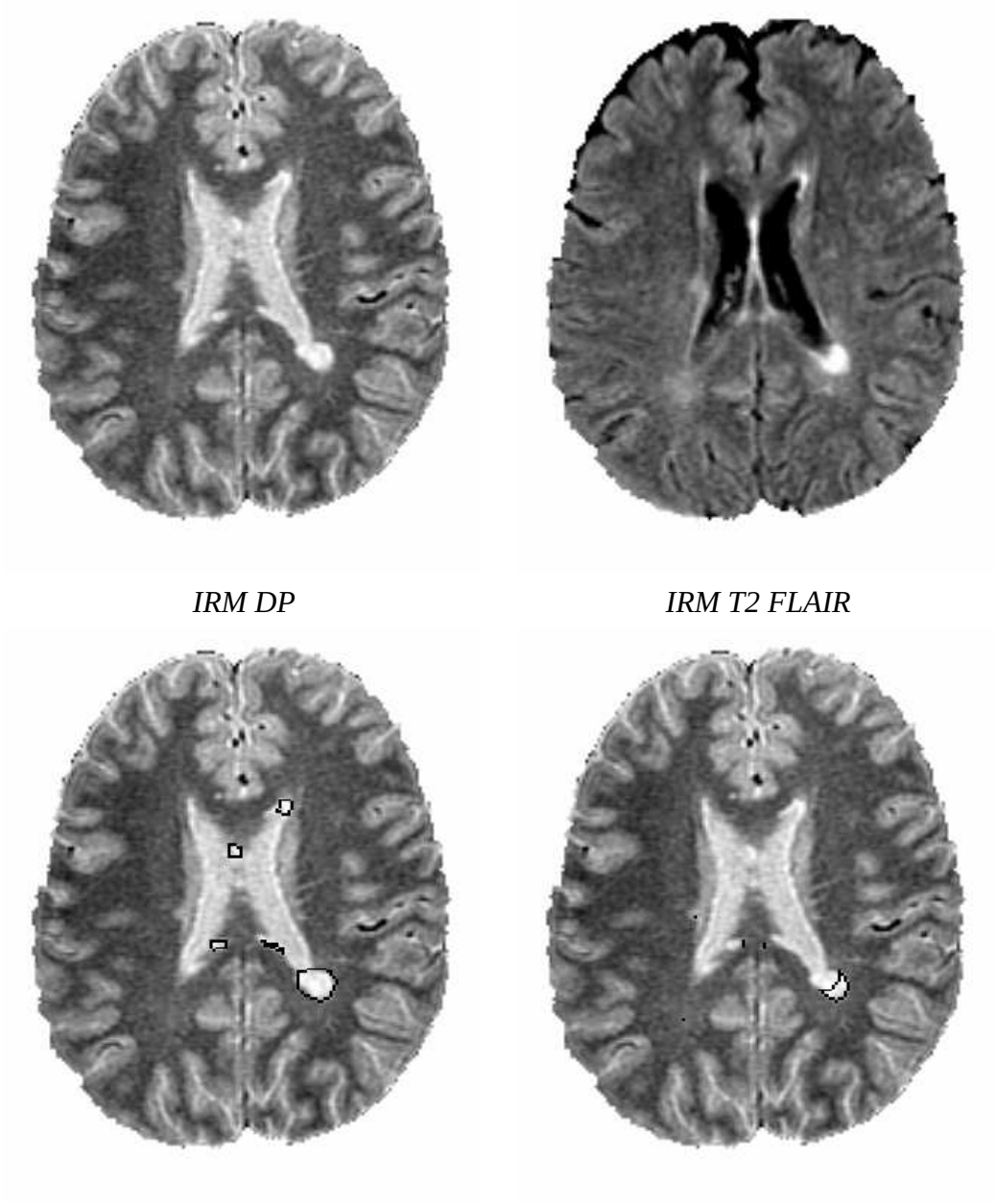


FIG. 7.15 – À cause de la correction des artefacts de flux par une zone de sécurité (figure 7.14), certaines lésions périventriculaires sont simplement omises, comme celle présente entre les deux ventricules, sur le plan médian sagittal. Il est à noter que cette lésion est très difficile à différencier des artefacts de flux, car elle se trouve dans une zone dans laquelle les artefacts de flux sont courants.

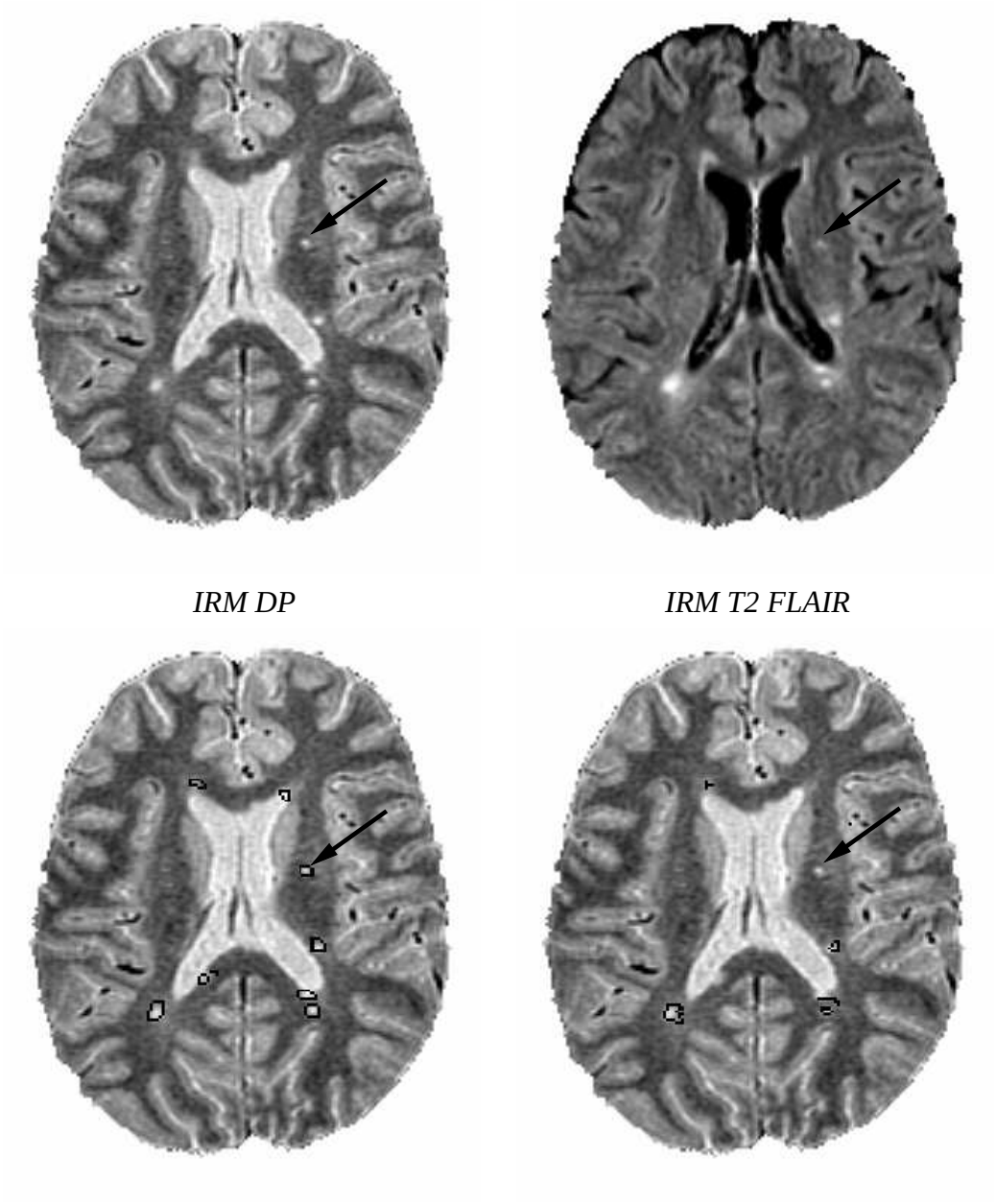


FIG. 7.16 – La plupart des lésions présentes sur cette coupe sont ici correctement traitées. La lésion périventriculaire présente sur la partie antérieure du ventricule droit n'est pas simple à détecter de part sa taille et sa localisation, alors qu'une petite lésion juxta-corticale à droite des ventricules (flèche) est omise, par manque de contraste en FLAIR.

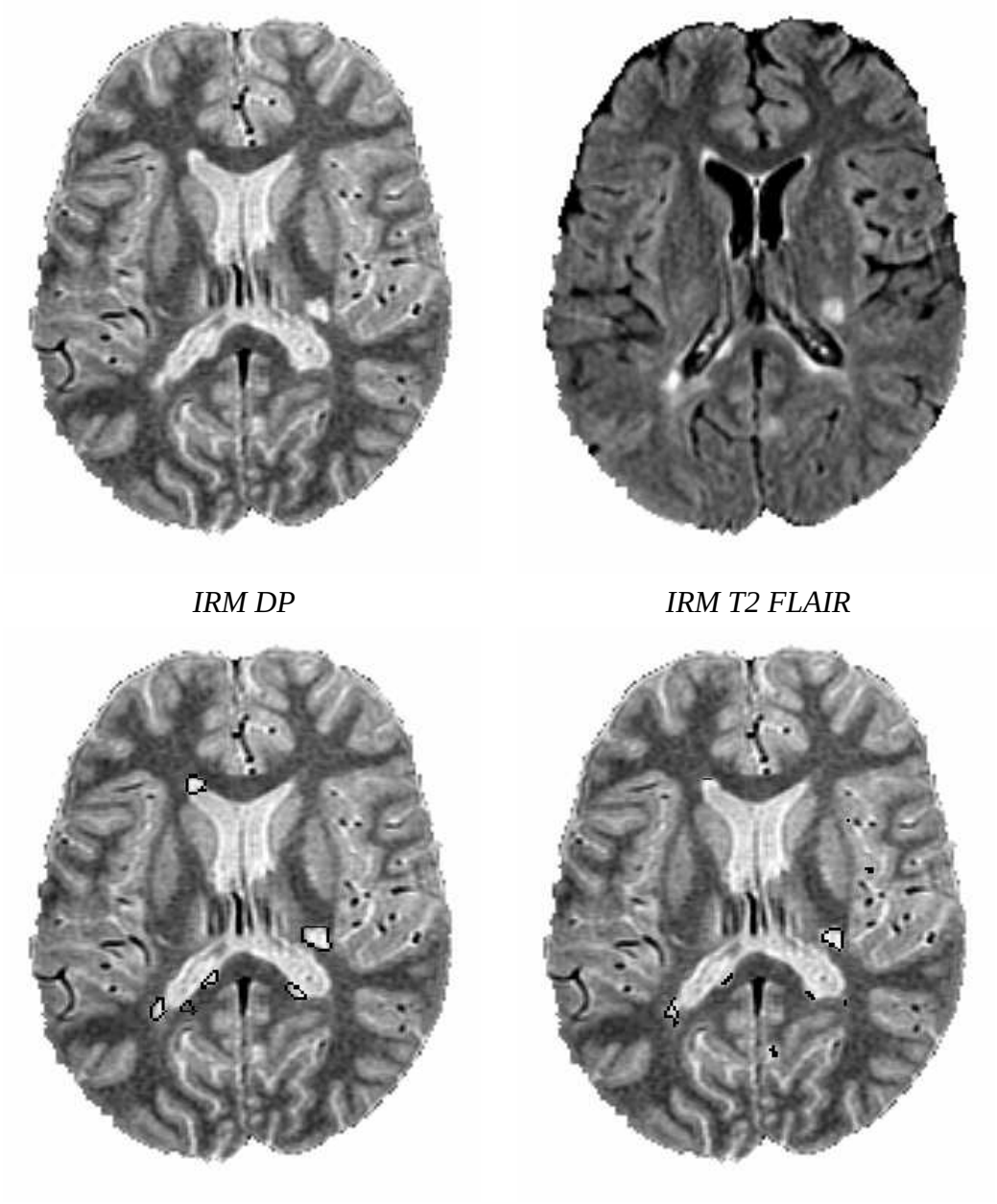


FIG. 7.17 – Les lésions proches de la partie postérieure des ventricules ne sont pas facile à différencier des artefact de flux, particulièrement sur cette coupe. Le FLAIR offre cependant un bon contraste, et le problème serait beaucoup plus difficile avec l'IRM T2/DP seule.

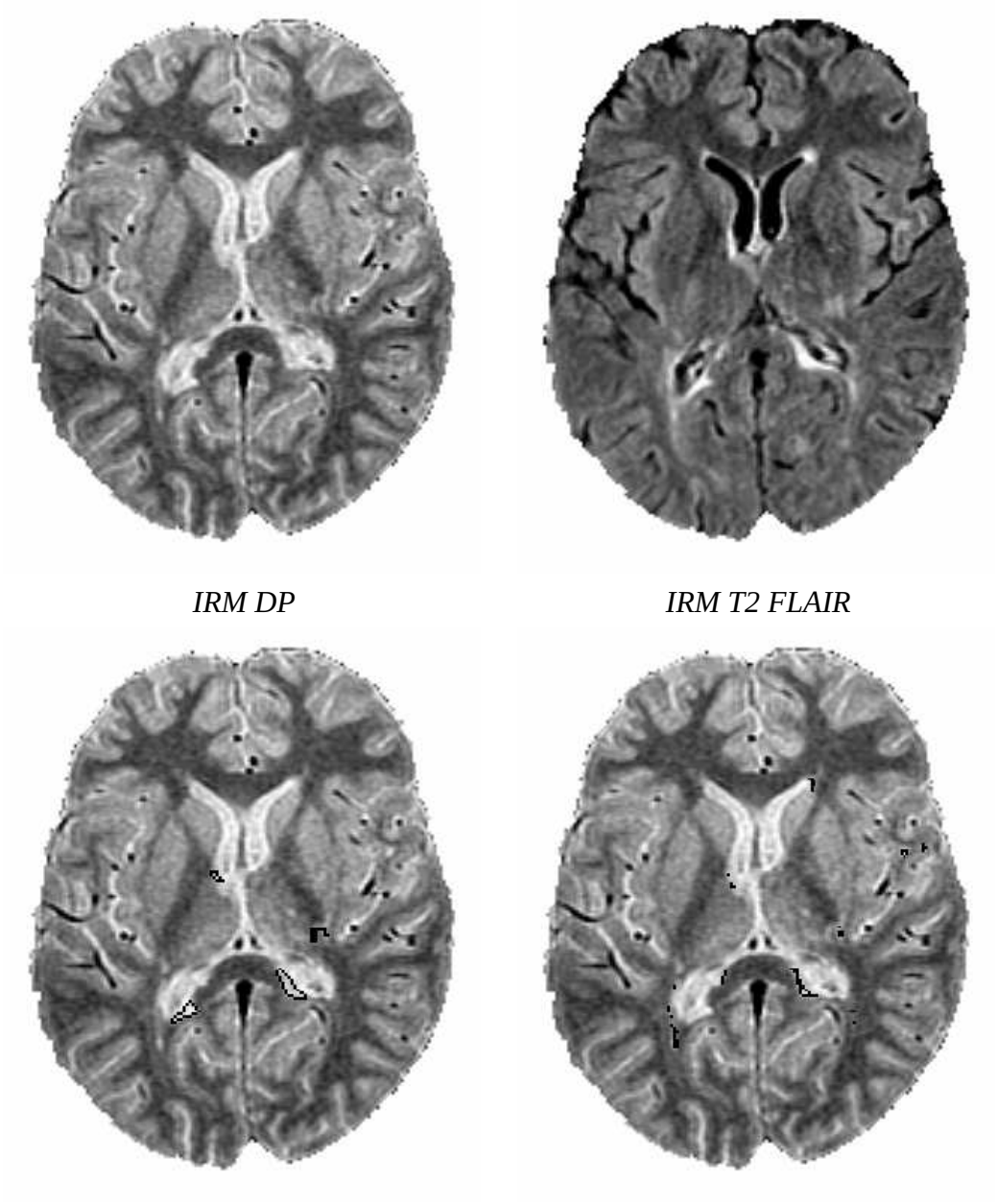


FIG. 7.18 – Lorsque les cornes sont fines, leur signal IRM est difficile voire impossible à différencier du signal d'une lésion de SEP. C'est le cas ici pour la naissance de la corne postérieure gauche, sur laquelle la chaîne de traitements crée une lésion qui n'existe pas.

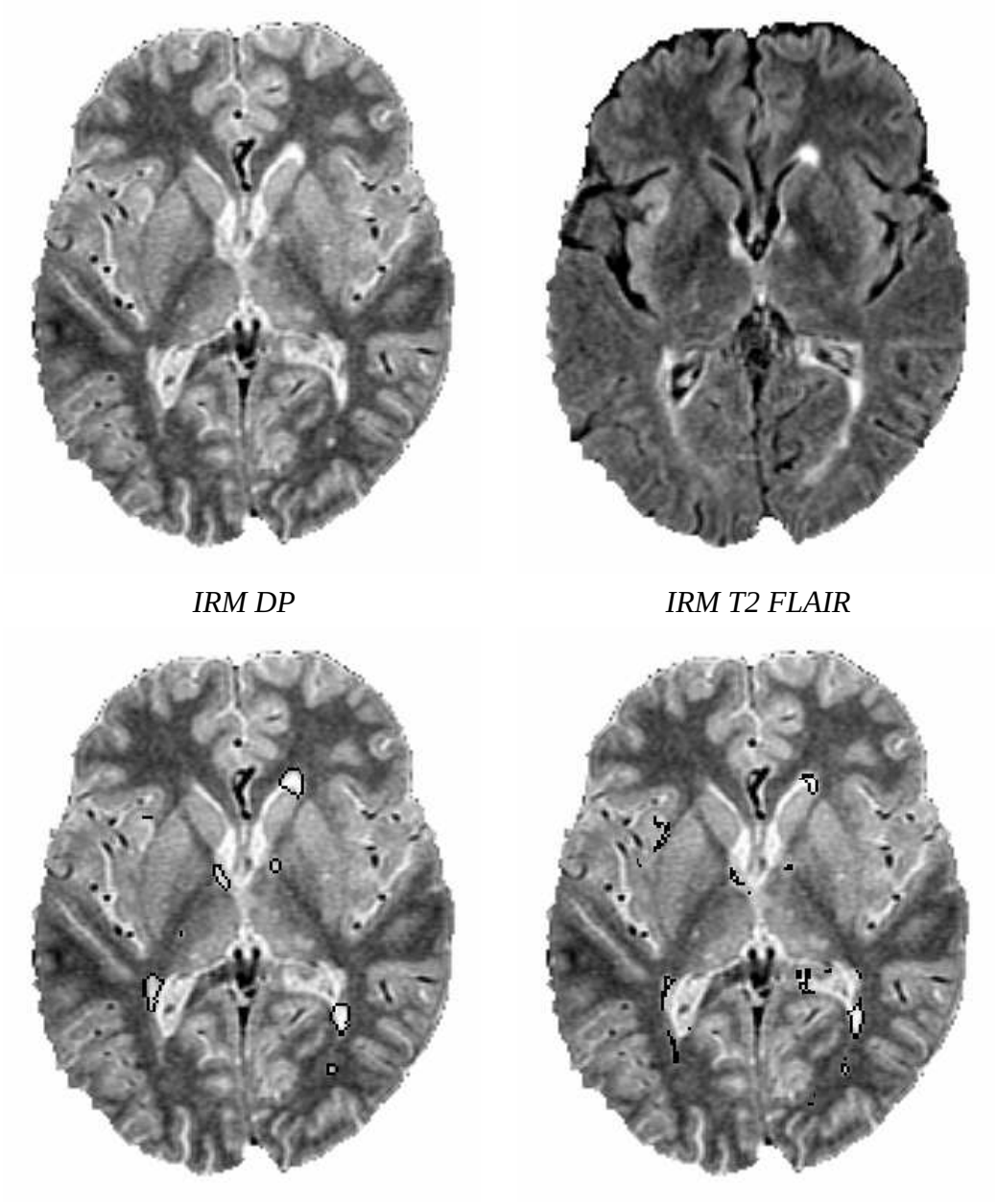


FIG. 7.19 – Le système anticipe ici l'apparition d'une lésion corticale à gauche. Sur la partie antérieure droite des ventricules, un signal suspect a été détecté, car le T1 indiquait que ce signal était hors des ventricules.

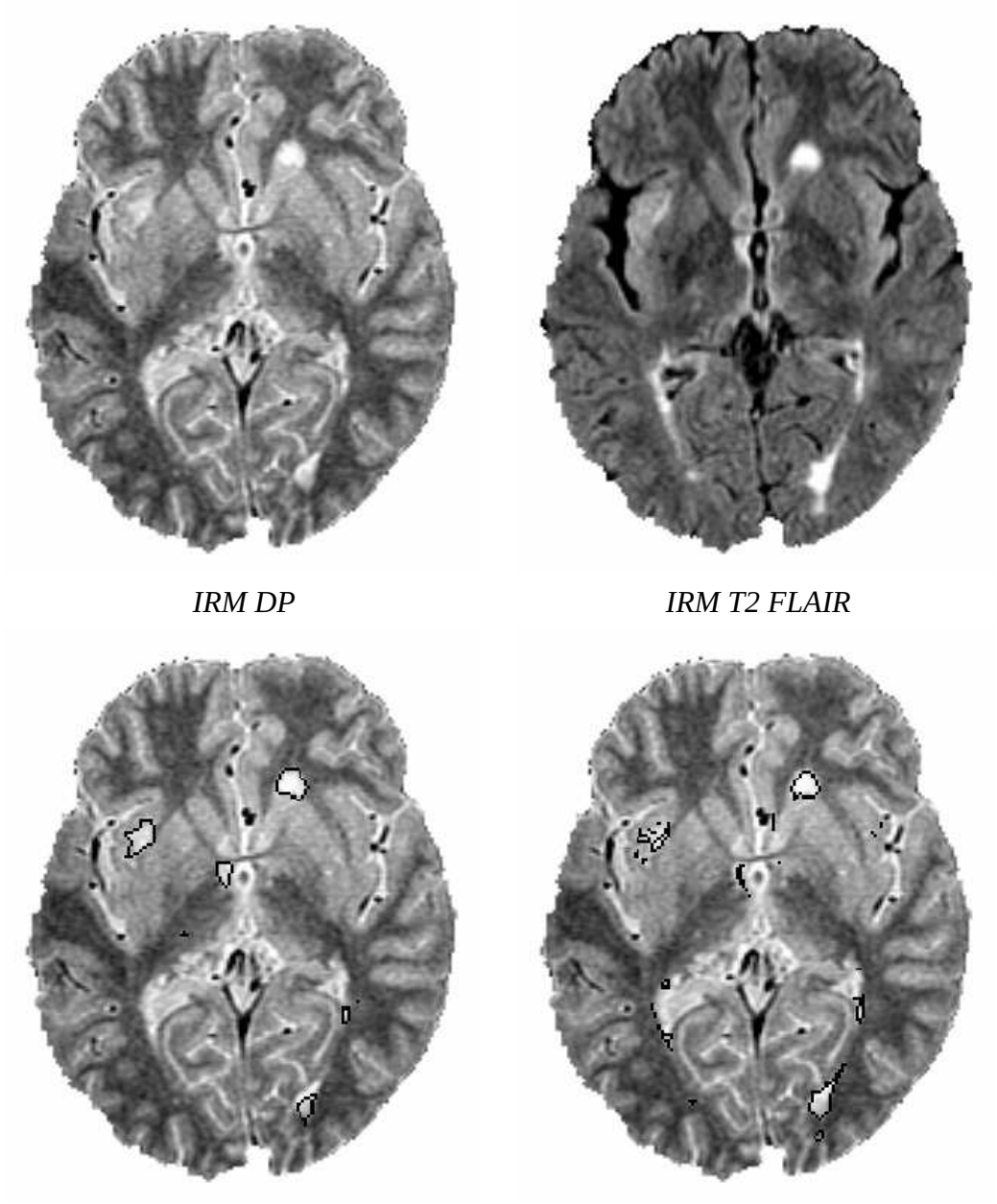


FIG. 7.20 – Au niveau de la corne postérieure droite, seul un examen attentif du signal en IRM DP permet de différentier ce qui semble être le LCR ventriculaire de la lésion de SEP. Le système a tendance à labéliser comme lésions certaines parties des cornes, quand celles-ci sont trop fines.

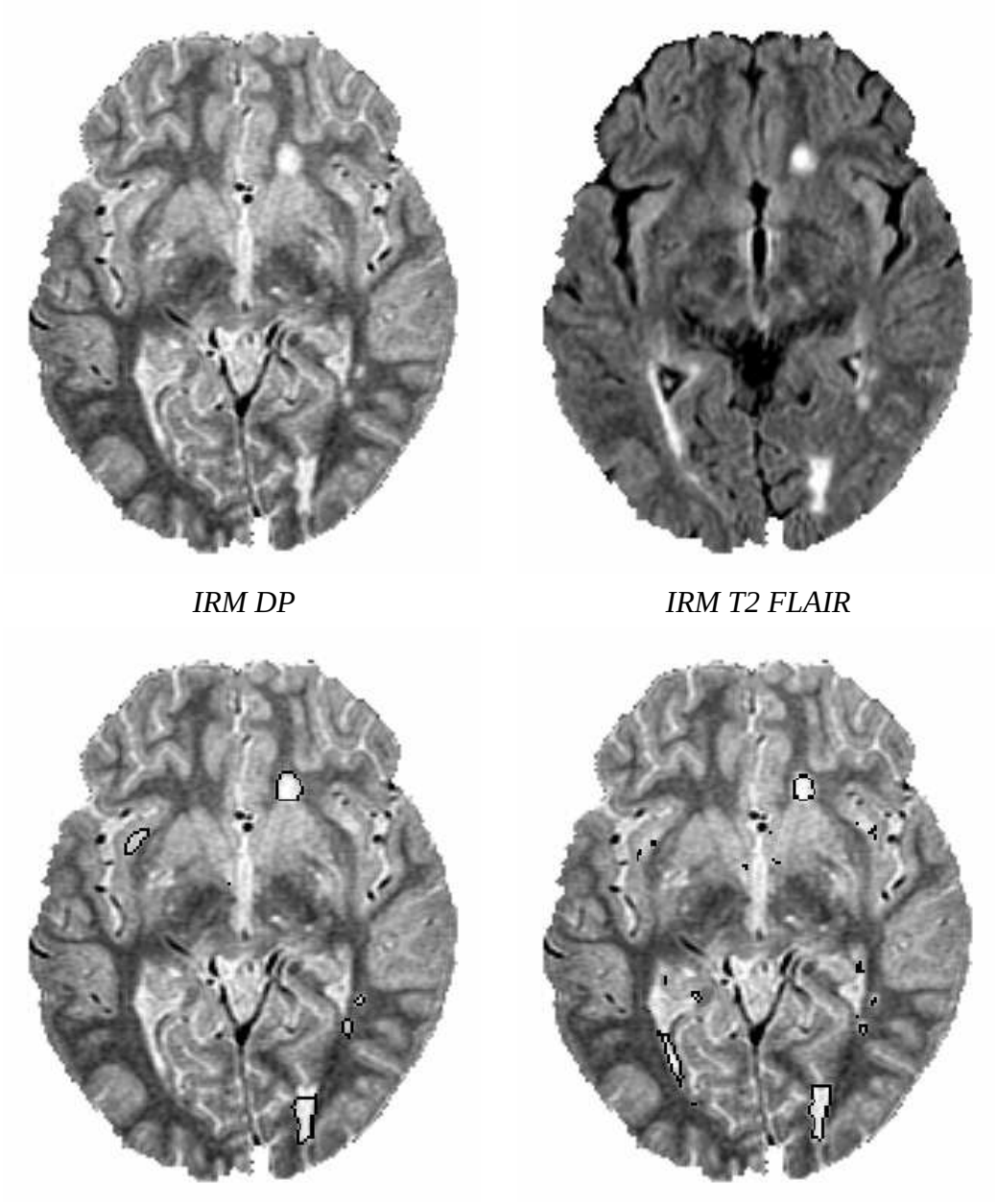


FIG. 7.21 – Deux lésions sont ici détectées au niveau des cornes postérieures. Une est sur-segmentée par rapport à la segmentation manuelle, l'autre n'existe pas. Le problème est ici subtil et difficile, et aucune solution immédiate n'est apparue pour séparer les cornes des ventricules des lésions de SEP périventriculaires.

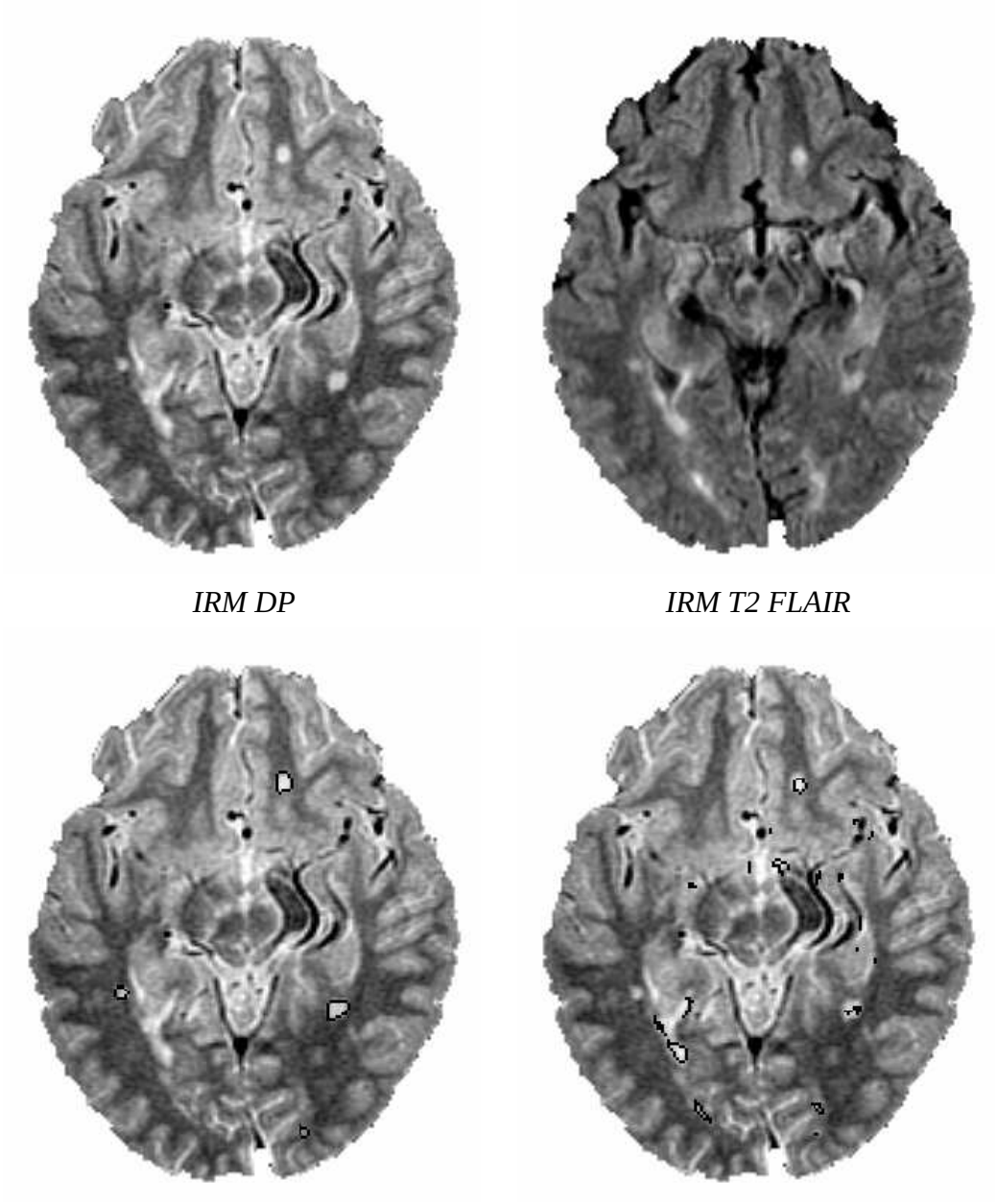


FIG. 7.22 – Comme sur la figure 7.21, l'hypersignal en T2 FLAIR sur la corne postérieure gauche des ventricules apparaît comme une lésion.

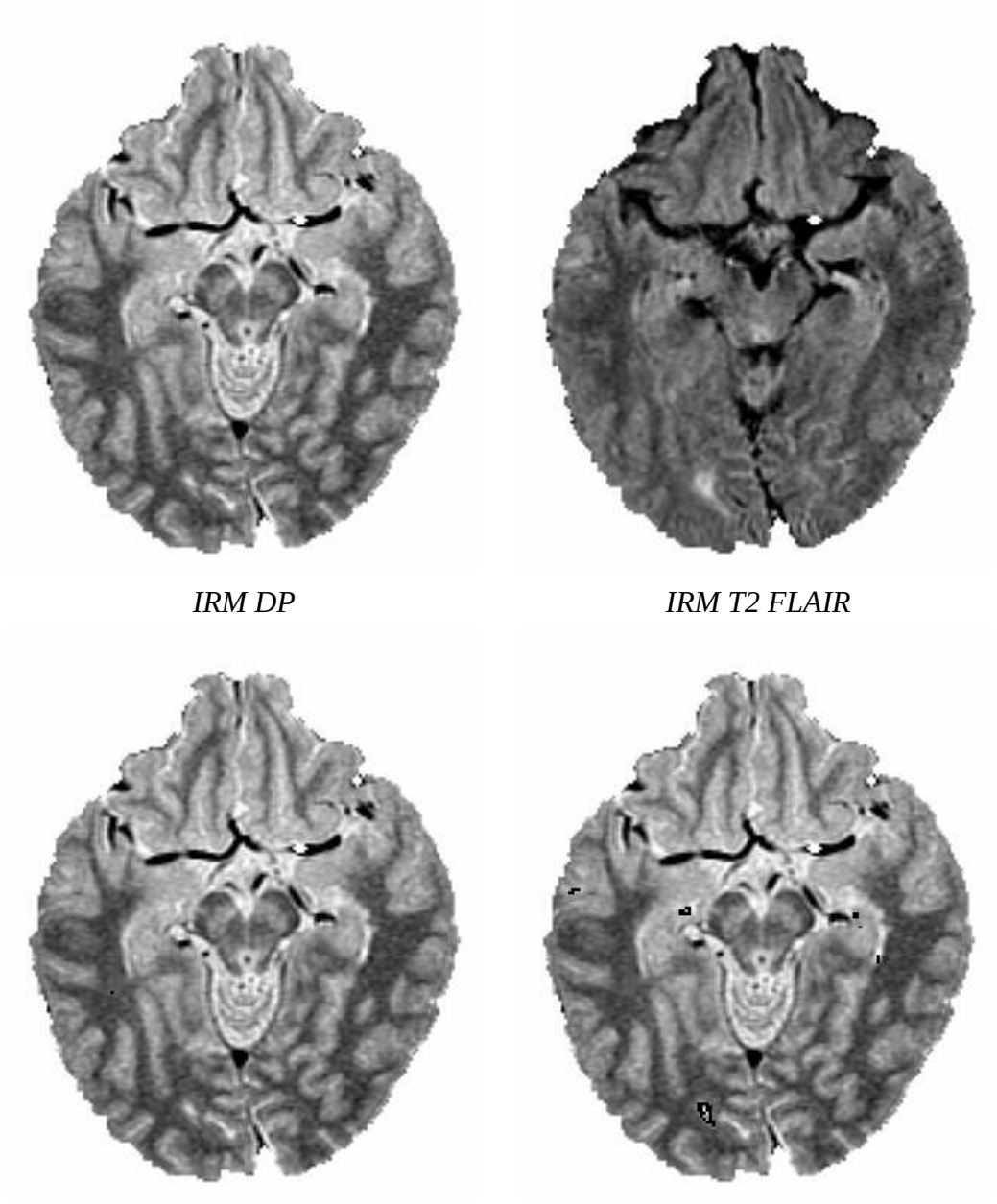


FIG. 7.23 – C'est généralement dans la région temporo-basale que le T2 FLAIR, avant toute correction de biais, donne le plus de faux positifs isolés de toute lésion. Ici, la correction de biais a bien fonctionné, et aucun faux positif n'est présent dans cette zone.

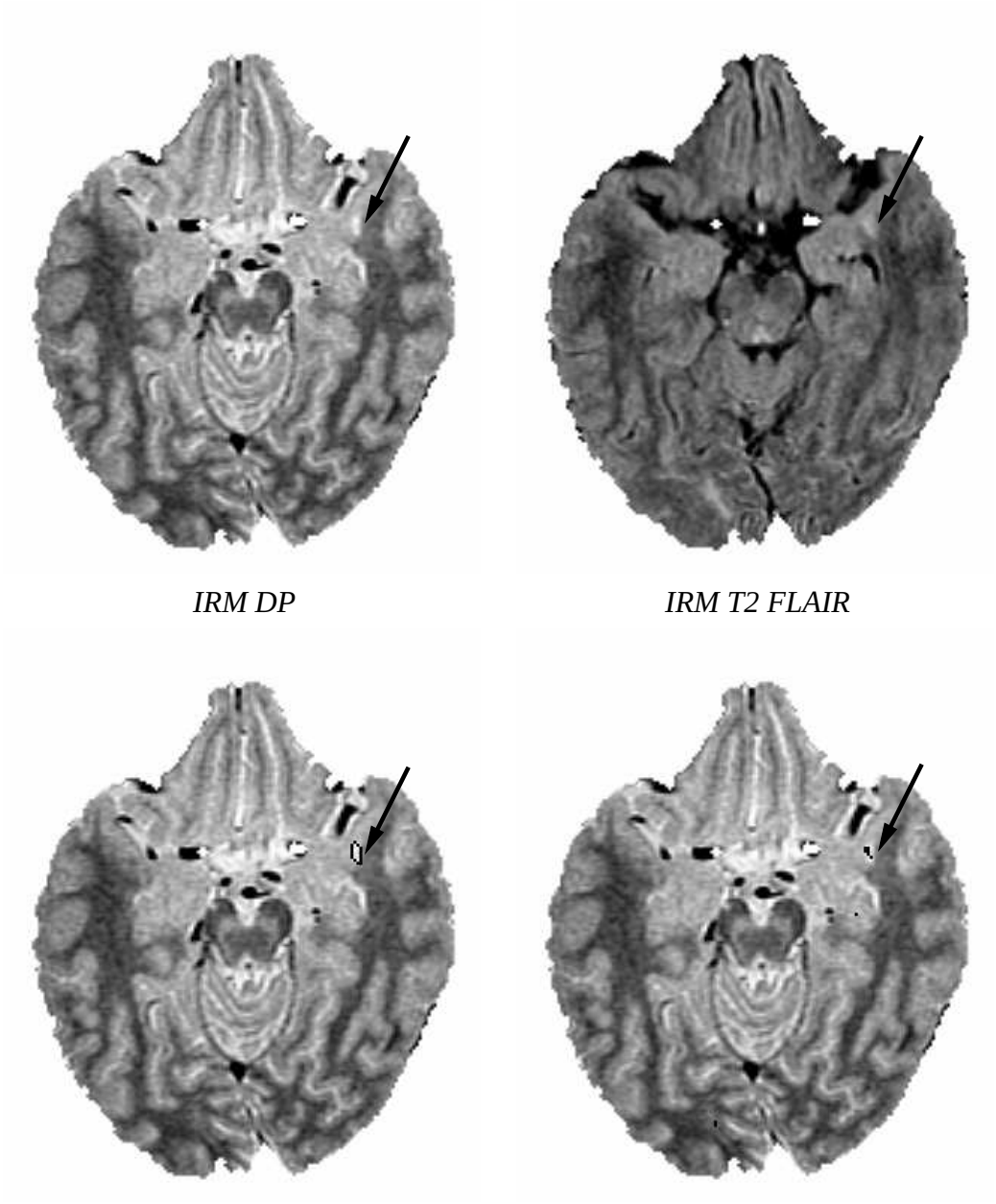


FIG. 7.24 – La lésion corticale (flèche) est ici mal contrastée en T2 FLAIR, ce qui conduit à un contourage médiocre, même si la lésion a tout de même été détectée.

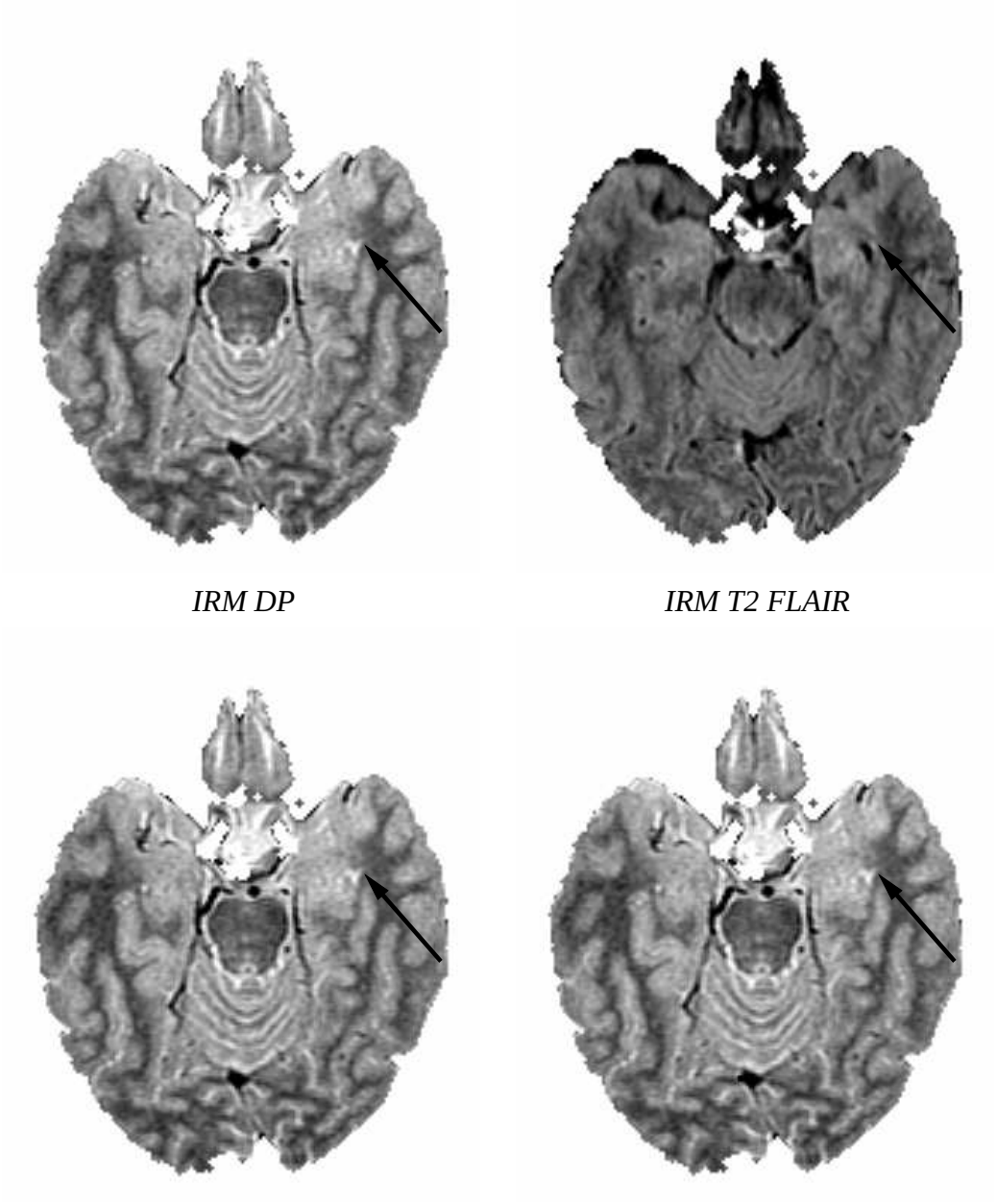


FIG. 7.25 – Malgré tous les problèmes inhérents au T2 FLAIR – artefact de flux, sur-segmentation des lésions péri-ventriculaires, sous-estimation des lésions corticales – cette séquence permet de différencier le signal du LCR (flèche) du signal des lésions, ce qui est impossible sur cette coupe à partir de la seule IRM DP.

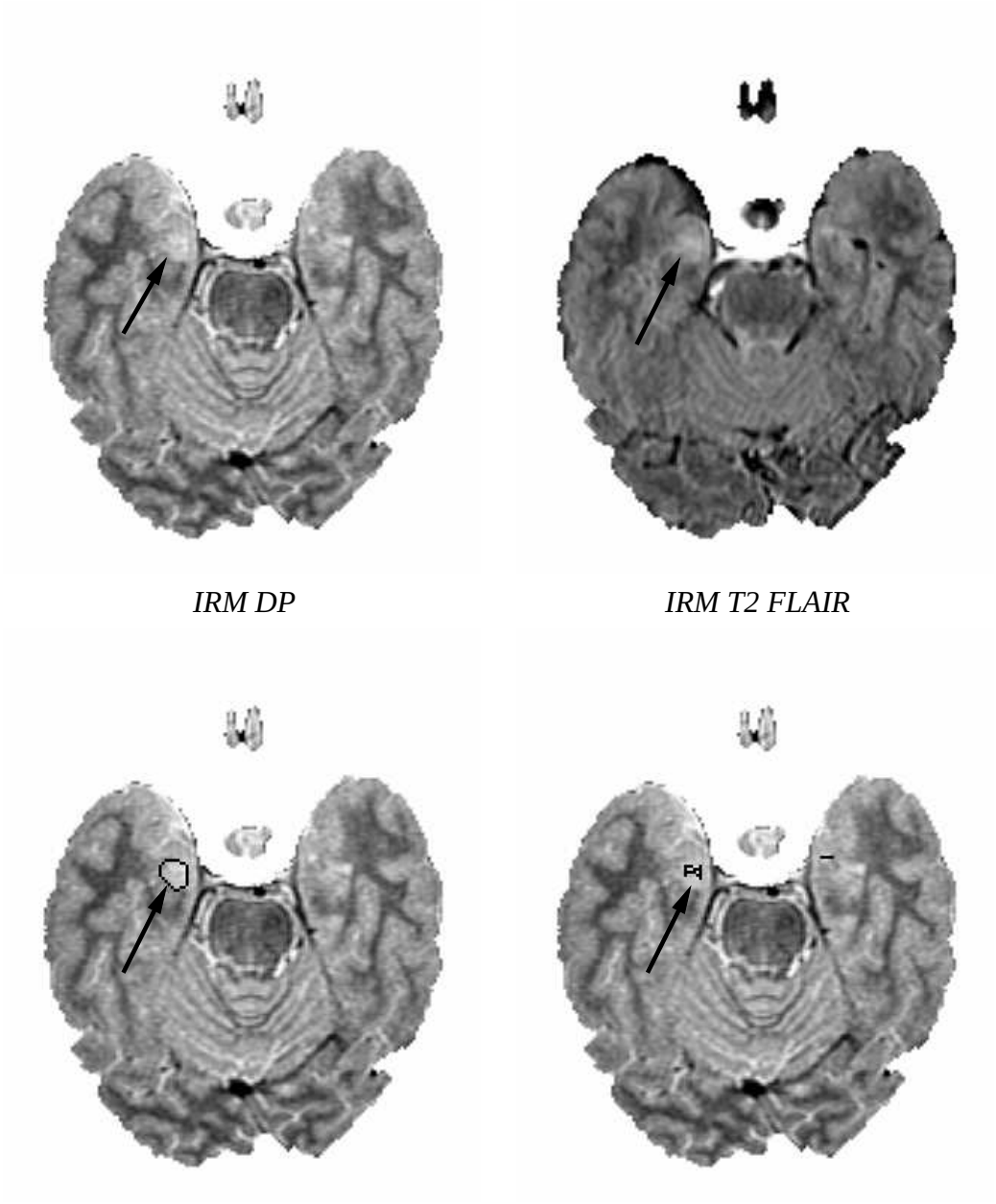


FIG. 7.26 – On voit ici un autre exemple de lésion corticale (flèche) sous-estimée par la chaîne de traitements, à cause du manque de contraste en IRM T2 FLAIR. Il est à noter que le contraste en IRM DP est également assez médiocre, et il ne serait pas facile de corriger le contourage, avec un algorithme de croissance de région par exemple.

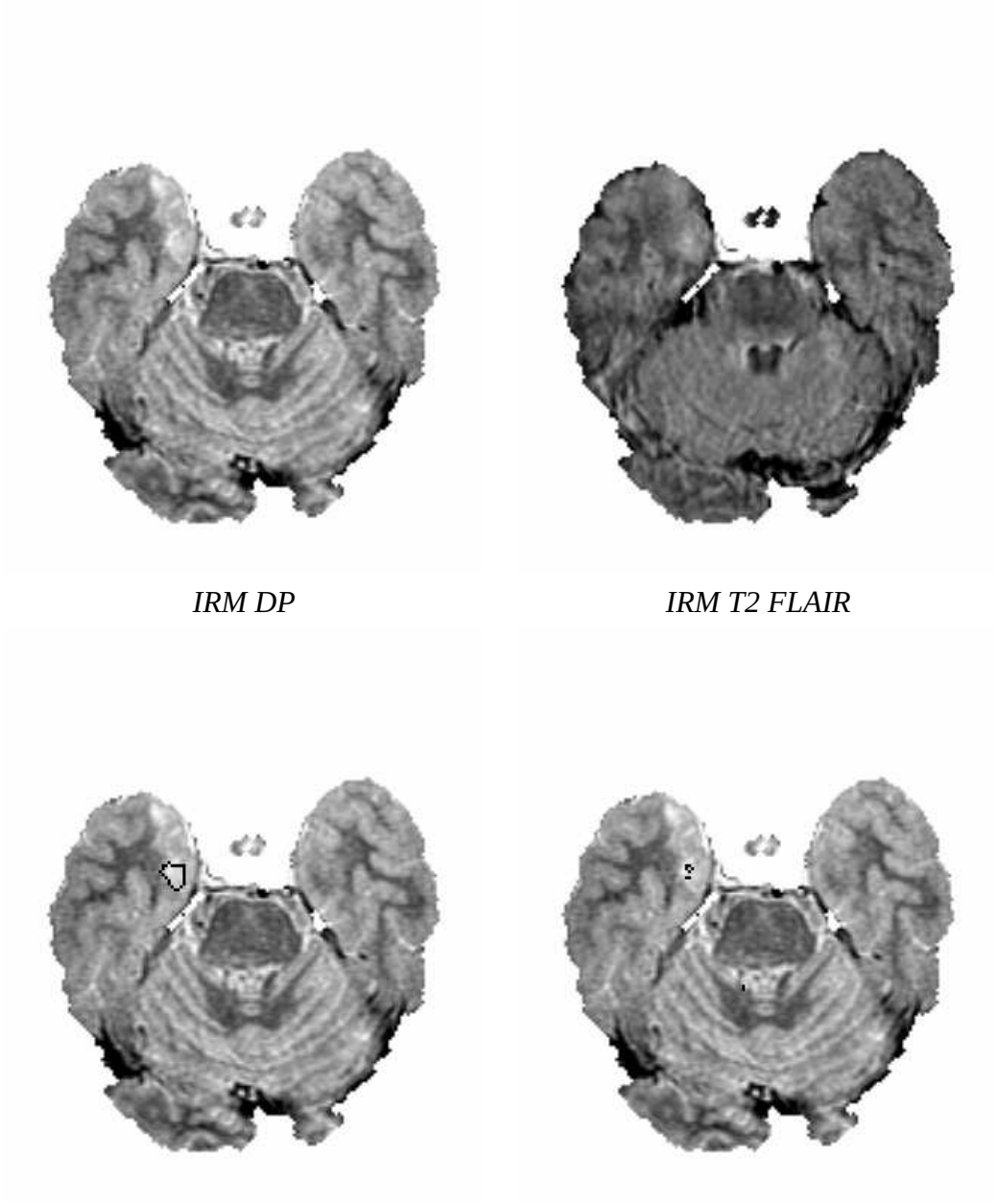


FIG. 7.27 – Comme sur la figure 7.27, la lésion corticale présente ici est sous-estimée par la chaîne de traitement

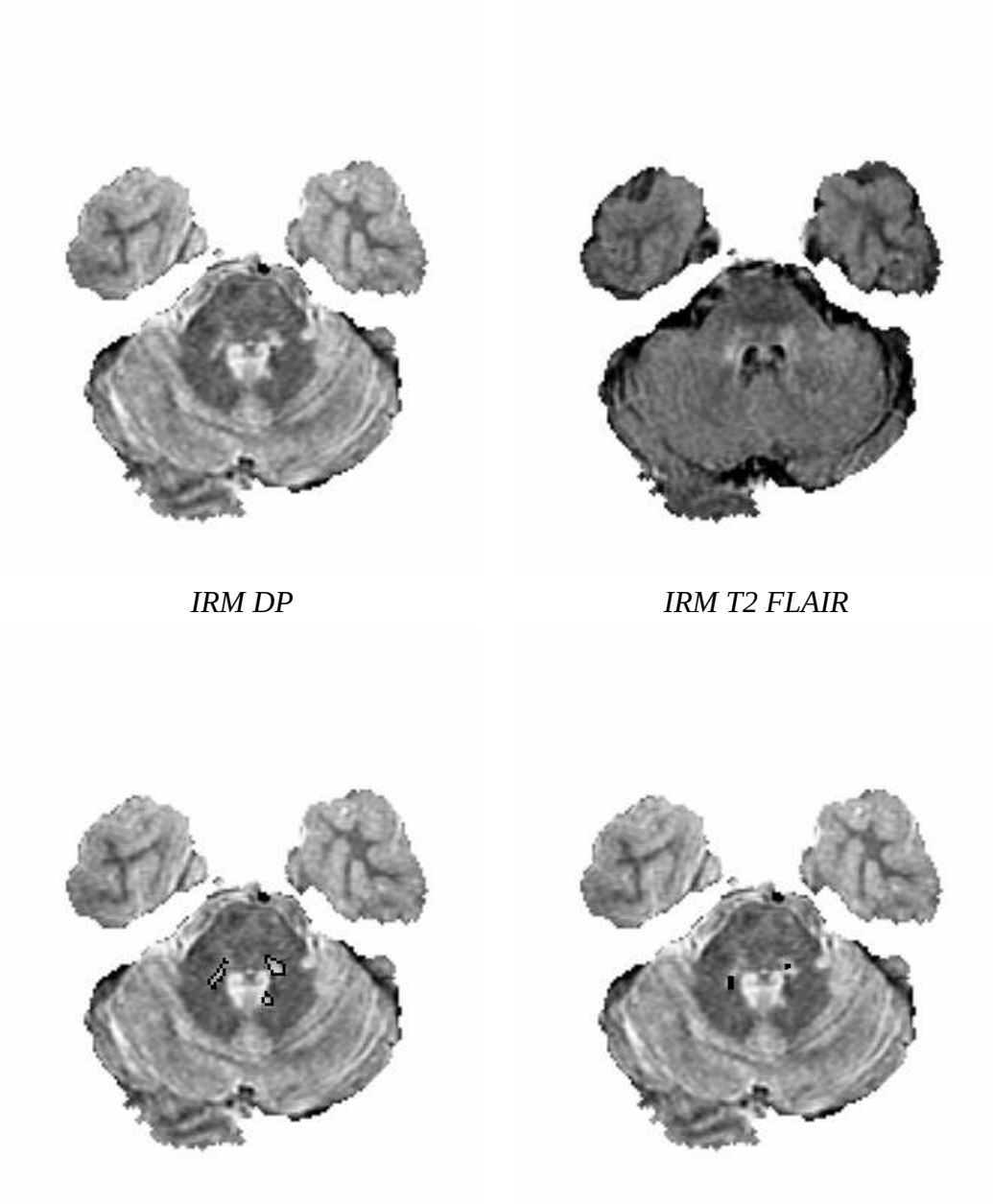


FIG. 7.28 – Sur cette coupe, les lésions de la fosse postérieure commencent à apparaître. Si on peut les deviner en T2 FLAIR, leur contraste n'est pas suffisant pour un contourage correct.

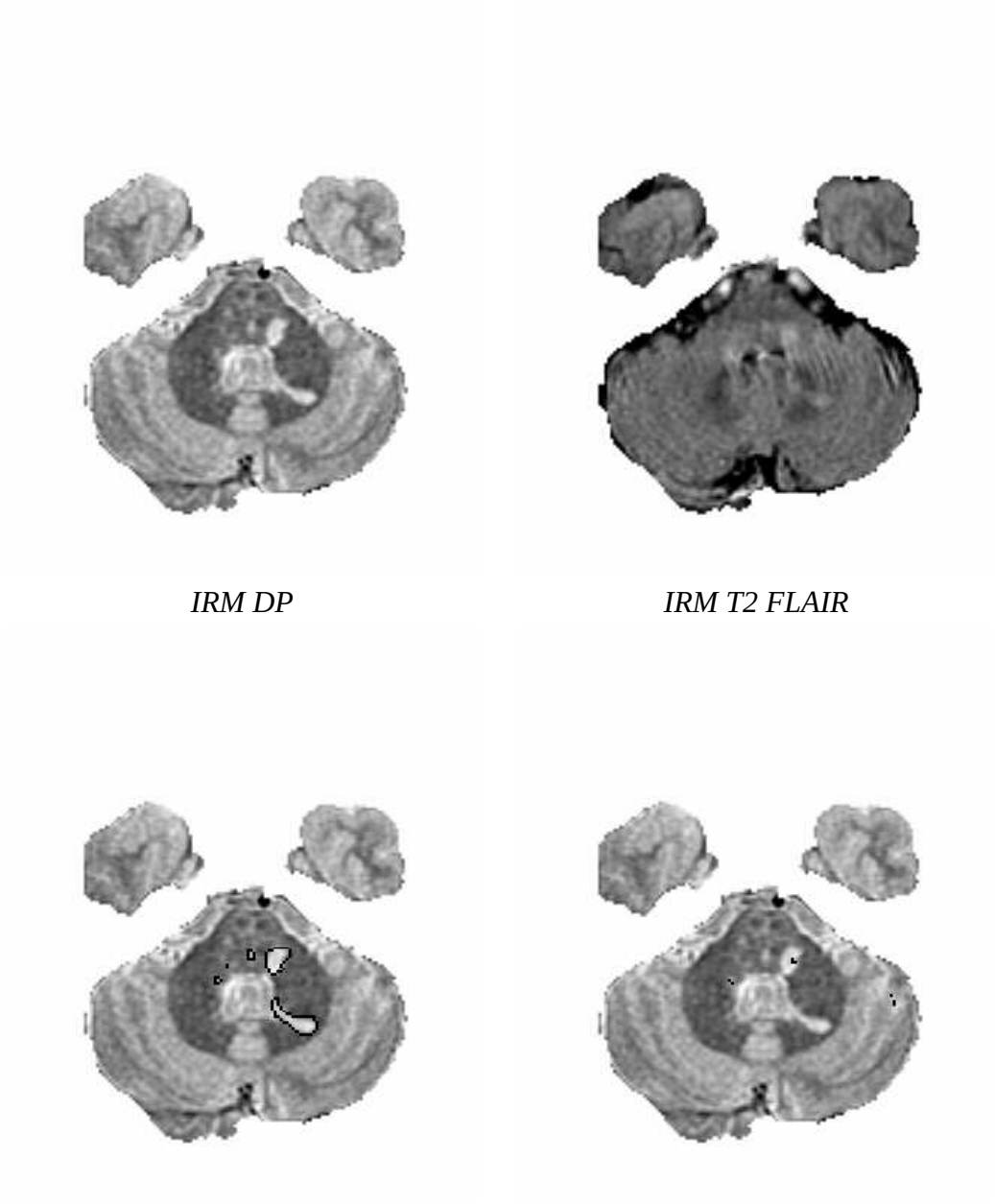


FIG. 7.29 – Les coupes axiales sous la fosse postérieure souffrent généralement d'un hyper-signal en FLAIR. La correction de ce signal permet de deviner les lésions du cervelet, mais pas d'améliorer leur contraste. Il y a donc une grande quantité de faux négatifs dans cette zone, sur l'ensemble des patients.

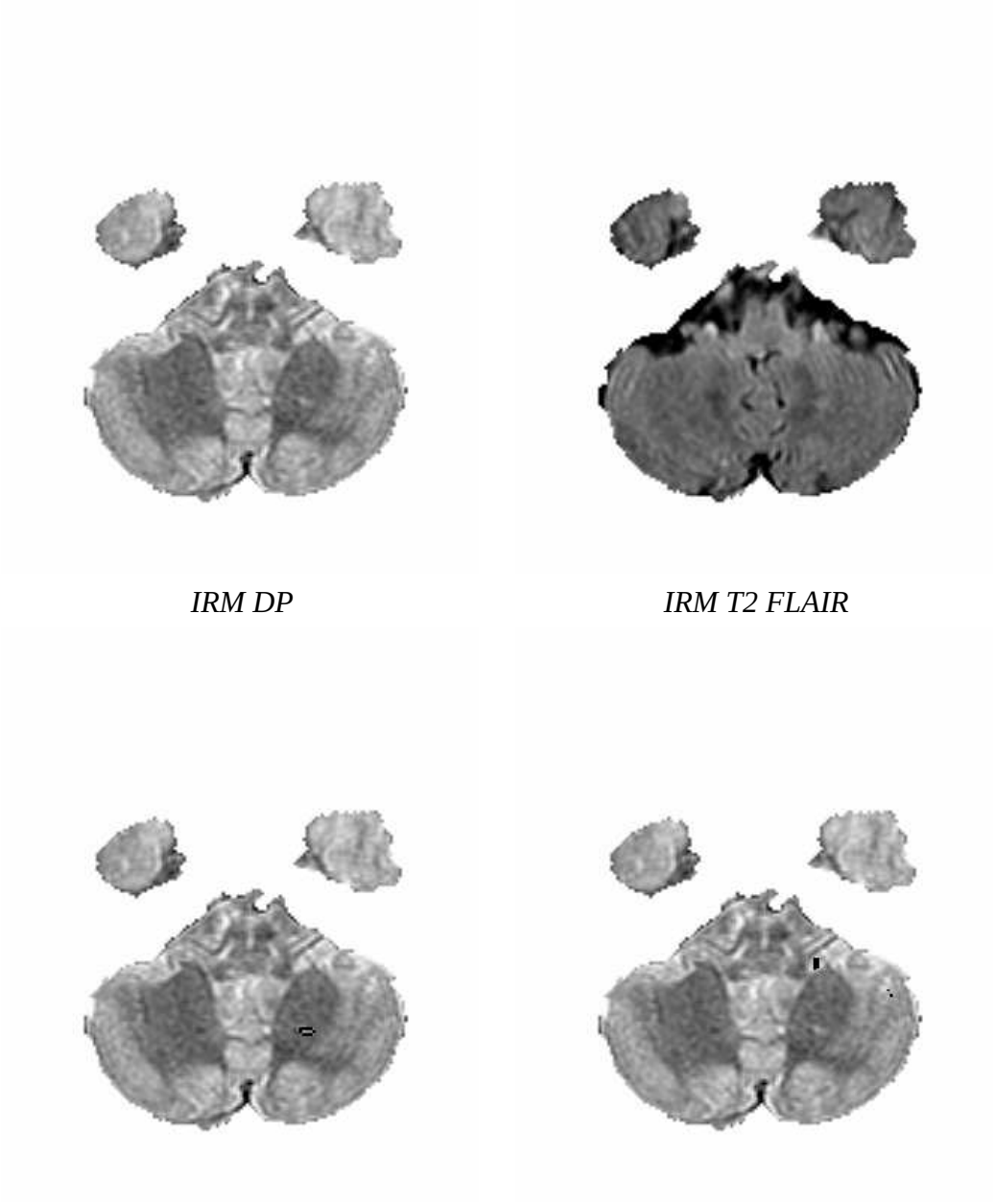


FIG. 7.30 – Ces coupes sont difficiles à segmenter, en général, car les lésions de SEP ont souvent un signal proche de la matière grise. En outre, les IRM souffrent parfois d'échos aux extrémités de l'axe Z – perpendiculaire à la coupe axiale – qui rendent difficile la lecture des images.

Chapitre 8

Discussion et perspectives

Les résultats obtenus sont intéressants par la qualité de la détection : la sensibilité moyenne est de 74.9 % et les lésions détectées représentent 95.6 % de la charge lésionnelle. La plupart des lésions sont détectées, et celles qui ne le sont pas ne représentent qu'un faible pourcentage de la charge lésionnelle globale, hors lésions de la fosse postérieure. Par contre, le contourage laisse à désirer : la plupart des lésions sont sous-segmentées, certains artefacts de flux sont labélisés comme lésions et sont donc une source de faux positifs.

Certes, la variabilité inter-expert n'a pas été prise en compte dans cette étude, et elle est très importante pour la segmentation des lésions de SEP. Lors de l'évaluation du système INSECT, la charge lésionnelle moyenne μ est de 381.2 mm^3 pour un écart type std de 169.1 mm^3 entre les différents experts disponibles [181, 180]. Le système automatique INSECT fournit un volume $V \in [\mu - std,\ \mu + std]$ pour 9 patients sur 10 : un écart entre la segmentation automatique et le contourage manuel n'est donc pas surprenant. Il reste néanmoins beaucoup à faire avant de pouvoir utiliser cette chaîne de traitements pour un calcul automatique de charge lésionnelle. Nous allons d'abord discuter de l'influence des prétraitements, puis aborder plus spécifiquement les différents types de lésion avant de conclure. Nous profiterons de cette discussion pour faire apparaître, au fil du texte, plusieurs directions de recherche qui restent à explorer.

8.1 Influence des prétraitements

Lorsqu'un algorithme de segmentation est en fait une succession d'étapes relativement simples – excepté la segmentation en tissus et le modèle de volume partiel – il est important de savoir quelles sont les étapes les plus cruciales pour le fonctionnement de l'algorithme global. Plus précisément, que se passe-t-il si une étape échoue, ou est tout simplement supprimée ? Il convient alors d'évaluer l'influence de chaque étape sur le résultat. L'importance des étapes spécifiques à la segmentation des lésions de SEP – détection en IRM T2 FLAIR, région d'intérêt calculée avec l'IRM T1 – a été étudiée dans la section précédente, et il faut maintenant regarder les prétraitements présentés au chapitre 3. Ceux-ci auront bien sûr une influence sur le résultat final : ainsi, la correction des inhomogénéités en IRM T2 FLAIR est capitale pour une bonne détection des lésions. Mais dans la plupart des cas, l'influence des prétraitements sur la segmentation en tissus est plus directe et plus intéressante.

8.1.1 Mise en correspondance des images

Une bonne mise en correspondance des images est cruciale à toute étude multi-séquences. Même lors de l'utilisation d'algorithmes moins "myopes" que les modèles voxelliques présentés dans ce manuscrit, il est important que chaque voxel de coordonnées identiques corresponde au même élément de volume physique dans toutes les séquences. Pour ce faire, le recalage rigide entre les images doit fonctionner parfaitement. Nous avons déjà visualisé dans la figure 3.5 du chapitre 3 l'influence d'un mauvais recalage intra-patient des différentes séquences sur l'histogramme conjoint. Lorsque les images ne sont pas parfaitement recalées, la signature en intensité des structures fines est particulièrement touchée : c'est le cas du LCR cortical, par exemple. La signature du LCR est beaucoup moins visible quand le recalage est mauvais, et on observe dans l'histogramme l'apparition d'une nouvelle structure n'ayant aucune signification anatomique.

Il est intéressant, en partant de ce point de vue, de chercher à évaluer les dégâts effectués par un mauvais recalage des séquences en effectuant une estimation

des paramètres de classes avec l'algorithme présenté au chapitre 5. Reprenons l'exemple de la figure 3.5 du chapitre 3. Dans cet exemple, au lieu de choisir le couple IRM T2/DP, les 2 séquences sont les IRM T2 et T1. Le T2 est toujours nécessaire pour différencier la classe "points aberrants" de la classe LCR, et l'algorithme donne des résultats corrects. Deux expériences sont menées : en recalant correctement l'IRM T2 sur l'IRM T1 et en ajoutant une rotation de 1 degré à la transformation optimale (figure 8.1). Les résultats obtenus correspondent à ce qu'on pouvait attendre après un examen qualitatif des histogrammes : même une faible erreur sur les paramètres du recalage dégrade fortement l'estimation des paramètres de classes, et les variances sont particulièrement touchées.

Cependant, deux manières de représenter ce bon recalage sont à notre disposition. Il faut en effet choisir l'image de référence pour construire l'histogramme et visualiser la segmentation. L'IRM T2 (taille du voxel : 1*1*2 mm) a une résolution deux fois inférieure à l'IRM T1 (1*1*0.8 mm). La figure 8.2 représente les deux choix possibles : T2 recalé sur T1 et l'inverse. Plusieurs remarques sont à faire sur cette expérience. Tout d'abord, les paramètres de classes de la matière matière blanche, dont le rapport surface / volume est faible, ne sont que faiblement modifiés. On note simplement une légère diminution de la variance sur l'axe de l'IRM T2 quand celle ci est rééchantillonnée sur le T1, ce qui correspond parfaitement aux effets lissants d'un rééchantillonnage. Mais l'effet le plus marquant est visible sur la matière grise et le LCR, très sensibles aux volumes partiels. Si les paramètres de classes sur l'axe de l'IRM T1 restent pratiquement inchangés, les paramètres sur l'axe de l'IRM T2 varient sensiblement, particulièrement la variance du LCR. L'explication est simple : les volumes partiels du LCR au niveau des sillons s'ajoutent aux volumes partiels du rééchantillonnage lors du recalage sur l'IRM T1.

Dans le chapitre 5, la segmentation en tissus a été réalisée sur les IRM T2/DP – taille de voxel : 1*1*2 mm^3. Il est donc possible d'ajouter l'IRM T1 – taille de voxel : 1*1*0.8 mm^3– recalée sur les images T2 pour augmenter la robustesse de l'algorithme. Par contre, intégrer l'IRM T2 FLAIR – taille de voxel : 1*1*4 mm^3– n'est pas recommandé, car le rééchantillonnage va créer des volumes par-

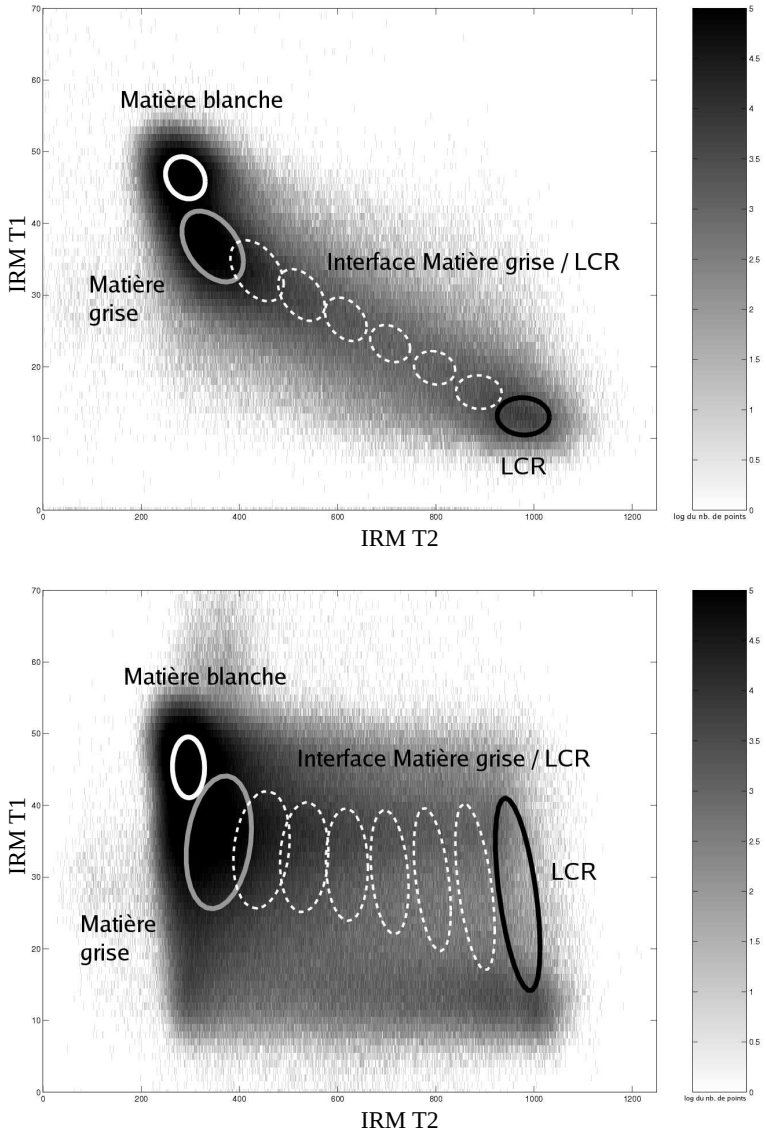


FIG. 8.1 – Histogrammes conjoints de deux séquences IRM (T2 FSE et T1). L'IRM T1 a été recalée sur l'IRM T2. Alors qu'en haut, les deux séquences ont été correctement recalées, en bas, une rotation de 1 degré a été ajoutée à la transformation optimale. Ceci est alors parfaitement visible dans l'histogramme conjoint, et l'estimation des paramètres de classes en est alors fortement altérée. La matière blanche, dont le rapport surface / volume est faible, est peu touchée par ce phénomène. Par contre, l'effet sur la segmentation du LCR est très important : la variance de cette classe a été multipliée par 10.

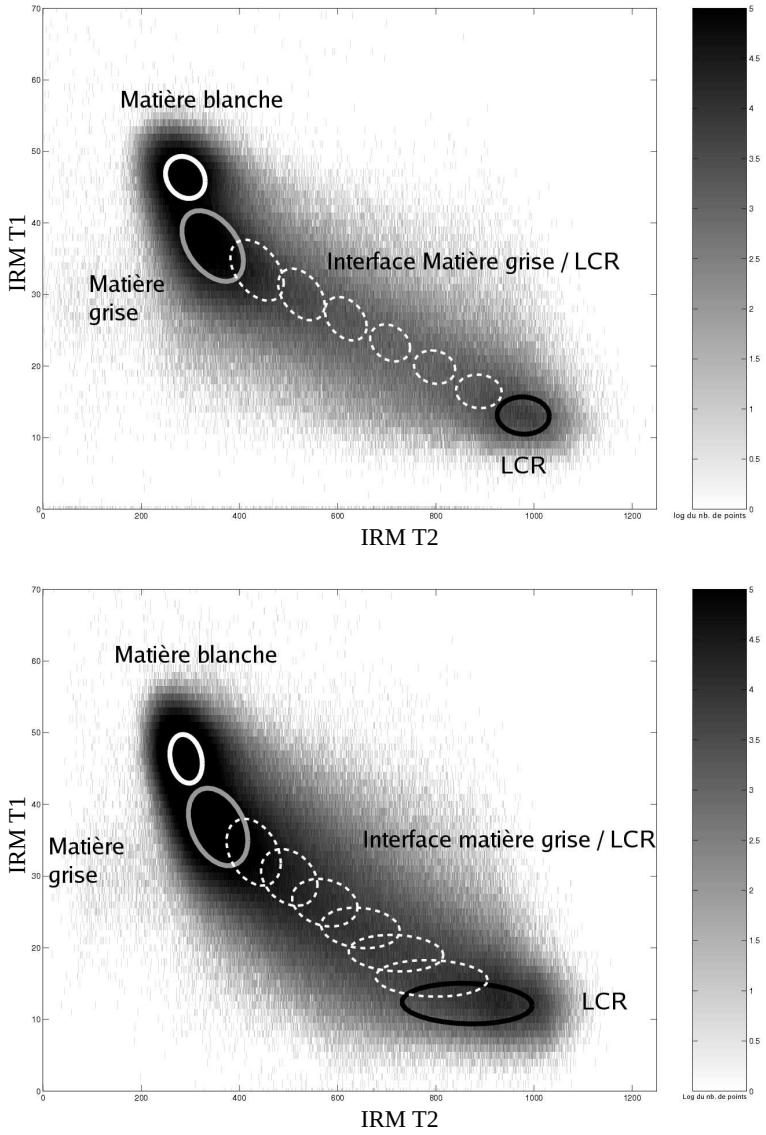


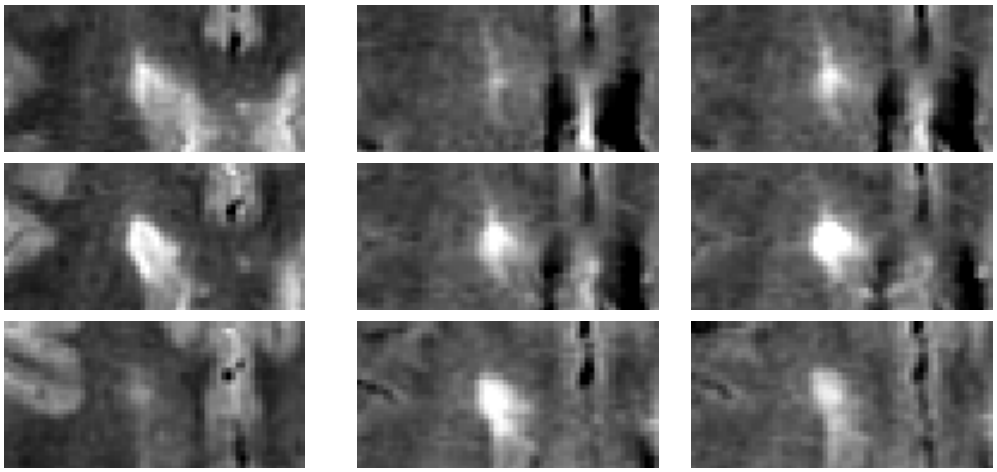
FIG. 8.2 – Histogrammes conjoints de deux séquences IRM (T2 FSE et T1). L'histogramme en haut est le résultat quand le T1 (taille du voxel : 1*1*0.8 mm^3) est recalé sur le T2 (1*1*2 mm^3). En bas, le T2 a été recalé sur le T1. Lors du recalage de deux séquences dont les résolutions sont différentes, il faut choisir une image de référence. Le choix de l'image de plus haute résolution est attractif pour limiter la perte d'information. Ce choix augmente par contre les effets de volumes partiels par une interpolation supplémentaire sur l'image de plus faible résolution : cet effet est visible par une augmentation significative de la variance du LCR en IRM T2.

tiels supplémentaires par rapport à l'information apportée par cette séquence. La résolution des images et le choix de l'algorithme de recalage sont donc des points très importants pour une bonne segmentation.

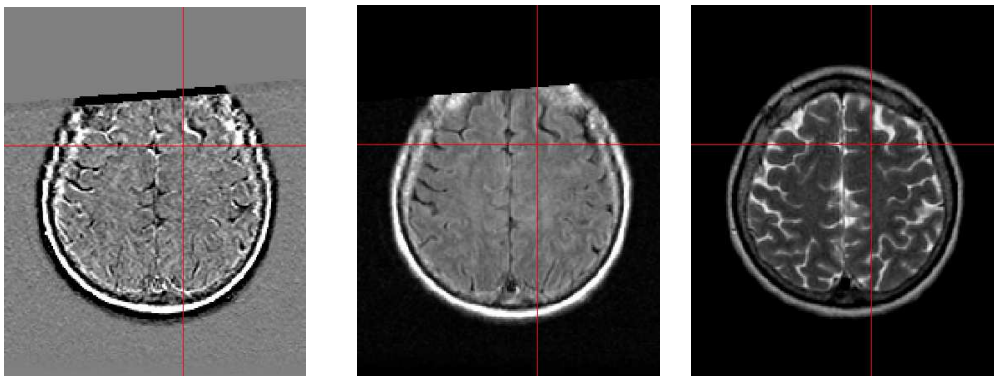
8.1.2 Correction de l'inclinaison de l'axe.

Un bon décodage des images, et particulièrement la correction de l'inclinaison de l'axe Z est indispensable : la détection est faite sur l'IRM T2 FLAIR ; le masque de la région d'intérêt est calculé à partir du T1, et la segmentation finale utilise la segmentation en tissus obtenue à partir des IRM T2/DP. L'inclinaison de l'axe, si elle n'est pas corrigée, peut aboutir à une erreur de recalage de 2 voxels entre les séquences, après recalage rigide. Une des conséquences a été visible lors de l'évaluation : certaines lésions en IRM DP ne correspondaient pas aux hypersignaux en IRM T2 FLAIR. Ceci était dû à une mauvaise reconstruction des images, comme le montre la figure 8.3 : la correction de l'inclinaison de l'axe Z corrige cet effet de décalage d'apparition des lésions entre les deux modalités. Les séquences dérivées du T2 – T2 FSE, densité de protons, T2 FLAIR – sont les plus touchées.

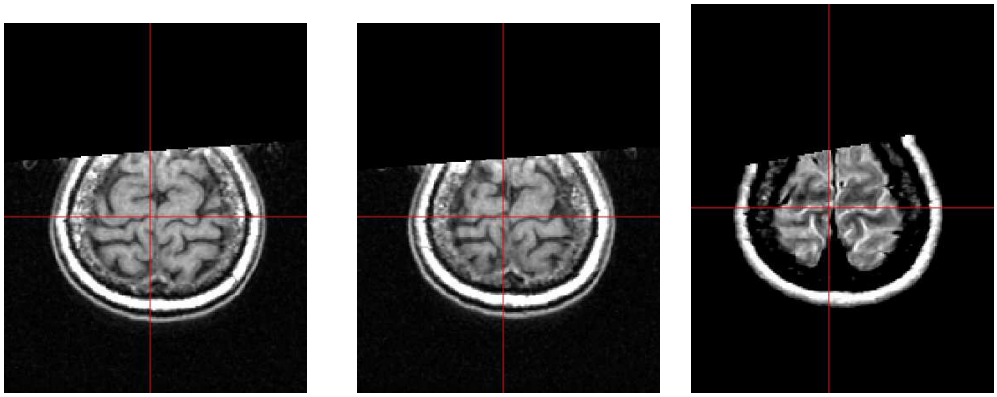
Ce problème concerne de la même manière le masque de la région d'intérêt obtenu à partir du T1. Le recalage rigide global a tendance à placer correctement le centre de l'image, alors que les mauvais alignements sont plus visibles sur les bords (figure 8.3). Comme l'intérêt majeur de cette région d'intérêt est de limiter les artefacts de flux au niveau des ventricules, il est difficile de se représenter l'influence de cette correction sur le résultat final. Il est pourtant important de corriger cette inclinaison : si un système ultérieur, analogue au travail fourni dans Brainvisa [36, 136, 135, 95, 93], permet une segmentation fine des sillons corticaux sans être perturbé par les lésions de SEP, un bon placement des sillons dans toutes les séquences est important pour éviter de faire sortir de la région d'intérêt certaines lésions juxtacorticales et corticales.



Lésion (DP) | *Lésion (T2 FLAIR)* | *Lésion (T2 FLAIR corrigé)*



T2 FLAIR : différence avant / après correction | *T2 FLAIR avant correction* | *IRM T2*



T1 non corrigé | *T1 corrigé* | *Densité de Protons*

FIG. 8.3 – Une mauvaise reconstruction des images peut avoir des conséquences lors de la visualisation des lésions en IRM DP et en T2 FLAIR. Avant correction de l'inclinaison de l'axe, il y a une nette impression de décalage d'environ une coupe entre les deux images, ce qui n'est plus vrai après correction (trois premières lignes). Le positionnement des sillons peut également être compromis (deux dernières lignes)

8.1.3 Correction du biais

Le biais est une source d'erreur pour tout système de segmentation basé sur l'intensité, et il est très important de le prendre en compte. Il peut cependant avoir de nombreuses origines, même si les conséquences sont souvent les mêmes : une mauvaise labélisation des tissus. La correction du biais utilisée est, dans une certaine mesure, optimale pour notre problème, car elle maximise l'uniformité de l'intensité au sein de chaque classe. Ceci permet de corriger un biais dû, par exemple aux inhomogénéités du champ magnétique principal B_0, mais également un biais tissulaire : dans le cas du LCR, le LCR ventriculaire est moins affecté par les effets de volume partiel que le LCR cortical. En IRM T2/DP, une correction du biais peut donc avoir comme effet la réduction de la variance de la classe LCR. Dans ce cas, le modèle de bruit gaussien étant le but final, la variance du LCR se rapproche de la variance de la matière blanche, et l'estimation des paramètres de classes a été améliorée (figure 8.4).

Mais les algorithmes de correction de biais peuvent être aussi utilisés pour éliminer d'autres artefacts. L'algorithme utilisé dans ce manuscrit utilisant une segmentation, laquelle peut être fournie par les résultats du chapitre 5, il est alors facile de corriger l'artefact osseux en IRM T2 FLAIR par une correction du biais. Cet artefact est important à corriger, car il induit des faux positifs dans la région temporo-basale. Il faut néanmoins faire attention à ce que la fréquence du biais spatial recherché soit suffisamment basse, pour ne pas réduire le contraste des grandes lésions.

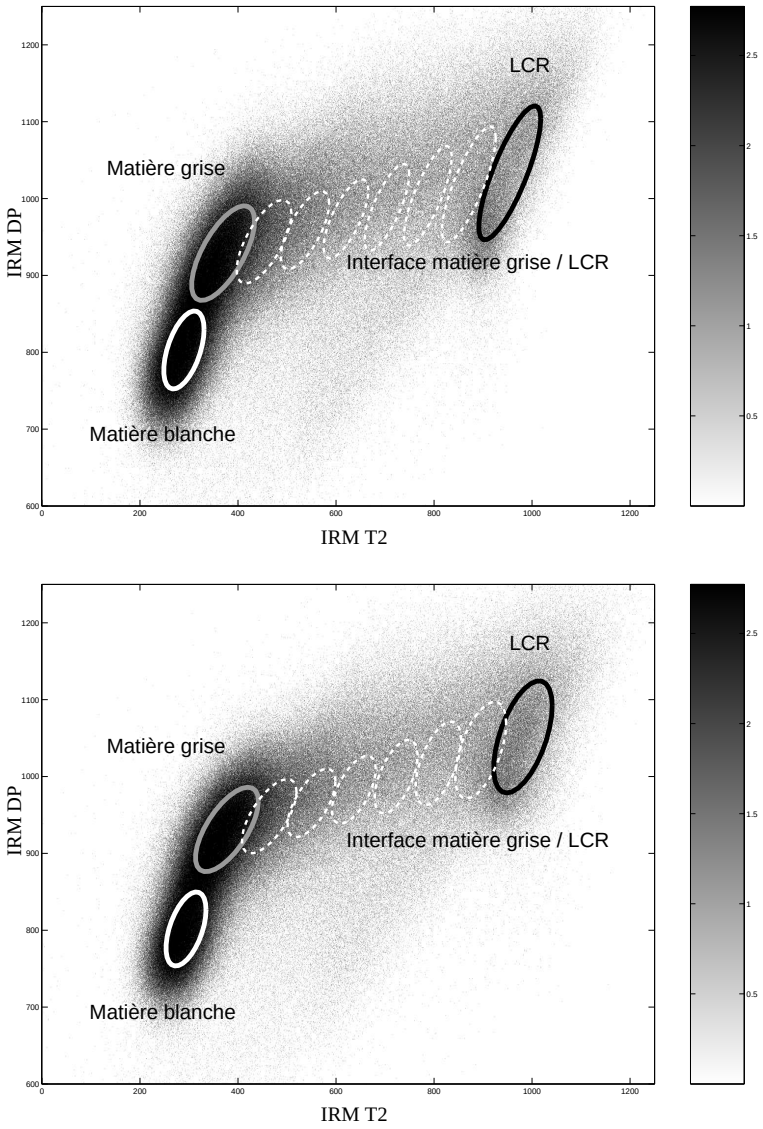


FIG. 8.4 – Histogrammes conjoints des IRM T2 / DP, et visualisation des paramètres de classes par les ellipses de Mahalanobis pour la matière blanche (blanc), la matière grise (gris), l'interface matière grise / LCR (pointillés) et le LCR (noir). L'histogramme du haut visualise le résultat avant correction du biais, en bas, après correction de biais. C'est ici un biais tissulaire entre le LCR cortical et le LCR ventriculaire qui a été corrigé, ce qui explique une diminution de la variance de cette classe.

8.2 Analyse par type de lésion

Les lésions de SEP ont une signature très variable, tant en termes de localisation que de contraste ou de taille. Pour pouvoir mieux cibler les succès et échecs de la chaîne de traitements, les catégories de lésions – périventriculaires, juxtacorticales, corticales, lésions de la fosse postérieure, et les lésions nécrotiques – doivent être regardées séparément.

8.2.1 Lésions périventriculaires : contourage et faux positifs

Les lésions périventriculaires sont les plus caractéristiques de la SEP et les plus nombreuses. Un bon traitement de ces lésions est donc un gage de bon fonctionnement pour un système d'analyse automatique d'IRM pour la SEP. Le choix de l'IRM T2 FLAIR pour la détection facilite grandement le travail : cette séquence assure en effet un très bon contraste entre le LCR, sombre, et les lésions qui apparaissent comme des hypersignaux forts : cette information est notre critère de détection. Par contre, essentiellement à cause des artefacts de flux, des hypersignaux supplémentaires apparaissent et altèrent la qualité de la segmentation.

- Des faux positifs sont ajoutés, par exemple entre les deux ventricules, mais aussi en général à l'interface parenchyme/LCR, comme par exemple autour des sillons corticaux (figure 8.5).
- Les hypersignaux de la substance blanche en T2 FLAIR sont généralement une sur-segmentation des lésions périventriculaires de SEP (figure 8.5).

La méthode choisie permet, la plupart du temps, d'écarter les faux positifs et de corriger le contourage, mais seulement d'une manière approximative. Nous allons voir pourquoi dans les paragraphes suivants, et prendrons deux exemples plus précis : le traitement des cornes ventriculaires et des artefacts de flux au niveau du plan médian sagittal, entre les deux ventricules.

La solution choisie est de tracer la frontière entre le LCR et le parenchyme cérébral : tout hypersignal suspect à un distance inférieure à 2 mm de cette frontière est un artefact de flux ou un volume partiel, et doit donc être considéré comme

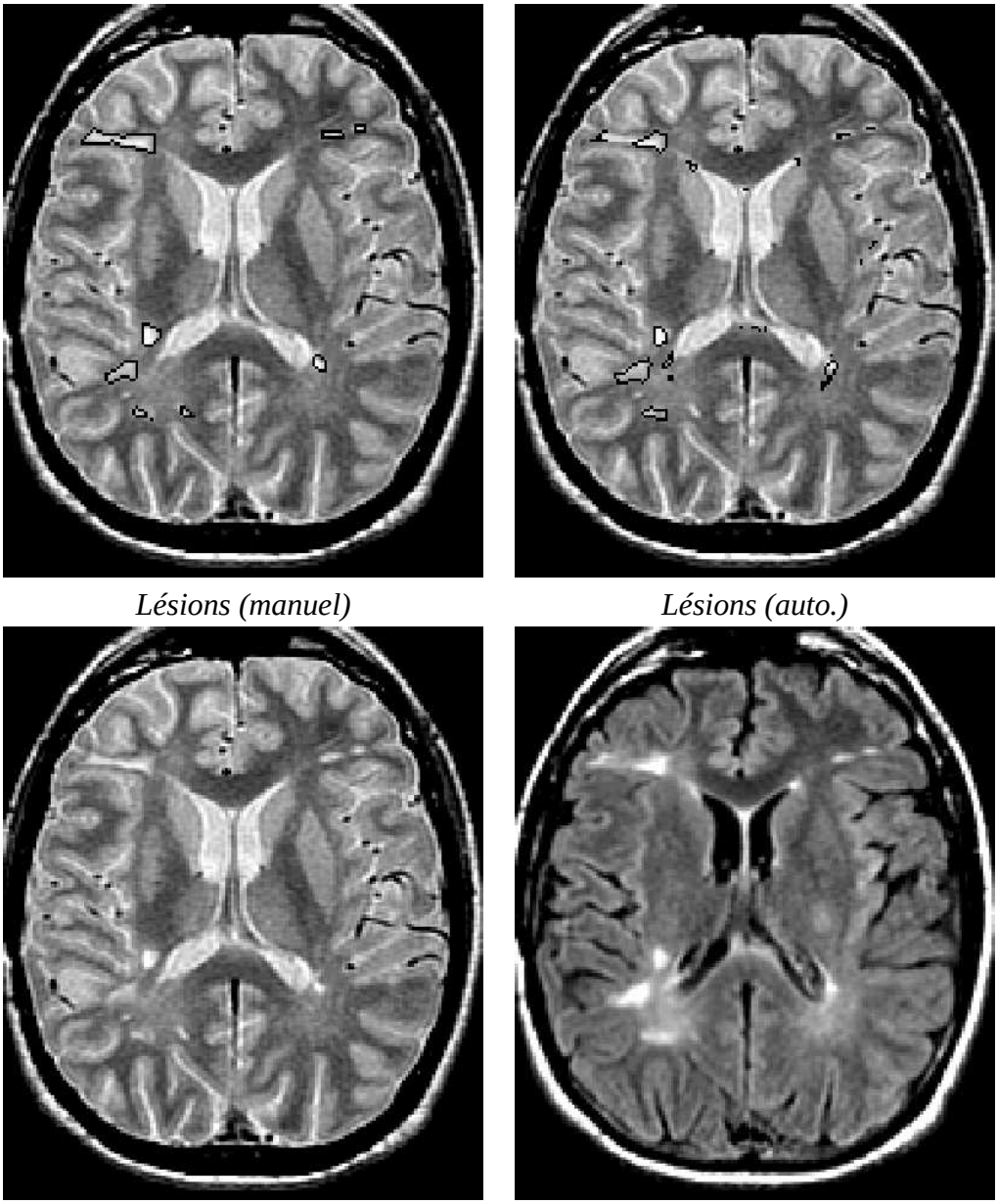


FIG. 8.5 – Les cornes des ventricules (corne antérieure droite sur cette image) et l'espace entre les deux ventricules sur cette coupe souffrent souvent d'artefacts de flux (hypersignaux) en T2 FLAIR, qui conduisent à des faux positifs. La bande de sécurité calculée à partir du T1 permet de limiter ce problème.

un faux positif. Ceci nécessite d'obtenir une segmentation du LCR – des ventricules dans le cas présent – à l'aide d'une image haute résolution : c'est la raison pour laquelle le T1 a été choisi. Le choix d'un seuillage simple sur l'IRM T1 accompagné d'opérations de morphologie mathématique est discutable, mais il permet aisément d'éliminer à coup sûr les artefacts de flux quand les ventricules ont une taille suffisante (figure 8.6). Par contre, des problèmes surviennent quand les volumes partiels gênent la segmentation de cette interface liquide/solide ; c'est le cas pour la majorité des sillons corticaux, mais aussi dans le cas où les ventricules sont trop fins pour être segmentés par un simple seuillage (figure 8.7). Le système échoue à positionner la frontière, et un artefact de flux sera segmenté comme lésion. Pour améliorer la qualité du contourage du système, deux points sont essentiels.

Meilleure segmentation de l'interface parenchyme LCR : une bonne segmentation de la frontière parenchyme/LCR est nécessaire, car c'est là que se situe la quasi-totalité des artefacts de flux. D'autres systèmes de segmentation des IRM T1 existent (figure 6.7). Les techniques de segmentation développées dans le logiciel Brainvisa [36, 136, 135, 95, 93] sont un très bon point de départ, mais le système doit être rendu plus robuste à la présence de lésions de SEP. Cette direction de recherche inclut une bonne gestion des volumes partiels, très importants dans la frontière parenchyme/LCR ; une connaissance *a priori* de l'anatomie cérébrale – variabilité des sillons corticaux [21, 131, 62, 37, 38, 96], par exemple – serait également profitable.

Introduction d'une connaissance *a priori* sur les artefacts de flux : cette première localisation des artefacts de flux n'est pas suffisante ; et il convient d'introduire une connaissance *a priori* sur leur localisation et leur intensité. Ceci nécessite, bien sûr, une base de donnée de témoins et non plus de patients, pour pouvoir effectuer des statistiques plus précises sur l'influence des artefacts de flux dans les IRM dérivées du T2 : T2, densité de protons, T2 FLAIR. Il reste ensuite à utiliser cette connaissance pour améliorer la segmentation des lésions, en particulier le contourage des lésions périventriculaires. Prenons deux exemples.

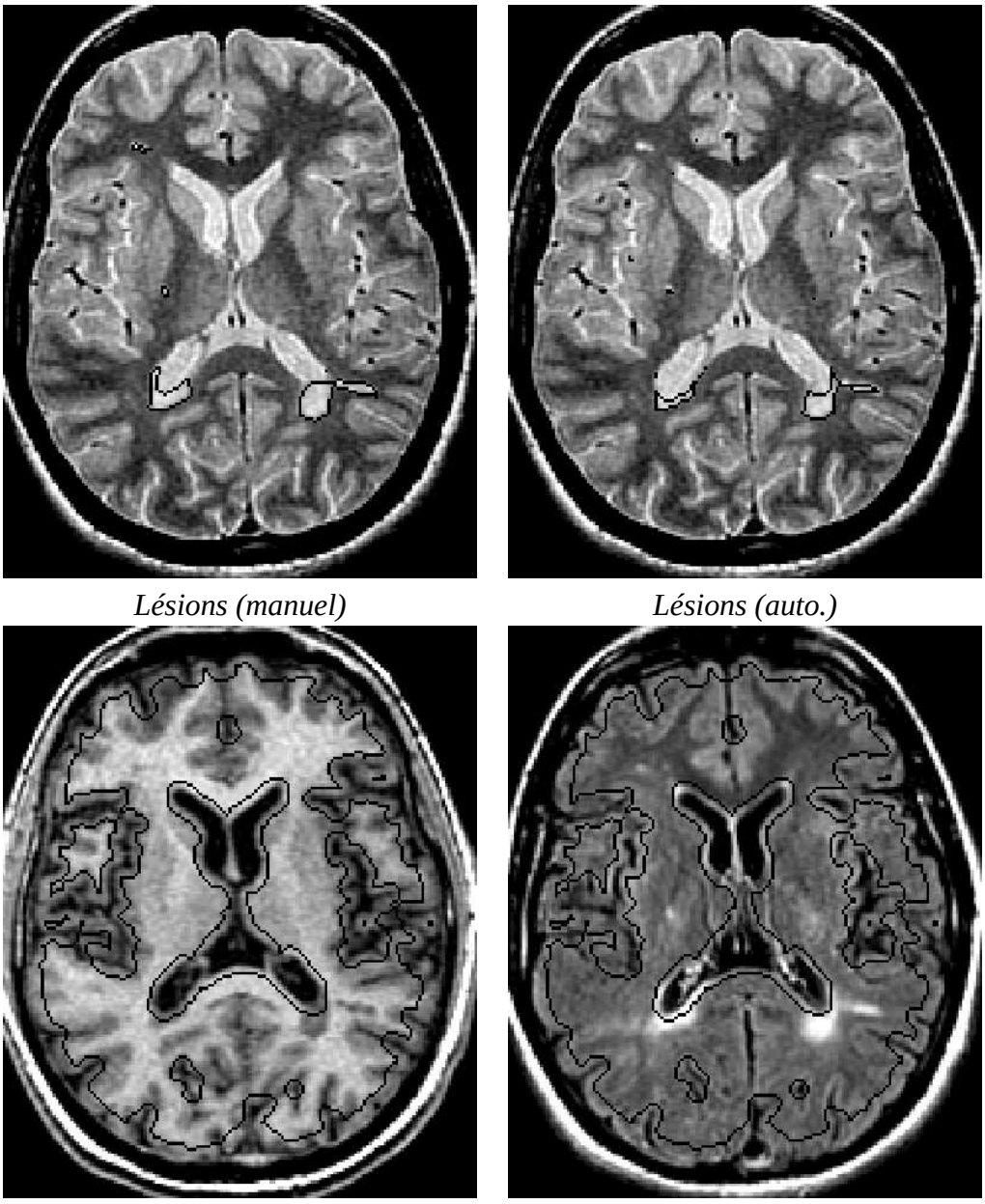


FIG. 8.6 – La région d'intérêt calculée à l'aide de l'IRM T1, visualisée en surimpression sur les deux coupes du bas, permet d'éliminer les artefacts de flux à la frontière entre les ventricules et le parenchyme cérébral, sans perturber la détection des lésions périventriculaires.

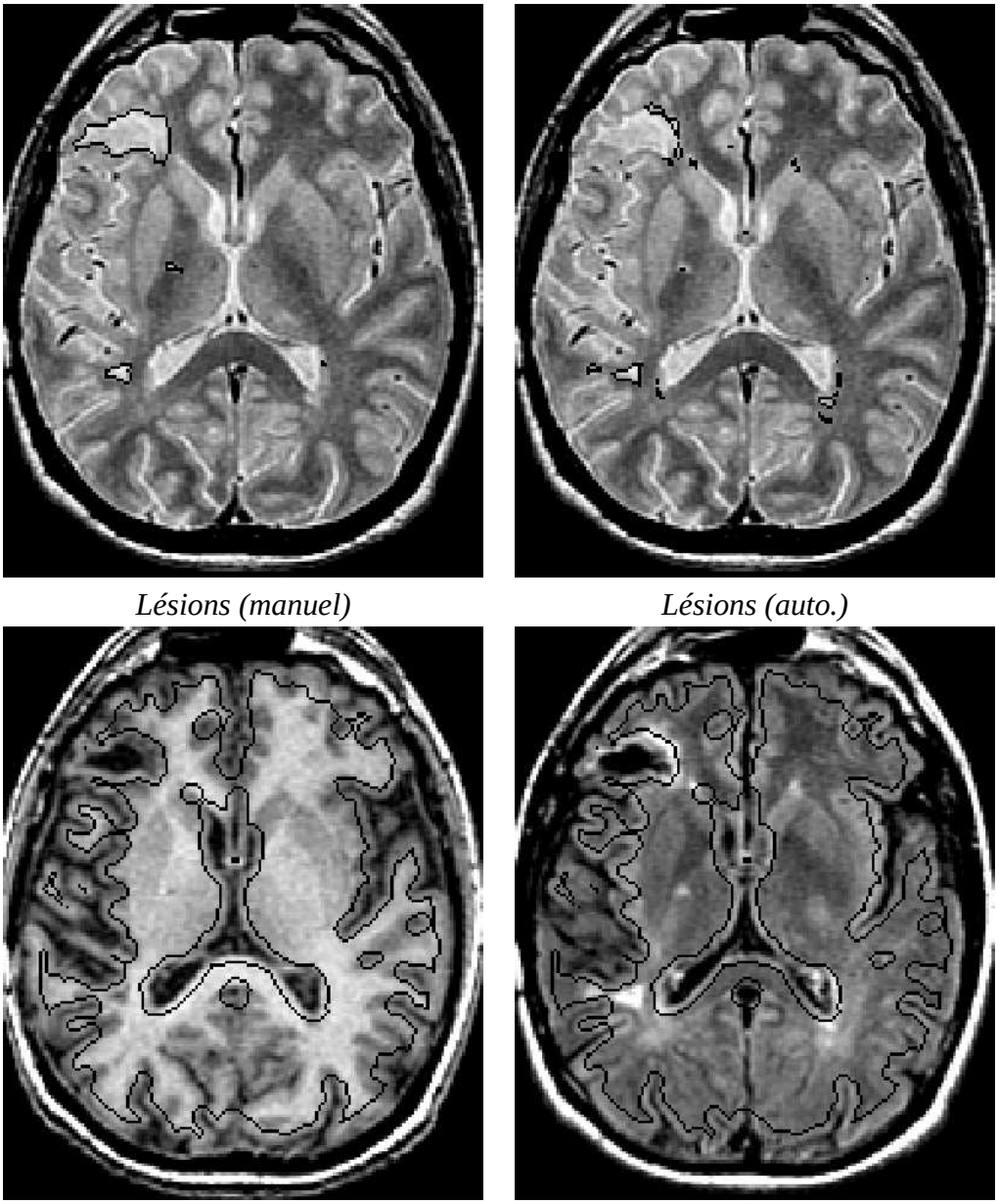


FIG. 8.7 – La finesse des cornes antérieures des ventricules perturbe la construction de la région d'intérêt issue du T1 : des faux positifs sont visibles à l'extrémité de ces cornes. De plus, le *trou noir* présente dans la partie antérieure gauche de cette image est très mal contouré par le système : la lésion étant connexe au LCR cortical, il est en effet impossible, sans critères morphologiques supplémentaires, de différencier la nécrose du LCR cortical.

8.2.1.1 Contourage des cornes ventriculaires

Le problème cité au paragraphe précédent est d'autant plus délicat pour la segmentation des cornes ventriculaires, très touchées par les volumes partiels et facilement confondues avec des lésions de SEP. Dans la figure 8.8, la corne postérieure gauche est segmentée comme une lésion, car l'hypersignal en FLAIR est particulièrement clair.

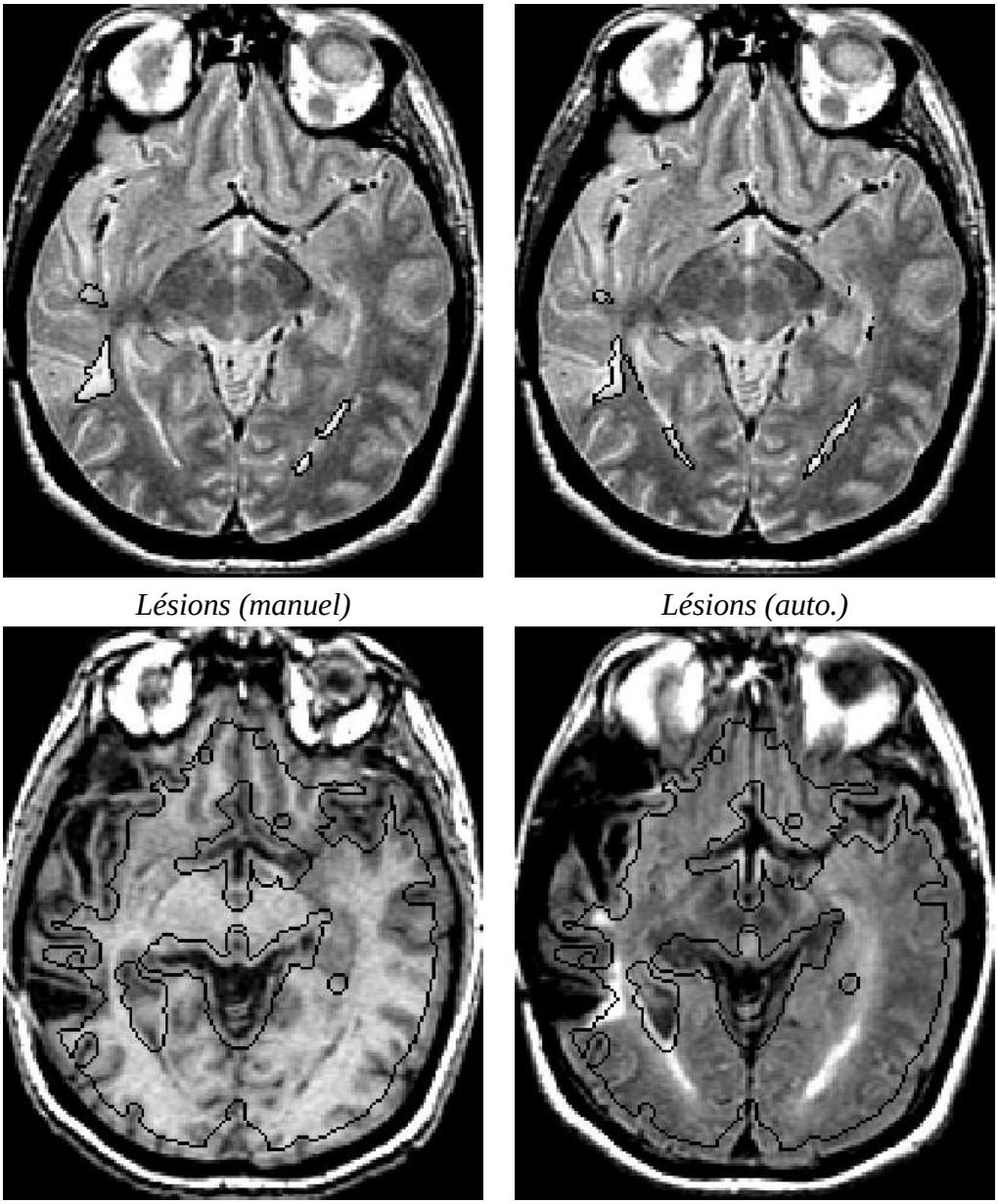
Les cornes ne sont finalement qu'une partie des ventricules ; leur segmentation est pourtant un problème encore plus délicat, même sur les sujets sains. Elle doit être couplée à un traitement adapté des volumes partiels – l'épaisseur des cornes est fréquemment inférieure à 1 mm – et nécessite une base importante de sujets sains pour être sûr de ne pas confondre ces cornes avec d'éventuelles lésions de SEP. Il reste ensuite à analyser les lésions de SEP et à les différencier des cornes, avec l'aide d'un expert, dans le cadre d'une évaluation plus spécifique.

8.2.1.2 Artefacts de flux entre les deux ventricules

Parfois, la position des artefacts de flux est prévisible. Dans le cas des artefacts présents entre les deux ventricules principaux, comme sur la figure 8.5, une segmentation correcte des ventricules permettrait d'anticiper la présence de ce faux positif. Ceci a été pris en compte dans la chaîne de traitements grâce au T1, et le faux positif a été supprimé. Par contre, certaines lésions périventriculaires sont alors sous-estimées, comme sur la figure 8.9. Dans ce cas, la naissance de la lésion est confondue avec un artefact de flux par le système, et la lésion est sous-segmentée. Il semble donc que l'intensité des artefacts de flux varie avec leur localisation. Une étude plus précise de ces artefacts sur une base de sujets sains serait préférable pour améliorer la qualité du résultat.

8.2.2 Lésions juxtacorticales et corticales

Les lésions juxtacorticales et corticales ont leur importance dans les critères de Barkhof, et ne doivent pas être laissées de côté. Elles sont plus touchées que les lésions périventriculaires par les volumes partiels à cause de leur taille : elles

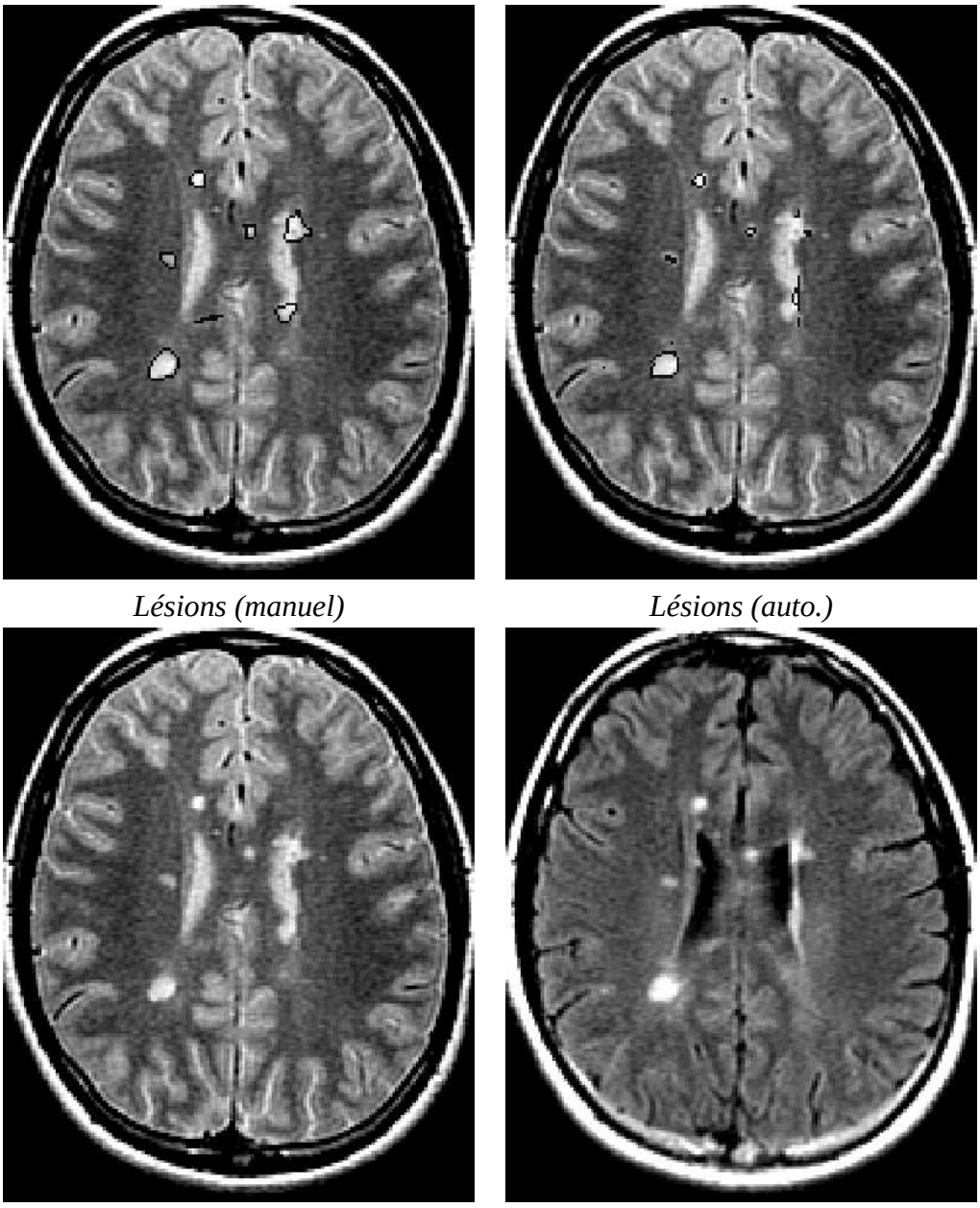


Lésions (manuel) *Lésions (auto.)*

IRM T1 + masque *IRM T2 FLAIR+ masque*

Patient : MCI02

FIG. 8.8 – La sclérose en plaques n'est pas le seul problème chez ce patient : il faut différentier les lésions de SEP de ce qui semble être une résection chirurgicale. Le système a ici un comportement raisonnable malgré la complexité du problème. On note par contre la détection comme lésion (en bas à gauche de l'image) d'un hypersignal en T2 FLAIR qui n'est en fait qu'un effet de flux sur les cornes de ventricules postérieurs.



Lésions (manuel) *Lésions (auto.)*

IRM DP *IRM T2 FLAIR*

Patient : MCI01

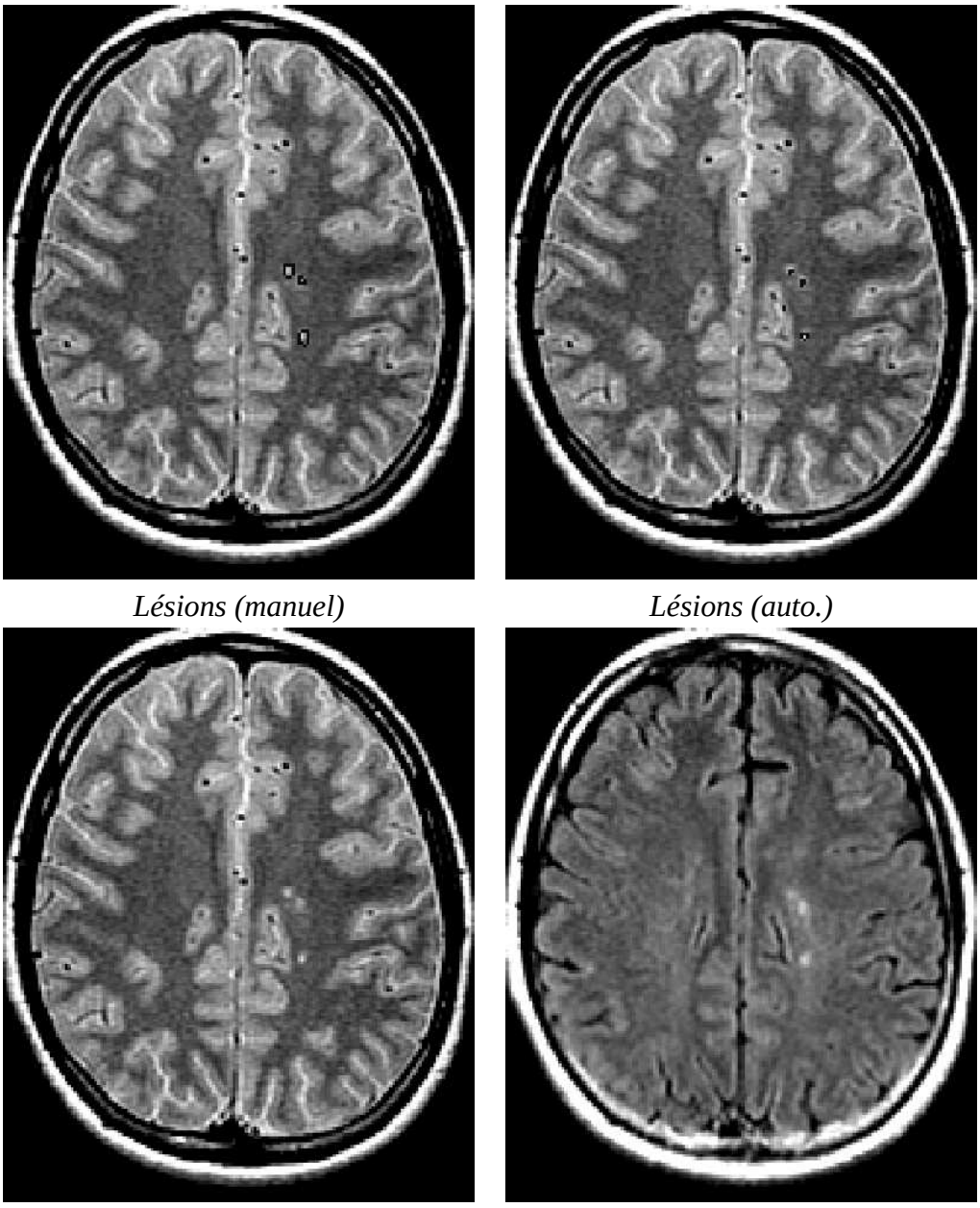
FIG. 8.9 – Les lésions périventriculaires sont ici sous-estimées par le système par rapport à la segmentation manuelle. Dans le cas des lésions péri-ventriculaires, il est toujours difficile de savoir où s'arrêtent précisément les ventricules, et la distance de sécurité autour des ventricules calculée à partir de l'IRM T1 contribue à cette sous-estimation.

sont donc plus difficilement détectées.

8.2.2.1 Lésions juxtacorticales

Les lésions juxtacorticales sont particulières : souvent plus petites que les lésions périventriculaires. Moins contrastées, elles se confondent plus facilement avec de la matière grise. Les plus petites font moins de 1 millimètre de diamètre, ce qui les rend très difficiles à détecter en IRM. Même les lésions de quelques millimètres sont délicates à détecter à cause des phénomènes de volumes partiels et quand elle sont détectées, elle sont généralement sous-segmentées (figure 8.10). L'épaisseur de coupe de cette séquence (taille du voxel : 1*1*4 mm^3) est un facteur limitant très important : toute lésion dont l'épaisseur est inférieure à 4 millimètres souffrira d'un manque de contraste et sera plus difficile à détecter. Plusieurs directions sont possibles pour résoudre ce problème.

- Dans un premier temps, la solution la plus simple serait de restreindre la zone de recherche pour choisir un seuillage plus sensible en IRM T2 FLAIR. Ceci nécessite une localisation plus précise des faux positifs en IRM T2 FLAIR, donc une bonne segmentation de l'interface parenchyme/LCR.
- Ensuite, il est toujours possible de raffiner l'analyse de l'IRM T2 FLAIR par des techniques plus fines que le seuillage, comme des seuillages adaptatifs ; mais ceci présuppose d'avoir une idée sur la taille de la lésion à segmenter. Cette taille étant très variable pour les lésions de SEP, ce genre de méthode est difficile à appliquer.
- Une autre possibilité est de détecter les lésions juxtacorticales en utilisant d'autres séquences – comme l'IRM DP – au lieu de l'IRM T2 FLAIR. Par contre, si l'hypersignal des lésions est en général très visible en T2 FLAIR, leur signal est proche de celui de la substance grise en IRM DP. Il faut donc rajouter des critères géométriques ou topologiques pour différencier, dans le cas des lésions juxtacorticales, le cortex des lésions. Des critères de connexité ont déjà été abordés dans le cadre d'études sur la connexité floue [162, 163]. Nous avons tenté de les appliquer, sans résultats concluants, mais il semble intéressant et prometteur de poursuivre dans cette voie.



Lésions (manuel) *Lésions (auto.)*

IRM DP *IRM T2 FLAIR*

Patient : MCI01

FIG. 8.10 – Dans le cas de petites lésions, l'épaisseur de coupe de l'IRM T2 FLAIR est un important facteur limitant, car les effets de volume partiel sur la lésion diminuent le contraste de l'image T2 FLAIR, donc les capacités de détection de celle-ci. Une segmentation directement sur la densité de protons nécessiterait des critères géométriques et topologiques supplémentaires.

8.2.2.2 Lésions corticales

La détection des lésions corticales présente les mêmes difficultés que celle des lésions juxtacorticales, mais la localisation de ces lésions rend difficile l'application de critères géométriques et topologiques : la littérature médicale confirme que la plupart des lésions corticales ne sont pas détectées en IRM conventionnelle [57]. La détection est d'autant plus délicate qu'en T2 FLAIR, les sillons corticaux sont sujets aux artefacts de flux : il est délicat de différencier l'extrémité d'un sillon d'une lésion corticale de SEP (figure 8.11). Un meilleur modèle de volume partiel incluant les artefacts de flux est ici nécessaire pour une bonne segmentation des sillons corticaux. Une autre piste est d'utiliser l'épaisseur du cortex : un hypersignal suspect détecte la lésion, et cette détection est confirmée par la présence d'un renflement dans le cortex. L'épaisseur et la structure du cortex sont un sujet d'étude pour de nombreuses maladies [159, 32, 13, 93, 96]. Il reste à élaborer un système qui demeure robuste à la présence de lésions de SEP.

8.2.3 Lésions nécrotiques

Les lésions nécrotiques – ou *trous noirs* – sont plus délicates à segmenter que de "simples" lésions périventriculaires : la partie nécrosée de la lésion est en fait liquide, et a donc en IRM un signal de LCR. Si la "couronne" autour de la lésion est généralement bien contrastée et permet une bonne détection à partir de l'IRM T2 FLAIR (figure 8.12), le centre de la lésion, hyposignal en FLAIR, est labélisé comme du LCR. Le problème n'est pas pour autant facile à résoudre, car lorsque le centre nécrotique de la lésion est connexe avec le LCR, rien hormis un critère de forme ne permet de les différencier (figure 8.13). Ce critère de forme est difficile à appliquer, car la structure des sillons corticaux et de la fosse postérieure est très complexe ; contourer une lésion de SEP dans cette structure reste probablement un problème très difficile.

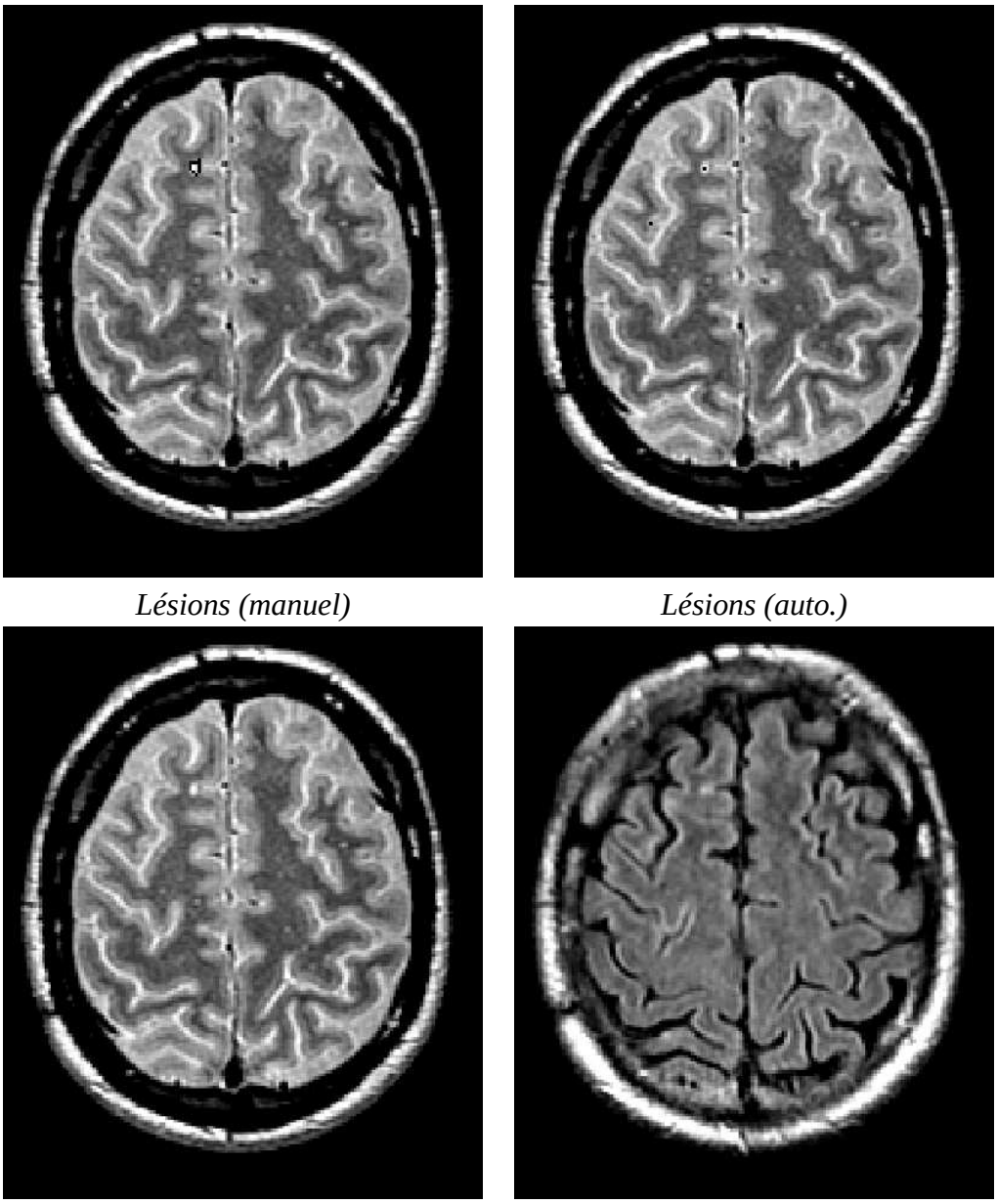
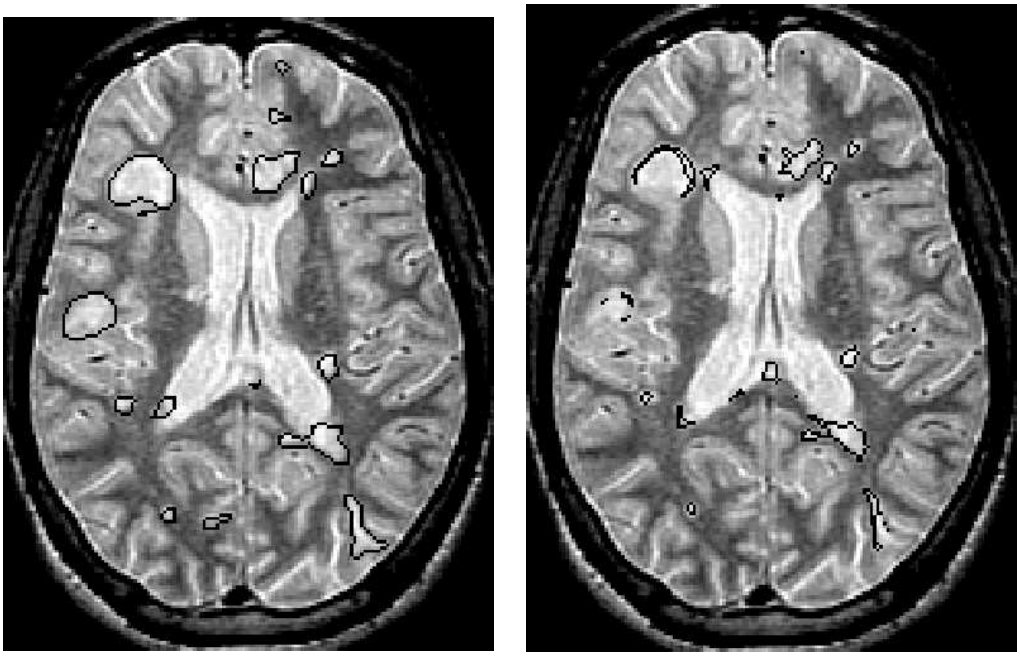
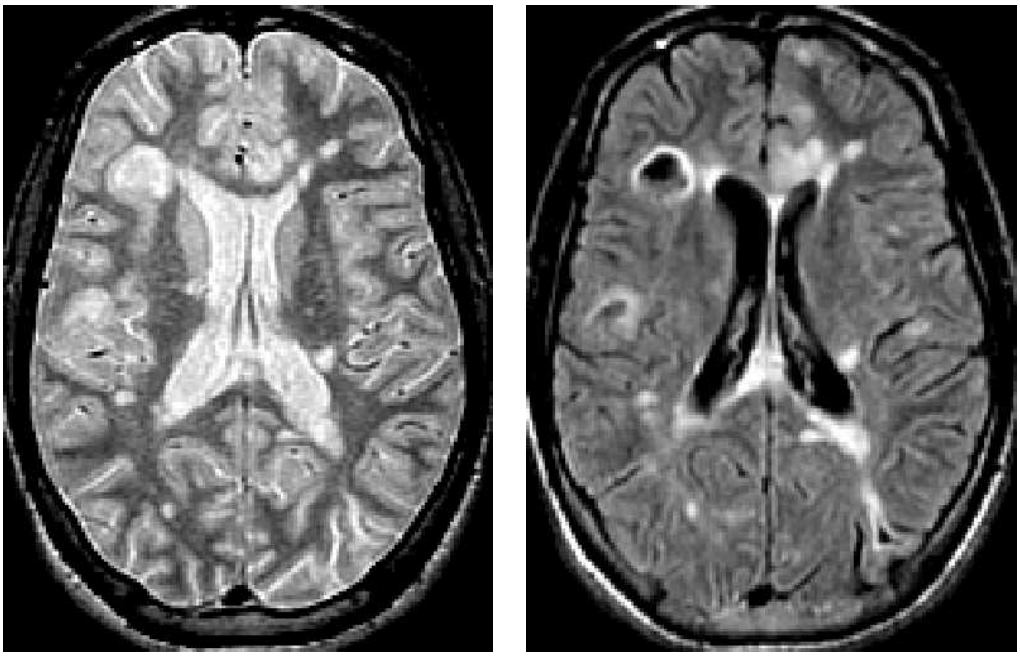


FIG. 8.11 – Toute lésion dont l'épaisseur est inférieure à l'épaisseur de coupe de l'IRM T2 FLAIR (4 millimètres dans le protocole d'acquisition utilisé) souffre d'une mauvaise détection à cause des volumes partiels entre la lésion et les tissus environnants.



Lésions (manuel) *Lésions (auto.)*



IRM DP *IRM T2 FLAIR*

Patient : MCI07

FIG. 8.12 – Les lésions de la matière grise, moins contrastées en IRM T2 FLAIR, et les *trous noirs* – hypo-signaux dans cette modalité – posent problème sur cette exemple : seuls les voxels sur la périphérie des *trous noirs* sont détectés comme pathologiques par la chaîne de traitement.

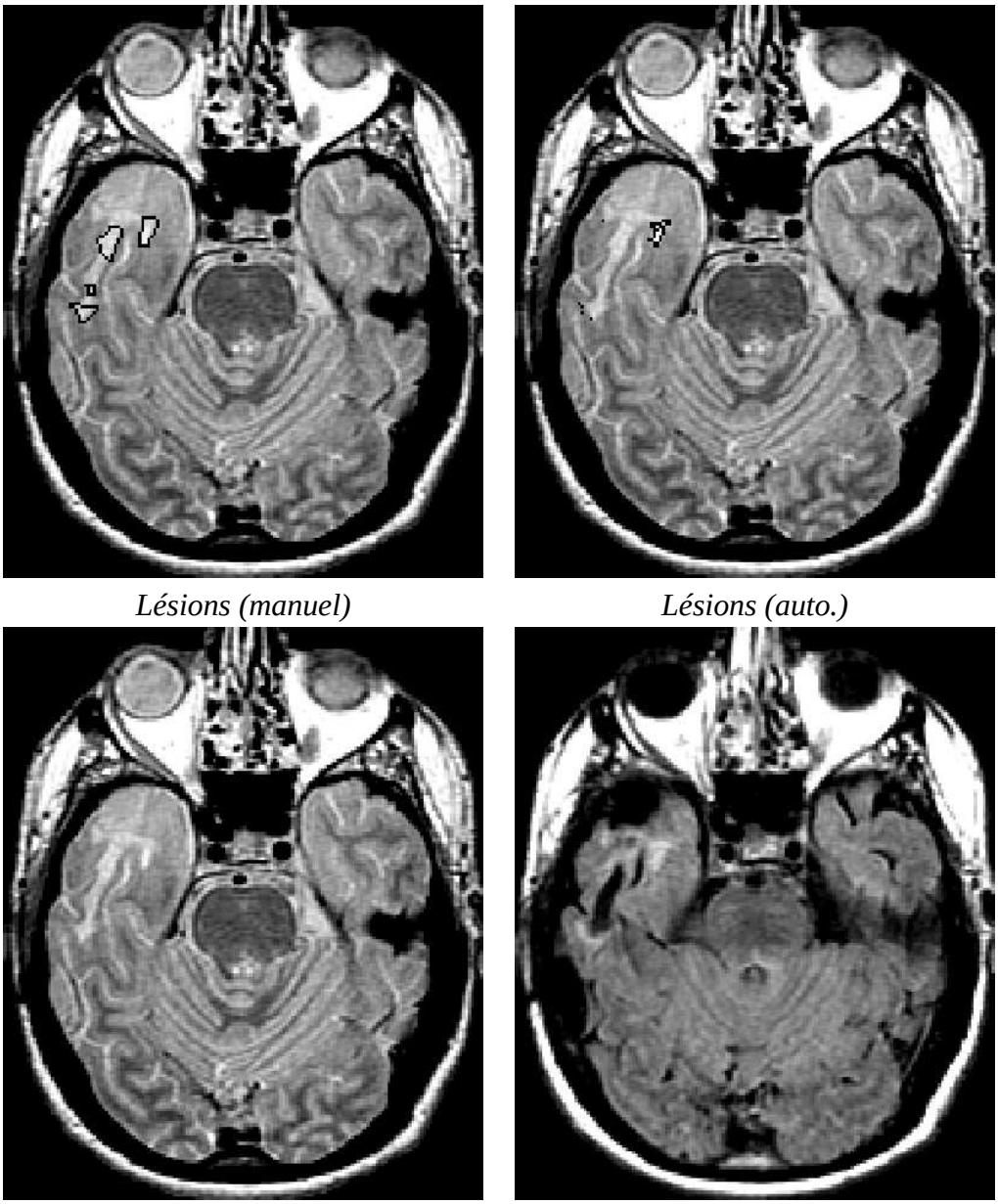


FIG. 8.13 – La lésion de SEP présente sur la partie antérieure gauche de ces coupes est connexe au LCR cortical ; elle est donc considérée comme faisant partie de l'extérieur du cerveau, ce qui explique les faux négatifs sur cet image. La présence de ce qui semble être une résection chirurgicale ne facilite pas le contourage de cette zone.

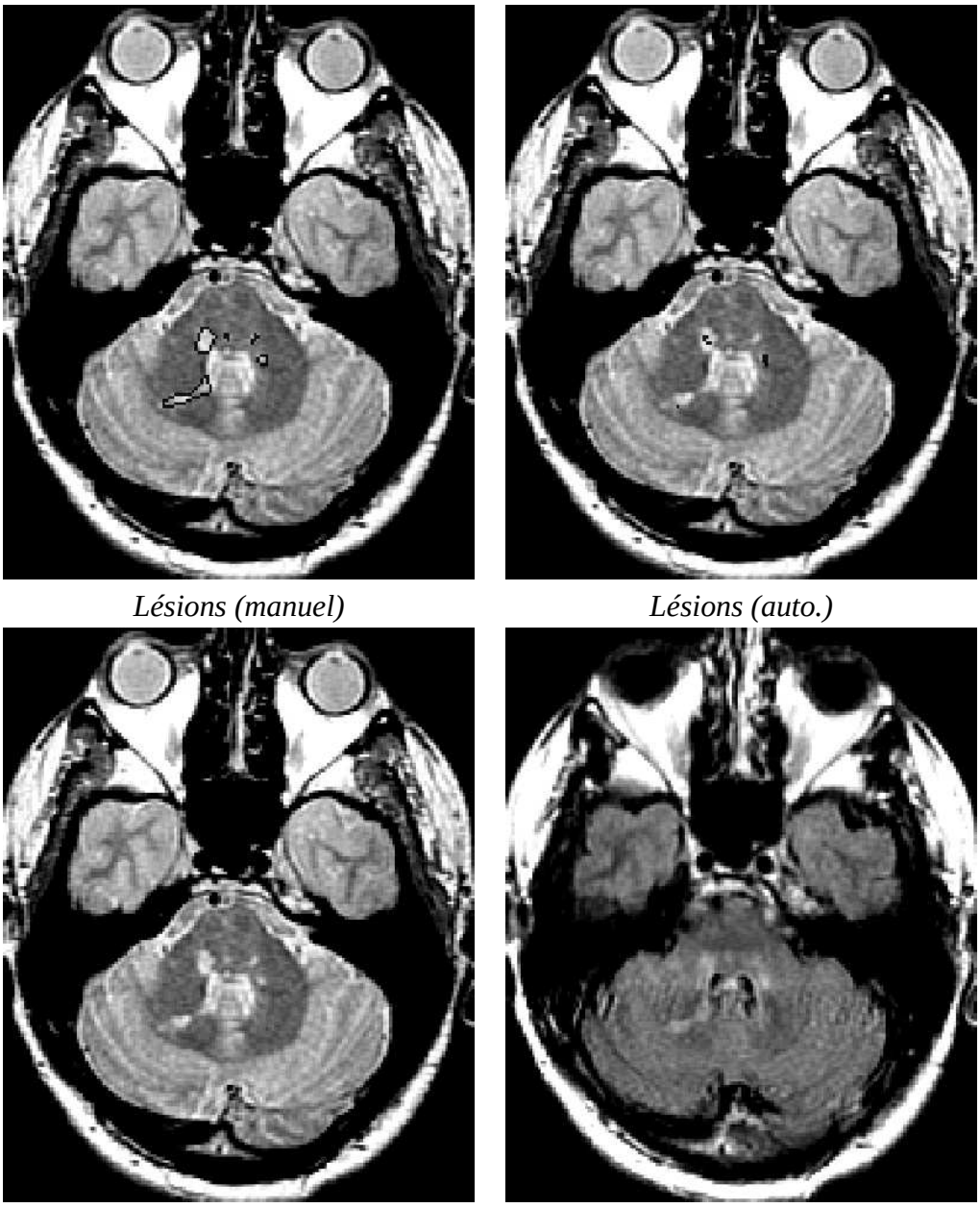
8.2.4 Lésions de la fosse postérieure

Lors de cette étude et de la présentation des résultats au chapitre précédent, les lésions de la fosse postérieure ont été écartées. La chaîne de traitements n'est ici pas appropriée, car ces lésions sont détectables en IRM DP (figure 8.14) mais pas en IRM T2 FLAIR. Les lésions ont par contre un signal proche de celui de la matière grise en IRM DP ; l'intensité ne peut pas être utilisée seule comme critère de détection. Dans ce cas, une piste pourrait être d'obtenir dans un premier temps un masque de ce que serait la matière blanche s'il n'y avait pas de lésions. Un recalage non-rigide ou l'application d'un modèle déformable sur la matière blanche du cervelet sont plusieurs pistes à explorer : il suffit alors d'extraire les lésions comme des points aberrants dans cette zone ne devant contenir que de la matière blanche.

Des lésions sont également visibles dans le tronc cérébral : ces lésions sont importantes car comme les lésions dans le cervelet, elles sont souvent associées à une atteinte physique définitive. Ces lésions sont très difficiles à voir, et la non-uniformité du signal dans cette zone ne facilite pas la lecture des images.

8.3 Conclusion

Nous avons ici regardé les lésions en fonction de leur type : périventriculaire, juxtacorticale, corticale et fosse postérieure avec le cas particulier des *trous noirs*. Ce classement correspond à une réalité médicale, car ces différentes catégories de lésions sont utilisées dans les critères de diagnostic et dans l'observation de la maladie. Dans le cadre de cette évaluation, la catégorisation de ces lésions a été effectuée à la main et de manière arbitraire. Il serait intéressant de définir une caractérisation précise de ces catégories à l'aide de quantificateurs – signal, taille, position, localisation par rapport à certaines structures cérébrales. L'automatisation de cette caractérisation permettrait l'extraction automatique des critères de Barkhof, mais aussi de pouvoir explorer de nouveaux algorithmes de segmentation spécifiques à chaque catégorie : segmenter chaque type de lésion de manière différente semble indispensable pour obtenir des résultats plus homogènes.



Lésions (manuel) *Lésions (auto.)*

IRM DP *IRM T2 FLAIR*

Patient : MCI01

FIG. 8.14 – Il est admis que la segmentation des lésions de SEP en IRM T2 FLAIR dans la zone sous-tentorielle conduit à l'apparition de faux positifs. La correction du biais opérée en FLAIR permet de limiter les dégâts, mais le problème est lié à la modalité elle-même. Les lésions sous-tentorielles doivent donc être segmentées autrement, à l'aide de l'IRM DP.

Cette étude aura montré que le choix d'algorithmes simples prenant pour base l'intensité des voxels donne des premiers résultats qui peuvent servir de socle à des traitements postérieurs. La poursuite d'une telle étude et le perfectionnement de la chaîne de traitements nécessitent néanmoins une évaluation plus approfondie. L'intégration de plusieurs experts pour une évaluation de plus grande envergure est recommandée, tant la pathologie se manifeste de manière diverse en IRM. La définition d'une méthode précise de segmentation manuelle des lésions de SEP permettrait de réduire cette variabilité en précisant le problème médical associé au problème d'analyse d'image.

Regarder uniquement les lésions ne suffit pourtant pas. Pour mieux segmenter les lésions de SEP, il semble nécessaire de mieux comprendre l'anatomie cérébrale et le processus d'acquisition IRM. Comme il est expliqué dans le chapitre 8, les artefacts de flux, très importants en IRM T2 FLAIR, gagneraient à être étudiés ; une bonne segmentation des ventricules cérébraux permettrait un meilleur contourage des lésions périventriculaires ; une étude des sillons corticaux et de l'épaisseur du cortex améliorerait la détection et le contourage des lésions juxta-corticales et corticales. Mais cette étude de l'anatomie cérébrale doit se faire en présence de lésions de sclérose. C'est là toute la difficulté du problème : obtenir des informations sur la structure générale du cerveau sans être perturbé par ces hypersignaux de la substance blanche.

D'autres quantificateurs sont enfin à explorer. L'atrophie cérébrale, l'évolution de la taille des ventricules, la matière blanche d'apparence normale sont autant de sujets d'étude intéressants et ouverts, avec des applications directes pour la sclérose en plaques, tant pour l'aide au diagnostic que pour le suivi des patients dans le cadre d'études cliniques. La base de données obtenue grâce au partenariat avec les docteurs Lebrun et Chanalet au CHU Pasteur à Nice continue d'être renforcée, et il reste encore beaucoup à faire.

Chapitre 9

Annexes

9.1 Quantificateurs, biomarqueurs et marqueurs cliniques

Lors de la définition de quantificateurs issus de l'imagerie médicale, il est intéressant de savoir comment définir le travail issu des techniques d'analyse automatique d'images. Dans [64, 174], une définition des termes majeurs pour les quantificateurs est proposée. Les termes, définitions et caractéristiques suivantes furent proposées pour décrire des mesures biologiques dans le développement thérapeutique.

Biomarqueur : Une caractéristique qui est mesurée objectivement et évaluée comme un indicateur de processus biologiques normaux, de processus pathogènes ou de réponse à une prise d'un médicament.

Les biomarqueurs ont une grande valeur dans les évaluations d'efficacité et de sécurité comme lors d'études *in vitro* sur des échantillons de tissus, lors d'études *in vivo* sur animaux et dans les premières phases d'essais cliniques pour établir une preuve de fonctionnalité. Les biomarqueurs peuvent avoir d'autres applications dans la détection de processus pathogènes et le suivi de l'état général du patient. Ces applications incluent les points suivants.

– Un outil diagnostique pour l'identification des patients ou sujets dont l'état

général laisse présager qu'ils sont malades (forte concentration de glucose dans le sang pour le diagnostic du diabète, par exemple).

- Un outil pour déterminer le stade actuel de la maladie ou spécification de la maladie (forte concentration d'un antigène spécifique à la prostate dans le sang pour le diagnostic d'une extension de la croissance de tumeur ou d'une métastase).
- Un indicateur pronostique pour une maladie (mesure de réduction de la taille de la tumeur sur certains cancers).
- Un suivi de la réponse clinique à une intervention externe (taux de cholestérol afin de déterminer le risque d'une maladie cardio-vasculaire).

Marqueur clinique : Un marqueur clinique est une caractéristique ou variable qui reflète comment le patient se sent, fonctionne ou survit. Les marqueurs cliniques sont des mesures distinctes ou des analyses de caractéristiques spécifiques à une maladie, observées lors d'une étude ou un test clinique, qui reflètent l'effet d'une prise médicamenteuse. Les marqueurs cliniques sont la plus crédible des caractéristiques utilisées pour l'évaluation de la prise d'un médicament en termes de risque et d'amélioration dans les tests cliniques.

Substitut de marqueur clinique : Un substitut de marqueur clinique est un biomarqueur conçu pour se substituer à un marqueur clinique. Il est donc sensé prédire une amélioration clinique (ou une détérioration) basée sur des études épidémiologiques, thérapeutiques, psycho-pathologiques ou d'autres preuves scientifiques solides.

Plusieurs remarques sont à faire par rapport à ces définitions. Lors de l'étude de maladies chroniques, les tests cliniques utilisant des marqueurs cliniques sont souvent longs, nécessitent un grand nombre de patients et sont donc extrêmement chers. Les substituts de marqueurs cliniques sont donc un ajout appréciable et diminuent de façon significative la longueur et les coûts de ces tests. En outre, les substituts de marqueurs cliniques sont donc un sous-ensemble des biomarqueurs. Si ces substituts peuvent être considérés comme des biomarqueurs, très peu de biomarqueurs atteignent le stade de substitut de marqueur clinique : la plupart des

quantificateurs issus de l'imagerie sont donc "évalués" (évaluation du facteur de risque) plutôt que "validés" (preuve du pouvoir prédictif). Typiquement, l'évaluation de ces biomarqueurs issus de l'imagerie comprend un certain nombre de points comme la reproductibilité des résultats, une forme d'invariance à un changement d'échelle (changement d'appareil d'acquisition, étude multi-centre, etc.), la comparaison avec des données plus fiables (histopathologie, par exemple) et la corrélation avec un marqueur clinique (qui peut potentiellement mener à une validation et donc à un substitut de marqueur clinique).

Il est intéressant de diviser les biomarqueurs issus de l'imagerie en trois catégories : morphologique, fonctionnel, et moléculaire. Les biomarqueurs morphologiques sont des variables *extensives* : volume, épaisseur, surface, nombre de lésions, etc. Il sont typiquement mesurés sur des images de type IRM ou scanner à l'aide d'une chaîne de traitements à base de segmentation. Les biomarqueurs fonctionnels sont des variables *intensives* : présence d'œdème, perfusion, ou activation d'une certaine région. En plus d'une segmentation, ces biomarqueurs nécessitent un regard plus approfondi sur les images qui demande des données biologiques ou moléculaires, généralement issues des techniques de type spectroscopie par résonance magnétique (*Magnetic Resonance Spectroscopy, MRS*), tomographie par émission de positons (*Positron Emission Tomography, PET*), tomographie par émission monophotonique (*Single Proton Emission Computerized Tomography, SPECT*), ou IRM avec produit de contraste. Les biomarqueurs moléculaires sont enfin acquis de manière similaire aux marqueurs fonctionnels, mais donnent des mesures précises sur un composant spécifique à l'aide des imageries PET ou SPECT.

En général et particulièrement dans le cadre d'études sur la sclérose en plaques, les biomarqueurs issus de l'imagerie sont particulièrement utiles dans la phase II des tests cliniques : le choix des dosages et du patient doit se faire dans des petites études à court-terme, et les essais en eux-mêmes sont souvent réalisables dans un petit nombre de centres disposant de technologies avancées.

9.2 Inclinaison de l'axe, lecture des DICOMS

Rappelons les données fournies par la lecture des images DICOM, pour chaque coupe.

- *Image Position (Patient)* I_p donne la position, dans le repère IRM, de l'origine du repère de la coupe 2D.
- *Image Orientation (Patient)* $(\overrightarrow{O_x}, \overrightarrow{O_y})$ donne, dans le repère IRM, les 2 vecteurs du repère de la coupe 2D.
- *Pixel Spacing* (V_x, V_y) donne, dans le repère IRM, la taille des deux vecteurs du repère de la coupe 2D.
- *Slice Thickness* V_z donne l'épaisseur de la coupe, ce qui permet de voir si les coupes sont jointives ou non.
- Un ensemble de voxels, dont la taille est $V_x * V_y * V_z$ mm^3.

L'ensemble des points I_p forme un axe, dont le vecteur de progression est $\overrightarrow{\delta I_p}$. La normale au plan de coupe est déduite des vecteurs unitaires du repère 2D : $\overrightarrow{N} = \overrightarrow{O_x} \wedge \overrightarrow{O_y}$. Lorsque $\overrightarrow{\delta I_p}$ est colinéaire à $\overrightarrow{N}$, un simple empilement des coupes suffit pour reconstruire l'image. Par contre, si $\overrightarrow{\delta I_p}$ et $\overrightarrow{N}$ ne sont pas colinéaires, une telle reconstruction n'est plus possible sans un rééchantillonnage supplémentaire.

La reconstruction correcte des images IRM peut se faire de deux manières. La plus directe est de se fixer un repère image, par exemple $(\overrightarrow{O_x}, \overrightarrow{O_y}, \overrightarrow{N})$ de reconstruire l'image coupe à coupe en plaçant le repère de chaque coupe dans le repère image, et de travailler directement sur les images reconstruites (section 3.2). Cette solution est très simple à mettre en œuvre mais nécessite un rééchantillonnage avant tout traitement sur les images (figure 9.1).

Une solution est de reconstruire l'image en considérant que le vecteur de progression de l'origine du repère 2D $\overrightarrow{\delta I_p}$ est colinéaire à $\overrightarrow{N}$, ce qui permet de reconstruire l'image directement à partir des données DICOM sans rééchantillonnage. En égalisant la taille du voxel image et la taille du voxel fournie par les données DICOM ($V_x * V_y * V_z$), aucun rééchantillonnage n'est nécessaire (figure 9.2).

Il reste ensuite à calculer la matrice de transformation affine permettant de finaliser la reconstruction de l'image. Le repère $(V_x\overrightarrow{O_x}, V_y\overrightarrow{O_y}, V_z\overrightarrow{N})$ est transformé

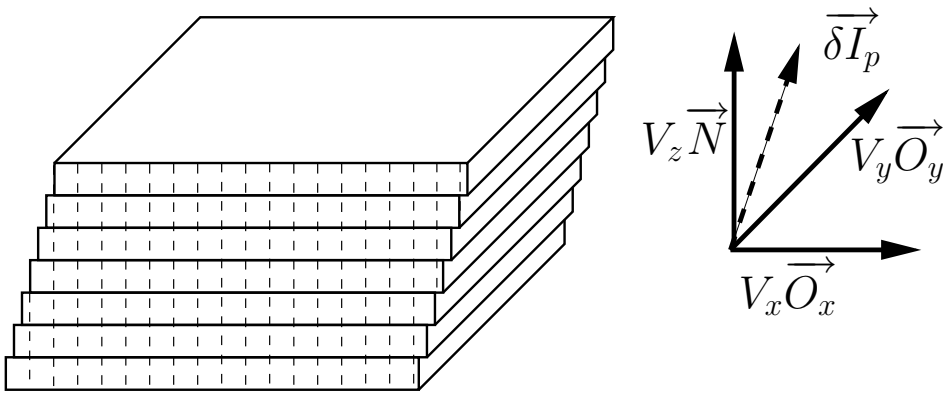


FIG. 9.1 – La reconstruction correcte des images IRM issues des données DICOM nécessite un rééchantillonnage : dans le cas général, $\overrightarrow{\delta I_p}$, le vecteur de progression de l'origine du repère 2D de chaque coupe I_p, n'est pas colinéaire à la normale au plan de coupe $\overrightarrow{N}$.

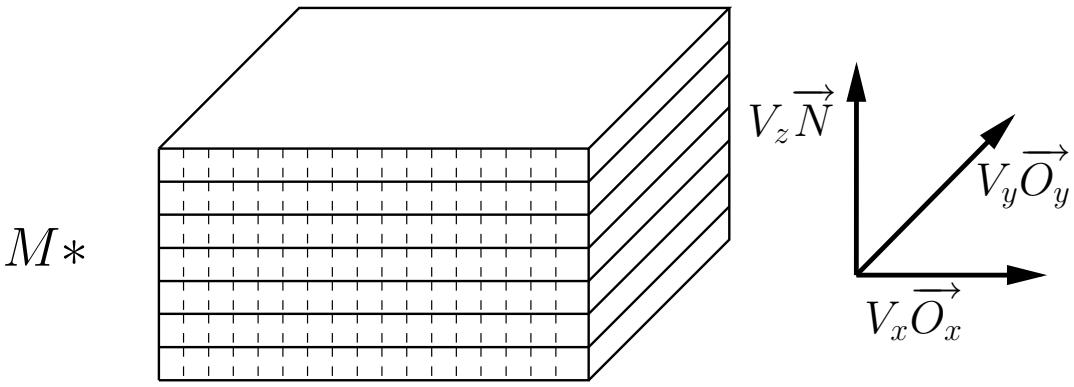


FIG. 9.2 – Pour éviter un rééchantillonnage lors de la reconstruction de image, il suffit de séparer cette dernière en deux étape.

– Construire une image 3D en faisant comme si l'axe de progression de l'origine du repère 2D était colinéaire à la normale au plan de coupe : les voxels de l'image sont alors les mêmes que ceux fournis par données DICOM, et aucun rééchantillonnage n'est nécessaire.
– Calculer une matrice de transformation affine M pour obtenir l'image finale.

en $(V_x\overrightarrow{O_x}, V_y\overrightarrow{O_y}, \overrightarrow{\delta I_p})$ dans le repère final $(\overrightarrow{O_x}, \overrightarrow{O_y}, \overrightarrow{N})$:

$$M = \begin{pmatrix} 1 & 0 & \frac{<\overrightarrow{\delta I_p}, \overrightarrow{O_x}>}{V_x} \\ 0 & 1 & \frac{<\overrightarrow{\delta I_p}, \overrightarrow{O_y}>}{V_y} \\ 0 & 0 & \frac{<\overrightarrow{\delta I_p}, \overrightarrow{O_z}>}{V_z} \end{pmatrix}$$

Ceci s'avère particulièrement utile, par exemple, dans le cas du recalage multi-séquences intra-patient : un algorithme de recalage par appariement par bloc peut prendre en entrée une image avec une matrice de transformation. En composant la matrice de recalage avec la matrice M, il est possible de n'obtenir qu'un seul rééchantillonnage au lieu de deux, ce qui est important, rappelons-le, lors de la gestion des problèmes de volumes partiels.

Bibliographie

[1] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 21(3) :193–199, 2002.

[2] L. S. Aït-Ali, S. Prima, P. Hellier, B. Carsin, G. Edan, and C. Barillot. STREM : A Robust Multidimensional Parametric Method to Segment MS Lesions in MRI. In *MICCAI*, volume 3749 of *Lecture Notes in Computer Science*, pages 409–416. Springer, 2005.

[3] A. Alpérovitch. Epidémiologie de la sclérose en plaques. *Médecine Thérapeutique*, 1 :545–548, 1995.

[4] P. Anbeek, K. Vincken, M. van Osch, B. Bisschops, M. Viergever, and J. van der Grond. Automated White Matter Lesion Segmentation by Voxel Probability Estimation. In Randy E. Ellis and Terry M. Peters, editor, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, volume 2878 of *LNCS*, pages 610–617, Montréal, Canada, November 2003. Springer.

[5] P. Anbeek, K. L. Vincken, G. S. van Bochove, M. J. P. van Osch, and J. van der Grond. Probabilistic segmentation of brain tissue in MR imaging. *NeuroImage*, 27(4) :795–804, 2005.

[6] P. Anbeek, K. L Vincken, M. J. P. van Osch, R. H. C. Bisschops, and J. van der Grond. Probabilistic segmentation of white matter lesions in MR imaging. *Neuroimage*, 21(3) :1037–44, March 2004.

[7] E. Ardizzone and R. Pirrone. An architecture for the recognition and classification of multiple sclerosis lesions in MR images. In Y. Shahar and S. Miksch, editors, *Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP'99)*, pages 1–12, 1999.

[8] M. S. Atkins and B. T. Mackiewich. Fully automatic segmentation of the brain in MRI. *IEEE Transactions on Medical Imaging*, 17(1) :98–107, February 1998.

[9] C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3D MR images using level sets and dense registration. *Medical Image Analysis*, 5 :185–194, 2001.

[10] I. N. Bankman, editor. *Handbook of medical imaging*. Academic Press, Inc., Orlando, FL, USA, 2000.

[11] F. Barkhof, M. Filippi, D. H. Miller, P. Scheltens, A. Campi, C. H. Polman, G. Comi, H. J. Ader, N. Losseff, and J. Valk. Comparison of MRI criteria at first presentation to predict conversion to clinically definite multiple sclerosis. *Brain*, 120 (Pt 11) :2059–69, November 1997.

[12] V. Barra and J. Y. Boire. Automatic segmentation of subcortical brain structures in MR images using information fusion. *IEEE Transactions on Medical Imaging*, 20(7) :549–58, July 2001.

[13] P. Barta, M. I. Miller, and A. Qiu. A stochastic model for studying the laminar structure of cortex from MRI. *IEEE Trans. Med. Imaging*, 24(6) :728–742, 2005.

[14] S. Bastianello. Magnetic resonance imaging of MS-like disease. *Neurol Sci*, 22 Suppl 2, November 2001.

[15] S. Beucher and C. Lantuéjoul. Use of watersheds in contour detection. In *Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*, pages 17–21, Rennes, France,, Sept. 1979.

[16] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[17] J. A. Bilmes. A gentle tutorial for the EM algorithm and its application to parameter estimation for gaussian mixture and hidden Markov models. Technical report, ICSI, Berkeley, 1998.

[18] F. Bloch. Nuclear induction. *Physical review*, 70 :460–474, 1946.

[19] P. A. Brex, J. I. O'Riordan, K. A. Miszkiel, I. F. Moseley, A. J. Thompson, G. T. Plant, and D. H. Miller. Multisequence MRI in clinically isolated syndromes and the early development of MS. *Neurology*, 53(6) :1184–90, October 1999.

[20] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325–376, 1992.

[21] A. Cachia, J.-F. Mangin, D. Rivière, F. Kherif, N. Boddaert, A. Andrade, D. Papadopoulos-Orfanos, J.-B. Poline, I. Bloch, M. Zilbovicius, P. Sonigo, F. Brunelle, and J. Régis. A primal sketch of the cortex mean curvature : a morphogenesis based approach to study the variability of the folding patters. *IEEE Trans. Med. Imaging*, 22(6) :754–765, 2003.

[22] P. Cachier, J.-F. Mangin, X. Pennec, D. Rivière, D. Papadopoulos-Orfanos, J. Régis, and N. Ayache. Multisubject non-rigid registration of brain MRI using intensity and geometric features. In W.J. Niessen and M.A. Viergever, editors, *4th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 734–742, Utrecht, The Netherlands, October 2001.

[23] G. Calmon and J.-P. Thirion. Calcul de variation de volume de lésions dans des images médicales tridimensionnelles. In *Traitement du Signal et des Images (GRETSI'97)*, Grenoble, September 1997.

[24] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *ICCV*, pages 694–699, 1995.

[25] J. E. Cates, R. T. Whitaker, and G. M. Jones. Case study: an evaluation of user-assisted hierarchical watershed segmentation. *Medical Image Analysis*, 9(6) :566–578, 2005.

[26] J. M. Charcot. Histologie de la sclérose en plaques. *Gazette des Hôpitaux*, 141 :554–558, 1868.

[27] J. M. Charcot. *Leçons sur les maladies du système nerveux*, volume 1, pages 237–238. Bourneville, Paris, 1898.

[28] J. T. Chen, D. L. Collins, M. S. Freedman, H. L. Atkins, and D. L. Arnold. Local magnetization transfer ratio signal inhomogeneity is related to subsequent change in MTR in lesions and normal-appearing white-matter of multiple sclerosis patients. *Neuroimage*, 25(4) :1272–8, May 2005.

[29] J. P. Chiverton and K. Wells. Estimation of partial volume mixtures with a confidence measure combining intensity and gradient magnitude information. In *Medical Image Understanding and Analysis*, University of Bristol, July 2005.

[30] H. S. Choi, D. R. Haynor, and Y. M. Kim. Partial volume tissue classification of multichannel magnetic resonance images – a mixel model. *IEEE Transactions on Medical Imaging*, 10(3) :395–407, 1991.

[31] G. E. Christensen and H. J. Johnson. Consistent image registration. *IEEE Transactions on Medical Imaging*, 20(7) :568–582, 2001.

[32] M. K. Chung, S. Robbins, and A. C. Evans. Unified statistical approach to cortical thickness analysis. In *IPMI*, pages 627–638, 2005.

[33] C. Ciofolo and C. Barillot. Brain segmentation with competitive level sets and fuzzy control. In *International Conference on Information Processing in Medical Imaging*, volume 3565, pages 333–344, 2005.

[34] M. Clanet, B. Fontaine, and C. Vuillemin. La susceptibilité génétique à la sclérose en plaques. *Revue Neurologique*, 152(1) :149–152, 1996.

[35] C. A. Cocosco, A. P. Zijdenbos, and A. C. Evans. A fully automatic and robust brain MRI tissue classification method. *Medical Image Analysis*, 7(4) :513–528, 2003.

[36] Y. Cointepas, J.-F. Mangin, L. Garnero, J.-B. Poline, and H. Benali. BrainVISA: Software platform for visualization and analysis of multi-modality brain data. *Neuroimage*, 13(6) :98, 2001.

[37] I. Corouge and C. Barillot. Statistical Modeling of Pairs of Sulci in the Context of Neuroimaging Probabilistic Atlas. In *MICCAI (2)*, volume 2489 of *Lecture Notes in Computer Science*, pages 655–662. Springer, 2002.

[38] I. Corouge, P. Hellier, B. Gibaud, and C. Barillot. Inter-individual functional mapping : a non linear local approach. *Neuroimage*, 19(4) :1337–1348, 2003.

[39] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood for incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1) :1–38, 1977.

[40] G. Dugas-Phocion, M. Á. González Ballester, G. Malandain, C. Lebrun, and N. Ayache. Improved EM-Based Tissue Segmentation and Partial Volume Effect Quantification in Multi-sequence Brain MRI. In *MICCAI*, volume 1496 of *Lecture Notes in Computer Science*, pages 26–33. Springer, 2004.

[41] G. Dugas-Phocion, M. A. González Ballester, C. Lebrun, S. Chanalet, C. Bensa, M. Chatel, N. Ayache, and G. Malandain. Automatic segmentation of white matter lesions in T2 FLAIR MRI of relapsing-remitting multiple sclerosis patients. In *20th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS)*, Vienna, Austria, October 2004.

[42] G. Dugas-Phocion, M. A. González Ballester, C. Lebrun, S. Chanalet, C. Bensa, G. Malandain, and N. Ayache. Hierarchical segmentation of multiple sclerosis lesions in multi-sequence MRI. In *International Symposium on Biomedical Imaging : From Nano to Macro (ISBI'04)*, Arlington, VA, USA, April 2004. IEEE.

[43] G. Dugas-Phocion, C. Lebrun, S. Chanalet, M. Chatel, N. Ayache, and G. Malandain. Automatic segmentation of white matter lesions in multi-sequence MRI of relapsing-remitting multiple sclerosis patients. In *21st Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS)*, Thessaloniki, September 2005.

[44] M. K. Erskine, L. L. Cook, K. E. Riddle, J. R. Mitchell, and S. J. Karlik. Resolution-dependent estimates of multiple sclerosis lesion loads. *Can J Neurol Sci.*, 32(2) :205–212, 2005.

[45] A. C. Evans, D. L. Collins, and B. Milner. An MRI-based stereotactic brain atlas from 300 young normal subjects. In *22nd Symposium of the Society for Neuroscience*, page 408, Anaheim, 1992.

[46] A. C. Evans, D. L. Collins, P. Neelin, D. MacDonald, M. Kamber, and T. S. Marrett. Three-Dimensional Correlative Imaging: Applications in Human Brain Mapping. *Functional Neuroimaging*, pages 145–162, 1994.

[47] A. C. Evans, J. A. Frank, J. Antel, and D. H. Miller. The role of MRI in clinical trials of multiple sclerosis: comparison of image processing techniques. *Ann Neurol.*, 41(1) :125–132, 1997.

[48] A. C. Evans, M. Kamber, D. L. Collins, and D. MacDonald. *Magnetic Resonance Scanning and Epilepsy*, chapter 48, pages 263–274. NATO ASI Series A, Life Sciences. Plenum Press, S. D. Shorvon et al. edition, 1994.

[49] F. Fazekas, H. Offenbacher, S. Fuchs, R. Schmidt, K. Niederkorn, S. Horner, and H. Lechner. Criteria for an increased specificity of MRI interpretation in elderly subjects with suspected multiple sclerosis. *Neurology*, 38(12) :1822–1825, 1988.

[50] M. Filippi. Linking structural, metabolic and functional changes in multiple sclerosis. *Eur. J. Neurol*, 8 :291–297, 2001.

[51] M. Filippi, A. Falini, D. L. Arnold, F. Fazekas, O. Gonen, J. H. Simon, V. Dousset, M. Savoiardo, and J. S. Wolinsky. Magnetic resonance techniques for the in vivo assessment of multiple sclerosis pathology: consensus report of the white matter study group. *J Magn Reson Imaging*, 21(6) :669–675, 2005.

[52] M. Filippi and R. I. Grossman. MRI techniques to monitor MS evolution: the present and the future. *Neurology*, 58(8) :1147–1153, April 2002.

[53] M. Filippi, M. A. Rocca, G. Mastronardo, and G. Comi. Lesion load measurements in multiple sclerosis: the effect of incorporating magneti-

zation transfer contrast in fast-FLAIR sequence. *Magn Reson Imaging*, 17(3) :459–61, April 1999.

[54] G. Flandin. *Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle*. PhD thesis, Université de Nice-Sophia Antipolis, April 2004.

[55] A. Gass, M. Filippi, M. E. Rodegher, A. Schwartz, G. Comi, and M. G. Hennerici. Characteristics of chronic MS lesions in the cerebrum, brainstem, spinal cord, and optic nerve on T1-weighted MRI. *Neurology*, 50(2) :548–50, February 1998.

[56] M. Gastaud, M. Barlaud, and G. Aubert. Combining shape prior and statistical features for active contour segmentation. *IEEE Trans. Circuits Syst. Video Techn.*, 14(5) :726–734, 2004.

[57] J. J. Geurts, L. Bo, P. J. Pouwels, J. A. Castelijns, C. H. Polman, and F. Barkhof. Cortical lesions in multiple sclerosis: combined postmortem MR imaging and histopathology. *American Journal of Neuroradiology*, 26(3) :572–577, 2005.

[58] J. Gohagan, E. Spitznagel, W. Murphy, and M. Vannier. Multispectral Analysis of MR Images of the Breast. *Radiology*, 163 :703–707, 1987.

[59] M. A. González Ballester. *Morphometric Analysis of Brain Structures in MRI*. PhD thesis, University of Oxford, 1999.

[60] M. A. González Ballester, A. Zisserman, and M. Brady. Segmentation and measurement of brain structures in MRI including confidence bounds. *Medical Image Analysis*, 4(3) :189–200, 2000.

[61] M. A. González Ballester, A. Zisserman, and M. Brady. Estimation of the partial volume effect in MRI. *Medical Image Analysis*, 6(4) :389–405, December 2002.

[62] G. Le Goualher, E. Procyk, D. L. Collins, R. Venugopal, C. Barillot, and A. C. Evans. Automated Extraction and Variability Analysis of Sulcal Neuroanatomy. *IEEE Trans. Med. Imaging*, 18(3) :206–217, 1999.

[63] J. Grimaud, Y.-M. Zhu, and M. Rombaut. Mise au point : Les techniques d'analyse quantitative des IRM cérébrales : application à la sclérose en plaques. *Revue Neurologique*, 158(3) :381–389, 2002.

[64] NIH Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: prefered definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, 69 :89–95, 2001.

[65] A. Guimond, J. Meunier, and J.-P. Thirion. Average Brain Models: A Convergence Study. *Computer Vision and Image Understanding*, 77(2) :192–210, 2000.

[66] J. V. Hajnal, D. L. G. Hill, and D. Hawkes, editors. *Medical Image Registration*. D. J. Hawkes, 2001.

[67] L. O. Hall, A. M. Bensaid, L. P. Clarke, R. P. Velthuizen, M. S. Silbiger, and J. C. Bezdek. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. *IEEE Transactions on Neural Networks*, 3 :672–682, 1992.

[68] R. J. Hathaway. Another interpretation of the EM algorithm for mixture distributions. *Journal of Statistics and Probability Letters*, 4 :53–56, 1986.

[69] P. Hellier and C. Barillot. Coupling dense and landmark-based approaches for non rigid registration. *IEEE Transactions on Medical Imaging*, 22(2) :217–227, 2003.

[70] A. Hoover, V. Kouznetsova, and M. H. Goldbaum. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3) :203–210, 2000.

[71] X. Hu, N. Alperin, D. N. Levin, K. K. Tan, and M. Mengeot. Visualization of MR angiographic data with segmentation and volume-rendering techniques. *J Magn Reson Imaging*, 1(5) :539–46, Sep-Oct 1991.

[72] B. Jahne. *Practical Handbook on Image Processing for Scientific Applications*. Boca Raton, 1997.

[73] K. Jellinger. Einige morphologische aspekte der multiplen sklerose. *Nervenheilk*, pages 12–37, 1969. Suppl II.

[74] M. Jordan and C. Bishop. *An introduction to Graphical Models*. MIT Press, 2002.

[75] L. Kappos, D. Moeri, E. W. Radue, A. Schoetzau, K. Schweikert, F. Barkhof, D. Miller, C. R. Guttmann, H. L. Weiner, C. Gasperini, and M. Filippi. Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. Gadolinium MRI Meta-analysis Group. *Lancet*, 353(9157) :964–9, March 1999.

[76] T. Kapur, W. E. Grimson, W. M. Wells III, and R. Kikinis. Segmentation of brain tissue from magnetic resonance images. *Medical Image Analysis*, 1(2) :109–27, June 1996.

[77] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1 :321–331, 1987.

[78] S. J. Kisner, T. M. Talavage, and J. L. Ulmer. Testing a model for MR imager noise. In *The 2nd Joint Meeting of the IEEE EMBS & the BMES*, Houston, Texas, October 2002.

[79] J. F. Kurtzke. Rating neurologic impairment in multiple sclerosis and the disability status scale. *Neurology*, 33 :1444–1452, 1983.

[80] D. H. Laidlaw, K. W. Fleischer, and A. H. Barr. Partial-volume bayesian classification of material mixtures in MR volume data using voxel histograms. *IEEE Transactions on Medical Imaging*, 17(1) :74–86, February 1998.

[81] E. G. Learned-Miller and P. Ahammad. Joint MRI bias removal using entropy minimization across images. *Neural Information Processing System 17*, pages 761–768, 2005.

[82] C Lebrun, D Rey, S Chanalet, V Bourg, C Bensa, M Chatel, N Ayache, and G Malandain. The contribution of automatic anatomical matching of sequential brain MRI scans in the monitoring of multiple sclerosis lesions. *Rev Neurol (Paris)*, 160(8-9) :805–10, September 2004. in French.

[83] M. M. J. Letteboer, O. F. Olsen, E. B. Dam, P. W. A. Willems, M. A. Viergever, and W. J. Niessen. Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm. *Academic Radiology*, 11(10) :1125–1138, 2004.

[84] B. S. Li, J. Regal, B. J. Soher, L. J. Mannon, R. I. Grossman, and O. Gonen. Brain metabolite profiles of T1-hypointense lesions in relapsing-remitting multiple sclerosis. *AJNR Am J Neuroradiol*, 24(1) :68–74, January 2003.

[85] Z. Liang, J. R. MacFall, and D. P. Harrington. Parameter estimation and tissue segmentation from multispectral MR images. *IEEE Transactions on Medical Imaging*, 13(3) :441–449, September 1994.

[86] A. W.-C. Liew and H. Yan. Magnetic Resonance Imaging - An Adaptive Spatial Fuzzy Clustering Algorithm for 3-D MR Image Segmentation. *IEEE Transactions on Medical Imaging*, 22(9) :1063–1075, 2003.

[87] M.G. Linguraru, J.M. Brady, and M. Yam. Detection of microcalcifications using SMF. In H.O. Peitgen, editor, *Digital Mammography IWDM'02*, pages 342–346. Springer, 2002.

[88] O. Lyon-Caen and M. Clanet. *La sclérose en plaques*. John Libbey, 1998.

[89] M. Maddah, K. H. Zou, W. M. Wells III, R. Kikinis, and S. K. Warfield. Automatic optimization of segmentation algorithms through simultaneous truth and performance level estimation (staple). In *MICCAI (1)*, volume 3216 of *Lecture Notes in Computer Science*, pages 274–282. Springer, 2004.

[90] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1) :1–36, March 1998.

[91] R. Malladi, J. Sethian, and B. Vemuri. Shape modelling with front-propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2) :158–175, 1995.

[92] J. F. Mangin. Entropy minimization for automatic correction of intensity non uniformity. In *MMBIA (Math. Methods in Biomed. Image Analysis)*, pages 162–169. IEEE Press, 2000.

[93] J.-F. Mangin. *Une vision structurelle de l'analyse des images cérébrales*. Habilitation à diriger des recherches, Université Paris 11, 2005.

[94] J.-F. Mangin, O. Coulon, and V. Frouin. Robust brain segmentation using histogram scale-space analysis and mathematical morphology. In *MICCAI*, volume 1496 of *Lecture Notes in Computer Science*, pages 1230–1241. Springer, 1998.

[95] J.-F. Mangin, V. Frouin, I. Bloch, J. Régis, and J. López-Krahe. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *J. Math. Imaging Vis.*, 5(4) :297–318, 1995.

[96] J.-F Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, P. Scifo, and T. Ochiai. A framework to study the cortical folding patterns. *Neuroimage*, 23 :129–138, 2004.

[97] W. I. McDonald, A. Compston, G. Edan, D. Goodkin, H. P. Hartung, F. D. Lublin, H. F. McFarland, D. W. Paty, C. H. Polman, S. C. Reingold, M. Sandberg-Wollheim, W. Sibley, A. Thompson, S. van den Noort, B. Y. Weinshenker, and J. S. Wolinsky. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol.*, 50(1) :121–127, 2001.

[98] H. F. McFarland, F. Barkhof, J. Antel, and D. H. Miller. The role of MRI as a surrogate outcome measure in multiple sclerosis. *Multiple Sclerosis*, 8(1) :40–51, February 2002.

[99] G. McLachlan and D. Peel. Finite mixture models. *Wiley Series in Probability and Statistics*, 2000.

[100] D. H. Miller, A. J. Thompson, and M. Filippi. Magnetic resonance studies of abnormalities in the normal appearing white matter and grey matter in multiple sclerosis. *J Neurol.*, 250(12) :1407–1419, 2003.

[101] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, 2004.

[102] J. Montagnat and H. Delingette. Globally constrained deformable models for 3D object reconstruction. *Signal Processing*, 71(2) :173–186, 1998.

[103] N. Moon, E. Bullitt, K. Van Leemput, and G. Gerig. Model-based brain and tumor segmentation. In *ICPR (1)*, pages 528–531, 2002.

[104] P. A. Narayana, W. W. Brey, M. V. Kulkarni, and C. L. Sievenpiper. Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magn Reson Imaging*, 6(3) :271–4, May-Jun 1988.

[105] J. Newcombe, C. P. Hawkins, H. L. Henderson, H. A. Patel, M. N. Woodroofe, G. M. Hayes, M. L. Cuzner, D. MacManus, E. P. du Boulay, and W. I. McDonald. Histopathology of multiple sclerosis lesions detected by magnetic resonance imaging in unfixed postmortem central nervous system tissue. *Brain*, 114 :1013–1023, 1991.

[106] W. J. Niessen. *Multiscale Medical Image Analysis*. PhD thesis, University of Utrecht, 1997.

[107] A. Noe and J. Gee. Efficient partial volume tissue classification in MRI scans. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2488 of *LNCS*, pages 698–705, Tokyo, September 2002. Springer.

[108] R. Nowak. Wavelet-based Rician noise removal for magnetic resonance imaging . *IEEE Transactions on Image Processing*, 8(10), October 1999.

[109] S. D. Olabarriaga and A. W. M. Smeulders. Interaction in the segmentation of medical images: A survey. *Medical Image Analysis*, 5 :127–142, 2001.

[110] J. I. O'Riordan, A. J. Thompson, D. P. Kingsley, D. G. MacManus, B. E. Kendall, P. Rudge, W. I. McDonald, and D. H. Miller. The prognostic value of brain MRI in clinically isolated syndromes of the CNS. a 10-year follow-up. *Brain*, 121(3) :495–503, 1998.

[111] S. Osher and J. A. Sethian. Fronts Propagating with curvature-dependant speed: algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79 :12–49, 1988.

[112] S. Ourselin, A. Roche, S. Prima, and N. Ayache. Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In A.M. DiGioia and S. Delp, editors, *Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI 2000)*, volume 1935 of *Lectures Notes in Computer Science*, pages 557–566, Pittsburgh, Pennsylvanie USA, octobre 11-14 2000. Springer.

[113] C. Pachai, Y. M. Zhu, J. Grimaud, M. Hermier, A. Dromigny-Badin, A. Boudraa, G. Gimenez, C. Confavreux, and J. C. Froment. A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI. *Computerized Medical Imaging and Graphics*, 22(5) :399–408, September 1998.

[114] N. Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *International Journal of Computer Vision*, 50(3) :345–362, 2002.

[115] D. W. Paty, J. J. Oger, L. F. Kastrukoff, S. A. Hashimoto, J. P. Hooge, A. A. Eisen, A. K. Eisen, S. J. Purves, M. D. Low, and V. Brandejs. MRI in the diagnostic of MS: a prospective study with comparison of clinical evaluation, evoked potential, oligoclonal banding and CT. *Neurology*, 38(2) :180–185, 1988.

[116] S. T. Pendlebury, M. A. Lee, A. M. Blamire, P. Styles, and P. M. Matthews. Correlating magnetic resonance imaging markers of axonal injury and demyelination in motor impairment secondary to stroke and multiple sclerosis. *Magn Reson Imaging*, 18(4) :369–78, May 2000.

[117] G. Peters. Multiple sclerosis. pages 821–843, New York : McGraw Hill, 1968.

[118] D. L. Pham. Robust fuzzy segmentation of magnetic resonance images. In *IEEE Symposium on Computer-Based Medical Systems*, pages 127–131, 2001.

[119] D. L. Pham and J. L. Prince. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Transactions on Medical Imaging*, 18(9) :737–52,

September 1999.

[120] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annu Rev Biomed Eng*, 2 :315–37, 2000.

[121] A. Pitiot. *Segmentation automatique des structures cérébrales s'appuyant sur des connaissances explicites*. Thèse de sciences, École des mines de Paris, November 2003.

[122] A. Pitiot, P. M. Thompson, and A. W. Toga. Adaptive Elastic Segmentation of Brain MRI via Shape Model Guided Evolutionary Programming. *IEEE Transactions on Medical Imaging*, 21(8) :910–923, August 2002.

[123] J. P. W. Pluim and J. M. Fitzpatrick. Image registration. *IEEE Trans. Med. Imaging*, 22(11) :1341–1343, 2003.

[124] K. M. Pohl, W. E. L. Grimson, S. Bouix, and R. Kikinis. Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework. In *ISBI*, pages 81–84. IEEE, 2004.

[125] C. M. Poser, D. W. Paty, L. Scheinberg, W. I. McDonald, F. A. Davis, G. C. Ebers, K. P. Johnson, W. A. Sibley, D. H. Silberberg, and WW Tourtellotte. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol*, 13 :227–231, 1983.

[126] C.M. Poser. The epidemiology of multiple sclerosis: a general overview. *Ann Neurol*, 36 :180–193, 1994.

[127] M. Prastawa, J. Gilmore, W. Lin, and G. Gerig. Automatic segmentation of neonatal brain MRI. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, volume 3216 of *LNCS*, pages 10–17. Springer, 2004.

[128] S. Prima, D. L. Arnold, and D. L. Collins. Multivariate Statistics for Detection of MS Activity in Serial Multimodal MR Images. In *MICCAI (1)*, volume 2878 of *Lecture Notes in Computer Science*, pages 663–670. Springer, 2003.

[129] S. Prima, N. Ayache, Tom Barrick, and Neil Roberts. Maximum Likelihood Estimation of the Bias Field in MR Brain Images: Investigating Different

Modelings of the Imaging Process. In W.J. Niessen and M.A. Viergever, editors, *4th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, volume 2208 of *LNCS*, pages 811–819, Utrecht, The Netherlands, October 2001.

[130] S. Prima, N. Ayache, A. Janke, S. Francis, D. Arnold, and L. Collins. Statistical Analysis of Longitudinal MRI Data: Application for detection of Disease Activity in MS. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2488 of *LNCS*, pages 363–371, Tokyo, September 2002. Springer.

[131] J. Régis, J.-F. Mangin, T. Ochiai, V. Frouin, D. Rivière, A. Cachia, M. Tamura, and Y. Samson. "Sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns. *Neurol Med Chir (Tokyo)*, 45 :1–17, 2005.

[132] M. A. Reidel, C. Stippich, S. Heiland, B. Storch-Hagenlocher, O. Jansen, and S. Hahnel. Differentiation of multiple sclerosis plaques, subacute cerebral ischaemic infarcts, focal vasogenic oedema and lesions of subcortical arteriosclerotic encephalopathy using magnetisation transfer measurements. *Neuroradiology*, 45(5) :289–294, 2003.

[133] D. Rey. *Détection et quantification de processus évolutifs dans des images médicales tridimensionnelles : application à la sclérose en plaques*. Thèse de sciences, Université de Nice Sophia-Antipolis, October 2002.

[134] D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis. *Medical Image Analysis*, 6(2) :163–179, June 2002.

[135] D. Rivière, J.-F. Mangin, D. Papadopoulos-Orfanos, J.-M. Martinez, V. Frouin, and J. Régis. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis*, 6(2) :77–92, 2002.

[136] D. Rivière, J. Régis, Y. Cointepas, D. Papadopoulos-Orfanos, A. Cachia, and J.-F. Mangin. A freely available anatomist/brainvisa package for struc-

tural morphometry of the cortical sulci. In *Proc. 9th HBM*, Neuroimage 19(2), page 934, 2003.

[137] J. B. T. M. Roerdink and A. Meijster. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies.. *Fundam. Inform.*, 41(1-2) :187–228, 2000.

[138] M. Rousson, N. Paragios, and R. Deriche. Implicit Active Shape Models for 3D Segmentation in MR Imaging. In *MICCAI (1)*, volume 3216 of *Lecture Notes in Computer Science*, pages 209–216. Springer, 2004.

[139] M. Rovaris and M. A. Rocca et al. Reproductibility of of brain MRI lesion volume measurements in multiple sclerosis using a local thresholding technique: effects of formal operator training. *European Journal of Neurology*, 41(4) :226–230, 1999.

[140] M Rovaris, M Filippi, L Minicucci, G Iannucci, G Santuccio, F Possa, and G Comi. Cortical/subcortical disease burden and cognitive impairment in patients with multiple sclerosis. *AJNR Am J Neuroradiol*, 21(2) :402–8, February 2000.

[141] S. Ruan, C. Jaggi, J. Xue, J. Fadili, and D. Bloyet. Brain tissue classification of magnetic resonance images using partial volume modeling. *IEEE Transactions on Medical Imaging*, 19(12) :1179–87, December 2000.

[142] A. Ruf, H. Greenspan, and J. Goldberger. Tissue Classification of Noisy MR Brain Images Using Constrained GMM. In *MICCAI (2)*, volume 3750 of *Lecture Notes in Computer Science*, pages 790–797. Springer, 2005.

[143] O. Sabouraud and G. Edan. *Sclérose en plaques*. Service de neurologie, CHU de Rennes, 2 rue Henri Le Guilloux 35033 Rennes Cedex, 1998. http ://www.med.univ-rennes1.fr/etud/neuro/SEP.htm.

[144] O. Salvado, D. L. Wilson, J. S. Suri, C. Hillenbrand, and S. Zhang. MR signal inhomogeneity correction for visual and computerized atherosclerosis lesion assessment. In *ISBI*, pages 1143–1146. IEEE, 2004.

[145] J. A. Sethian. Curvature and evolution of fronts. *Commun. Math. Phys.*, 101 :487–499, 1985.

[146] J. A. Sethian. A review of recent numerical algorithms for hypersurfaces moving with curvature dependent speed. *Journal of Differential Geometry*, 31 :131–161, 1989.

[147] A. Shahar and H. Greenspan. Probabilistic spatial-temporal segmentation of multiple sclerosis lesions. In *ECCV Workshops CVAMIA and MMBIA*, volume 3117 of *Lecture Notes in Computer Science*, pages 269–280. Springer, 2004.

[148] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Shaper, D. A. Rottenberg, and R. M. Leahy. Magnetic Resonnance Image Tissue Classification Using a Partial Volume Model. *NeuroImage*, 13 :856–876, 2001.

[149] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck. Maximum Likelihood Estimation of Rician Distribution Parameters. *IEEE Transactions on Medical Imaging*, 17(3) :357–361, 1998.

[150] J. H. Simon. MRI in multiple sclerosis. *Phys Med Rehabil Clin N Am.*, 16(2) :383–409, 2005.

[151] H. R. Singleton and G. M. Pohost. Automatic cardiac MR image segmentation using edge detection by tissue classification in pixel neighborhoods. *Magnetic Resonance in Medicine*, 37 :418–424, 1997.

[152] J. G. Sled and G. B. Pike. Standing-wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI. *IEEE Trans. Med. Imaging*, 17(4) :653–662, 1998.

[153] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. PWS Publishing, 1999.

[154] G. Subsol. Crest lines for curve based warping. In A. W. Toga, editor, *Brain Warping*, chapter 13, pages 225–246. Academic Press, 1998.

[155] J. S. Suri, S. K. Setarehdan, and S. Singh, editors. *Advanced algorithmic approaches to medical image segmentation: state-of-the-art application in cardiology, neurology, mammography and pathology*, chapter 5. Springer-Verlag New York, Inc., 2002.

[156] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3) :243–260, 1998.

[157] J.-P. Thirion and G. Calmon. Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences. *IEEE Transactions on Medical Imaging*, 18(5) :429–441, 1999.

[158] J.-P. Thirion, G. Subsol, and D. Dean. Cross validation of three inter-patients matching methods. In *Visualization in Biomedical Computing, VBC'96*, Hamburg (D), September 1996.

[159] P. M. Thompson and A. D. Lee. Abnormal Cortical Complexity and Thickness Profiles Mapped in Williams Syndrome The Journal of Neuroscience. *The Journal of Neuroscience*, 25(16) :4146–4158, 2005.

[160] A. Tourbah and I. Berry. Contribution des techniques de résonance magnétique à la SEP. *Pathol Biol*, 48 :151–161, 2000.

[161] A. Tourbah and O. Lyon-Caen. IRM et SEP: intérêt dans le diagnostic et la connaissance de l'histoire naturelle. *Revue Neurologique*, 157(8–9) :750–760, 2001.

[162] J. K. Udupa, L. G. Nyul, Y. Ge, and R. I. Grossman. Multiprotocol MR image segmentation in multiple sclerosis : Experience with over 1000 studies. *Academic Radiology*, 8(11) :1116–1126, November 2001.

[163] J. K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M. A. van Buchem, and R. I. Grossman. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Transactions on Medical Imaging*, 16(5) :598–609, 1997.

[164] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Transactions on Medical Imaging*, 20(8) :677–688, 2001.

[165] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(10) :885–896, October 1999.

[166] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classsification of MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(10) :897–908, October 1999.

[167] K. Van Leemput, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. A unifying framework for partial volume segmentation of brain MR images. *IEEE Transactions on Medical Imaging*, 22(1) :105–119, 2003.

[168] J. H. van Waesberghe, W. Kamphorst, C. J. De Groot, M. A. van Walderveen, J. A. Castelijns, R. Ravid, G. J. Lycklama a Nijeholt, P. van der Valk, C. H. Polman, A. J. Thompson, and F. Barkhof. Axonal loss in multiple sclerosis lesions : magnetic resonance imaging insights into substrates of disability. *Ann Neurol*, 46(5) :747–54, November 1999.

[169] M. A. van Walderveen, L. Truyen, B. W. van Oosten, J. A. Castelijns, G. J. Lycklama a Nijeholt, J. H. van Waesberghe, C. Polman, and F. Barkhof F. Development of hypointense lesions on T1-weighted spin-echo magnetic resonance images in multiple sclerosis: relation to inflammatory activity. *Archives of Neurology*, 56(3) :345–351, 1999.

[170] M. A. A. van Walderveen, F. Barkhof, O. R. Hommes, C. H. Polman, H. Tobi, S. T. F. M. Frequin, and J. Valk. Correlating MRI and clinical disease activity in multiple sclerosis: relevance of hypointense lesions on short-TR / short-TE T1-weighted spin-echo images. *Neurology*, 45 :1684–1690, 1995.

[171] S. Warfield, J. Dengler, J. Zaers, C. R. G. Guttmann, W. M. Wells III, G. J. Ettinger, J. Hiller, and R. Kikinis. Automatic identification of grey matter structures from MRI to improve the segmentation of white matter lesions. *The Journal of Image Guided Surgery*, 1(6) :326–338, 1995.

[172] S. Warfield, K. Zou, and W. Wells. Validation of image segmentation and expert quality with an expectation-maximization algorithm. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2488 of *LNCS*, pages 298–306, Tokyo, September 2002. Springer.

[173] Simon K. Warfield, Kelly H. Zou, and William M. Wells III. Simultaneous truth and performance level estimation (staple) : an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging*, 23(7) :903–921, 2004.

[174] J. C. Waterton. What does the drug developer need from medical imaging ? challenges for image analysis. In *Medical Image Understanding and Analysis 2005*, pages 2–6, University of Bristol, July 2005. Invited talk.

[175] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Adaptive Segmentation of MRI Data. In *CVRMed '95 : Proceedings of the First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *Lecture Notes in Computer Science*, pages 59–69, London, UK, 1995. Springer-Verlag.

[176] Z. Wu, H. W. Chung, and F. W. Wehrli. A bayesian approach to subvoxel tissue classification in NMR microscopic images of trabecular bone. *Magn Reson Med*, 31(3) :302–8, March 1994.

[177] Y. Yang, P. Lin, and C. Zheng. An Efficient Statistical Method for Segmentation of Singe-Channel Brain MRI. In *2004 International Conference on Computer and Information Technology (CIT 2004)*, pages 149–154, 2004.

[178] L. A. Zadeh. Fuzzy sets. *Inf. and Control*, 8 :338–353, 1965.

[179] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer. Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *IEEE Transactions on Medical Imaging*, 13(4) :716–724, December 1994.

[180] A. P. Zijdenbos, R. Foghani, and A. C. Evans. Automatic 'Pipeline' Analysis of 3D MRI Data for Clinical Trials : Application to Multiple Sclerosis. *IEEE Trans. Med. Imaging*, 21(10) :1280–1291, 2002.

[181] A. P. Zijdenbos, R. Forghani, and A. C. Evans. Automatic Quantification of MS Lesions in 3D MRI Brain Data Sets : Validation of INSECT. In *MICCAI*, volume 1496 of *Lecture Notes in Computer Science*, pages 439–448. Springer, 1998.

Dans ce manuscrit, nous nous sommes intéressés à l'analyse d'IRM cérébrales dans le cadre notamment du suivi de patients souffrant de sclérose en plaques (SEP). L'extraction automatique de quantificateurs pour la SEP a de nombreuses applications potentielles, tant dans le domaine clinique que pour des tests pharmaceutiques. Dans un premier temps, nous nous sommes consacrés à l'étude générale de la maladie, afin de situer le rôle de l'imagerie. Nous présentons ensuite les différents prétraitements nécessaires à la robustesse du système, puis la segmentation des tissus sains via un modèle probabiliste des volumes partiels, ainsi qu'une nouvelle version de l'algorithme EM doté d'un traitement spécifique des points aberrants. L'utilisation des différentes séquences du protocole d'acquisition permet de spécialiser le processus pour la détection des lésions de sclérose en plaques. Une évaluation quantitative des résultats ainsi qu'un ensemble de perspectives sont enfin détaillées.

In this thesis, we present our work on the analysis of human brain MRI, in particular brain MRI of multiple sclerosis (MS) patients. The automatic extraction of quantifiers for MS has many potential applications, as well in the clinical field as for clinical trials. Initially, we devoted ourselves to the general study of the disease, in order to precise the role of the MRI in the diagnosis process. We present then the various pretreatments necessary to the robustness of the system, then the segmentation of healthy tissues via a probabilistic model of partial volume effect, as well as a new version of the EM algorithm with a specific treatment of outliers. The use of the various sequences of the acquisition protocol makes it possible to specialize the process for MS lesion detection. A quantitative evaluation of the results as well as a whole of prospects are finally detailed.