

ÉCOLE DOCTORALE 269 MSII

Laboratoire ICube UMR 7357 CNRS - Université de Strasbourg

THÈSE présentée par :

Franck VALENTINI

soutenue le : **31 mars 2025**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Signal, Image, Robotique, Automatique

Analyse d'IRM cérébrales de souris par apprentissage profond dans le cadre des études de groupes

THÈSE dirigée par :

Mme HARSAN Laura
M. NOBLET Vincent

MCU-PH, Université de Strasbourg, ICube
Ingénieur de recherche, CNRS, ICube

RAPPORTEURS :

Mme MATEUS Diana
M. ROUSSEAU François

Professeure des universités, Ecole Centrale de Nantes, LS2N
Professeur, IMT Atlantique, LATIM

AUTRES MEMBRES DU JURY :

M. LALLEMENT Alex

Maître de conférences, Université de Strasbourg –
Encadrant

M. PASSAT Nicolas

Professeur, Université de Reims Champagne-Ardenne, CRESTIC –
Examineur

Financement

Je souhaitais remercier Vincent Noblet et Laura Harsan d'avoir dirigé ma thèse. Vincent pour sa présence, ses idées et sa motivation pour avancer mes travaux de recherche. Laura pour ses conseils et son point de vue biologique.

Merci à Alex Lallement d'avoir encadré mes travaux, d'avoir toujours su les encourager, dans certains cas les limiter, et d'avoir partagé plusieurs heures d'enseignement avec moi.

Merci à Benjamin Lemasson du GIN et à Sylvain Faisan d'ICube d'avoir fait partie du comité de thèse.

Merci aux membres du jury d'avoir accepté d'en être membres, de m'avoir accordé du temps et d'avoir activement participé à ma soutenance de thèse ainsi qu'à la correction de ce manuscrit.

Merci au laboratoire ICube, notamment à l'équipe IMAGEs et à l'équipe IMIS, qui m'ont hébergé pendant ces trois années et demie. Par extension, merci aux personnes qui ont partagé mon bureau pendant mes rares moments de présence, notamment à Arnaud.

Merci à nouveau aux équipes IMAGEs et IMIS du laboratoire ICube ainsi qu'au laboratoire GIN d'avoir fourni les données nécessaires à mes recherches.

Merci à mes proches de m'avoir changé les idées, et contrairement aux idées reçues, à ne pas m'avoir posé tant de questions que ça sur mes recherches.

Merci à ma meilleure amie et conjointe Arina de m'avoir supporté pendant tout ce temps, dans les bons moments comme dans les mauvais. Je te serai toujours reconnaissant.

Financement

Ce travail a été soutenu par le projet ArtIC "Artificial Intelligence for Care" (programme ANR-20-THIA-0006-01) et cofinancé par la Région Grand Est, l'Inria

Nancy - Grand Est, l'IHU de Strasbourg, l'Université de Strasbourg et l'Université de Haute-Alsace.

Acronymes

ANTs	Advanced Normalization Tools
BET	Brain Extraction Tool
CE	Cross-Entropy
CNN	Convolutional Neural Network
CSF	Cerebrospinal Fluid
CT-scan	Computed Tomography-scan
DARTEL	Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra
DBM	Deformation-Based Morphometry
DL	Deep Learning
DSC	Dice Similarity Coefficient
FCN	Fully Convolutional Network
FFD	Free-Form Deformation
FSL	FMRIB Software Library
GAN	Generative Adversarial Network
GD	Gradient Descent
GM	Grey Matter
GPU	Graphics Processing Unit
HMMRF	Hidden Markov Random Field
IRM	Imagerie par Résonance Magnétique
LDDMM	Large Deformation Diffeomorphic Metric Mapping
MI	Mutual Information
MSE	Mean Square Error

NCC Normalized Cross Correlation
NLP Natural Language Processing
ROI Region Of Interest
SGD Stochastic Gradient Descent
SPM Statistical Parametric Mapping
SSIM Structural Similarity Index Measure
STN Spatial Transformer Network
SVF Stationary Velocity Field
SyN Symmetric Image Normalization
TBM Tensor-Based Morphometry
VBM Voxel-Based Morphometry
ViT Vision Transformer
WM White Matter

Sommaire

Financement	iii
Financement	iii
Acronymes	v
Sommaire	vii
Liste des tableaux	xi
Table des figures	xiii
Introduction générale	1
Organisation	2
Contributions	3
Publications	4
1 Études de groupes en imagerie préclinique	7
1.1 Études précliniques	8
1.1.1 Modèles murins	9
1.1.2 Perspectives et imagerie	10
1.2 Imagerie par résonance magnétique	10
1.2.1 Principes physiques de l'IRM	11
1.2.2 Modalités d'acquisition d'images en IRM	12
1.2.3 Limites techniques et méthodologiques	13
1.3 Morphométrie basée voxel	14
1.3.1 Principes de l'analyse VBM	14
1.3.2 Applications de l'analyse VBM	17
1.4 Présentation des jeux de données	18
1.4.1 IRIS	18

1.4.2	GIN	20
1.4.3	Feminad	21
1.4.4	Atlas	23
1.4.5	Prétraitement commun	23
1.5	Limites des méthodes conventionnelles	24
1.5.1	Limites des méthodes d'extraction du cerveau	24
1.5.2	Limites des méthodes de segmentation en tissus	25
1.5.3	Limites des méthodes de recalage	26
1.5.4	Impact sur l'analyse VBM	28
1.6	Conclusion	28
2	Extraction du cerveau et segmentation des tissus cérébraux	31
2.1	Méthodes de segmentation par apprentissage profond	33
2.1.1	Architectures	35
2.1.2	Fonctions de coût	37
2.1.3	Augmentation des données	40
2.1.4	Métriques de performances	40
2.2	Etat de l'art des méthodes de segmentation pour les cerveaux de souris	42
2.3	Évaluation de méthodes d'extraction et de segmentation sur nos jeux de données	45
2.3.1	Expériences et validation	46
2.3.1.1	Jeu de données	46
2.3.1.2	Implémentation	46
2.3.2	Résultats et discussion	47
2.3.2.1	Extraction du cerveau	47
2.3.2.2	Segmentation en tissus	50
2.3.3	Conclusion	53
3	Recalage d'images médicales	55
3.1	Introduction	56
3.2	Méthodes conventionnelles pour le recalage d'images médicales	57
3.2.1	Types de déformations	59
3.2.1.1	Transformations linéaires	60
3.2.1.2	Transformations non linéaires	60
3.2.2	Régularisation du champ de déformation	62
3.2.3	Fonctions de similarité	63
3.2.4	Algorithmes d'optimisation	65
3.2.5	Validation des méthodes de recalage	66
3.2.6	Etat de l'art des méthodes conventionnelles	67
3.3	Méthodes de recalage d'images par apprentissage profond	69

3.3.1	Approche supervisée	70
3.3.2	Approche non supervisée	71
3.3.3	Approches hybrides	72
3.3.4	État de l'art des méthodes de recalage par apprentissage profond pour l'imagerie médicale	73
3.4	Conclusion	74
4	Recalage déformable dans le cadre d'études de groupe	77
4.1	Introduction	78
4.2	Travaux relatifs	79
4.3	Recalage d'images pour l'étude de groupes	81
4.3.1	Formulation globale du recalage de groupes	81
4.3.2	Paradigmes d'entraînement ou d'optimisation	82
4.3.2.1	Optimisation pour étude de groupes	82
4.3.2.2	Entraînement sur jeu de données externe	82
4.3.2.3	Entraînement k-fold	83
4.3.2.4	Optimisation par paire	83
4.3.2.5	Paramétrisation du champ de déformation	84
4.4	Expériences et validation	85
4.4.1	Jeu de données	85
4.4.2	Méthodes	86
4.4.3	Implémentation	88
4.4.4	Métriques d'évaluation	89
4.5	Résultats et discussions	89
4.6	Conclusion	94
5	Améliorations pour le transfert d'atlas et l'inversibilité du champ	95
5.1	Introduction	96
5.2	Travaux relatifs	96
5.3	Méthode de recalage inverse	99
5.3.1	Métriques de validation	100
5.3.2	Expériences et validation	101
5.3.3	Résultats et discussions	101
5.4	Transfert d'atlas pour la segmentation des tissus cérébraux	106
5.5	Conclusion	107
6	Application à l'analyse VBM	109
6.1	Introduction	110
6.2	Etat de l'art des pipelines VBM	111
6.2.1	Pipeline VBM	111
6.2.2	Intégration de méthodes par apprentissage profond	112

6.3 Application au jeu de données Feminad	113
6.3.1 Intérêt de l'étude	113
6.3.2 Pipelines comparés	116
6.3.3 Évaluation du recalage sur Feminad	116
6.3.4 Résultats	117
6.4 Conclusion	119
Conclusion	121
Bibliographie	125
Résumé	159
Résumé en anglais	159

Liste des tableaux

1.1	Tableau récapitulatif des jeux de données IRIS, GIN et Feminad.	22
2.1	Dice moyen par jeu de données par modèle entraîné pour l'extraction du cerveau.	48
2.2	Comparaison de la distance de Hausdorff moyenne en voxels par jeu de données par modèle entraîné pour l'extraction du cerveau.	50
2.3	Comparaison du Dice moyen par classe pour la segmentation du cerveau en tissus.	51
4.1	Tableau récapitulatif des méthodes implémentées	86
4.2	Scores de Dice en pourcentage pour les 9 ROI de GIN. Une erreur de segmentation sur MAP6 a conduit à la suppression de la 10ème ROI sur l'ensemble de GIN. Les meilleurs scores par colonne sont indiqués en gras.	90
4.3	Score de Dice moyen et temps total de recalage des méthodes comparées sur le jeu de données GIN	91
4.4	Score de Dice en pourcentage pour 3 ROI spécifiques : GP, HY et OLF représentent respectivement le globus pallidus, l'hypothalamus et les zones olfactives. Les meilleurs scores par colonne sont indiqués en gras.	92
5.1	Comparatif des métriques pour différentes valeurs de λ_2	102
5.2	Comparaison du Dice de 3 régions d'intérêt par méthode avec et sans modifications pour le recalage inverse	104
5.3	Comparaison des scores de Dice par classe pour différentes méthodes de segmentation	107
6.1	Régions cérébrales avec dimorphisme sexuel chez les souris	114

Table des figures

1.1	Pipeline pour l'étude de groupes préclinique en neuro-imagerie.	11
1.2	Exemples d'images IRM de cerveaux de souris de modalité T1, T2 et T2*	12
1.3	Pipeline de traitement d'image individuelle pour l'analyse VBM.	15
1.4	Photographie de l'IRM 7T Bruker Biospec de la plateforme d'imagerie IRIS	19
1.5	Défauts des méthodes classiques d'extraction	20
1.6	Exemple de la segmentation en 10 régions sur un cerveau provenant du jeu de données GIN.	21
1.7	Présentation d'une coupe de cerveau du jeu de données IRIS avec un biais d'intensité visible et son atténuation par l'application de l'algorithme N4.	24
1.8	Coupe d'un cerveau provenant d'IRIS et extraction obtenue par la méthode BET (Brain Extraction Tool).	25
1.9	Coupe d'un cerveau et segmentation en tissus cérébraux obtenue par SPMmouse	26
1.10	Cerveau recalé par la méthode conventionnelle SyN sur un atlas cible.	27
2.1	Architecture U-Net	34
2.2	Pipeline classique d'un problème de segmentation dans le cadre des CNN en apprentissage supervisé.	34
2.3	Développements récents autour de l'architecture U-Net avec les Transformers et les State Space Models pour la segmentation d'image médicale.	37
2.4	Exemple visuel du calcul de la distance de Hausdorff	42
2.5	Deux exemples d'extraction de cerveau.	49
2.6	Exemple de segmentation du cerveau	52

3.1	Types de déformations en 2D	59
3.2	Approche supervisée.	70
3.3	Approche non-supervisée.	71
4.1	Schéma des paradigmes majeurs d'entraînement ou d'optimisation	84
4.2	Méthode de recalage direct	87
4.3	Moyenne des log-Jacobiens des champs de déformation sur le jeu de données GIN. (a) SyN , (b) SyN-DL , (c) DL-IRIS , (d) Pair , (e) OP .	93
5.1	Méthode de recalage simultanée direct et inverse	100
5.2	Comparaison des moyennes des log-Jacobiens des déformations pour chaque méthode avec et sans modifications	105
5.3	Comparaison entre une vérité terrain, une segmentation par UNETR et une segmentation par transfert d'atlas avec la méthode OP-sym	108
6.1	Zones d'intérêt du cerveau pour l'étude Feminad identifiées sur l'atlas Allen	115
6.2	Images moyennes du jeu de données Feminad recalé sur l'atlas Allen par deux méthodes de recalage	118
6.3	Dimorphisme sexuel localisé dans le cerveau de souris pour le jeu de données Feminad	119

Introduction générale

L'Imagerie par Résonance Magnétique (IRM) a permis de nombreuses avancées dans le domaine des neurosciences pour la compréhension du fonctionnement normal et pathologique du cerveau [1]. Son utilisation dans le contexte de la recherche préclinique est en plein essor [2], offrant la possibilité de caractériser des modèles murins *in vivo* et de suivre leur évolution au cours du temps.

La Morphométrie Basée Voxel (Voxel-Based Morphometry, VBM) est une méthode d'analyse d'IRM qui consiste à étudier les différences de volume, de forme ou de densité des tissus cérébraux entre différents groupes de sujets. Cette méthode repose principalement sur une étape de recalage, permettant de placer plusieurs images dans un référentiel commun afin de les comparer voxel par voxel, l'équivalent d'un pixel pour une image 3D. Une étape facultative de segmentation peut être utilisée pour isoler certaines régions ou tissus cérébraux, permettant une analyse plus localisée. L'analyse VBM permet ainsi de détecter des différences subtiles dans la structure cérébrale, qui pourraient être associées à des maladies, des troubles neurologiques ou d'autres conditions médicales.

Les approches traditionnelles de recalage et de segmentation peuvent s'avérer laborieuses à mettre en œuvre, en matière de temps et de qualité. Il a été démontré qu'elles sont sujettes à des erreurs, notamment en présence de variations anatomiques importantes entre les sujets ou sur certaines modalités d'imagerie [3, 4], tandis que les méthodes par apprentissage profond ont démontré une grande efficacité dans le recalage et la segmentation d'images médicales ces dernières années [5, 6]. Les travaux présentés ci-après visent à explorer l'intérêt

de l'intégration de méthodes par apprentissage profond dans le processus de l'analyse VBM dans le cadre spécifique des études précliniques.

Organisation

Dans le chapitre 1, nous soulignons l'importance des études précliniques et le rôle déterminant que joue l'IRM pour étudier le cerveau de la souris, un modèle largement utilisé en raison de sa proximité génétique avec l'humain [7]. Nous détaillons succinctement les principes et les limites de l'IRM, puis nous introduisons l'analyse VBM comme méthode pour détecter et localiser les changements dans le cerveau. Pour illustrer son application aux études précliniques, nous présentons les jeux de données (IRIS, GIN et Feminad) utilisés dans le cadre de cette thèse. Enfin, nous mettons en lumière les difficultés pratiques liées à la segmentation et au recalage, étapes essentielles à l'analyse VBM, ouvrant ainsi la voie à des approches plus performantes.

Le chapitre 2 aborde le problème de la segmentation et propose un état de l'art. Nous évaluons ensuite une méthode de segmentation par apprentissage supervisé, issue de l'état de l'art, pour l'extraction automatique du cerveau ainsi que pour sa segmentation en Matière grise (Grey Matter, GM), Matière blanche (White Matter, WM) et Liquide Céphalo-Rachidien (Cerebrospinal Fluid, CSF).

Le chapitre 3 introduit le principe de recalage et le formalise en tant que problème d'optimisation. Un état de l'art des méthodes conventionnelles et des méthodes basées sur l'Apprentissage Profond (Deep Learning, DL) est ensuite dressé, mettant en évidence les similarités entre ces approches, renforcées par les développements récents, ainsi que leurs limites.

Le chapitre 4 développe une méthode de recalage adaptée à notre cadre d'étude de groupes. Cette approche, spécifiquement conçue pour les études de groupes, exploite à la fois la capacité de représentation des Réseaux Neuraux Convolutifs (Convolutional Neural Networks, CNN) pour modéliser la déformation et la capacité d'apprentissage à partir des données. En tirant parti de l'apprentissage non supervisé, le modèle est entraîné sur l'ensemble du jeu de données. Cette méthode est comparée à d'autres paradigmes d'optimisation et à d'autres modèles de représentation du champ de déformation, notamment

la méthode conventionnelle SyN (Symmetric Image Normalization). L'objectif est d'obtenir des résultats de recalage plus performants, tout en maintenant un temps de calcul inférieur à celui des méthodes traditionnelles.

La méthode décrite dans le chapitre 4 permet d'aligner les individus sur un atlas, mais ne propose pas de recalage inverse (aligner l'atlas sur chaque individu). Disposer de ce recalage inverse est particulièrement utile, notamment pour réaliser la segmentation par transfert d'atlas (transporter la segmentation de l'atlas sur chaque individu). En outre, le recalage inverse contribue à garantir des propriétés spécifiques du champ de déformation, assurant ainsi des transformations plus régulières et réalistes. Le chapitre 5 présente ainsi les modifications apportées à la méthode pour obtenir à la fois la transformation directe et la transformation inverse. La segmentation par transfert d'atlas devient ainsi possible, et ses performances sont comparées à celles obtenues par les méthodes de segmentation supervisées présentées dans le chapitre 2.

Enfin, le chapitre 6 examine l'impact des modifications apportées au recalage sur une analyse VBM consistant à comparer une cohorte de souris mâles avec une cohorte de souris femelles. L'analyse VBM est réalisée à la fois avec notre méthode et une méthode conventionnelle. Les résultats sont ensuite comparés à ceux issus de la littérature. Seul l'impact de la méthode de recalage est évalué. Cependant, cette évaluation reste qualitative, bien qu'expertisée, soulignant ainsi le besoin de cadres d'évaluation et de métriques pour comparer différents pipelines d'analyse VBM.

Contributions

Dans cette thèse, nous apportons plusieurs contributions :

1. Nous formulons un cadre unifié pour la tâche de recalage, intégrant à la fois les méthodes conventionnelles et les approches par apprentissage profond, en explorant divers paradigmes d'entraînement et d'optimisation. La contribution majeure de cette thèse est le développement d'un paradigme d'entraînement spécifique à notre cadre d'étude des groupes, intitulé **OP**. Cette approche se situe à mi-chemin entre un apprentissage

non supervisé et une optimisation par paires issue des méthodes conventionnelles. En adoptant une architecture U-Net classique comme référence commune, nous comparons les principales approches étudiées et validons les performances d'**OP** par rapport aux approches par paires (**Pair**) et aux méthodes d'apprentissage profond classiques (**DL**). Ce travail est en cours de publication à la conférence ISBI, qui se tiendra en avril 2025.

2. La méthode proposée présente des différences notables par rapport aux méthodes conventionnelles dues à la régularisation implicite induite par l'optimisation sur les groupes. Pour les minimiser, nous améliorons le modèle U-Net utilisé, afin d'obtenir à la fois le recalage inverse et des champs de déformation nettement plus réguliers. Nous étudions l'impact de cette régularisation sur les performances du recalage, tant direct qu'inverse, pour les trois principales approches examinées (**OP**, **Pair** et **DL**). Au-delà des performances, nous analysons qualitativement son effet sur la régularisation du champ, en comparaison avec les méthodes conventionnelles. Grâce à la proximité de notre méthode **OP-sym** avec celles-ci, nous l'appliquons à une analyse VBM spécifique et validons qualitativement les résultats obtenus. En identifiant des régions plus spécifiques du dimorphisme sexuel chez la souris, notre approche surpasse la méthode conventionnelle SyN pour cette analyse VBM.
3. Enfin, nous implémentons et évaluons des méthodes issues de la recherche pour les tâches d'extraction et de segmentation du cerveau, afin de confirmer leur efficacité dans notre cadre spécifique d'étude sur les cerveaux de souris. Les approches de recalage développées précédemment sont également évaluées sur cette tâche de segmentation en utilisant des cartes statistiques spécifiques aux souris. Ce travail permet ainsi de valider l'ensemble de ces méthodes sur ces tâches.

Publications

Un article intitulé : *Bridging the Gap Between Conventional and CNN-Based Registration: Application to Mouse Brain Morphometry* a été accepté à la conférence

ISBI 2025 qui aura lieu du 14 au 17 avril 2025. D'autres communications locales ont été réalisées durant la thèse, comme la présentation des travaux lors des journées FMTS (*Fédération de Médecine Translationnelle de Strasbourg*) en juillet 2022, lors d'une journée SAIN (*Small Animal Imaging Network*) en novembre 2022 ou pendant une journée workshop de l'axe transverse DSAI (*Data Science and Artificial Intelligence*) du laboratoire ICube en avril 2023.

CHAPITRE 1

Études de groupes en imagerie préclinique

Sommaire

1.1 Études précliniques	8
1.1.1 Modèles murins	9
1.1.2 Perspectives et imagerie	10
1.2 Imagerie par résonance magnétique	10
1.2.1 Principes physiques de l'IRM	11
1.2.2 Modalités d'acquisition d'images en IRM	12
1.2.3 Limites techniques et méthodologiques	13
1.3 Morphométrie basée voxel	14
1.3.1 Principes de l'analyse VBM	14
1.3.2 Applications de l'analyse VBM	17
1.4 Présentation des jeux de données	18
1.4.1 IRIS	18
1.4.2 GIN	20
1.4.3 Feminad	21

1.4.4 Atlas	23
1.4.5 Prétraitement commun	23
1.5 Limites des méthodes conventionnelles	24
1.5.1 Limites des méthodes d'extraction du cerveau	24
1.5.2 Limites des méthodes de segmentation en tissus . . .	25
1.5.3 Limites des méthodes de recalage	26
1.5.4 Impact sur l'analyse VBM	28
1.6 Conclusion	28

Ce chapitre présente le contexte des études de groupes en imagerie pré-clinique, en soulignant l'importance des modèles murins et de l'IRM. Nous introduisons succinctement les principes de l'IRM et ses différentes modalités d'acquisition. Ensuite, nous détaillons la morphométrie basée voxel, une approche employée pour identifier des variations structurelles cérébrales. Enfin, nous décrivons les jeux de données exploités dans cette thèse et discutons des limitations des méthodes conventionnelles, ouvrant ainsi la voie à des approches innovantes basées sur l'apprentissage profond.

1.1 Études précliniques

Les études précliniques constituent l'un des piliers fondamentaux de la recherche biomédicale, car elles précèdent et conditionnent le passage aux essais cliniques chez l'humain [8]. Grâce à l'utilisation de modèles animaux, les chercheurs peuvent étudier en détail les mécanismes sous-jacents à une pathologie, tester des approches thérapeutiques et optimiser les protocoles expérimentaux avant toute application clinique. Cette rigueur permet ainsi de limiter les risques encourus par les participants lors des phases ultérieures.

L'un des enjeux majeurs des études précliniques réside dans la validation de cibles thérapeutiques et la sélection de candidats médicaments prometteurs. Dans le domaine des neurosciences, il s'agit par exemple d'évaluer la capacité d'une molécule à ralentir la progression d'une maladie neurodégénérative ou à atténuer des manifestations pathologiques. Les données recueillies (mesures

biologiques, observations comportementales, imagerie spécifique) fournissent des preuves essentielles pour anticiper l'efficacité chez l'humain et écarter les interventions trop risquées ou inefficaces. Ces études constituent donc un filtre scientifique critique, guidant les choix stratégiques et financiers de la recherche médicale.

1.1.1 Modèles murins

Le recours à des modèles animaux s'impose comme un levier essentiel des études précliniques [9]. Parmi la grande variété d'espèces disponibles, la souris occupe une place privilégiée pour l'exploration du système nerveux central [10]. Sa facilité d'élevage, son coût relativement modéré et la diversité des lignées génétiquement modifiées offrent une vaste palette expérimentale. Il est ainsi possible de reproduire, sous une forme simplifiée, de nombreuses pathologies humaines, comme la maladie d'Alzheimer ou la maladie de Parkinson, et d'examiner la progression de ces maladies dans des conditions maîtrisées.

Le potentiel de la souris en tant que modèle de recherche repose sur la proximité génétique qu'elle partage avec l'humain [7]. Cette similarité permet de cibler précisément certains gènes candidats dans l'investigation des maladies, tout en bénéficiant d'une rétroaction rapide et fiable sur l'efficacité d'un traitement expérimental. Les techniques de manipulation génétique, qu'il s'agisse de knock-in (l'insertion ou la modification d'un gène), de knock-out (l'inhibition ou la suppression d'un gène) ou d'approches plus récentes fondées sur l'édition du génome, autorisent une grande finesse dans la création de modèles adaptés à des questions spécifiques.

Toutefois, il demeure des différences anatomiques et physiologiques non négligeables entre la souris et l'humain, qui peuvent limiter la validité des conclusions. Les structures cérébrales ne sont pas identiques dans leur organisation, et certaines réponses immunitaires ou hormonales peuvent différer sensiblement. Ces contraintes nécessitent une interprétation nuancée des résultats, laquelle devra impérativement être confirmée par d'autres approches *in vitro*, *ex vivo* ou cliniques. L'élaboration d'un modèle murin pertinent est donc un équilibre délicat entre la complexité nécessaire à la transposition humaine et

la simplicité indispensable à une analyse rigoureuse et contrôlable.

1.1.2 Perspectives et imagerie

Au fil des avancées technologiques, les études précliniques se sont progressivement enrichies d’approches d’imagerie de plus en plus performantes. En neurosciences, l’IRM [11] offre une méthode non invasive permettant d’observer le cerveau murin avec une résolution croissante. Elle est particulièrement utile pour caractériser les modifications anatomiques, repérer l’apparition de lésions ou évaluer la réponse à un protocole thérapeutique. L’IRM s’inscrit pleinement dans la règle des 3R (Remplacer, Réduire et Raffiner), une approche instaurée pour répondre aux préoccupations éthiques liées à l’expérimentation animale [8]. Ces principes visent à minimiser la souffrance des animaux en remplaçant les méthodes invasives par des techniques non invasives, en réduisant le nombre d’animaux nécessaires grâce à des données plus précises et en raffinant les protocoles expérimentaux pour améliorer le bien-être animal.

L’intégration de l’IRM dans les protocoles précliniques, présentée par la figure 1.1, a ainsi transformé la recherche en neurosciences en fournissant une vision quantitative et dynamique des processus pathologiques [12]. L’exploitation de ces données requiert, néanmoins, un soin tout particulier dans l’acquisition, la calibration et le traitement des images, afin de fiabiliser les comparaisons et d’éviter les biais induits. Les chercheurs peuvent ensuite confronter les résultats issus de l’imagerie à ceux provenant d’analyses comportementales, histologiques ou moléculaires. Cette mise en relation multidimensionnelle participe à l’émergence d’une compréhension plus fine de la physiopathologie des maladies du système nerveux et de la réponse à l’intervention thérapeutique envisagée.

1.2 Imagerie par résonance magnétique

L’IRM est une modalité d’imagerie non ionisante pour l’acquisition d’images de haute résolution spatiale et de fort contraste, particulièrement adaptée à l’observation de la structure [13] et de la fonction du cerveau [14], ainsi que d’autres organes. Le principe repose essentiellement sur la résonance magnétique

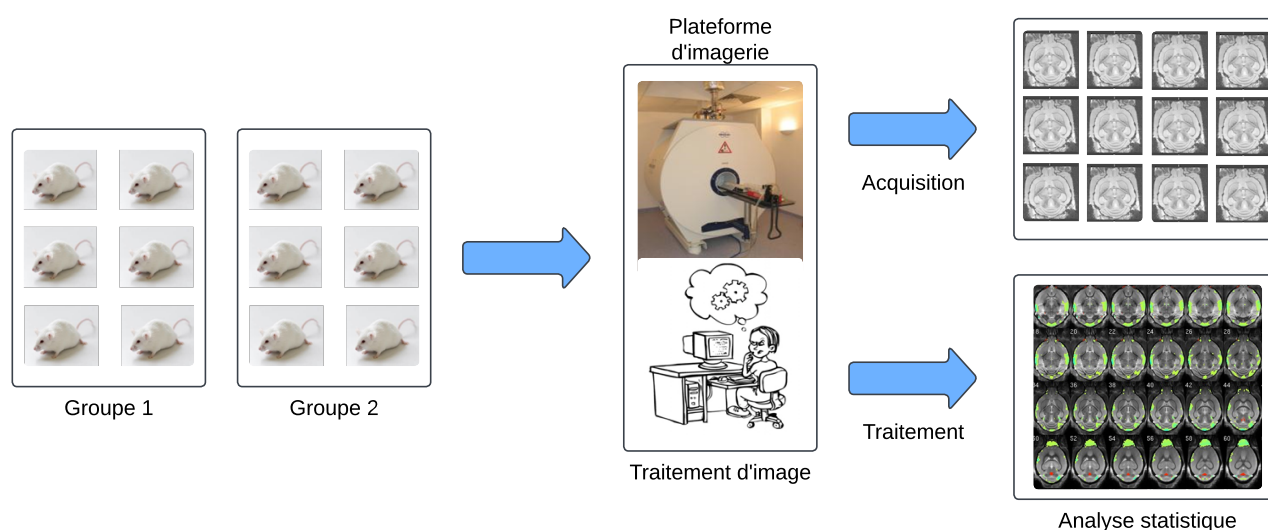


FIGURE 1.1 – Pipeline pour l'étude de groupes préclinique en neuro-imagerie.

nucléaire. Sans entrer dans des détails superflus par rapport à l'enjeu de cette thèse, nous présentons ici les principes de l'IRM [15] de manière succincte.

1.2.1 Principes physiques de l'IRM

L'IRM repose sur la propriété de certains noyaux atomiques, tels que le proton de l'hydrogène, de posséder un spin et un moment magnétique intrinsèque. Lorsque ces protons sont placés dans un champ magnétique statique intense, ils s'orientent préférentiellement selon l'axe de ce champ. Au niveau microscopique, chaque proton effectue un mouvement de précession à la fréquence de Larmor, directement proportionnelle à l'intensité du champ magnétique. C'est cet alignement partiel des spins et leur réponse à des perturbations radiofréquence qui forment le socle de la détection du signal nécessaire à la production des images IRM [16].

Pour caractériser ce signal, on applique des impulsions radiofréquences à la fréquence de Larmor, ce qui déplace temporairement la magnétisation hors de son axe d'équilibre. La magnétisation revient ensuite à l'équilibre selon deux mécanismes de relaxation : la relaxation longitudinale (T_1), qui caractérise la

vitesse de retour de la composante de la magnétisation dans l'axe du champ principal, et la relaxation transversale (T2), qui décrit la perte de cohérence de la magnétisation dans le plan perpendiculaire au champ. À ce phénomène s'ajoutent les inhomogénéités de champ magnétique locales, qui accélèrent le déphasage et sont prises en compte par la relaxation $T2^*$, généralement plus rapide que T2. Les mesures T1, T2 et $T2^*$ diffèrent d'un tissu à l'autre, ce qui confère à l'IRM sa capacité de contraste et de discriminer les diverses structures anatomiques. Les gradients de champ magnétique permettent de coder spatialement l'information, et une transformée de Fourier convertit ensuite les signaux recueillis en images anatomiques ou fonctionnelles.

1.2.2 Modalités d'acquisition d'images en IRM

L'acquisition des images en IRM repose sur l'ajustement des paramètres d'impulsions et de collecte du signal pour exploiter les propriétés distinctes des tissus.

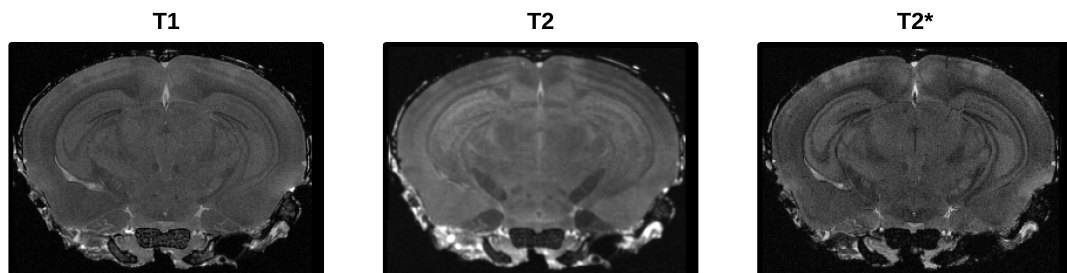


FIGURE 1.2 – Exemples d'images IRM de cerveaux de souris de modalité T1, T2 et $T2^*$. L'image est l'atlas Waxholm [17] sous différentes modalités d'acquisition d'IRM.

La séquence pondérée en T1 est souvent utilisée pour observer la structure générale du cerveau, car elle fait ressortir la substance blanche et la sépare nettement de la substance grise. Le liquide qui entoure et traverse certaines cavités du cerveau apparaît généralement sombre sur ce type d'images, ce qui aide à bien distinguer les contours et l'anatomie.

Avec une séquence pondérée en T2, ce liquide devient au contraire très clair.

La substance grise apparaît plus lumineuse, et la substance blanche un peu plus sombre, ce qui peut aider à repérer certains changements dans les tissus. Une autre modalité, FLAIR (FLuid Attenuated Inversion Recovery) [18], atténue le signal du liquide cébrospinal pour mieux mettre en évidence les anomalies cérébrales, notamment dans les zones difficiles à distinguer sur les images T2 classiques.

Les séquences dites T2* sont sensibles aux inhomogénéités du champ magnétique dans le corps. Ce type d'acquisition, souvent réalisé avec des échos de gradient, peut révéler de petites irrégularités ou des variations liées à la présence de différents composants dans le sang ou dans les tissus.

La figure 1.2 illustre les différentes modalités d'acquisition T1, T2 et T2* sur l'atlas Waxholm [17]. Naturellement, il existe d'autres types d'IRM, comme l'IRM de diffusion ou l'IRM de perfusion, par exemple, mais les présenter dépasse le cadre des travaux de cette thèse.

1.2.3 Limites techniques et méthodologiques

Les études précliniques menées chez l'animal tirent grand profit de l'IRM pour caractériser la structure et la fonction de divers organes, notamment le cerveau, sans recourir à la dissection ou à des méthodes invasives. Cependant, cette technique présente des limites liées à la configuration complexe du matériel et à la nécessité d'adapter l'équipement à la taille et à la morphologie des animaux étudiés. Les machines d'IRM à haut champ dédiées aux modèles rongeurs exigent des champs très intenses, ce qui engendre un coût d'acquisition et de maintenance particulièrement élevé, souvent hors de portée des laboratoires ne disposant pas d'infrastructures dédiées.

La question de l'anesthésie constitue un autre point critique, notamment pour l'IRM fonctionnelle. La majorité des expériences IRM chez l'animal requiert une immobilité stricte pour éviter les artefacts de mouvement, d'où l'usage quasi systématique de l'anesthésie générale et de la contention. Or, l'anesthésie peut modifier la physiologie de l'animal, affecter ses paramètres circulatoires, métaboliques et respiratoires, et donc influencer la qualité et la pertinence des données recueillies [2]. Les protocoles doivent être précisément standardisés

et surveillés (température, ventilation) pour limiter l'impact de l'anesthésie et éviter des variations qui pourraient biaiser les résultats.

Dépassant le cadre de cette thèse, des artefacts dus aux variations du champ magnétique proches des cavités remplies d'air et aux mouvements involontaires comme la respiration ou les battements cardiaques, peuvent engendrer des distorsions ou du flou de mouvement [19], malgré l'utilisation de techniques de synchronisation. De plus, l'hétérogénéité inter-espèce impose des ajustements spécifiques des bobines et des protocoles pour chaque modèle, rendant la reproductibilité des expériences plus complexe et demandant un calibrage rigoureux avant chaque session pour garantir la fiabilité des résultats.

1.3 Morphométrie basée voxel

L'analyse VBM [20] est une méthode quantitative de neuro-imagerie permettant d'évaluer de manière automatique et statistique les différences de volume ou de concentration tissulaire dans l'ensemble du cerveau, voxel par voxel. Développée à la fin des années 1990, elle s'est imposée comme un outil majeur pour détecter et localiser des variations structurelles fines, par exemple liées à la neurodégénérescence, au vieillissement ou à diverses manipulations expérimentales [21, 22]. Contrairement à des approches axées sur des régions d'intérêt prédéfinies, l'analyse VBM explore le cerveau entier sans *a priori*, ce qui s'avère particulièrement pertinent lorsqu'on ne dispose pas d'hypothèses fortes sur la localisation des modifications anatomiques. Son utilisation s'est par la suite étendue au domaine préclinique, notamment aux modèles murins [23, 24] ou à d'autres espèces animales, comme le chien [25] ou le macaque [26].

1.3.1 Principes de l'analyse VBM

L'analyse VBM se déroule en plusieurs étapes clés, chacune impliquant des choix méthodologiques qui peuvent influencer sur la sensibilité et la validité des résultats. Bien que présentées ici de manière séquentielle, les trois étapes de correction du biais, de segmentation et de recalage sont très souvent effectuées simultanément [27]. Le pipeline formé par ces étapes pour le traitement d'une

image est brièvement schématisé en figure 1.3.

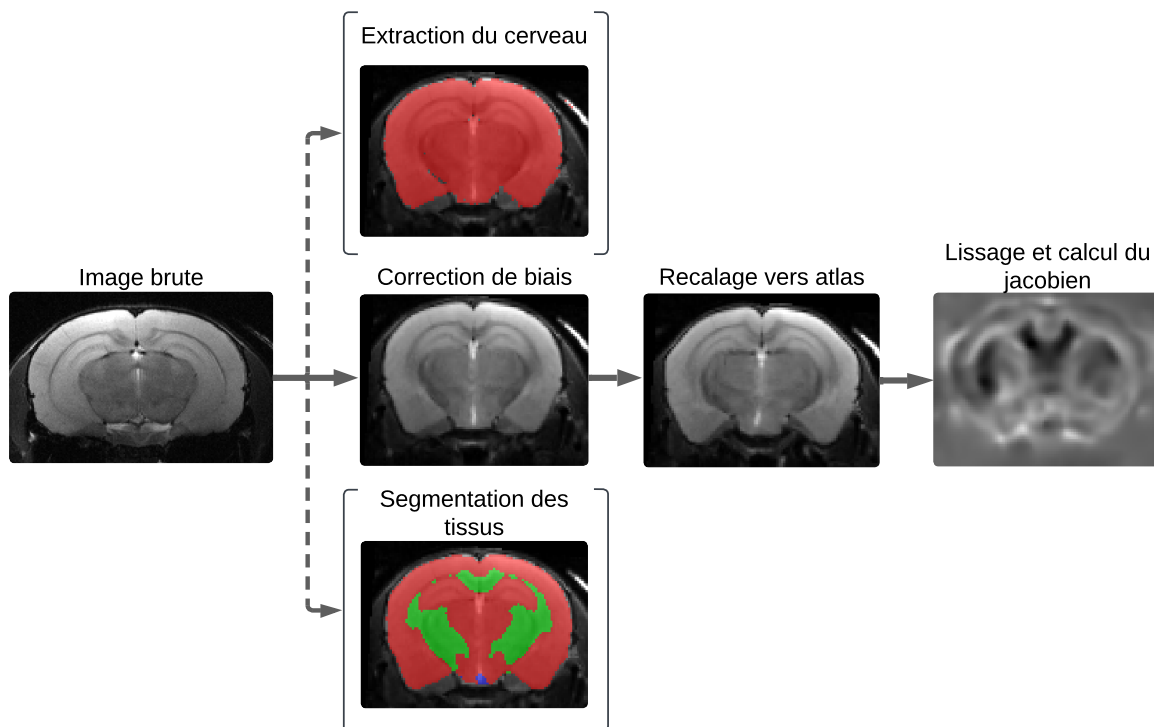


FIGURE 1.3 – Pipeline de traitement d’une image pour l’analyse VBM. L’extraction du cerveau et la segmentation des tissus sont des étapes facultatives en fonction du type d’analyse souhaitée. La correction de biais est toujours effectuée pour compenser les inhomogénéités dues à l’acquisition. Pour la segmentation, GM, WM et CSF sont respectivement représentés en rouge, vert et bleu.

1. Prétraitement et correction des inhomogénéités

Les images IRM subissent une correction de l’hétérogénéité du champ magnétique, afin de limiter leur impact sur les performances des méthodes de segmentation et de recalage ultérieures. L’algorithme N4, par exemple, ajuste localement l’intensité pour réduire les variations parasites dues aux bobines de réception ou à des inhomogénéités de l’aimant. En cas de faible rapport signal/bruit, on peut aussi appliquer un algorithme de débruitage afin d’améliorer la qualité de l’image. Cette étape est particulièrement importante dans les études précliniques de par l’utilisation

d'antennes surfaciques provoquant une grande inhomogénéité de champ lors de l'acquisition IRM. La plupart du temps, pour améliorer les performances des étapes suivantes, le cerveau peut être isolé du crâne et des tissus environnants.

2. **Segmentation des tissus**

Chaque voxel de l'image est classé selon sa probabilité d'appartenir à la GM, à la WM ou au CSF. Cette étape repose sur des approches statistiques, ou d'autres méthodes de segmentation, et peut être affinée par l'utilisation d'atlas probabilistes de l'espèce étudiée. Cette étape est facultative, mais elle permet de se concentrer uniquement sur l'analyse d'un tissu cérébral spécifique.

3. **Recalage spatial**

Les cartes de probabilités résultantes sont alignées sur un espace de référence (atlas) via une transformation d'abord linéaire (rigide ou affine), puis non linéaire pour corriger les déformations locales. L'objectif est de faire correspondre autant que possible les structures ou tissus homologues entre les différents sujets. La variabilité morphologique est ainsi encodée dans le champ de déformation recalant l'individu sur l'atlas. Chez la souris, on peut utiliser un atlas dédié, comme par exemple le WHS2012 [17] ou l'Allen Brain Atlas [28], en ajustant soigneusement les paramètres de l'algorithme de recalage.

4. **Lissage spatial**

Les images segmentées et recalées sont ensuite lissées à l'aide d'un noyau gaussien. Le choix de la taille du noyau dépend d'un compromis entre sensibilité et précision anatomique. Ce lissage normalise la distribution des intensités sur le cerveau et compense de faibles erreurs de recalage. Chez la souris, un lissage trop important peut masquer des changements de petite taille, un noyau plus restreint qu'en clinique humaine est souvent choisi.

5. **Analyse statistique**

Enfin, l'analyse se termine par une comparaison voxel par voxel des cartes de densité ou de volume de matière grise (ou blanche) entre différents

groupes, ou en fonction de variables d'intérêt telles que l'âge, le traitement ou le génotype. Ces cartes sont modulées par les jacobiens des déformations, qui représentent les variations locales des transformations induites par les champs de déformation appliqués pour recalcr les images. Pour d'autres études, comme celles présentées dans cette thèse, les jacobiens sont directement utilisés pour représenter les changements locaux de volumes dans l'entièreté du cerveau. Cette étape repose généralement sur des approches statistiques, paramétriques ou non, mises en œuvre à l'aide de logiciels comme SPM (Statistical Parametric Mapping) ou FSL (FMRIB Software Library). Pour visualiser les résultats, les cartes statistiques sont alors superposées sur le cerveau de référence utilisé pour le recalage, facilitant l'interprétation anatomique et la localisation précise des régions affectées.

Toutes ces étapes constituent l'analyse VBM. Cependant, on différencie plusieurs analyses lorsque certaines étapes ne sont pas réalisées. Si la segmentation des tissus cérébraux n'est pas effectuée, on parle alors de TBM (Tensor-Based Morphometry). Si l'objet étudié par l'analyse statistique n'est pas le jacobien de la déformation mais directement le champ de déformation, on parle alors de DBM (Deformation-Based Morphometry) [29].

1.3.2 Applications de l'analyse VBM

Au fil des années, l'analyse VBM s'est imposée dans l'investigation de multiples processus neurobiologiques. Elle occupe une place majeure dans l'étude de la neurodégénérescence (maladies d'Alzheimer [30], de Parkinson [31], de Huntington [32]), en identifiant les zones d'atrophie les plus précoces. Elle s'applique aussi aux troubles psychiatriques (schizophrénie, dépression, troubles du spectre autistique) [33], offrant des pistes pour détecter des différences structurales parfois discrètes. Dans le cadre de la plasticité cérébrale, des variations de densité de substance grise peuvent être associées à l'apprentissage ou à la répétition intensive d'habiletés motrices ou cognitives. Chez la souris, l'analyse VBM s'est révélée particulièrement utile pour étudier des modèles transgéniques conçus pour modéliser des pathologies humaines. Par exemple, dans le cadre

de la maladie d'Alzheimer, l'analyse VBM peut détecter, dès un stade précoce, une atrophie de l'hippocampe et de certaines régions du cerveau [34, 35]. Ces changements peuvent ensuite être validés et mis en parallèle avec des tests comportementaux, consolidant la pertinence du modèle. Toutefois, les analyses VBM présentent certaines limites, notamment la variabilité des résultats en fonction des pipelines utilisés, un point exploré plus en détail à la fin de ce chapitre.

1.4 Présentation des jeux de données

Dans cette section, nous présentons les trois jeux de données utilisés tout au long de ces travaux. Les jeux de données publics pour les cerveaux de souris étant très limités, nous avons dû nous appuyer sur l'utilisation de jeux de données privés. Un tableau récapitulatif et comparatif 1.1 des trois jeux de données présentés ci-après se trouve à la suite de cette section.

1.4.1 IRIS

Le jeu de données IRIS, nommé d'après la plateforme d'imagerie du laboratoire ICube (Strasbourg) dans laquelle il a été acquis, est en réalité un regroupement de plusieurs jeux de données acquis au fur et à mesure des années. Il regroupe une dizaine de jeux de données pour des études précliniques dont les objectifs initiaux sont omis. Le but d'IRIS est d'obtenir un volume conséquent d'images de cerveaux de souris pour entraîner et ou tester des méthodes. Il contient ainsi 412 IRM de cerveaux de souris pondérées T2. Son hétérogénéité est caractérisée par la présence de cerveaux *ex vivo* et *in vivo* et par l'existence de multiples déformations dues à des traitements ou non. Cependant, toutes les images ont été obtenues avec le même IRM 7 Tesla Bruker Biospec 70/30 USR présenté en figure 1.4, le rendant ainsi homogène dans son acquisition.

Comme des annotations manuelles sont complexes à obtenir sur ce grand nombre d'IRM, les masques pour l'extraction du cerveau pour ce jeu de données proviennent de méthodes conventionnelles automatiques. Trois méthodes conventionnelles différentes ont été appliquées aux 412 IRM du jeu de données. La première méthode consiste à recalcr plusieurs images sur l'image à segmenter



FIGURE 1.4 – Photographie de l'IRM 7T Bruker Biospec de la plateforme d'imagerie IRIS

et à obtenir une représentation probabiliste de l'image afin d'en extraire une segmentation du cerveau. La deuxième méthode consiste à exécuter ce processus de manière itérative. La troisième méthode est une méthode de transfert d'atlas qui consiste à recalcr une image segmentée sur l'image à segmenter, tirant profit de la segmentation initiale sur l'image à segmenter. Visuellement, la meilleure des trois méthodes a été retenue. Cependant, lorsqu'aucune des méthodes ne donne de résultat convenable, il a été nécessaire de rectifier la segmentation à la main à l'aide d'un outil de segmentation. Sur ces 412 images, 204 masques ont nécessité une retouche manuelle pour être satisfaisants visuellement. Deux exemples de défauts sont présentés en figure 1.5 : en haut, un phénomène de sur-segmentation (la segmentation déborde sur tous les contours), et en bas, une partie du cerveau n'est pas segmentée. Pour IRIS, nous disposons donc seulement d'annotations pour l'extraction.

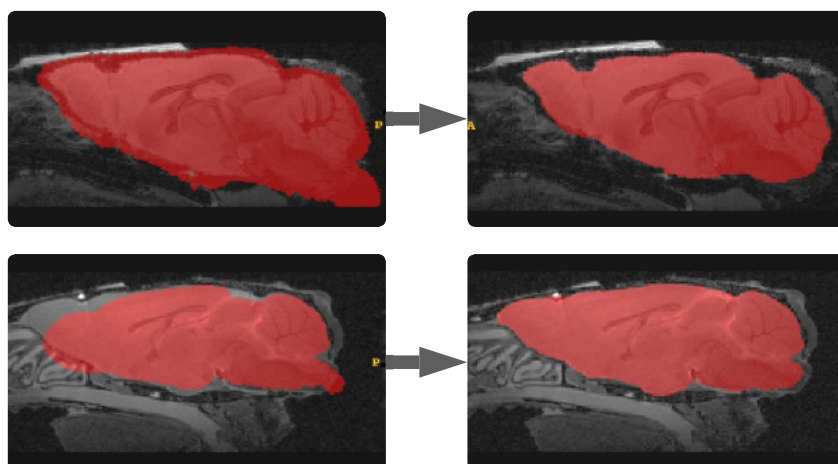


FIGURE 1.5 – Deux défauts présentés par les méthodes classiques d’extraction de cerveau sur le jeu de données IRIS. Les images de droites présentent les masques corrigés manuellement à partir des masques des images de gauche.

1.4.2 GIN

GIN est le deuxième jeu de données utilisé dans cette thèse et comprend 27 IRM cérébrales de souris. Les acquisitions IRM ont été réalisées grâce à un IRM Bruker Biospec Avance III 9,4T de la plateforme IRMaGe du laboratoire GIN (Grenoble). Ce jeu de données est la concaténation de deux jeux de données : SVBP et MAP6. MAP6 est le regroupement de plusieurs acquisitions IRM pondérées T1 de cerveaux de souris *ex vivo* réalisées pour une étude de la connectivité cérébrale [36] dont les souris sont issues d’un modèle murin MAP6 knockout [37]. MAP6 est ainsi constitué de 5 souris MAP6 (KO, knock-out) et 6 souris wild-type (WT, individus sains). De même, le jeu de données SVBP a été acquis pour une étude sur les anomalies structurelles cérébrales avec déficience cognitive chez les souris [38]. SVBP est constitué de 8 souris SVBP (KO) et 8 souris WT. Comme pour MAP6, il s’agit d’acquisitions IRM pondérées T1 de cerveaux de souris *ex vivo*. Ainsi, GIN est constitué de 27 cerveaux de souris. Les spécificités biologiques et éthiques de chacune des acquisitions sont indiquées dans les

articles [36, 38]. L'intérêt du jeu de données GIN, contrairement aux autres jeux de données, réside dans sa segmentation manuelle. En effet, il a été segmenté manuellement par des experts de deux manières : une segmentation unique représentant l'ensemble du cerveau (masque pour l'extraction du cerveau) et une segmentation multiple représentant 10 structures anatomiques cérébrales (labels) suivantes : Hypothalamus, Midbrain, Hippocampal Formation, Striatum, Cerebral Cortex, Globus Pallidus, Thalamus, Cerebellum, Olfactory Areas, Brainstem. Cette segmentation, plus fine que celle en tissus cérébraux (GM, WM et CSF), servira ainsi à évaluer les performances de futures méthodes de recalage. Un exemple de segmentation d'un cerveau provenant de ce jeu de données est illustré en figure 1.6.

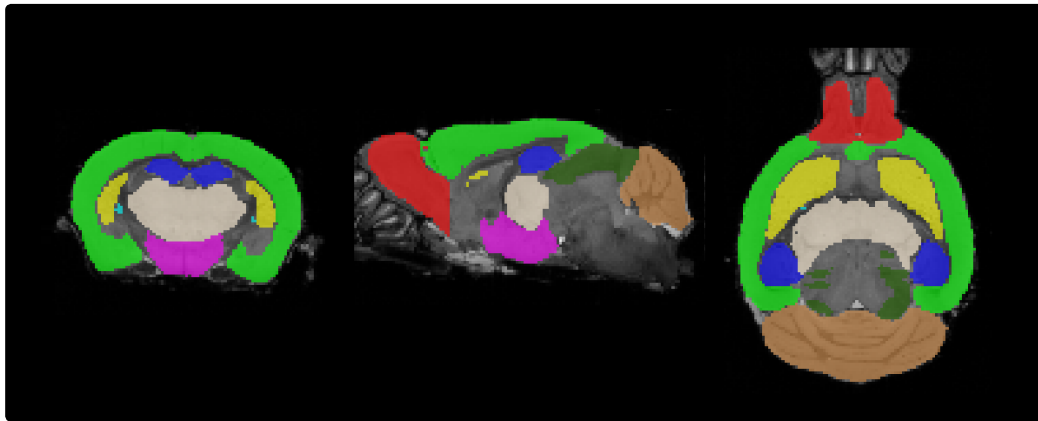


FIGURE 1.6 – Exemple de la segmentation en 10 régions sur un cerveau provenant du jeu de données GIN. En rouge, l'hypothalamus ; en vert, le colliculi ; en bleu, l'hippocampe ; en jaune, le striatum ; en cyan, le cortex cérébral ; en rose, le globus pallidus ; en beige, le thalamus ; en kaki, le cervelet ; en blanc, le bulbe olfactif ; en orange, le tronc cérébral.

1.4.3 Feminad

Le dernier jeu de données utilisé est dénommé Feminad. Tout comme IRIS, il a été obtenu grâce à la plateforme IRIS et à l'IRM 7 Tesla Bruker Biospec 70/30 USR. Ce jeu de données a été acquis pour l'exploration du dimorphisme sexuel

dans la maladie d'Alzheimer grâce au modèle murin dKI. Les données sont donc divisées selon le sexe, l'âge et le modèle des individus. Il s'agit d'un grand jeu de données représentant 67 individus, 34 dKI (dont 17 mâles et 17 femelles) et 33 WT (dont 16 mâles et 17 femelles), imagés à l'âge de 2 mois et de 4 mois. Cela correspond donc à un total de 134 IRM de cerveaux de souris pondérées T2. Pour ce jeu de données, les masques d'extraction du cerveau et la segmentation en tissus cérébraux ont été obtenus grâce à SPMmouse [39] sous forme de cartes de probabilité et binarisées par la suite. Ces segmentations, bien qu'obtenues avec une méthode conventionnelle, sont considérées comme des vérités terrain. Ce jeu de données est particulièrement intéressant pour l'analyse VBM, car les souris WT utilisées ici appartiennent à la souche C57BL/6, largement employée dans les études précliniques [40]. Cette caractéristique facilite la comparaison de nos résultats d'analyse VBM sur le dimorphisme sexuel avec ceux issus de la littérature scientifique.

TABLEAU 1.1 – Tableau récapitulatif des jeux de données IRIS, GIN et Feminad.

Caractéristiques	IRIS	GIN	Feminad
Type d'acquisition	T2	T1	T2
Nombre total d'IRM	412	27	134
Condition	<i>Ex et in vivo</i>	<i>Ex vivo</i>	<i>In vivo</i>
Segmentation	Automatique + Retouches	Manuelle (experts)	Automatique (SPMmouse)
Structures segmentées	Cerveau	Cerveau + 10 structures anatomiques	Cerveau + 3 tissus cérébraux
Particularité	Volume conséquent pour tests	Segmentation manuelle précise	Étude sur le dimorphisme sexuel
Pertinence	Entraînement de modèles	Validation du recalage	Analyse VBM

1.4.4 Atlas

Un atlas est un ensemble standardisé de données d'IRM représentant la structure et la fonction du cerveau ou d'autres organes. Il sert de référence pour comparer, segmenter et analyser les images provenant de différents individus. En d'autres termes, l'atlas fournit un modèle de base qui facilite l'identification, la localisation et l'étude des différentes régions anatomiques dans les examens d'imagerie. L'atlas employé tout au long de ces travaux est l'atlas Allen [28]. En effet, l'Allen Mouse Brain Atlas offre un référentiel anatomique de très haute résolution du cerveau de la souris, accompagné de cartes d'expression génique détaillées et d'une segmentation très précise, ce qui en fait un atlas très utilisé par la communauté [41].

1.4.5 Prétraitement commun

Le tableau 1.1 récapitule les trois jeux de données présentés ci-dessus. Afin de permettre leur utilisation commune, un pipeline de prétraitements a été appliqué à tous les jeux de données. Tout d'abord, toutes les images ont été ré-échantillonnées suivant les dimensions $128 \times 128 \times 128$. Ce choix est arbitraire mais intéressant car à la fois cubique, une puissance de deux pour éviter les problèmes dus à l'utilisation de convolutions, suffisamment petit pour tenir en mémoire sur la carte graphique à disposition (NVIDIA RTX 3080 Ti avec 12 Go de VRAM) et suffisamment grand pour que les images présentent des détails fins. En parallèle, la résolution des images est modifiée afin de les rendre isotropes pour standardiser les calculs. En raison de l'utilisation d'une antenne surfacique lors de l'acquisition, un biais d'intensité marqué est observable sur les images. Pour atténuer cet effet, l'algorithme N4 [42] a été appliqué. La figure 1.7 illustre l'impact de cette correction, comparant une coupe cérébrale avant et après application de N4.

Après la correction du biais, les intensités supérieures au 99ème percentile ont été tronquées au niveau de ce seuil, puis une normalisation linéaire a été appliquée pour ramener les intensités dans l'intervalle unité. Cette étape vise à harmoniser les profils d'intensité entre les jeux de données. Le même prétraitement a été appliqué à l'atlas pour garantir sa compatibilité avec les images

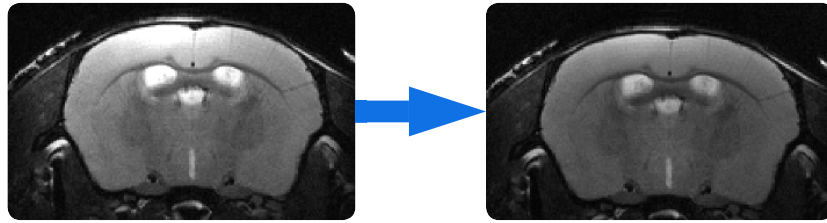


FIGURE 1.7 – Présentation d’une coupe de cerveau du jeu de données IRIS avec un biais d’intensité visible et son atténuation par l’application de l’algorithme N4. À gauche l’image initiale et à droite l’image corrigée par N4.

traitées. Enfin, toutes les images ont été recalées sur l’atlas à l’aide d’un recalage rigide et affine, servant d’initialisation au recalage déformable ultérieur. Cette étape a été réalisée en utilisant le framework ANTs (Advanced Normalization Tools). Les deux étapes sont optimisées par une fonction de coût d’information mutuelle pendant 2000 itérations dans une approche pyramidale sur trois résolutions d’images.

1.5 Limites des méthodes conventionnelles

Il est nécessaire d’identifier les limites des méthodes conventionnelles afin d’évaluer si les approches par apprentissage profond peuvent les surmonter. Cette analyse se concentre sur les problèmes rencontrés en études précliniques à partir d’images IRM pondérées T1, T2 ou T2*. D’autres modalités d’IRM peuvent présenter des défis distincts sur chacune des tâches.

1.5.1 Limites des méthodes d’extraction du cerveau

L’extraction du cerveau (*skull-stripping*) consiste à isoler le cerveau des structures adjacentes comme le crâne et les tissus extra-cérébraux. Bien que de nombreux algorithmes existent pour le cerveau humain (par exemple BET [43], HWA [44], BEaST [45] ou encore BrainMAPS [46]), leur application aux souris est

problématique. Les différences d'acquisition, provoquant de fortes disparités en matière de contraste, de résolution spatiale, de champ de vision, de rapport signal/bruit et de profils d'intensité entre les jeux de données, entraînent une baisse conséquente des performances si les paramètres ne sont pas correctement ajustés aux données [47]. Un exemple d'extraction du cerveau obtenu par la méthode BET est présenté sur la figure 1.8. Cette méthode automatique d'extraction du cerveau utilise un *a priori* de forme spécifique au cerveau humain, nécessitant une adaptation pour les souris. Bien que ces erreurs d'extraction affectent les analyses ultérieures, leur impact reste généralement modéré comparativement aux étapes suivantes [48].

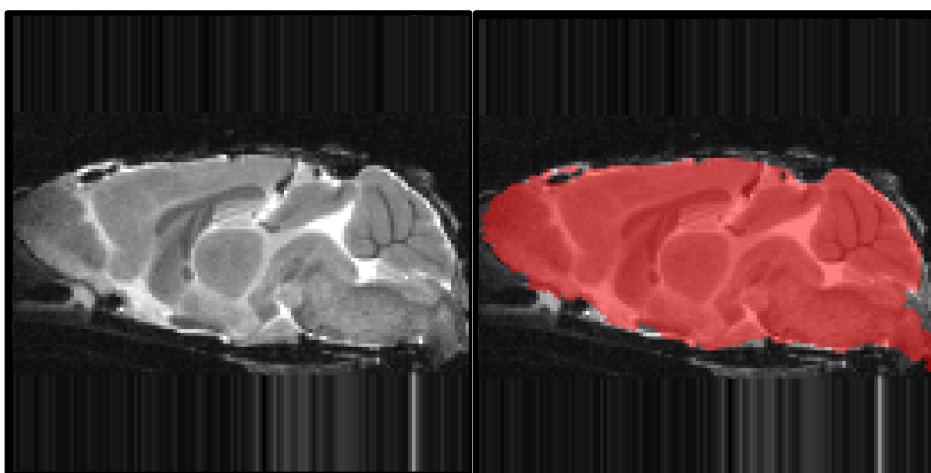


FIGURE 1.8 – Présentation d'une coupe d'un cerveau de souris provenant du jeu de données IRIS et de l'extraction du cerveau associée obtenue par BET [43].

1.5.2 Limites des méthodes de segmentation en tissus

La plupart des analyses VBM s'effectuent en limitant l'étude aux structures de la substance grise, bien que cela ne soit pas systématique comme nous pourrions le voir dans l'étude de Feminad. Ainsi, il est intéressant de segmenter les tissus cérébraux pour chaque image. Des algorithmes, comme RATS [49], PCNN [50] ou des logiciels comme SPMmouse [39], sont couramment utilisés pour obtenir cette segmentation chez les souris. En général, des cartes de probabilité pour les

trois classes sont obtenues pour chaque voxel du cerveau puis sont binarisées en segmentation. Comme pour l'extraction, les algorithmes sont développés pour segmenter le cerveau humain et leur performance chute drastiquement sur les cerveaux de souris [51]. Pour les études précliniques, les méthodes les plus évoluées sont basées sur le transfert d'atlas, rendant ainsi la segmentation intégralement dépendante de l'algorithme de recalage et de l'atlas utilisé [52]. La segmentation obtenue avec l'algorithme de SPMmouse sur un cerveau de souris provenant de Feminad, présentée en figure 1.9, présente de nombreux artefacts : une partie du lobe à gauche n'est pas segmentée, les zones rouges à l'intérieur du cerveau sont très dissociées et le liquide cérébrospinal en bleu n'est que partiellement segmenté.

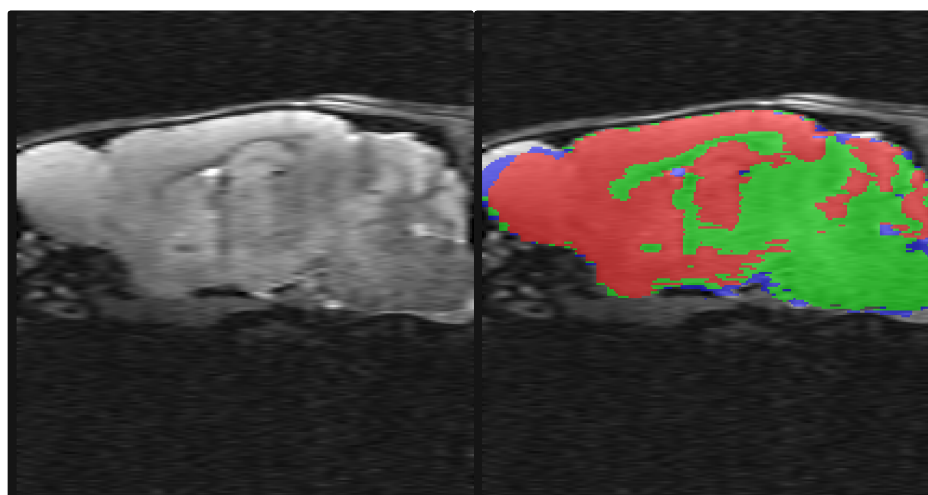


FIGURE 1.9 – Présentation d'une coupe d'un cerveau de souris et de la segmentation en tissus cérébraux GM, WM et CSF, respectivement en rouge, vert et bleu sur l'image, obtenue par SPMmouse [39].

1.5.3 Limites des méthodes de recalage

Le recalage d'images, en particulier dans sa composante non linéaire, constitue une étape cruciale pour aligner les cerveaux des souris sur un atlas ou un modèle de référence. Cette étape est d'autant plus cruciale qu'elle est identifiée

comme la plus influente dans l'analyse VBM [48]. Bien que des algorithmes comme SyN [53] ou Demons [54] soient performants, leur efficacité dépend du choix des paramètres et de la similarité entre l'image source et l'image cible [55]. Lorsque ces paramètres ne sont pas correctement ajustés, le recalage peut produire des déformations inexactes, comme le montre la figure 1.10. Les flèches rouges mettent en évidence des zones anatomiques similaires qui ne sont pas correctement superposées, entraînant des erreurs de recalage qui affectent également les régions voisines.

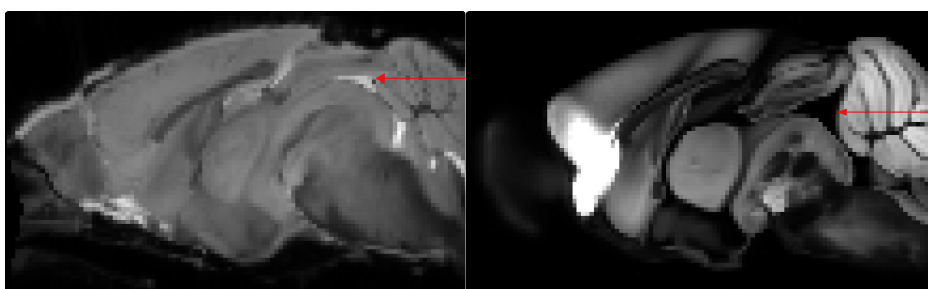


FIGURE 1.10 – Coupe d'un cerveau de souris recalé sur un atlas par la méthode SyN. À gauche : résultat du recalage par SyN. À droite : atlas de référence. Les flèches rouges indiquent une zone anatomique commune soulignant les imperfections du recalage.

De plus, ces méthodes présentent un coût computationnel élevé, dépassant fréquemment plusieurs dizaines de minutes par image [56], limitant ainsi leur applicabilité aux études à grande échelle. L'atlas utilisé peut aussi introduire des biais dans la méthode. L'atlas Allen présente notamment certaines limitations, telles que des troncutures sur les extrémités du cerveau et des variations de contraste, pouvant engendrer des erreurs supplémentaires dans le pipeline d'analyse. Par ailleurs, il peut introduire un biais lié au sexe [57]. Pour limiter ces biais, une approche couramment adoptée consiste à générer un *template* spécifique à l'étude, bien que cela implique un surcoût en ressources [58]. Brièvement, parallèlement au recalage, un template est construit de manière itérative en recalant chaque individu sur l'image moyenne des individus.

1.5.4 Impact sur l'analyse VBM

L'analyse VBM, considérée dans sa globalité, présente des limites intrinsèques liées à la variabilité des pipelines utilisés. Chez l'humain, des études comparatives ont mis en évidence d'importantes disparités selon les choix méthodologiques. Par exemple, les effets de l'âge sur le cerveau sont tributaires du choix des méthodes de segmentation et de recalage employés [59]. Une analyse systématique de 10 pipelines a montré que chaque étape, en particulier le recalage, influence fortement les résultats, notamment pour la prédiction de l'âge [48]. Ces variations sont également observées entre les logiciels SPM et FSL, dont les implémentations distinctes de l'analyse VBM produisent des résultats VBM divergents [60]. Plus récemment, des travaux ont évalué l'impact du choix de l'atlas et de la méthode de segmentation sans changer la méthode de recalage, confirmant l'hétérogénéité provoquée par le choix des méthodes [61]. Une étude complémentaire a révélé que les régions cérébrales associées à l'âge et au sexe varient selon le pipeline utilisé [62].

Chez la souris, l'étude de l'impact du pipeline VBM est très limitée. Une étude a identifié des zones d'intérêt divergentes selon que le recalage était effectué avec DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra) [63] ou via des cartes statistiques de segmentation [64, 65]. Une analyse des paramètres du recalage SyN a par ailleurs démontré de grandes disparités dans les résultats de l'analyse VBM obtenus dans un contexte murin [66].

Ces résultats soulignent que chaque étape du pipeline VBM introduit une variabilité significative dans les résultats. Cependant, cette influence n'a été évaluée que sur des données humaines ou sur des études très limitées pour la souris, sur un choix restreint de pipelines et dans des contextes spécifiques.

1.6 Conclusion

Ce chapitre a introduit les concepts fondamentaux et le contexte de cette thèse, en mettant l'accent sur les études précliniques, l'IRM pour l'acquisition d'images et l'analyse VBM comme outil pour étudier les changements structuraux du cerveau. Nous avons également présenté les caractéristiques distinctes

et la pertinence des trois jeux de données, IRIS, GIN et Feminad, utilisés par la suite dans cette thèse. Les limites et faiblesses des méthodes conventionnelles employées dans les différentes étapes de l'analyse VBM ont été exposées. Par ailleurs, les algorithmes conçus pour l'humain nécessitent souvent des adaptations complexes pour être appliqués au petit animal, impliquant souvent la création de nouveaux *a priori* spécifiques à la souris.

Ces limitations ouvrent des perspectives pour des approches innovantes visant à améliorer l'efficacité, la précision et la reproductibilité des pipelines VBM. L'émergence de l'apprentissage profond offre donc des opportunités pour résoudre ces problèmes. En effet, les méthodes basées sur l'apprentissage profond se sont imposées comme les méthodes de référence pour l'analyse d'images médicales, notamment en neuro-imagerie, que ce soit pour la segmentation des tissus cérébraux ou pour le recalage. Cependant, leur adoption par la communauté scientifique reste limitée par rapport aux pipelines traditionnels. Ces méthodes, bien que performantes, sont souvent très spécifiques et peinent à s'adapter à de nouveaux jeux de données, une situation fréquente en neuro-imagerie avec la diversité des scanners IRM et des protocoles d'acquisitions [67].

Cette thèse vise ainsi à tirer profit à la fois des méthodes conventionnelles et des méthodes basées sur l'apprentissage profond pour améliorer le pipeline VBM en répondant aux limites identifiées. Un pipeline VBM entièrement fondé sur l'apprentissage profond a récemment été proposé pour les études cliniques, mettant particulièrement l'accent sur l'optimisation du temps d'exécution des méthodes [68]. Ce travail illustre l'intérêt croissant pour les méthodes DL dans l'analyse VBM. Les chapitres suivants s'intéressent plus particulièrement aux tâches de segmentation et de recalage, en explorant comment exploiter le contexte spécifique de l'étude de groupes pour chacune des étapes de l'analyse VBM.

CHAPITRE 2

Extraction du cerveau et segmentation des tissus cérébraux

Sommaire

2.1 Méthodes de segmentation par apprentissage profond	33
2.1.1 Architectures	35
2.1.2 Fonctions de coût	37
2.1.3 Augmentation des données	40
2.1.4 Métriques de performances	40
2.2 Etat de l’art des méthodes de segmentation pour les cer- veaux de souris	42
2.3 Évaluation de méthodes d’extraction et de segmentation sur nos jeux de données	45
2.3.1 Expériences et validation	46
2.3.1.1 Jeu de données	46
2.3.1.2 Implémentation	46
2.3.2 Résultats et discussion	47

2.3.2.1 Extraction du cerveau	47
2.3.2.2 Segmentation en tissus	50
2.3.3 Conclusion	53

L'extraction du cerveau consiste à isoler le cerveau des structures adjacentes comme le crâne et les tissus extra-cérébraux. Par ailleurs, elle sert de prétraitement à la segmentation des tissus cérébraux [69, 70, 71]. Au-delà de l'analyse VBM, la segmentation d'images médicales permet, entre autres, d'améliorer la précision des diagnostics et d'aider à la planification d'interventions médicales [72, 73]. Il s'agit donc d'un domaine majeur de l'analyse d'images médicales, ce qui explique l'intérêt de nombreux travaux de recherche pour la segmentation d'images médicales [74, 75].

Les méthodes classiques de segmentation s'appuient sur des algorithmes basés sur le seuillage [76], l'apprentissage machine [77, 78, 79], les contours actifs [80], la croissance de régions [81] ou encore les formulations d'optimisation énergétique, telles que les méthodes graph-cut [82], les champs aléatoires de Markov [83] et les méthodes level-sets [84, 85]. Ces méthodes classiques ont déjà fait l'objet d'études de revue [86, 87] et de comparaisons [88, 89], mettant en avant leurs bonnes performances dans certains cas d'utilisation.

Dans le cadre spécifique de la segmentation du cerveau, les méthodes conventionnelles les plus performantes [3] sont les méthodes basées sur les atlas comme la segmentation par transport d'atlas ou multi-atlas [90, 91]. Dans ces approches, un ou plusieurs atlas et leurs segmentations associées sont recalés sur l'image à segmenter, puis les segmentations sont fusionnées pour obtenir la segmentation finale. De manière générale, ces méthodes sont désormais surpassées dans de nombreux domaines d'application par les approches basées sur l'apprentissage profond [92].

L'objectif de ce chapitre est d'évaluer des architectures DL de l'état de l'art dans le contexte spécifique de l'extraction cérébrale et de la segmentation des tissus cérébraux chez la souris. Ce chapitre présente ainsi les méthodes d'apprentissage profond pour la segmentation, ainsi que leurs stratégies d'évaluation, avant de proposer un état de l'art des techniques d'extraction du cerveau et de segmentation des tissus pour les IRM. Enfin, nous avons implanté et comparé les

performances de plusieurs architectures sur les problématiques spécifiques de l'extraction du cerveau et de la segmentation des tissus cérébraux à partir d'IRM de souris.

2.1 Méthodes de segmentation par apprentissage profond

L'émergence de l'apprentissage profond a marqué un tournant dans la segmentation des IRM [93, 94]. Pour une introduction au DL, le lecteur pourra se référer à [95]. Contrairement aux méthodes traditionnelles d'apprentissage automatique, les CNN apprennent directement des caractéristiques pertinentes à partir des données brutes, éliminant ainsi le besoin d'extraction manuelle. Ces modèles sont particulièrement adaptés aux tâches de segmentation, car ils exploitent leur capacité à apprendre des structures hiérarchiques, depuis des motifs locaux jusqu'aux représentations globales de l'image.

Les premières architectures appliquées à la segmentation au niveau du pixel sont les Réseaux Entièrement Convolutifs (Fully Convolutional Networks, FCN) [97]. Un encodeur extrait les informations en compressant les images d'entrée, tandis que le décodeur consiste alors en une déconvolution unique entraînée pour sur-échantillonner l'information codée et obtenir la carte de segmentation. Pour améliorer la partie décodeur, [98] introduit les réseaux de déconvolution avec plusieurs couches de déconvolution, ou convolution transposée, afin d'obtenir le symétrique des encodeurs pour décoder l'information. Ainsi, la combinaison d'un encodeur de convolution et d'un décodeur de déconvolution a permis de créer l'architecture emblématique U-Net pour le domaine médical [96] en ajoutant aussi des *skip-connections* qui sont des connexions directes entre des couches correspondantes de l'encodeur et du décodeur pour préserver des informations de localisation pour une segmentation plus précise. Cette architecture U-Net est schématisée en figure 2.1.

Dans le cadre de l'apprentissage supervisé, comme c'est généralement le cas pour la segmentation en apprentissage profond, le modèle est entraîné à partir d'un ensemble de données d'entraînement, où chaque image est accompagnée

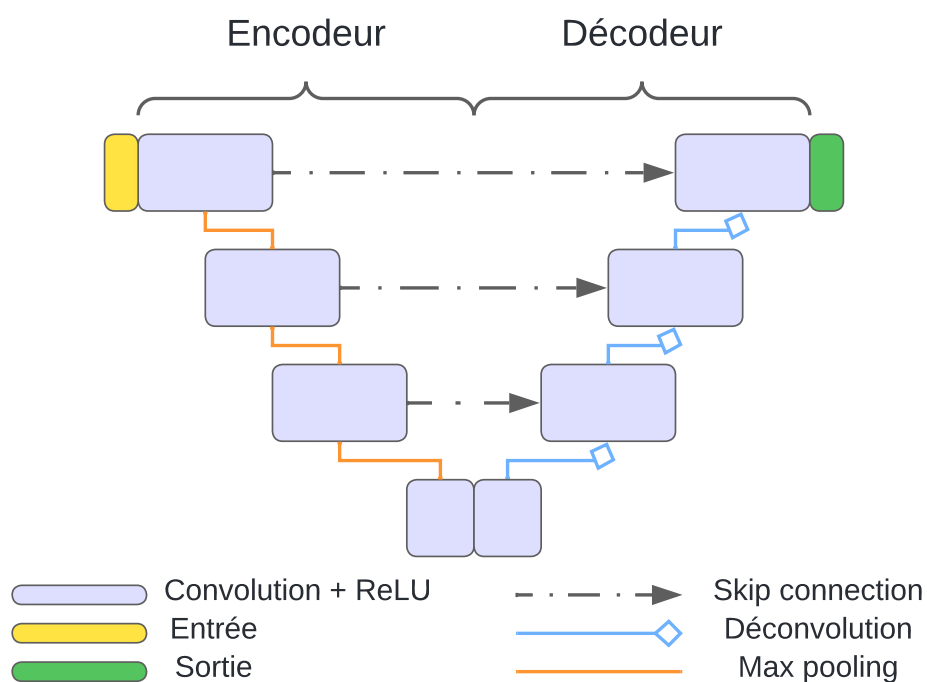


FIGURE 2.1 – Architecture U-Net classique avec convolution, ReLU, maxpooling et déconvolution comme présentée dans [96].

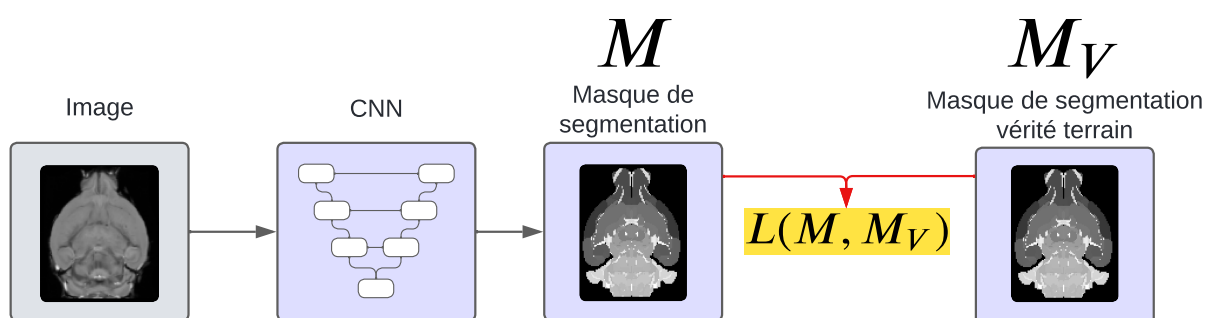


FIGURE 2.2 – Pipeline classique d'un problème de segmentation dans le cadre des CNN en apprentissage supervisé.

d'annotations de référence. Ces annotations, ou labels, fournissent des informations explicites sur les structures cérébrales permettant au réseau d'associer les caractéristiques des images aux régions annotées. Pour la segmentation cérébrale, chaque voxel de l'image IRM est étiqueté comme appartenant à une région spécifique du cerveau, en fonction des besoins de l'étude. Le modèle estime un label pour chaque voxel en étant contraint par une fonction de coût associée à la supervision. Plus concrètement, il s'agit d'un problème d'apprentissage dans lequel le modèle est entraîné à segmenter le cerveau en cherchant à minimiser une fonction objectif, qui calcule très souvent l'erreur entre la segmentation obtenue par le modèle et la segmentation réelle donnée par l'annotation. La figure 2.2 illustre le problème de segmentation supervisé tel qu'il est formulé dans le cadre des CNN, avec une fonction de coût L calculée entre les deux annotations M et M_V .

L'apprentissage supervisé appliqué aux CNN nécessite cependant un ensemble de données annotées de haute qualité, ce qui représente un défi majeur : les experts en imagerie médicale doivent créer ces annotations détaillées en délimitant les structures cérébrales sur chaque IRM. Ce processus est coûteux et chronophage, mais il est essentiel pour fournir des labels précis pour l'entraînement du modèle. Très souvent, pour pallier le manque de données annotées, des techniques d'augmentation des données sont employées. Ces techniques génèrent de nouvelles images à partir d'images existantes en appliquant, par exemple, des transformations affines. Cela accroît la diversité des données d'entraînement et améliore la robustesse du modèle.

Ainsi, la conception d'un pipeline de segmentation ou d'extraction du cerveau en apprentissage profond requiert la définition de l'architecture du modèle, le choix des fonctions de coût, l'élaboration d'une stratégie d'augmentation des données et la sélection des métriques de performance pour la validation.

2.1.1 Architectures

La majorité des travaux de recherche récents ont repris l'architecture U-Net dans le but de l'améliorer. On peut citer par exemple, son extension aux images 3D avec V-Net [99], l'ajout de couches résiduelles avec ResUnet [100]

ou MultiResUNet [101], ainsi que d'autres modifications d'architectures comme UNet++ [102] et RU-Net [103], permettant d'améliorer les performances de l'architecture.

Plus récemment, avec les avancées en Traitement automatique des langues (Natural Language Processing, NLP), les Vision Transformers (ViT) [104], des modèles basés sur l'attention [105], ont été introduits pour le traitement d'images, notamment pour la segmentation d'IRM. Contrairement aux CNN, les Transformers capturent les relations à longue distance entre les voxels grâce à une attention globale sur l'image entière. Cela a conduit à l'émergence de nombreuses architectures combinant U-Net et Transformers, comme Swin-Unet [106], Transformer-Unet [107] et UNETR [108] pour la segmentation 3D. [109] et [110] proposent un inventaire complet ainsi qu'une comparaison, jusqu'à fin 2022, des améliorations pouvant être apportées à l'architecture U-Net, avec ou sans Transformers.

Dans le domaine du NLP, Mamba [111], basé sur les modèles en espace d'état (State Space Models), rivalise avec les Transformers tout en réduisant significativement le coût en calcul. De la même manière que pour les Transformers, des variantes de Mamba adaptées à la vision par ordinateur ont émergé, telles que Vision Mamba [112], MambaVision [113] ou VMamba [114] qui combinent Mamba et Transformers. Inévitablement, des architectures hybrides combinant Mamba, Transformers et U-Net ont été développées, à l'image de U-Mamba [115], Mamba-UNet [116], VM-UNet [117] et Swin-UMamba [118], ainsi que d'autres variantes comme HMT-UNet [119] et MLLA-UNet [120], prônant toutes des performances accrues.

Dernièrement, les architectures SegKAN [121] et KM-UNet [122], qui combinent Mamba et les réseaux de Kolmogorov-Arnold [123], surpassent les performances des autres architectures, notamment TransUNet et U-Mamba, sur plusieurs tâches de segmentation. Cependant, la puissance de calcul requise par ces nouvelles architectures constitue un défi, tant pour leur entraînement que pour leur déploiement, en particulier dans le cadre de l'imagerie médicale en 3D.

Au-delà du cadre strict de la segmentation, certaines approches intègrent le recalage affine afin de réorienter l'image sur des axes standards avant de la

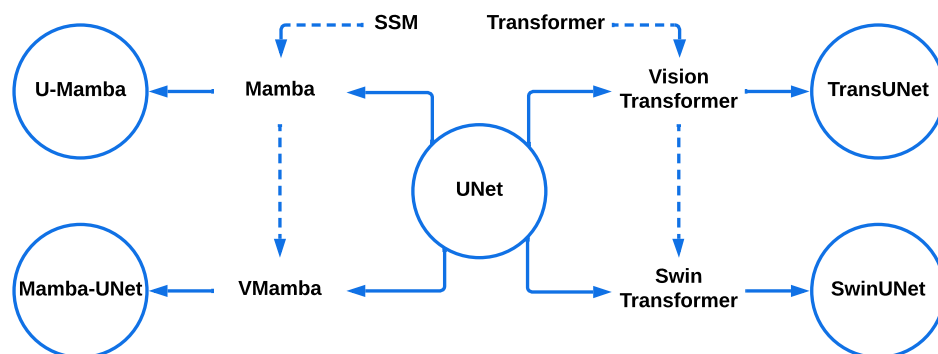


FIGURE 2.3 – Développements récents autour de l’architecture U-Net avec les Transformers et les State Space Models pour la segmentation d’image médicale. Figure reprise de [116].

segmenter [124]. Par ailleurs, des modèles multitâches ont émergé et démontré leur efficacité en apprenant simultanément plusieurs tâches : la segmentation et le recalage [125], la segmentation et l’extraction [52], ou encore les trois combinées [126] dans le but d’améliorer la précision globale de chaque étape.

2.1.2 Fonctions de coût

Les fonctions de coût jouent un rôle essentiel dans l’apprentissage profond, car elles déterminent la manière dont un modèle ajuste ses paramètres afin de distinguer correctement les différentes classes de pixels (ou voxels). Ce critère reflète non seulement la capacité du modèle à classifier les régions pertinentes, mais aussi sa robustesse face à la variabilité anatomique, au bruit et à l’inhomogénéité des images. Le choix d’une fonction de coût adaptée peut considérablement améliorer la qualité finale de la segmentation, en particulier dans des contextes où les jeux de données présentent un fort déséquilibre entre les classes, une situation fréquente dans les applications médicales [127].

Dans notre cas d’application, la segmentation des tissus cérébraux chez la souris, les classes sont naturellement déséquilibrées : l’atlas Allen du cerveau de souris indique des proportions respectives de 23 %, 70 % et 7 % pour les volumes de WM, GM et CSF. Nous notons ainsi, pour chaque voxel x , $p(x)$ la probabilité

prédite par le modèle et $y(x) \in \{0, 1\}$ la vérité terrain.

Cross-Entropy Loss Parmi les fonctions de coût les plus couramment utilisées figure l'Entropie Croisée (Cross-Entropy, CE), ainsi que sa version binaire, la *Binary Cross-Entropy* (BCE). La BCE s'écrit, pour un lot de N voxels, comme :

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N \left[y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \right]. \quad (2.1)$$

Cette fonction de coût est simple à implémenter et constitue un standard pour de nombreux problèmes de segmentation. Toutefois, elle peut se révéler sous-optimale en présence d'un déséquilibre important entre les classes ou lorsque l'on souhaite accorder une importance particulière à la précision des frontières.

Focal Loss Pour remédier aux problèmes de déséquilibre des classes, des variantes pondérées de la CE comme la *Weighted Cross-Entropy* ainsi que des approches telles que la *Focal Loss* [128] sont utilisées. Cette dernière introduit un terme supplémentaire qui renforce la pénalisation des exemples mal classés et réduit l'impact des exemples facilement classés :

$$\mathcal{L}_{\text{Focal}} = -\frac{1}{N} \sum_{i=1}^N \left[\alpha (1 - p_i)^\gamma y_i \log p_i + (1 - \alpha) p_i^\gamma (1 - y_i) \log(1 - p_i) \right], \quad (2.2)$$

où α est un hyperparamètre permettant de contrôler l'importance de la classe minoritaire, et γ ajuste le degré d'accentuation des exemples difficiles. Dans les contextes médicaux, où la classe d'intérêt (par exemple, une tumeur) peut représenter moins de 1 % de l'image, ces stratégies s'avèrent particulièrement efficaces.

Dice Loss Un autre groupe de fonctions de coût adaptées à la segmentation repose sur le Coefficient de Similarité de Dice (Dice Similarity Coefficient, DSC), une métrique de comparaison des segmentations. La *Dice Loss* [129] exploite cette métrique en l'intégrant directement dans la fonction de coût. Elle se définit

comme l'opposé du coefficient de Dice :

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_i p_i y_i}{\sum_i p_i + \sum_i y_i + \epsilon}, \quad (2.3)$$

où ϵ est un petit terme ajouté pour éviter la division par zéro. L'avantage de la *Dice Loss* réside dans sa capacité à prendre directement en compte le recouvrement entre les régions prédites et réelles, tout en gérant efficacement les problèmes de déséquilibre de classes. En effet, la similarité de Dice dépend du rapport entre les volumes prédits et les volumes réels. Pour les problèmes multi-classes, la *Dice Loss* est calculée séparément pour chaque classe.

La *Generalized Dice Loss* [127] propose d'utiliser un score unique pour les problèmes multi-classes. Pour N classes, elle se définit comme suit :

$$\mathcal{L}_{\text{GDice}} = 1 - \frac{2 \sum_{l=1}^N w_l \sum_i p_{li} y_{li}}{\sum_{l=1}^N w_l (\sum_i p_{li} + \sum_i y_{li})}, \quad (2.4)$$

où $w_l = \frac{1}{(\sum_i y_{li})^2}$ est un facteur de pondération qui tient compte du volume de chaque classe afin de mieux gérer les déséquilibres.

Fonctions de coût hybrides Enfin, des fonctions de coût hybrides ont émergé pour combiner les avantages de plusieurs métriques. Par exemple, en associant la *Cross-Entropy Loss* et la *Dice Loss*, on bénéficie à la fois d'une bonne stabilité lors de l'entraînement (apportée par la *Cross-Entropy*) et d'un alignement plus direct avec la métrique d'évaluation (grâce à la *Dice Loss*). Certains auteurs proposent également la *Tversky Loss* [130], qui généralise la *Dice Loss* en introduisant des coefficients de pondération pour gérer plus finement les faux positifs et les faux négatifs.

Il existe aussi plusieurs variantes de la *Dice Loss* comme la *Generalized Dice Loss* [127], présentée précédemment, ou encore la *Wasserstein Dice Loss* [131]. Ainsi, selon les objectifs cliniques et la nature des données, le choix de la fonction de coût vise à maximiser à la fois la pertinence anatomique de la segmentation et la robustesse de l'algorithme lors de l'inférence.

D'autres variantes moins répandues existent, mais en dresser un inventaire exhaustif dépasse le cadre de ce chapitre.

2.1.3 Augmentation des données

Dans le cadre de l'imagerie médicale, où la quantité de données est relativement faible, l'usage de l'augmentation de données est primordial pour éviter le sur-apprentissage dans un contexte supervisé, phénomène où le modèle est performant uniquement sur les données d'entraînement. Les stratégies d'augmentation les plus courantes incluent les transformations géométriques (rotations, translations, symétries, mises à l'échelle), la variation d'intensité ou encore l'ajout de bruit.

Au-delà des stratégies d'augmentation utilisées dans cette thèse, l'augmentation des données est devenue un domaine de recherche à part entière, notamment avec les Réseaux Antagonistes Génératifs (Generative Adversarial Networks, GAN) [132] et les nouvelles méthodes basées sur l'occlusion de l'image [133] ou le mélange d'images [134, 135].

2.1.4 Métriques de performances

Nous présentons ici les différents critères d'évaluation des performances d'une méthode de segmentation. Si l'on obtient une segmentation A par la méthode choisie et que l'on souhaite la comparer à la segmentation B , considérée comme vérité terrain (obtenue par annotation d'experts, par exemple), plusieurs métriques peuvent être utilisées pour évaluer quantitativement la qualité de la segmentation A par rapport à la segmentation B .

Au-delà des métriques classiques utilisées en classification et en segmentation, telles que la sensibilité, la spécificité, la précision ou l'exactitude, d'autres sont particulièrement courantes dans les applications de segmentation d'images, notamment le DSC, défini comme suit :

$$\text{DSC} = \frac{2 \cdot |A \cap B|}{|A| + |B|} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}}, \quad (2.5)$$

où A désigne l'ensemble des voxels de la prédiction, B l'ensemble des voxels de la vérité terrain, TP (True Positive) le nombre de voxels correctement prédits comme appartenant à la classe d'intérêt, FP (False Positive) le nombre de voxels incorrectement prédits comme appartenant à la classe d'intérêt alors qu'ils n'en

font pas partie, TN (True Negative) le nombre de voxels correctement prédits comme n'appartenant pas à la classe d'intérêt, et FN (False Negative) le nombre de voxels incorrectement prédits comme n'appartenant pas à la classe d'intérêt alors qu'ils en font partie. Un coefficient de Dice proche de 1 indique une forte similarité entre la prédiction et la référence. Grâce à sa simplicité de calcul et à sa bonne interprétabilité, le DSC est largement adopté dans la communauté scientifique pour évaluer la qualité globale d'une segmentation [136], en particulier pour les structures dont les volumes sont de taille intermédiaire par rapport à la taille de l'image. Dans le cas extrême où la région recouvre presque intégralement l'image, choisir l'image entière comme segmentation permet d'obtenir des résultats élevés en termes de Dice malgré une mauvaise segmentation. À l'inverse, si la région est de seulement quelques voxels, une faible sur-segmentation entraînera un mauvais résultat en matière de score de Dice.

Une autre mesure couramment utilisée est l'index de Jaccard ou Intersection over Union (IoU) :

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{TP + FP + FN}. \quad (2.6)$$

L'index de Jaccard pénalise davantage les erreurs de segmentation que le DSC. Mathématiquement, les deux métriques sont liées par :

$$DSC = \frac{2 \cdot IoU}{1 + IoU}. \quad (2.7)$$

On note ainsi que l'IoU est toujours strictement inférieur au DSC, sauf aux extrémités, ce qui indique que l'IoU est une mesure plus stricte que le Dice. Étant donné que ces deux métriques moyennent les intersections par le volume global des objets, elles sont sensibles à la taille des volumes segmentés. Ainsi, pour un petit volume à segmenter, une erreur de quelques voxels aura un impact proportionnellement plus important que pour un volume plus grand.

Enfin, la distance de Hausdorff est une métrique orientée sur les contours, particulièrement utile pour évaluer la précision spatiale d'une segmentation, mais très sensible aux valeurs aberrantes. Elle se calcule entre X et Y comme suit :

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}, \quad (2.8)$$

où $d(x, y)$ représente la distance euclidienne entre les points x et y . Cette mesure évalue la plus grande des distances minimales entre chaque point d'un ensemble et l'autre ensemble, quantifiant ainsi leur écart spatial. Un exemple visuel de cette métrique est présenté en figure 2.4.

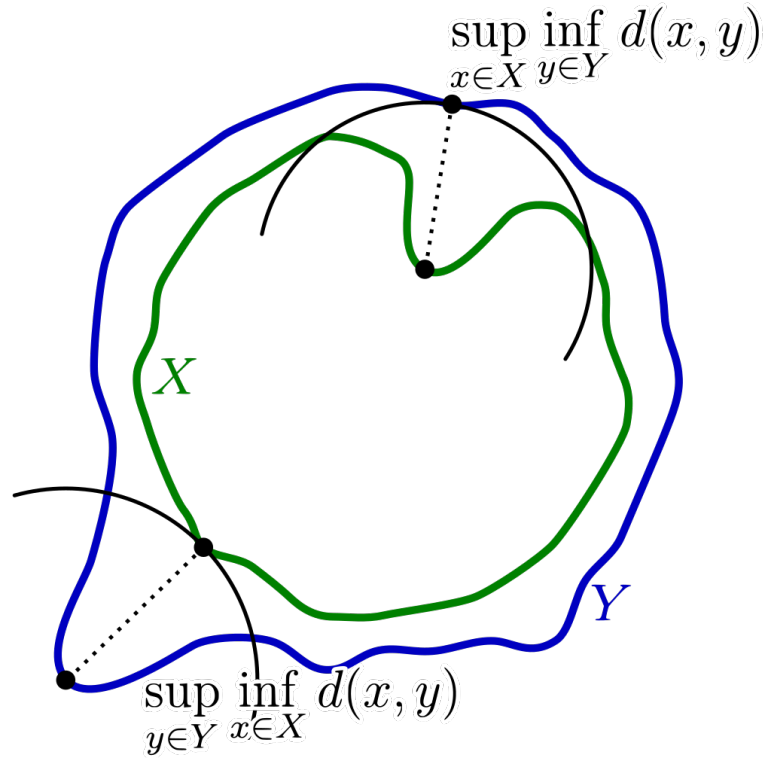


FIGURE 2.4 – Exemple visuel du calcul de la distance de Hausdorff. La figure provient directement de https://en.wikipedia.org/wiki/Hausdorff_distance.

2.2 Etat de l'art des méthodes de segmentation pour les cerveaux de souris

Au-delà des architectures dont l'état de l'art a déjà été présenté dans la section précédente 2.1, nous nous intéressons ici aux méthodes de segmentation

effectivement appliquées dans des études précliniques, notamment celles intégrées dans les logiciels ou frameworks d'analyse d'images médicales tels que FSL [137], SPM [27], ANTs [53] ou encore BrainSuite [138].

Dans FSL, l'algorithme BET [43], utilisé pour l'extraction du cerveau humain, commence par un seuillage grossier afin de localiser le centre du crâne. Un modèle déformable de sphère est ensuite ajusté itérativement pour paramétrer la surface du cerveau et permettre son extraction. Toutefois, l'extraction du cerveau par BET sur des images IRM pondérées T1 est souvent problématique : elle peut soit inclure une fraction seulement du cerveau, soit incorporer trop de tissus non cérébraux [139].

Dans le même logiciel, l'algorithme FAST (*FMRIB's Automated Segmentation Tool*) [70] de FSL assure à la fois la segmentation des tissus cérébraux et la correction du biais d'intensité. De manière simplifiée, il repose sur une approche HMRF (Hidden Markov Random Fields) optimisée par un algorithme itératif d'espérance-maximisation. La première étape de l'algorithme estime les probabilités d'appartenance des voxels aux différentes classes de tissus en tenant compte des relations locales définies par le modèle statistique. La seconde étape met à jour les paramètres des distributions gaussiennes. L'algorithme requiert en entrée une image IRM dont le cerveau a été préalablement extrait, généralement à l'aide de BET, et recalée en affine dans l'espace d'intérêt.

Dans le logiciel SPM, l'algorithme nommé *Unified Segmentation* [27, 140] réalise simultanément la segmentation en tissus cérébraux, la correction du biais d'intensité et le recalage spatial. Le champ de biais est représenté par une fonction basse fréquence, le recalage est paramétré par une combinaison linéaire de fonctions cosinus discrètes et la distribution d'intensité de chaque classe de tissu est représentée par un mélange de gaussiennes. Un algorithme d'espérance-maximisation est employé pour estimer et mettre à jour les paramètres des trois processus jusqu'à convergence. Le recalage affine permet ainsi d'aligner l'image avec les cartes statistiques de segmentation d'un atlas fourni, qui servent de contraintes pour guider la segmentation.

D'autres méthodes existent, comme l'algorithme *Brain Surface Extractor* (BSE) pour l'extraction du cerveau et *Partial Volume Classifier* (PVC) pour la segmentation [69], tous deux intégrés dans BrainSuite [138]. Le logiciel FSL propose

également la méthode FIRST (FMRIB's Integrated Registration and Segmentation Tool) [141], qui effectue à la fois la segmentation et le recalage des images IRM de cerveaux humains à l'aide d'une base de données d'images manuellement segmentées.

Ces méthodes, développées pour le traitement d'images cérébrales humaines, ont déjà fait l'objet de comparaisons [142, 143, 144]. Cependant, leur application aux modèles murins pose un problème majeur : chaque algorithme repose soit sur l'hypothèse d'un cerveau humain (BET, BSE), soit sur un recalage vers un atlas statistique humain fourni (*Unified Segmentation*, FAST), soit encore sur l'utilisation de jeux de données de cerveaux humains pour la segmentation (PVC, FIRST). Il est donc nécessaire d'adapter ces algorithmes aux cerveaux de souris en modifiant leurs paramètres ou en fournissant des atlas et des bases de données spécifiques aux modèles murins. De plus, les images IRM de cerveaux de souris présentent généralement des contrastes plus faibles et des structures moins définies, malgré une variabilité anatomique inférieure à celle des cerveaux humains [145].

Ainsi, certaines méthodes ont été adaptées aux cerveaux de souris et de rats, comme BET avec *Rat Brain Deformable* (RBD) [146] ou, plus récemment, *bet4animal* dans FSL. De nouvelles approches ont également été développées, notamment RATS [49]. Par ailleurs, des atlas, y compris des atlas statistiques, ont été spécifiquement conçus pour les modèles murins, facilitant l'adaptation de certains algorithmes mentionnés précédemment, tels que *Unified Segmentation* de SPM dans SPMmouse [39].

Les méthodes de recalage, en particulier les recalages affines, sont généralement indépendantes des spécificités des espèces, ce qui permet d'utiliser des approches de segmentation par transfert d'atlas [145, 147] pour les modèles murins. Toutefois, la qualité de la segmentation dépend directement de celle du recalage.

Avec l'émergence de plusieurs atlas dédiés aux modèles murins, les méthodes de segmentation basées sur la fusion multi-atlas ont également été explorées [148, 149, 150]. Ces approches se distinguent principalement par leur stratégie de fusion des annotations. Une revue des limites de ces méthodes appliquées aux modèles murins est proposée dans [151], soulignant notamment le temps

de calcul élevé de ces méthodes (jusqu'à deux heures pour le recalage d'un cerveau humain ou murin) et les difficultés d'adaptation des algorithmes conçus initialement pour les cerveaux humains.

Ces défis mettent en évidence les limites des algorithmes de segmentation classiques pour les modèles murins. L'essor des CNN a permis le développement de modèles capables de réaliser une segmentation sans dépendre d'*a priori* ou de cartes statistiques, mais en s'appuyant uniquement sur des données annotées nécessaires à l'entraînement.

Concernant les méthodes de DL appliquées spécifiquement aux souris, plusieurs approches ont été proposées, notamment DeepBrainIPP [152], SHERM [47], un modèle basé sur 3D U-Net [153], MU-Net [52] et MouseGan++ [51]. La plupart de ces méthodes sont des adaptations directes des architectures de l'état de l'art, appliquées à la segmentation du cerveau de souris, obtenant des résultats comparables ou meilleurs qu'avec les méthodes conventionnelles. Des revues de la littérature plus complètes sur l'évolution des solutions DL pour la segmentation d'images médicales sont consultables dans [154] ou [75]. Toutefois, la revue la plus récente confirme l'utilisation des Transformers pour obtenir les meilleures performances de segmentation [155].

2.3 Évaluation de méthodes d'extraction et de segmentation sur nos jeux de données

Dans ce chapitre, nous appliquons puis évaluons le potentiel de plusieurs architectures DL issues de l'état de l'art pour l'extraction du cerveau et la segmentation en tissus dans des IRM acquises chez la souris. Notre objectif n'est pas d'optimiser les architectures ou les fonctions de coût présentées ci-avant, mais plutôt d'appliquer des méthodes déjà existantes pour obtenir des performances au moins meilleures que celles obtenues avec les méthodes conventionnelles de segmentation du cerveau.

L'extraction constitue traditionnellement une étape préalable et essentielle pour la majorité des algorithmes de segmentation des tissus cérébraux, voire pour certaines méthodes de recalage. Pour autant, les tâches d'extraction du

cerveau et de segmentation en tissus cérébraux sont ici explorées de manière indépendante. Ainsi, les modèles de segmentation sont entraînés sur les images originales, après correction du biais par N4, afin d'éliminer toute dépendance de la segmentation vis-à-vis de l'extraction.

2.3.1 Expériences et validation

2.3.1.1 Jeu de données

L'apprentissage étant supervisé, chaque jeu de données est divisé en trois ensembles : entraînement, validation et test. La répartition classiquement adoptée est de 60 % du jeu de données pour l'entraînement, 20 % pour la validation et 20 % pour le test.

Pour l'extraction du cerveau, seuls les jeux de données IRIS et Feminad sont utilisés pour l'entraînement. Les modèles d'extraction du cerveau sont ensuite évalués sur les ensembles de test respectifs d'IRIS et de Feminad afin d'illustrer leur capacité à généraliser sur des données non vues pendant l'entraînement. Une dernière évaluation sur le jeu de données GIN démontre la robustesse de ces modèles et valide leur utilisation sur un jeu de données externe avec un autre système d'acquisition IRM.

Concernant la segmentation des tissus cérébraux, les cartes de segmentation n'étant disponibles que pour le jeu de données Feminad, l'entraînement des modèles est réalisé exclusivement sur ce dernier.

2.3.1.2 Implémentation

Dans l'ensemble de cette thèse, les méthodes deep learning sont implémentées à l'aide des frameworks PyTorch [156] et MONAI [157]. L'architecture sélectionnée est un réseau U-Net [96], plus précisément un *Residual U-Net* [100] sur 5 couches, implémenté dans MONAI. Cette architecture, à la fois compétitive et peu consommatrice en ressources, est facilement déployable. L'entraînement a été réalisé sur un ordinateur équipé d'une carte graphique NVIDIA RTX 3080 Ti (12 Go de VRAM) et d'un processeur Intel Xeon W-2225 (4.10 GHz), en utilisant l'optimiseur Adam [158] avec ses paramètres par défaut et un taux d'appren-

tissage de 10^{-4} . Concernant l'augmentation de données, la stratégie repose sur l'application aléatoire de chaque opération avec une probabilité de 0,5. Elle comprend une rotation sur chaque axe entre -45° et $+45^\circ$, une symétrie selon chaque axe ainsi qu'un décalage global d'intensité d'une valeur de 10 %.

Trois modèles d'extraction de cerveau ont été entraînés selon les jeux de données d'entraînement utilisés. Le modèle A est entraîné sur IRIS (248 IRM), le modèle B sur Feminad (81 IRM) et le modèle C sur la concaténation des deux jeux de données (329 IRM). L'extraction du cerveau étant une tâche de segmentation binaire, la fonction de coût choisie est la *Dice Loss* [129]. Le pipeline utilisé est identique à celui présenté en figure 2.2, mais appliqué à une segmentation binaire.

Plusieurs modèles de segmentation des tissus cérébraux ont été entraînés sur le jeu de données Feminad. Pour cette tâche, nous avons entraîné plusieurs instances du modèle U-Net, identique à celui utilisé pour l'extraction du cerveau, en faisant varier la fonction de coût parmi les suivantes : *Dice + CE*, *Generalized Dice* et *Generalized Dice + Focal* [127] en raison du déséquilibre entre les trois classes.

Afin d'évaluer l'impact d'une architecture Transformer sur nos jeux de données, un modèle UNETR [108] a été entraîné pour cette tâche en utilisant la meilleure des trois fonctions de coût précédemment testées. Le pipeline de segmentation du cerveau a déjà été présenté en figure 2.2. Nous nous limitons à cette architecture UNETR pour les mêmes raisons invoquées pour le choix d'U-Net, à savoir sa relative simplicité et son intégration constante dans les frameworks de DL.

2.3.2 Résultats et discussion

2.3.2.1 Extraction du cerveau

Le tableau 2.1 présente les résultats de l'extraction du cerveau sur chaque jeu de données (jeu de test d'IRIS, jeu de test de Feminad et l'ensemble du jeu de données GIN) pour les trois modèles entraînés, évalués selon la métrique de Dice.

Les scores particulièrement élevés s'expliquent en partie par le volume im-

TABLEAU 2.1 – Comparaison du Dice moyen $\pm$ l'écart-type par jeu de données pour chaque modèle entraîné pour l'extraction du cerveau. La flèche ascendante indique qu'un score plus élevé équivaut à une meilleure performance. Les meilleurs scores par colonne sont indiqués en gras.

Jeu de données de test	IRIS $\uparrow$	Feminad $\uparrow$	GIN $\uparrow$
A (IRIS)	0.969 $\pm$ 0.013	0.966 $\pm$ 0.005	0.965 $\pm$ 0.008
B (Feminad)	0.911 $\pm$ 0.060	0.975 $\pm$ 0.002	0.923 $\pm$ 0.035
C (IRIS+Feminad)	0.969 $\pm$ 0.013	0.972 $\pm$ 0.004	0.955 $\pm$ 0.009

portant du cerveau, qui conduit mécaniquement à des scores de Dice importants. Pour comparaison, des méthodes basiques de seuillages atteignent en moyenne 0.94 de score de Dice sur les trois jeux de données. IRIS, étant un jeu de données très hétérogène, permet au modèle A d'obtenir de bonnes performances sur Feminad et IRIS. L'ajout de Feminad à l'entraînement améliore très peu la segmentation sur celui-ci. À l'inverse, celui entraîné sur Feminad, plus homogène, ne permet pas d'obtenir de bonnes performances sur IRIS et GIN. Ces résultats confirment que la diversité des données d'entraînement joue un rôle clé dans la capacité de généralisation des modèles. Un point particulièrement intéressant est la performance des modèles sur GIN, un jeu de données extérieur à l'entraînement et issu d'un autre système d'imagerie, ce qui témoigne de leur robustesse. Le modèle A, entraîné uniquement sur IRIS, obtient également des résultats satisfaisants sur l'ensemble des jeux de données, renforçant ainsi l'intérêt d'un jeu de données hétérogène pour l'apprentissage supervisé.

La figure 2.5 illustre deux exemples d'extraction du cerveau : en haut, un cas provenant de GIN (jeu de données externe), et en bas, un cas issu du jeu de test d'IRIS (jeu de données interne). Sur l'exemple d'IRIS, ainsi que sur certaines zones de l'exemple de GIN, la sortie du modèle est qualitativement plus cohérente avec la définition de l'extraction du cerveau que la vérité terrain, tout en présentant des contours plus lisses, comme le soulignent les cercles jaunes. En revanche, sur l'exemple de GIN, une partie du cerveau n'est pas correctement extraite, mise en évidence par le cercle bleu.

Ces observations qualitatives remettent en question l'utilisation exclusive du score de Dice pour évaluer la qualité d'une segmentation lorsqu'un certain seuil

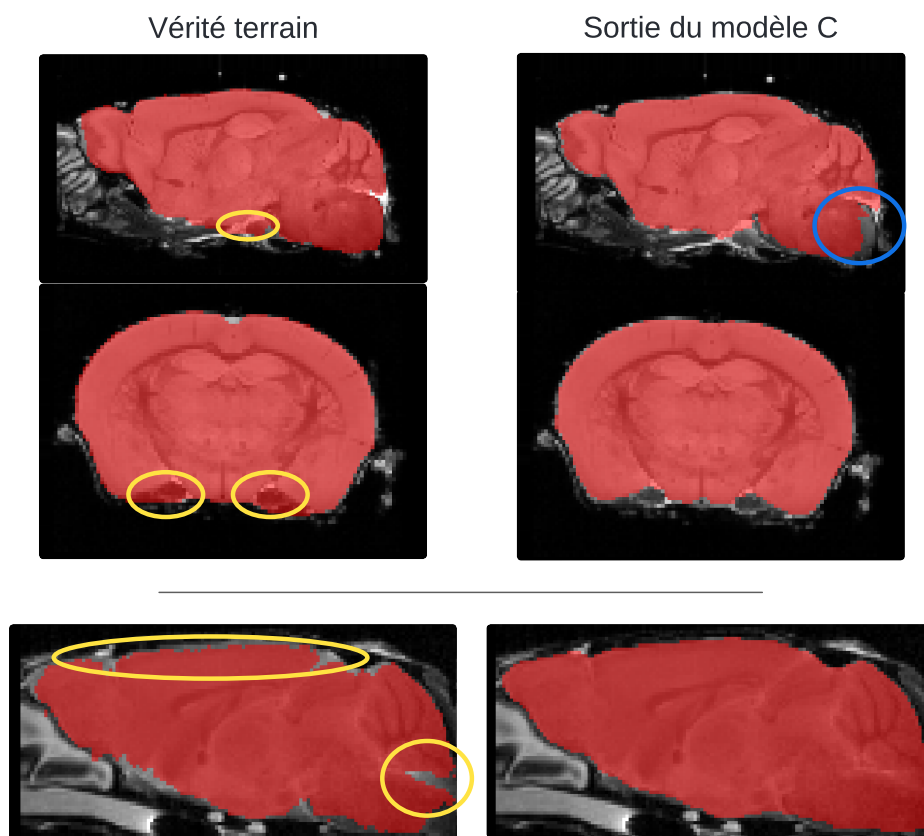


FIGURE 2.5 – Deux exemples d'extraction du cerveau : en haut, un cas issu de GIN; en bas, un cas issu du jeu de test d'IRIS. À gauche, la vérité terrain; à droite, l'extraction obtenue en sortie du modèle C. Les cercles jaunes indiquent des régions où la performance du modèle est meilleure, tandis que le cercle bleu met en évidence une erreur d'extraction du modèle.

est dépassé (supérieur à 0.9 notamment). En effet, cette métrique ne reflète pas toujours les erreurs localisées ni la fidélité des contours segmentés. Il est donc essentiel d'envisager des métriques plus discriminantes pour une évaluation plus fine de la segmentation.

TABLEAU 2.2 – Comparaison de la distance de Hausdorff moyenne $\pm$ l'écart-type par jeu de données pour chaque modèle entraîné pour l'extraction du cerveau. La flèche descendante indique qu'un score plus faible équivaut à une meilleure performance. Les meilleurs scores par colonne sont indiqués en gras.

Jeu de données de test	IRIS ↓	Feminad ↓	GIN ↓
A (IRIS)	17.42 $\pm$ 10.56	19.21 $\pm$ 1.80	12.91 $\pm$ 3.46
B (Feminad)	28.87 $\pm$ 11.76	19.36 $\pm$ 4.41	20.49 $\pm$ 3.17
C (IRIS+Feminad)	16.85 $\pm$ 10.10	18.61 $\pm$ 0.85	11.80 $\pm$ 4.59

Le tableau 2.2 présente les résultats de la distance de Hausdorff, exprimée en voxels, pour les modèles évalués sur les ensembles de test considérés. Alors que le score de Dice suggérerait une performance équivalente des modèles A et C sur IRIS, la distance de Hausdorff, plus discriminante, met en évidence une supériorité du modèle C sur l'ensemble des trois jeux de données. Ces résultats confirment ainsi la pertinence d'IRIS comme jeu d'entraînement pour l'apprentissage supervisé et l'intérêt de la concaténation de jeux de données pour obtenir de meilleures segmentations.

2.3.2.2 Segmentation en tissus

Les résultats présentés dans le tableau 2.3 permettent d'évaluer la performance de différentes architectures de réseaux de neurones et de fonctions de coût pour la segmentation des tissus cérébraux de souris. Les modèles évalués incluent les architectures UNet et UNETR. U-Net a été optimisé avec différentes combinaisons de fonctions de coût, comme présenté ci-dessus. Le modèle UNETR, quant à lui, a été entraîné en utilisant la meilleure fonction de coût identifiée pour UNet. Il est important de rappeler que ces modèles ont été entraînés exclusivement sur le jeu de données Feminad (81 IRM d'entraînement, 27 de validation et 27 de test). Les segmentations considérées comme vérités terrain ont été obtenues à l'aide de SPMmouse via un processus de recalage.

Pour permettre une comparaison directe avec les modèles d'extraction, nous ajoutons une colonne intitulée *Cerveau* dans le tableau 2.3 qui regroupe les trois tissus cérébraux en une seule région. On observe ainsi une performance équivalente sur la tâche d'extraction, avec une légère amélioration de l'écart-type. Ces résultats soulignent l'inutilité de recourir à deux modèles distincts pour ces tâches, contrairement aux méthodes conventionnelles.

TABLEAU 2.3 – Comparaison du Dice moyen $\pm$ l'écart-type par structure sur l'ensemble de test du jeu de données Feminad pour chacun des modèles entraînés. Les classes incluent la matière grise (GM), la matière blanche (WM) et le liquide cébrospinal (CSF). La dernière colonne, intitulée *Cerveau*, correspond à l'extraction du cerveau. La flèche ascendante indique qu'un score plus élevé équivaut à une meilleure performance et les meilleurs scores par colonne sont indiqués en gras.

Modèle	GM $\uparrow$	WM $\uparrow$	CSF $\uparrow$	Cerveau $\uparrow$
UNet, Dice + CE	0.929 $\pm$ 0.009	0.672 $\pm$ 0.017	0.292 $\pm$ 0.086	0.972 $\pm$ 0.009
UNet, GDice	0.898 $\pm$ 0.008	0.601 $\pm$ 0.020	0.266 $\pm$ 0.078	0.945 $\pm$ 0.009
UNet, GDice + Focal	0.893 $\pm$ 0.009	0.621 $\pm$ 0.020	0.236 $\pm$ 0.069	0.933 $\pm$ 0.012
UNETR, Dice + CE	0.940 $\pm$ 0.009	0.702 $\pm$ 0.021	0.353 $\pm$ 0.126	0.975 $\pm$ 0.010

La segmentation de la GM est réalisée avec une bonne précision par tous les modèles, les scores de Dice variant de 89.3 % à 94.0 %. Les modèles UNETR et UNet optimisés avec la fonction de coût *Dice* + *CE* affichent les meilleures performances, surpassant les autres combinaisons de fonctions de coût. Néanmoins, l'ensemble des modèles montre des résultats satisfaisants pour la segmentation de la matière grise.

La segmentation de la WM s'avère plus problématique pour tous les modèles. Une fois encore, UNETR obtient le meilleur score. Ces résultats suggèrent que la matière blanche est plus difficile à segmenter, possiblement en raison de son faible contraste ou de sa similarité avec les tissus environnants.

Les performances pour la segmentation du CSF sont médiocres pour tous les modèles, quelle que soit la fonction de coût utilisée lors de l'entraînement. Le modèle UNETR se distingue en améliorant considérablement le Dice. Ces faibles performances pourraient être attribuées à la complexité des caractéristiques du liquide cébrospinal ainsi qu'au déséquilibre présent entre les classes.

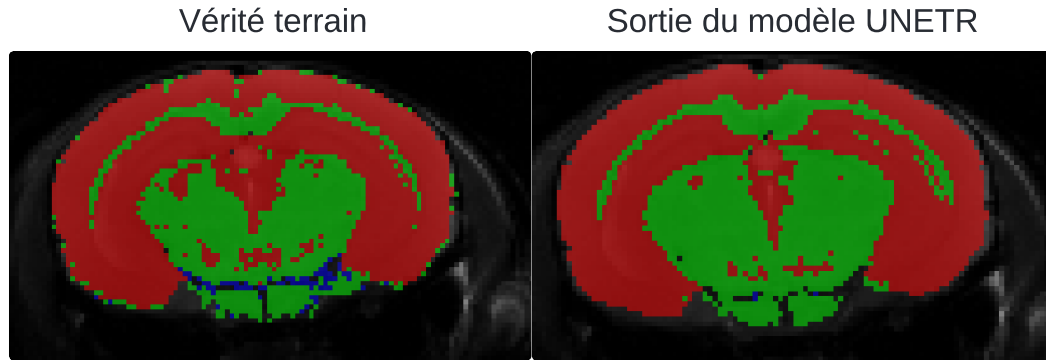


FIGURE 2.6 – Exemple de segmentation de cerveau sur l'ensemble de test. En rouge : la matière grise (GM); en vert : la matière blanche (WM); en bleu : le liquide cébrospinal (CSF). À gauche : la vérité terrain utilisée pour calculer le coefficient de Dice. À droite : la sortie du modèle de segmentation (UNETR, *Dice* + *CE*).

Contre toute attente, l'utilisation de fonctions de coût spécifiquement conçues pour gérer le déséquilibre entre les classes, comme *GDice* ou *GDice* + *Focal*, ne permet pas d'améliorer les résultats dans notre cas. Au contraire, les modèles entraînés avec *Dice* + *CE* restent systématiquement plus performants pour les trois tissus cérébraux.

Un autre facteur influençant la qualité des résultats est la méthode d'obtention des vérités terrain utilisées pour l'entraînement. Ces vérités terrain sont générées par recalage d'un atlas sur les individus du jeu de données, via la méthode intégrée dans SPMmouse [39]. Une piste d'amélioration consisterait à recourir à des annotations manuelles réalisées par des experts.

Au-delà des performances quantitatives des modèles, il est intéressant d'examiner un exemple de segmentation. Sur l'exemple présenté en figure 2.6, la segmentation semble globalement proche de la vérité terrain. Cependant, le CSF, représenté en bleu au bas de l'image, n'est pratiquement pas segmenté. La segmentation de la GM et de la WM apparaît convenable, mais des lacunes persistent pour le CSF, ce qui reflète les scores de Dice observés. Une amélioration potentielle consisterait à utiliser les cartes de probabilité dans les fonctions de coût, plutôt que des segmentations binaires, afin de mieux correspondre aux

annotations de référence. Enfin, bien que la segmentation soit légèrement plus lisse que la vérité terrain, elle présente toujours des artefacts, tels que des bords crénelés, qui résultent de la qualité insuffisante des vérités terrain.

2.3.3 Conclusion

À condition de disposer de vérités terrain fiables, la mise en œuvre de méthodes DL performantes semble relativement aisée pour ces tâches. Le problème d'extraction est très bien maîtrisé et apparaît relativement simple d'un point de vue DL : les modèles entraînés fournissent de très bons résultats, même sur des jeux de données externes provenant d'autres systèmes d'imagerie médicale. Cette performance s'explique par la forte différence d'intensité entre le cerveau et son environnement, facilitant la segmentation. Les contrastes marqués entre l'extérieur du cerveau et le tissu cérébral, ainsi qu'entre le cerveau et le crâne, permettent aux modèles d'isoler précisément le cerveau, rendant l'extraction à la fois fiable et généralisable. Cela est renforcé par la faible variabilité anatomique du cerveau de souris au sein de l'espèce.

En ce qui concerne la segmentation en tissus, les modèles entraînés produisent des segmentations cohérentes pour la matière blanche et la matière grise, mais présentent des difficultés sur le liquide cébrospinal. Ces limitations sont principalement dues à la qualité insuffisante des vérités terrain à disposition. Sans l'utilisation de segmentations annotées manuellement par des experts, il semble difficile d'améliorer significativement ces résultats. Toutefois, on remarque sur les exemples illustrés que la qualité des segmentations obtenues peut surpasser celles des vérités terrain utilisées. L'utilisation d'architectures plus évoluées, ainsi que l'utilisation d'annotations non binaires, peut certainement contribuer à améliorer les performances. Le chapitre présenté ici a ainsi permis d'analyser les performances et les limites des méthodes par apprentissage profond pouvant être rapidement mises en place dans un contexte préclinique.

Nous choisissons donc de mettre l'accent sur les méthodes de recalage, qui, si nécessaire, pourront également servir de méthode de segmentation grâce au transfert d'atlas. Le chapitre suivant formalise le problème du recalage d'image avant d'en établir un état de l'art.

CHAPITRE 3

Recalage d'images médicales

Sommaire

3.1 Introduction	56
3.2 Méthodes conventionnelles pour le recalage d'images médicales	57
3.2.1 Types de déformations	59
3.2.1.1 Transformations linéaires	60
3.2.1.2 Transformations non linéaires	60
3.2.2 Régularisation du champ de déformation	62
3.2.3 Fonctions de similarité	63
3.2.4 Algorithmes d'optimisation	65
3.2.5 Validation des méthodes de recalage	66
3.2.6 Etat de l'art des méthodes conventionnelles	67
3.3 Méthodes de recalage d'images par apprentissage profond	69
3.3.1 Approche supervisée	70
3.3.2 Approche non supervisée	71

3.3.3 Approches hybrides	72
3.3.4 État de l'art des méthodes de recalage par apprentis- sage profond pour l'imagerie médicale	73
3.4 Conclusion	74

3.1 Introduction

Le recalage d'images est un domaine important du traitement d'images, notamment en imagerie médicale mais aussi dans de nombreux autres domaines scientifiques. Ce processus consiste à aligner et superposer deux ou plusieurs images d'une même scène, capturées à différents moments ou de modalités différentes, afin de les rendre comparables.

Dans le contexte médical, le recalage d'images est utilisé dans de nombreuses applications cliniques. Par exemple, il est couramment utilisé en radiologie pour aligner des images CT-scan (Computed Tomography-scan) ou IRM prises à différents moments afin de suivre la progression des tumeurs, des lésions cérébrales ou des maladies dégénératives. Il est également essentiel en chirurgie assistée par ordinateur afin de superposer des images préopératoires sur des images per-opératoires pour guider les interventions chirurgicales avec précision.

Avant d'aborder les méthodes de recalage en détail, on distingue généralement trois cas d'utilisation du recalage en imagerie médicale :

- le recalage intra-patient mono-modal : l'intérêt est de suivre l'évolution d'une pathologie au cours du temps grâce à une seule et identique technique d'imagerie ;
- le recalage intra-patient multi-modal : il met en correspondance plusieurs techniques d'imagerie et donc différentes informations anatomiques et fonctionnelles ;
- le recalage inter-patients (et/ou vers un atlas) : il est utilisé pour faire correspondre les informations entre plusieurs patients afin de, soit créer un atlas, soit à l'inverse utiliser l'information de l'atlas pour obtenir des informations plus précises sur les individus.

Le concept de mettre en correspondance deux représentations d'un même

objet existe au moins depuis le début de l'imagerie médicale avec l'apparition de la radiographie. Dans les années 1970 et 1980, avec l'avènement du scanner à rayons X et de l'IRM, les chercheurs ont commencé à explorer des méthodes permettant d'aligner et de corrélérer ces images [159] afin de mieux comprendre les structures anatomiques et les processus pathologiques du corps humain. À cette époque, le recalage est effectué manuellement par des radiologues et des techniciens en imagerie, nécessitant non seulement une expertise mais aussi un investissement en temps important et un effort fastidieux.

Le recalage d'images a considérablement évolué grâce aux progrès de la technologie informatique et des algorithmes d'apprentissage automatique [160, 4, 161]. Les méthodes automatisées, basées sur des transformations géométriques linéaires ou déformables ainsi que sur les réseaux de neurones convolutionnels, ont permis d'améliorer la précision et l'efficacité du recalage d'images. Ces avancées ont ouvert de nouvelles perspectives en imagerie médicale, offrant des outils puissants pour l'analyse quantitative des images médicales et l'aide à la décision clinique.

Ce chapitre sert d'appui pour les prochains chapitres de la thèse, tant sur le plan scientifique que sur le contexte apporté par certaines sections. Dans une première partie, nous présentons le problème du recalage ainsi que les solutions existantes, sans inclure les approches basées sur l'apprentissage profond qui seront abordées dans la seconde partie. Évidemment, cette division n'est pas dichotomique car certains points de la première partie sont applicables aux méthodes par apprentissage profond, comme par exemple la validation des méthodes de recalage. Cette idée est renforcée par le fait que les méthodes par apprentissage profond peuvent être introduites de manière très similaire aux méthodes conventionnelles.

3.2 Méthodes conventionnelles pour le recalage d'images médicales

Le recalage d'images consiste à transformer une image (appelée image source) afin de l'aligner sur une autre image (image fixe ou image de référence), dans

le but de maximiser leur similarité. Soient M l'image source et F l'image fixe définies sur l'ensemble $\Omega \subset \mathbb{R}^3$ pour une image 3D. Le recalage de M sur F revient à trouver une transformation spatiale $\phi : \Omega \rightarrow \Omega$ telle que l'image recalée $R = M \circ \phi$ soit la plus similaire à F selon une fonction de coût à minimiser.

D'un autre point de vue, ϕ peut aussi se décomposer comme $\phi(x) = x + u(x)$ pour tout $x \in \Omega$. Ainsi, recalier M sur F revient à trouver le champ de déformation u tel que pour tout $x \in \Omega$, $R(x) = (M \circ \phi)(x) = M(x + u(x))$ soit le plus similaire à $F(x)$ selon une fonction de coût à minimiser. Soit L_{sim} la fonction de coût traduisant la dissimilarité entre deux images, le recalage peut alors être formulé comme un problème d'optimisation cherchant à trouver le champ optimal ϕ^* tel que :

$$\phi^* = \arg \min_{\phi} L_{sim}(F, M \circ \phi). \quad (3.1)$$

Étant donné que le recalage est un problème inverse où la solution n'est pas unique, il est mal posé au sens d'Hadamard. Pour pallier ce problème, un second terme de régularisation est ajouté afin de restreindre l'espace des solutions admissibles et de contraindre la transformation spatiale à des solutions physiquement plausibles. Ce terme de régularisation L_{reg} est pondéré par un hyperparamètre λ qui ajuste son importance. L'équation d'optimisation devient alors :

$$\phi^* = \arg \min_{\phi} L_{sim}(F, M \circ \phi) + \lambda L_{reg}(\phi). \quad (3.2)$$

Pour concevoir un algorithme de recalage, il est nécessaire de définir les termes cités précédemment, c'est-à-dire concevoir une fonction de dissimilarité cohérente pour nos deux images F et M ainsi qu'une fonction de régularisation. Le champ de déformation ϕ peut aussi être défini de plusieurs manières en suivant un type de déformation. Une fois les termes définis, il faut choisir et appliquer un algorithme d'optimisation pour minimiser cette fonction de coût. Le choix de tous ces paramètres dépend du type d'images traitées et de leur domaine d'application.

3.2.1 Types de déformations

Les types de déformations (ou modèles de transformation spatiale) peuvent être catégorisés en deux grandes catégories appelées linéaires (rigides ou affines) et non linéaires (voir figure 3.1). Le type de déformation est alors choisi en fonction du problème à résoudre. Pour les études de groupes portant sur la déformation de certaines zones du cerveau, les déformations non linéaires sont bien plus pertinentes à étudier. Cependant, les déformations linéaires restent utilisées comme prétraitement pour éliminer les différences globales avant d'appliquer une transformation non linéaire. Un modèle de recalage complet combine généralement un recalage affine suivi d'un recalage déformable, permettant d'augmenter progressivement la complexité du recalage et d'éviter de converger vers des minima locaux.

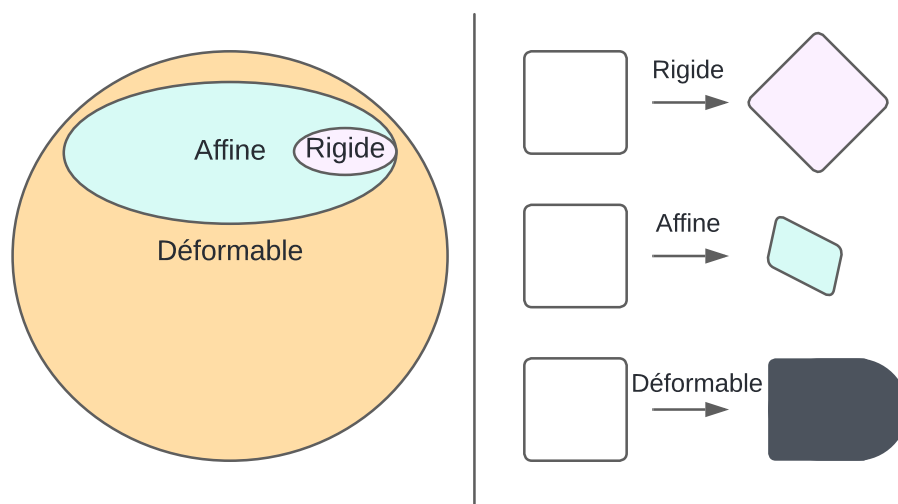


FIGURE 3.1 – Exemples de transformation par type en 2D. La transformation rigide comprend une rotation et une translation. La transformation affine comprend en plus une mise à l'échelle 2D et un cisaillement. La transformation déformable est non-linéaire.

3.2.1.1 Transformations linéaires

Les transformations rigides sont les plus simples : elles conservent les distances. Mathématiquement, si ϕ est une transformation rigide, alors $\forall x, y \in \Omega^3$, $\|x - y\| = \|\phi(x) - \phi(y)\|$. Il est alors facile de définir ϕ par six degrés de liberté correspondant à trois translations et trois rotations. Ce type de transformation compense uniquement les différences de positionnement.

Les transformations affines étendent les transformations rigides en ajoutant les mises à l'échelle et les cisaillements portant le nombre de degrés de liberté à douze. Les transformations affines sont souvent utilisées pour modéliser des effets globaux comme, par exemple, la croissance liée à l'âge ou la différence de taille entre plusieurs individus.

3.2.1.2 Transformations non linéaires

Alors que les déformations rigides et affines sont considérées comme globales, car les paramètres sont les mêmes pour tous les voxels de l'image, les déformations non linéaires sont généralement qualifiées de locales, car leurs paramètres varient en fonction de la position dans l'image. Les transformations déformables possèdent donc un nombre plus élevé de degrés de liberté, ce qui permet de représenter des déformations plus complexes et plus localisées. En imagerie médicale, ces transformations sont à la fois utilisées pour recalibrer des images entre des sujets différents ou vers un atlas de référence (recalage inter-patients) et pour le recalage intra-patient avec par exemple le suivi temporel pour l'imagerie pulmonaire ou l'imagerie cardiaque.

On distingue alors la déformation paramétrique et non paramétrique. La déformation est dite non paramétrique lorsqu'elle est définie implicitement par un modèle physique. En pratique, chaque voxel est défini par des paramètres qui lui sont propres. Comme $\Omega \subset R^3$ et $\phi(x) = x + u(x)$, le déplacement u est exprimé par 3 paramètres par voxel de l'image. À titre d'illustration, cela représente, pour une image de taille 128^3 , un peu plus de 6 millions de paramètres. La déformation dite paramétrique s'appuie sur un modèle de déformation, plus complexe que la déformation affine, basé sur des paramètres afin d'estimer le champ de déplacement. Plusieurs modèles de déformation non linéaires sont

présentés ci-dessous.

Parmi les modèles de déformation non linéaires, les plus couramment utilisés sont ceux fondés sur des points de contrôle, aussi appelés modèles de Déformations de Formes Libres (Free-Form Deformations, FFD). Dans ces approches, la déformation des points de contrôle constitue les paramètres du modèle, tandis que la déformation du reste de l'image est obtenue par interpolation à partir de ces points. Beaucoup de méthodes se sont basées sur ce modèle de déformation en interpolant le champ de déformation avec des splines [162], des fonctions à bases radiales [163] ou encore des B-splines [164]. Ces méthodes sont généralement utilisées pour représenter des déformations de faible amplitude.

D'autres méthodes génèrent le champ de déplacement en respectant certaines lois physiques. Les modèles dits élastiques [165, 166] considèrent l'image comme un solide élastique soumis à des forces qui correspondent aux mesures de similarité d'intensité. Les modèles fluides visqueux [167], quant à eux, expriment le champ de déformation à partir des équations de Navier-Stokes et sont adaptés à des déformations plus larges. Ces deux types de modèles relèvent des déformations non paramétriques, car le champ de déformation est défini voxel par voxel en fonction des contraintes physiques appliquées.

Après les modèles fluides, les modèles basés sur le flot optique [168, 169] considèrent l'image source et l'image cible comme deux échantillons d'un même flot optique à deux instants. Le modèle des Demons [170] considère le recalage comme un échange de flot depuis l'image source vers l'image fixe, faisant le lien avec les modèles de flot optique. Les variantes de ce modèle sont très utilisées car particulièrement efficaces pour représenter de larges déformations. Toutefois, sa principale limitation réside dans l'impossibilité d'obtenir le champ de déformation inverse, ce qui empêche de garantir la conservation de la structure topologique de l'image. Lorsque les déformations sont suffisamment petites, avec $\phi(x) = x + u(x)$, l'approximation $\phi^{-1}(x) \simeq x - u(x)$ est vraie. Cependant, avec les modèles Demons qui sont capables de modéliser de larges déformations, cela n'est plus applicable. Des modèles Demons difféomorphiques ont donc été introduits [171, 54]. Les modèles dits difféomorphiques sont devenus très courants grâce à leur conservation de la topologie, leur inversibilité et leur régularité. Le recalage difféomorphique consiste en l'expression de $\phi(x)$ par la composition de

plusieurs petites déformations difféomorphiques ϕ_i telles que $\phi = \phi_1 \circ \dots \circ \phi_N$ avec N un entier naturel. De cette manière, ϕ est aussi difféomorphique et son inverse s'exprime tel que $\phi^{-1} = \phi_N^{-1} \circ \dots \circ \phi_1^{-1}$.

Les modèles LDDMM (Large Deformation Diffeomorphic Metric Mappings) [172] sont un exemple de modèle de déformation difféomorphique. Les déformations sont obtenues en intégrant un champ de vitesse dépendant du temps, optimisé en minimisant l'énergie géodésique. Les modèles basés sur des SVF (Stationary Velocity Fields) [63] modélisent le champ de déformation en intégrant un champ de vitesse indépendant du temps. En pratique, le champ de déformation est obtenu itérativement grâce à l'algorithme *Scaling and Squaring* [173]. On a ainsi $\phi_{1/2^{t-1}} = \phi_{1/2^t} \circ \phi_{1/2^t}$ avec $t \in [1 : T]$ et $T = 2^N$. Après N étapes, $\phi_1 \simeq \exp(v)$ où v représente le champ de vitesse.

L'inventaire, complet ou non, des modèles de déformation ainsi que leurs variantes a déjà fait l'objet de plusieurs revues de la littérature [174, 4, 175]. D'un certain point de vue, les CNN présentés plus en détail dans la seconde partie de ce chapitre sont aussi des modèles de déformation dont le nombre de degrés de liberté peut dépasser celui du modèle non paramétrique.

3.2.2 Régularisation du champ de déformation

En recalage d'images, le problème étant mal posé au sens d'Hadamard, l'introduction de termes de régularisation du champ de déformation est essentielle pour tendre vers des transformations géométriques plus cohérentes et plus régulières. La régularisation peut être soit implicite, soit explicite. La régularisation est implicite lorsque la déformation est directement paramétrée par des fonctions régulières, comme c'est le cas des modèles de FFD. En revanche, elle est dite explicite lorsqu'il est nécessaire d'ajouter des contraintes externes pour compenser les propriétés du modèle de déformation utilisé, ce qui est typiquement le cas des modèles de déformation élastiques ou fluides.

L'un des objectifs majeurs de la régularisation est de tendre vers l'unicité de la solution afin que le problème soit correctement posé. Une approche couramment utilisée consiste à minimiser l'énergie de courbure, une technique directement inspirée des modèles de splines [162], qui limite les distorsions en pénalisant les

dérivées secondes de la déformation [164]. Le jacobien de la déformation fournit également beaucoup d'informations sur les propriétés locales du champ et est fréquemment employé pour détecter la présence de repliements [167]. À cet effet, un terme de régularisation basé sur le jacobien a été introduit [176, 177]. De plus, certaines fonctions de coût intègrent des contraintes supplémentaires, comme l'*inverse-consistency* [176, 178], qui peut être considérée comme une forme de régularisation.

3.2.3 Fonctions de similarité

Les fonctions de similarité sont définies comme les opposées des fonctions de dissimilarité : maximiser la similarité entre deux images revient à minimiser leur dissimilarité. Elles quantifient la qualité de l'alignement entre l'image source recalée et l'image fixe. L'optimisation est entièrement guidée par ces fonctions ; leur choix est déterminant. De nombreuses approches ont été proposées au fil des années [4], mais les principales peuvent être classées en deux grandes catégories : celles reposant sur des attributs géométriques et celles basées sur l'intensité des images.

Historiquement, les premières fonctions de similarité reposaient sur la mise en correspondance d'éléments géométriques présents dans les deux images, comme des points d'intérêt ou des lignes caractéristiques, pour imiter le processus naturel d'un humain pour recaler deux images. Le principe consiste à aligner ces attributs géométriques en optimisant les paramètres de la transformation. Ces points peuvent être identifiés automatiquement par des algorithmes ou placés manuellement.

Un exemple courant est l'utilisation de points de repère. Plusieurs points de repère, des amers, sont placés à des endroits précis sur les deux images. Un point de repère aura donc une position sur l'image source et une position correspondante sur l'image cible. La fonction de similarité cherche alors à minimiser la distance euclidienne entre ces points après recalage. En minimisant cette fonction, les points de repère appariés deviennent très proches, mais cela ne garantit pas un recalage correct sur le reste de l'image. L'utilisation de points de repère est intéressante en raison de leur signification anatomique et de leur

simplicité d'implémentation. Ils sont utilisés principalement pour des transformations rigides pour recaler des objets relativement simples. Avec $\|\cdot\|$ une norme définie sur Ω , N landmarks placés aux voxels x_i sur l'image M et aux voxels y_i sur l'image F , appariés entre x_i et y_i pour chaque $i \in [1 : N]$, la similarité sur les landmarks est définie par :

$$L_{landmarks} = \sum_{i=1}^N \|\phi(x_i) - y_i\|. \quad (3.3)$$

Contrairement aux méthodes géométriques, les approches basées sur l'intensité permettent de comparer directement l'image recalée et l'image cible sur l'ensemble des voxels, sans nécessiter d'identification préalable d'éléments caractéristiques. Elles sont particulièrement adaptées aux situations mono-modales, où des structures identiques doivent présenter des intensités similaires. Par la suite, d'autres mesures ont été développées pour mieux répondre aux besoins des cas multi-modaux.

L'Erreur Quadratique Moyenne (Mean Square Error, MSE) est un exemple de fonctions de similarité basées sur les distances. Elle s'obtient en additionnant à chaque voxel la différence d'intensité au carré entre les deux images, puis en en faisant la moyenne.

$$L_{MSE}(F, R) = \frac{1}{N} \sum_{x \in \Omega} (R(x) - F(x))^2. \quad (3.4)$$

Elle nécessite donc deux images possédant les mêmes domaines d'intensité pour établir un recalage cohérent. Due à l'utilisation de la norme L^2 , la MSE est sensible aux valeurs aberrantes. D'autres normes peuvent être employées pour réduire cette sensibilité aux valeurs aberrantes mais introduisent à la fois d'autres spécificités.

La Corrélation Croisée Normalisée (Normalized Cross-Correlation, NCC) provient de la théorie du signal et mesure la similarité entre deux images. Comparée à MSE, cette mesure est robuste aux variations d'échelle et de luminosité grâce à une normalisation par la moyenne et l'écart-type des intensités de l'image. Soient $\bar{F}$ et $\bar{R}$ les moyennes et σ_F et σ_R les écarts-type des images, elle s'exprime

comme :

$$L_{NCC}(F, R) = -\frac{1}{N} \sum_{x \in \Omega} \frac{(F(x) - \bar{F})(R(x) - \bar{R})}{\sigma_F \sigma_R}. \quad (3.5)$$

Pour employer la NCC dans un cas multimodal, sa variante locale est souvent utilisée : il s'agit de calculer la NCC sur une fenêtre glissante. Supposons qu'on dispose, pour chaque voxel x , d'une fenêtre W_x centrée en x définissant deux sous-images F_{W_x} et R_{W_x} de F et R , la corrélation croisée normalisée locale est définie par :

$$L_{LNCC}(F, R) = \sum_{x \in \Omega} L_{NCC}(F_{W_x}, R_{W_x}). \quad (3.6)$$

Dans le cas multi-modal, où la même structure n'est pas représentée par les mêmes intensités, on utilise plutôt l'Information Mutuelle (Mutual Information, MI) [179, 180]. Il s'agit d'une mesure de similarité basée sur la théorie de l'information. MI évalue l'interdépendance des intensités entre deux images. Sa définition statistique lui donne une certaine robustesse au bruit et aux variations d'intensité. Elle est définie ainsi :

$$L_{MI}(F, R) = - \sum_{i,j} p_{F,R}(i, j) \log \frac{p_{F,R}(i, j)}{p_F(i)p_R(j)}, \quad (3.7)$$

où $P_I(i)$ est la probabilité de l'intensité i dans l'image I , obtenue par l'histogramme des intensités de l'image, et $P_{I,J}(i, j)$ la probabilité conjointe des intensités i et j dans les images F et R à partir de l'histogramme joint.

3.2.4 Algorithmes d'optimisation

Comme vu auparavant, le recalage d'image se formule comme un problème d'optimisation, consistant à trouver les paramètres de la transformation qui minimisent la fonction objectif définie dans l'équation 3.2.

L'algorithme de Descente de Gradient (Gradient Descent, GD) est l'un des algorithmes d'optimisation les plus utilisés [167, 164]. L'algorithme ajuste de manière itérative les paramètres de la transformation en fonction du gradient

de la fonction objectif. Sa principale limite est sa faible vitesse de convergence, notamment lorsque l'espace de recherche est complexe. Le Gradient Conjugué est similaire à la GD mais accélère sa convergence en prenant en compte les directions précédentes. Le Descente de Gradient Stochastique (Stochastic Gradient Descent, SGD) est une variante du GD où l'optimisation a lieu sur un sous-ensemble aléatoire des données à chaque itération, plutôt que sur l'ensemble des données, améliorant grandement sa rapidité et son efficacité.

De nouveaux algorithmes d'optimisation plus efficaces ont été développés par la suite. L'algorithme Adam [158], très utilisé en DL, combine les avantages à la fois du SGD et de l'Adaptive Gradient Algorithm (AdaGrad) [181]. Il ajuste les taux d'apprentissage pour chaque paramètre afin de converger rapidement et de s'adapter à différents types de données. Très récemment, [182] adapte Adam pour son utilisation avec SyN, Greedy-SyN ainsi que pour le recalage rigide et affine, et réduit les temps de convergence.

Au-delà de l'algorithme d'optimisation, pour réduire le temps de calcul mais aussi pour guider progressivement le recalage, la majorité des méthodes [183] utilisent une stratégie de résolution dite pyramidale, ou multi-échelle, où l'optimisation est effectuée séquentiellement sur des images de plus en plus résolues, les premières étant effectuées sur des versions sous-échantillonnées de l'image.

3.2.5 Validation des méthodes de recalage

La validation des méthodes de recalage constitue une tâche complexe en l'absence de vérités terrain fiables. Pour évaluer l'efficacité d'un algorithme de recalage, plusieurs stratégies sont couramment employées.

Une première approche consiste à réutiliser directement les fonctions de coût sur l'intensité présentées plus haut comme méthodes de validation. Ces fonctions de coût peuvent inclure des métriques telles que la NCC ou la MSE. Cependant, bien que ces quantités soient fréquemment utilisées pour valider les performances des algorithmes, elles peuvent facilement être mises en défaut. Obtenir de bons scores sur ces métriques ne garantit pas nécessairement un recalage correct [184].

Les fonctions sur les propriétés géométriques peuvent aussi être réutilisées comme par exemple l'utilisation de landmarks anatomiques pour évaluer la précision du recalage. Ces landmarks, qui sont des points caractéristiques placés sur les images, servent de repères pour mesurer la correspondance spatiale entre les images avant et après recalage. Un grand nombre de points de repère est donc nécessaire pour obtenir une métrique de validation convenable.

Une autre approche repose sur l'utilisation de segmentations d'organes ou de structures anatomiques pour évaluer la qualité du recalage. En comparant les segmentations après le recalage avec des segmentations de référence, il est possible de quantifier la précision du recalage dans des régions spécifiques. Cela permet de détecter d'éventuelles déformations indésirables localisées. Pour valider le recalage de cette manière, il faut réutiliser les mesures de performance de la segmentation présentées dans le chapitre précédent.

En outre, l'analyse des propriétés de la déformation, telles que le jacobien de la déformation, constitue un autre critère de validation. Un jacobien négatif signale la présence de repliements dans l'image recalée, ce qui est souvent indésirable en imagerie médicale [167].

En résumé, la validation des méthodes de recalage repose sur l'utilisation de plusieurs critères, tels que la similarité des images, l'alignement des landmarks, la précision des segmentations et l'analyse des propriétés de la déformation. La combinaison de ces approches permet une évaluation exhaustive des performances des algorithmes de recalage et de leur fiabilité en contexte médical.

3.2.6 Etat de l'art des méthodes conventionnelles

Sans faire doublon avec les différents modèles de déformation présentés dans la section correspondante, cette section répertorie les algorithmes de recalage les plus couramment utilisés pour le recalage d'IRM cérébrales, tant pour les études de groupes que comme référence pour comparer et valider de nouvelles méthodes. Nous présentons ainsi les méthodes à ce jour majoritairement utilisées à des fins comparatives pour de nouvelles méthodes de recalage [185, 186, 187, 188, 189], notamment SyN, GreedySyN, NiftyReg et Symmetric Log Demons.

SyN [53] est un algorithme de recalage déformable implémenté dans le logi-

ciel ANTs [190]. SyN produit un recalage à la fois symétrique et difféomorphique. La symétrie provient du fait que les deux images, M et F , sont considérées dans le processus d'optimisation de manière symétrique. Plus précisément, les deux images sont recalées selon deux champs de déformation ϕ_1 et ϕ_2 tels que le problème d'optimisation devienne :

$$\phi_1^*, \phi_2^* = \arg \min_{\phi_1, \phi_2} L_{sim}(F \circ \phi_2, M \circ \phi_1). \quad (3.8)$$

La méthode SyN utilise cette approche symétrique en calculant les champs de vitesse dépendant du temps et contraint ϕ_1 à être l'inverse de ϕ_2 . La comparaison a lieu à l'instant $t = 0.5$ considéré comme image intermédiaire : chaque image a été recalée l'une vers l'autre. Le champ de déformation ϕ est alors obtenu par composition $\phi = \phi_1 \circ \phi_2^{-1}$.

La méthode Greedy, implémentée dans [191, 192], est une variante [193] de SyN où la partie symétrique est abandonnée afin d'obtenir des calculs plus rapides. NiftyReg [194] est une implémentation sur GPU (Graphics Processing Unit) d'une méthode de FFD avec des B-splines cubiques [164] dont le problème majeur était le coût computationnel. La méthode Symmetric Log Demons [54], implémentée dans ITK [195, 196], est une variante symétrique et difféomorphique des Demons. La méthode FireANTs [182], beaucoup plus récente, reprend les méthodes SyN et Greedy [191] avec une implémentation partielle sur GPU et l'utilisation d'un algorithme d'optimisation Adam [158].

Il existe d'autres méthodes dont deede [197, 198], DARTEL [63] dans SPM et adapté dans SPMmouse [39] ainsi que d'autres frameworks comme Elastix [199] dont la méthode principale de recalage déformable est équivalente à NiftyReg.

Les problèmes majeurs de ces méthodes, comme déjà présentés dans le chapitre 1, sont notamment leur application aux cerveaux de souris présentant d'autres contrastes et à différents atlas. Pour le temps de calcul, les implémentations sur GPU comme NiftyReg ou FireANTs ont permis de progresser en ce sens. FireANTs indique une amélioration moyenne d'un facteur 600 du temps d'exécution [182].

3.3 Méthodes de recalage d'images par apprentissage profond

Comme pour les méthodes de segmentation, les méthodes de recalage ont connu une révolution grâce à l'émergence des CNN [200] et en particulier de l'architecture U-Net [96] dans le domaine de l'imagerie médicale. Les CNN ont démontré une capacité remarquable à extraire des caractéristiques significatives des images, tandis que l'architecture U-Net s'est avérée particulièrement efficace pour l'imagerie médicale, notamment pour les tâches de segmentation puis pour le recalage d'images. Contrairement aux méthodes conventionnelles qui optimisent seulement sur une paire d'images pour effectuer le recalage, les CNN nécessitent un grand volume de données pour être entraînés. Afin d'éviter toute redondance avec le chapitre précédent, nous nous focalisons ici sur l'application de l'apprentissage profond aux tâches de recalage, en particulier au recalage déformable. Le choix de l'architecture U-Net comme modèle de base se justifie par son adoption majoritaire dans les méthodes récentes d'apprentissage profond.

D'un point de vue DL, le problème de recalage revient à prédire en sortie un champ de déformation ϕ avec en entrée une paire d'images (F, M) selon les paramètres θ du réseau :

$$\phi = f_{\theta}(F, M), \quad (3.9)$$

les paramètres θ du réseau étant optimisés au cours de l'apprentissage du réseau sur l'ensemble des images d'entraînement.

Déjà présentée en figure 2.2 du chapitre 2, l'architecture U-Net se compose d'un encodeur et d'un décodeur, lui conférant une structure en forme de U. L'encodeur est un réseau convolutionnel classique, composé d'applications répétées de convolutions, chacune suivie d'une unité linéaire rectifiée (ReLU) et d'une opération de *max-pooling*. Pendant cette partie, l'information spatiale est réduite tandis que l'information sur les caractéristiques est augmentée. Le décodeur combine ces deux informations à travers une séquence de convolutions inversées et de concaténations avec les caractéristiques haute résolution de l'encodeur pour reconstituer l'information spatiale.

3.3.1 Approche supervisée

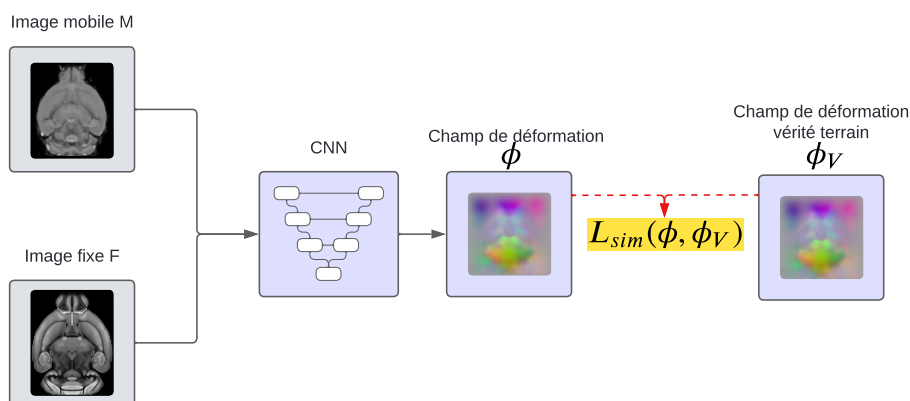


FIGURE 3.2 – Approche supervisée. L'entraînement se fait en comparant directement la sortie du modèle ϕ aux vérités terrain ϕ_V .

Pour les approches supervisées, les vérités terrain doivent être fournies. Le pipeline de l'approche supervisée pour le recalage est présentée en figure 3.2. En pratique, elles peuvent être obtenues grâce à des méthodes conventionnelles, comme chez [201], ou bien grâce à des simulations comme par [202]. Il est alors classique de comparer la sortie du modèle ϕ avec ces vérités terrain ϕ_V telles que la fonction coût devienne :

$$L = L_{sim}(\phi, \phi_V), \quad (3.10)$$

où L_{sim} est une fonction de dissimilarité basée sur la distance terme à terme. Les approches supervisées n'ont pas besoin de fonction de dissimilarité basée sur les intensités ou sur les propriétés géométriques comme les méthodes conventionnelles, car tout provient de la qualité des vérités terrain. En réalité, ces méthodes supervisées reproduisent des comportements des méthodes conventionnelles, d'où viennent les vérités terrain, et sont donc en deçà de leur performance en matière de qualité de recalage. De plus, l'obtention de vérités terrain pour les champs de déformation est la plupart du temps impossible ou alors extrêmement fastidieuse, comme l'était le recalage manuel.

3.3.2 Approche non supervisée

Les méthodes supervisées ont été peu à peu abandonnées dû au temps nécessaire à l'obtention des vérités terrain, que ce soit manuellement ou par méthode conventionnelle, pour laisser place aux approches non supervisées.

Le principe de l'approche non-supervisée est de s'affranchir des vérités terrain. En sortie du modèle, le champ de déformation ϕ est utilisé pour obtenir l'image recalée. Classiquement, en DL, les fonctions de coût sont calculées sur les sorties du modèle. Or, les critères de similarité vus pour les approches conventionnelles sont majoritairement calculés entre les images recalées et les images de référence, et non directement sur le champ de déformation.

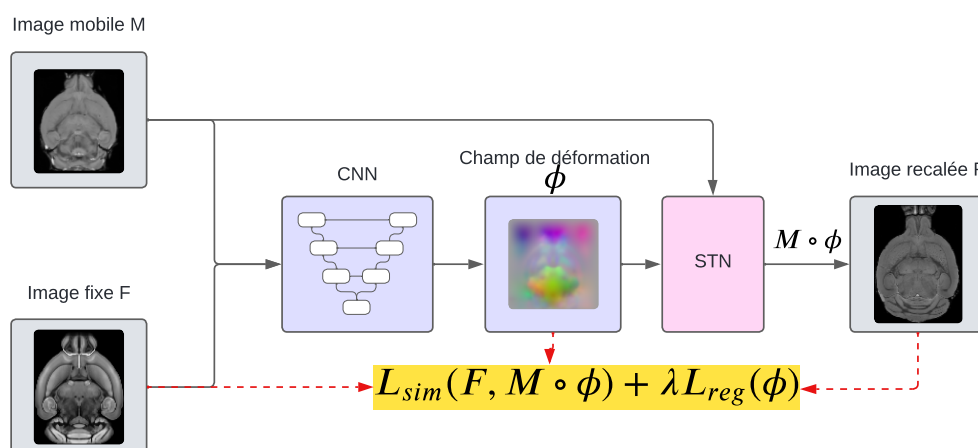


FIGURE 3.3 – Approche non-supervisée. L'image R est obtenue grâce au STN (Spatial Transformer Network). Un terme de régularisation contrôlé par λ est ajouté à la fonction de dissimilarité L_{sim} . L'entraînement se fait en comparant les paires d'images (F, R) de manière assez similaire aux méthodes conventionnelles.

Afin de permettre cette comparaison, la majorité des architectures actuelles intègrent un réseau supplémentaire appelé STN [203] qui applique le champ de déformation en sortie du réseau à l'image source pour obtenir l'image recalée directement au sein du réseau. Comme le STN est différentiable, il est utilisé pour rétropropager les critères de similarité calculés entre l'image recalée et l'image de référence sur les paramètres θ du réseau. Ainsi, les fonctions de

dissimilarité établies par les méthodes conventionnelles peuvent être réutilisées pour guider l'apprentissage du réseau. Comme pour ces dernières, un terme de régularisation est ajouté à cette fonction objectif pour les mêmes raisons que pour les méthodes conventionnelles. Par contre, la fonction de dissimilarité est optimisée sur l'ensemble du jeu de données plutôt que sur une paire d'images. Mathématiquement, on retrouve :

$$L = L_{sim}(F, M \circ \phi) + \lambda L_{reg}(\phi). \quad (3.11)$$

Cette approche est devenue l'approche classique en apprentissage pour le recalage. Les mêmes questionnements que pour les approches conventionnelles sur les fonctions de dissimilarité et les termes de régularisation sont repris. Ce pipeline classique est illustré en figure 3.3. Beaucoup utilisent les fonctions vues précédemment comme NCC [204, 205], MSE [206], MI [207], SSIM (Structural Similarity Index Measure) [208] ou plusieurs d'entre-elles [209, 210]. Il en est de même pour le terme de régularisation où la majorité des méthodes actuelles pénalisent soit les dérivées premières (diffusion) soit les dérivées secondes (énergie de courbure) du champ de déformation.

Comme les méthodes conventionnelles, les méthodes DL n'aboutissent pas à des solutions difféomorphiques par défaut. En imitant les méthodes LDDMM ou Demons, le CNN donne en sortie un champ de vitesse puis, grâce au *Scaling and Squaring* [173], ce champ est numériquement intégré en un champ de déformation difféomorphique [211].

3.3.3 Approches hybrides

D'autres approches sont par ailleurs utilisées en recalage d'image médicale, combinant les approches décrites ci-dessus.

Les méthodes faiblement supervisées [212, 213] sont des approches qui utilisent un niveau d'annotation plus limité que les méthodes entièrement supervisées, mais qui nécessitent tout de même une certaine supervision pour guider le processus de recalage. Ces approches se contentent d'annotations moins précises et qui ne concernent pas forcément la sortie directe du réseau. Au lieu d'annoter manuellement chaque paire d'images avec des points de correspondance, il est

possible d'utiliser des annotations globales telles que des étiquettes de classe ou des contours grossiers pour guider le recalage. Elles offrent donc un compromis entre l'efficacité des méthodes entièrement supervisées et la flexibilité des méthodes non supervisées. Soient $M_{labels}, F_{labels}, R_{labels}$ les segmentations ou les landmarks de l'image source, de l'image fixe et de l'image recalée tels que $R_{labels} = M_{labels} \circ \phi$ et L_{labels} une nouvelle fonction de dissimilarité pour ces segmentations pondérée par un hyper-paramètre λ_{labels} , le problème se formule ainsi :

$$L = L_{sim}(F, R) + \lambda L_R(\phi) + \lambda_{labels} L_{labels}(F_{labels}, R_{labels}). \quad (3.12)$$

Les méthodes doublement supervisées, comme utilisées par [214], combinent les termes supervisés et les termes non-supervisés afin que le problème s'énonce ainsi :

$$L = L_{sim}(F, R) + \lambda L_R(\phi) + \lambda_{sim_\phi} L_{sim_\phi}(\phi, \phi_V). \quad (3.13)$$

Ces méthodes duales sont souvent entraînées de manière consécutive en apprenant d'abord la partie supervisée puis la partie non-supervisée en faisant varier les poids de chaque terme contenu dans la fonction coût.

3.3.4 État de l'art des méthodes de recalage par apprentissage profond pour l'imagerie médicale

Après l'adoption majoritaire de l'approche non supervisée pour le recalage, et compte tenu de la similarité croissante entre ces méthodes et les approches conventionnelles, les techniques d'apprentissage profond ont principalement évolué en implémentant les améliorations déjà développées pour les méthodes conventionnelles. Nous notons par exemple les fonctions de régularisation pénalisant le jacobien [215, 205], les approches symétriques [205], cycliques [216] ou difféomorphiques [217] avec des fonctions de coût différentes, des modèles affines puis déformables [213, 218, 219], les approches pyramidales [188, 220, 221].

Afin de se rapprocher encore davantage des méthodes conventionnelles, certains travaux ont proposé d'utiliser les CNN non pas comme des modèles d'apprentissage en soi, mais comme des outils de représentation du champ de

déformation [222, 223, 224]. En optimisant les paramètres du réseau seulement sur une paire d'images [225] pour mieux représenter les déformations individuelles, le CNN est employé comme un modèle de déformation de manière similaire à une méthode conventionnelle. L'objectif de cette représentation est aussi de supprimer la nécessité d'avoir un grand jeu de données spécifique pour l'entraînement, difficilement obtainable pour certaines applications, par exemple pour différentes espèces animales.

Exactement comme pour la segmentation, l'évolution des réseaux avec l'arrivée des Transformers [105] et Mamba [111] a aussi eu un impact sur le recalage avec [226, 227, 228, 219] pour les Transformers ou [229, 230] pour Mamba, améliorant le recalage au prix d'une architecture plus lourde. D'autres approches multi-tâches, combinant recalage et segmentation simultanés, permettent d'améliorer la qualité du recalage comme le montrent [204, 125, 126] en étendant l'approche faiblement supervisée. De plus, l'utilisation de GAN dans le recalage d'image, entre autres [231, 232], a ouvert de nouvelles perspectives en générant des transformations réalistes entre les images de référence et cibles, avec des applications potentielles dans la création de données synthétiques pour l'entraînement des modèles.

Des inventaires complets des méthodes conventionnelles et DL pour le recalage d'image peuvent être consultés dans [233, 234] ou [4].

Pour finir sur les limites des méthodes DL, malgré des performances améliorées constatées par les articles les implémentant, d'autres articles sont plus critiques et n'indiquent aucune amélioration notable sur certains jeux de données lors de leur ré-implémentation par rapport aux méthodes classiques [185]. Ces affirmations sont par ailleurs soutenues par les performances des méthodes classiques lors de challenges de recalage comme par exemple [235] ou spécifiquement pour les souris le challenge RnR-ExM [236] avec FireANTS [182] et bigstream [237] comme méthodes les plus performantes.

3.4 Conclusion

Après avoir formalisé et présenté parallèlement les méthodes conventionnelles et celles basées sur l'apprentissage profond, il apparaît que ces deux

familles convergent progressivement. Les améliorations apportées au cours du temps pour les méthodes conventionnelles se retrouvent implémentées aujourd'hui dans les méthodes par apprentissage profond depuis l'adoption majoritaire de l'approche non supervisée. La non-supervision supprime la dépendance aux annotations, souvent manquantes pour ce type de données. Les différences principales résident dans l'utilisation des CNN comme modèle de déformation ou dans la stratégie d'apprentissage, mais les problèmes imposés par les fonctions de coût ou les régularisations sont les mêmes.

Une différence notable apportée par les méthodes par apprentissage profond est la possibilité d'utiliser des stratégies d'apprentissage faiblement supervisées, apportant des performances accrues en comparaison aux méthodes classiques seules et aux méthodes de recalage non supervisées [185]. Ce même article affirme que les changements d'architecture n'ont pas d'impact significatif sur les performances du recalage. Il met également en évidence l'incapacité des méthodes basées DL à s'adapter à différents jeux de données, même lorsqu'il s'agit d'une anatomie identique. Dans notre contexte spécifique de modèles murins, l'absence de grands jeux de données pour l'entraînement constitue un obstacle majeur. Bien que le jeu de données IRIS ait été conçu pour pallier ce manque, les méthodes d'apprentissage profond rencontrent encore des difficultés à généraliser lorsqu'un changement de domaine survient.

À part pour les nouvelles implémentations GPU de méthodes classiques, les temps d'exécution des méthodes restent longs. Pour les méthodes basées sur l'apprentissage profond au sens strict, il est souvent nécessaire de considérer le temps d'entraînement du modèle, surtout pour une application aux cerveaux de souris où l'accès à des jeux de données publics est restreint. Pour les méthodes DL utilisant un CNN comme représentation de la déformation, un temps considérable est requis pour optimiser les nombreux paramètres du réseau. Cette limite est exacerbée par l'utilisation de nouvelles architectures possédant de plus en plus de paramètres et occupant de plus en plus de mémoire sur GPU.

Le chapitre suivant fait ainsi la synthèse de ces observations en proposant une méthode de recalage adaptée à notre cadre d'études de groupes, où nous rappelons que l'objectif est de recaler un groupe d'images vers un atlas, pour des modèles murins.

CHAPITRE 4

Recalage déformable dans le cadre d'études de groupe

Sommaire

4.1 Introduction	78
4.2 Travaux relatifs	79
4.3 Recalage d'images pour l'étude de groupes	81
4.3.1 Formulation globale du recalage de groupes	81
4.3.2 Paradigmes d'entraînement ou d'optimisation	82
4.3.2.1 Optimisation pour étude de groupes	82
4.3.2.2 Entraînement sur jeu de données externe	82
4.3.2.3 Entraînement k-fold	83
4.3.2.4 Optimisation par paire	83
4.3.2.5 Paramétrisation du champ de déformation	84
4.4 Expériences et validation	85
4.4.1 Jeu de données	85
4.4.2 Méthodes	86

4.4.3 Implémentation	88
4.4.4 Métriques d'évaluation	89
4.5 Résultats et discussions	89
4.6 Conclusion	94

Avant d'aborder ce chapitre, il est important de noter que nous limitons le cadre au recalage déformable. L'étude de groupes, via l'analyse VBM, repose surtout sur les variations locales du cerveau qui sont représentées par le recalage déformable. Nous ne nions pas les difficultés pour obtenir un recalage affine de qualité, surtout dans le contexte des modèles murins, ni son impact sur l'analyse VBM. Cependant, nous considérons que, tant que les erreurs associées au recalage affine restent limitées, elles peuvent être compensées par le recalage déformable. Le recalage affine est donc considéré comme un prétraitement, comme indiqué lors de la présentation des jeux de données. De plus, le cadre préclinique n'impose pas de contrainte de temps stricte.

4.1 Introduction

À la suite des conclusions amenées par l'état de l'art des méthodes de recalage d'images médicales, ce chapitre propose une formulation globale du problème de recalage en regroupant les méthodes conventionnelles et celles basées sur l'apprentissage profond. Cette approche nous permet d'introduire différents paradigmes d'entraînement (ou d'optimisation selon le point de vue) et nous en présentons un spécifique au recalage d'images pour l'étude de groupes.

Dans ce paradigme, le champ de déformation est modélisé par un CNN et est directement optimisé sur les images à recaler, nous affranchissant d'utiliser des données d'entraînement. Ce paradigme est nommé **OP** pour Optimisation et relie, selon nous, les méthodes DL aux méthodes conventionnelles. Cette approche offre une meilleure qualité des déformations grâce à l'information partagée du recalage de groupe tout en conservant un temps d'exécution plus court que les méthodes conventionnelles. Nous la comparons aux autres paradigmes d'entraînement utilisés pour les méthodes par apprentissage profond et nous

l'évaluons sur le jeu de données GIN. Ce travail a été accepté à la conférence ISBI 2025.

4.2 Travaux relatifs

Le framework Deep Image Prior [238] a récemment été adapté au recalage. Cette approche utilise un CNN comme moyen de représenter le champ de déformation et les paramètres sont directement optimisés sur la paire d'images à recaler, supprimant la nécessité de l'entraînement préalable. Dans [239], cette technique est implémentée en remplaçant les images d'entrée du réseau par du bruit blanc. [225] propose une variante où seule l'image source est fournie en entrée du réseau. Il est aussi possible de représenter la déformation, non pas par un CNN, mais par une représentation neurale implicite [222] afin de passer à une représentation continue du champ de déformation plutôt que discrète.

Pour résoudre le problème des données d'entraînement, d'autres solutions ont été proposées. Les approches *One-shot domain adaptation* et *One-shot learning from scratch* [240] optimisent, après un entraînement ou non, les paramètres du réseau sur la paire d'images avant d'effectuer le recalage. Plus spécifique aux recalages de groupes et plus proche du travail présenté dans ce chapitre, le recalage simultané du jeu de données a déjà été considéré. Le concept de *One-shot learning* a ainsi été adapté pour le recalage global d'un ensemble d'images, en empilant les images du jeu de données en une seule image 4D de grande dimension. Pour pallier la contrainte mémoire du GPU utilisée par cette image de grande dimension, ce concept est implémenté en utilisant de petits patches au lieu des images entières [241]. Une approche similaire consiste à sous-échantillonner l'image avant son entrée dans le réseau, puis à sur-échantillonner le champ de déformation en sortie afin d'éviter l'utilisation de patches [242]. Dans ces deux implémentations, la forte consommation de mémoire GPU imposée par l'image 4D limite la complexité des architectures utilisées.

L'approche *Few-shot learning* consiste, quant à elle, à entraîner un modèle sur une très faible quantité de données. Cette approche est implémentée en pré-entraînant le modèle, puis en l'optimisant en quelques itérations sur la paire d'images à recaler [243], ce qui réduit le temps de calcul des approches *One-shot*

précédemment citées, le pré-entraînement n'étant pas pris en compte. S'agissant d'un recalage d'image 3D+t, les images 3D sont aussi empilées en entrée du réseau. L'entraînement est donc tout de même effectué sur 50 images 3D, puis l'optimisation sur 10 images 3D.

D'autres concepts partagent l'idée d'entraîner les modèles sur les données de test lorsque l'apprentissage est non supervisé. Le *Test-Time Training* consiste à prévoir la généralisation d'un modèle à différents domaines [244] avec une approche auto-supervisée. Cela est réalisé en adaptant le modèle pendant la phase de test avant d'émettre une prédiction [245]. L'idée commune est d'entraîner le modèle sur des données de test non annotées pour éviter un biais. Ce concept est adapté au recalage en intégrant un modèle auto-supervisé s'optimisant à la fois sur les données d'entraînement et sur les données de test [246]. Le *Training in the Loop* ou *On-the-fly Guidance* [247, 248] est un modèle en deux étapes qui s'approche du paradigme présenté ici. La première est basée apprentissage profond où un réseau prédit un champ de déformation, tandis que la seconde optimise le champ de déformation émis par le réseau comme une méthode conventionnelle. Ce concept combine en série une méthode apprentissage profond et une méthode conventionnelle. Le concept d'*Instance Optimization* consiste à raffiner le champ de déformation obtenu pour chaque paire d'images en suivant une optimisation classique de méthode conventionnelle. Cette approche consiste plutôt à modifier les paramètres du modèle en amont pour chaque paire d'images, en optimisant ainsi le modèle spécifiquement pour chaque cas en s'appuyant sur un modèle pré-entraîné [249].

À notre connaissance, l'approche **OP** proposée ci-dessous pourrait être vue comme une approche *Few-shot learning from scratch* ou de *Group Instance Optimisation*. Ce choix nous évite d'utiliser des données d'entraînement, tire profit du cadre d'études de groupe et nous rapproche des méthodes conventionnelles, tout en étant similaire aux approches de l'état de l'art comme vu précédemment. Nous considérons donc le CNN comme un moyen de représenter plusieurs champs de déformation, contrairement à un unique champ de déformation pour les méthodes de recalage par paire, en étant une fonction de l'image à recaler. Ainsi, notre problème d'apprentissage supervisé devient un problème de recalage classique consistant en une optimisation. Pour optimiser au mieux les

paramètres de cette représentation, elle est directement effectuée sur les images à recaler, comme pour une méthode conventionnelle.

4.3 Recalage d'images pour l'étude de groupes

4.3.1 Formulation globale du recalage de groupes

Pour mieux illustrer la similarité entre les méthodes conventionnelles et les méthodes basées sur l'apprentissage profond, il est nécessaire de reformuler le recalage d'une manière commune.

En reprenant les notations du chapitre 3 et en notant un jeu de données de paires d'images à recaler $D_T = \{(F_i, M_i)\}$ de taille N_T , le problème de recalage de groupe peut se formuler comme :

$$\theta^* = \arg \min_{\theta} \sum_i^{N_T} L_{sim}(F_i, M_i \circ \phi_{\theta}(F_i, M_i)) + \lambda L_{reg}(\phi_{\theta}(F_i, M_i)). \quad (4.1)$$

Déjà formulée de cette manière dans [185], elle est plus générale et contient les deux formulations selon que :

$$\phi_{\theta}(F_i, M_i) = \begin{cases} f_{\theta}(F_i, M_i), & \text{pour les réseaux de neurones,} \\ \phi_i, & \text{pour les méthodes conventionnelles,} \end{cases} \quad (4.2)$$

et les paramètres θ sont, dans un cas, les paramètres du réseau de neurones et, dans le cas des méthodes conventionnelles, l'ensemble des paramètres indépendants de chaque modèle de déformation optimisé tels que $\theta = \cup_i \{\phi_i\}$. Pour notre cas particulier d'études de groupes, toutes les images fixes F_i sont égales à l'atlas, qui peut être considéré comme un hyper-paramètre avec λ :

$$\theta^* = \arg \min_{\theta} \sum_i^{N_T} L_{sim}(F, M_i \circ \phi_{\theta}(F, M_i)) + \lambda L_{reg}(\phi_{\theta}(F, M_i)). \quad (4.3)$$

4.3.2 Paradigmes d'entraînement ou d'optimisation

Plusieurs paradigmes d'entraînement ou d'optimisation sont alors envisageables, les termes entraînement et optimisation étant interchangeables. Nous présentons en premier lieu le paradigme utilisé pour notre méthode de recalage avant de formuler les autres paradigmes considérés avec lesquels nous comparons notre méthode. Les 4 paradigmes majeurs sont schématisés sur la figure 4.1

4.3.2.1 Optimisation pour étude de groupes

Dans notre cas d'étude de groupes où l'on souhaite recalculer directement les N images du jeu de données D de taille limitée par rapport à un jeu de données d'entraînement, nous introduisons le paradigme suivant :

$$\theta^* = \arg \min_{\theta} \sum_i^N L_{sim}(F, M_i \circ \phi_{\theta}(F, M_i)) + \lambda L_{reg}(\phi_{\theta}(F, M_i)). \quad (4.4)$$

Ce paradigme sous-entend que le recalage de toutes les paires d'images améliore le recalage global du groupe d'images. Les N champs de déformation sont directement représentés par un unique CNN. Ce paradigme est dénommé **OP** par la suite. Il présente selon nous les avantages suivants : il ne nécessite pas d'annotations, il ne nécessite pas de nombreuses données d'entraînement car il s'adapte directement aux images d'intérêt pour une contrepartie en temps de calcul. Pour améliorer la rapidité de l'optimisation, nous pouvons utiliser un modèle DL pré-entraîné avant d'optimiser le modèle sur les images du groupe. Nous optimisons le modèle **DL-IRIS**, présenté ci-dessous, directement sur GIN avec l'approche **OP** pour obtenir le modèle **DL+OP**, de manière similaire à [249] mais adaptée pour l'étude de groupes. Ici, **DL-IRIS** correspond au modèle pré-entraîné et **DL+OP** au modèle optimisé.

4.3.2.2 Entraînement sur jeu de données externe

Le paradigme d'entraînement d'apprentissage profond classique où le jeu de données d'entraînement D_T de taille N_T sur lequel sont optimisés les paramètres

θ est différent de l'ensemble D que l'on souhaite recaler se formule :

$$\theta^* = \arg \min_{\theta} \sum_i^{N_T} L_{sim}(F, M_i \circ \phi_{\theta}(F, M_i)) + \lambda L_{reg}(\phi_{\theta}(F, M_i)), \quad (4.5)$$

puis l'unique modèle est évalué sur chaque paire d'images du jeu de données D . Généralement, la taille N_T est très grande pour tendre vers la généralisation du modèle. Dans notre cas d'utilisation, le jeu de données utilisé pour D_T est IRIS en raison de son volume et de son hétérogénéité. Ce paradigme est dénommé **DL-IRIS** par la suite.

4.3.2.3 Entraînement k-fold

Supposons le jeu de données à recaler D subdivisé en k partitions D_j de taille N_j . Le paradigme d'entraînement d'apprentissage profond "k-fold" où k modèles sont entraînés tels que le j -ième modèle soit entraîné sur $k - 1$ subdivisions D_n de taille N_n avec $n \neq j$ se formule :

$$\theta_j^* = \arg \min_{\theta_j} \sum_{\substack{n=1 \\ n \neq j}}^k \sum_{i=1}^{N_n} L_{sim}(F, M_i \circ \phi_{\theta_j}(F, M_i)) + \lambda L_{reg}(\phi_{\theta_j}(F, M_i)). \quad (4.6)$$

Les k modèles sont alors utilisés pour obtenir le recalage entier du jeu de données D , le j -ème modèle étant évalué sur D_j . Dans les méthodes par apprentissage, les valeurs $k = 3$, $k = 5$ ou $k = 10$ sont fréquemment utilisées. Dans un souci de temps de calcul et pour avoir des jeux de données de test tout de même conséquents, nous choisissons $k = 3$ et ainsi ce paradigme est dénommé **DL-3fold** par la suite.

4.3.2.4 Optimisation par paire

Chaque champ de déformation est directement optimisé sur une paire d'images du jeu de données D indépendamment des autres paires. Ainsi, pour la j -ème paire du jeu de données D :

$$\theta_j^* = \arg \min_{\theta_j} L_{sim}(F, M_j \circ \phi_{\theta_j}(F, M_j)) + \lambda L_{reg}(\phi_{\theta_j}(F, M_j)). \quad (4.7)$$

Les N modèles indépendants sont alors utilisés pour obtenir le recalage entier du jeu de données D . Il s'agit du paradigme utilisé par les méthodes effectuant des recalages par paire en représentant le modèle de déformation par un CNN comme dans [225] ou dans [222]. Ce paradigme est dénommé **Pair** par la suite.

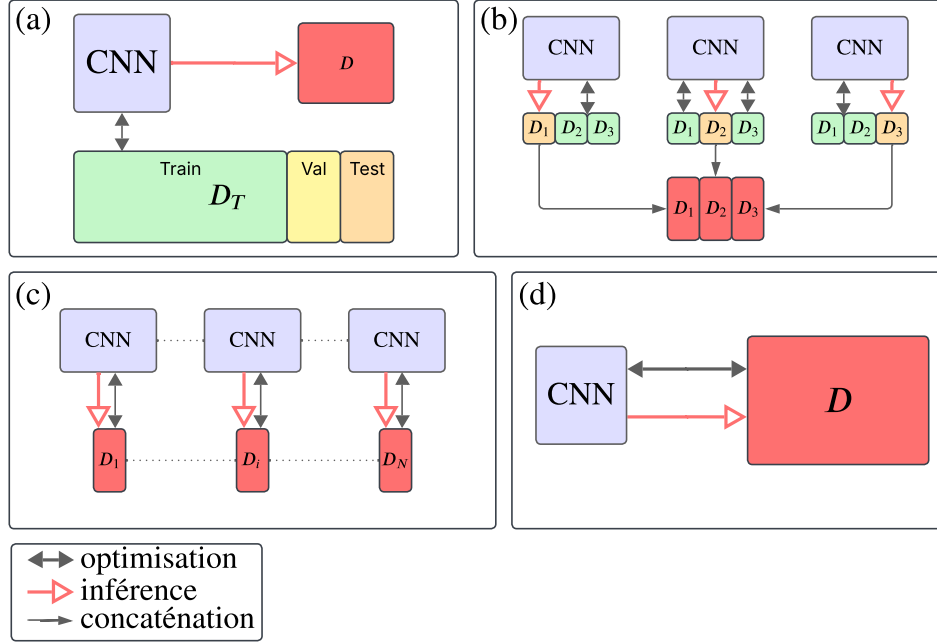


FIGURE 4.1 – Schéma présentant les 4 paradigmes majeurs d'entraînement ou d'optimisation. (a) correspond à **DL-IRIS**, (b) correspond à **DL-3fold**, (c) correspond à **Pair** et (d) correspond à **OP**.

4.3.2.5 Paramétrisation du champ de déformation

Dans l'optique d'évaluer l'impact de la paramétrisation du champ de déformation par un CNN, nous proposons d'ajouter les deux expérimentations suivantes :

Effet de la paramétrisation du modèle par une paire d'images Pour analyser l'influence de donner en entrée du réseau la paire d'images à recalier, nous introduisons ce paradigme. Il reprend donc le paradigme **Pair** en supprimant l'entrée du modèle telle que :

$$\phi_{\theta_i}(F, M_i) = f_{\theta_i},$$

où la fonction f du CNN ne prend plus F et M_i en entrée. Pour compenser, au lieu de donner en entrée du bruit [239, 225], nous donnons des images unitaires constantes à l'entrée du réseau, indépendantes de F et M_i .

$$\underset{\theta_i}{\operatorname{argmin}} L_{sim}(F, M_i \circ f_{\theta_i}) + \lambda L_{reg}(f_{\theta_i}). \quad (4.8)$$

Ce paradigme est dénommé **Pair-Zero**.

Effet de la paramétrisation du champ par un CNN Ce sous-paradigme de **Pair** retire complètement l'effet du réseau et cherche à représenter directement le champ de déformation par ses paramètres comme un modèle non paramétrique [225]. Ainsi, $\phi_{\theta_i} = \phi_i$ et le paradigme se formule alors :

$$\underset{\phi_i}{\operatorname{argmin}} L_{sim}(F, M_i \circ \phi_i) + \lambda L_{reg}(\phi_i). \quad (4.9)$$

Il s'agit en fait du paradigme d'optimisation des méthodes conventionnelles. Ce paradigme est dénommé **Pair-Null** et est équivalent aux méthodes conventionnelles, modulo les optimisations propres à chacune.

4.4 Expériences et validation

4.4.1 Jeu de données

Nous utilisons l'atlas Allen comme cible du recalage. Hormis **DL-IRIS**, toutes les méthodes sont optimisées/entraînées sur le jeu de données GIN qui présente l'avantage de disposer de cartes de segmentation précises, qui sont exploitées pour mesurer quantitativement les performances des méthodes. Comme indiqué en début de chapitre, le jeu de données est aussi pré-initialisé avec un recalage rigide et affine pour n'étudier que l'effet du recalage déformable.

TABLEAU 4.1 – Tableau récapitulatif des méthodes implémentées.

Méthodes	Équation	Architecture	Type	Entrées
SyN	4.9	Difféomorphisme symétrique	Pair	/
DL-IRIS	4.5	LocalNet	Apprentissage	Source & Atlas
DL-3fold	4.6	LocalNet	Apprentissage	Source & Atlas
Pair	4.7	LocalNet	Pair	Source & Atlas
Pair-Zero	4.8	LocalNet	Pair	Image constante
Pair-Null	4.9	DDF paramétrique	Pair	/
OP	4.4	LocalNet	Groupe	Source & Atlas
DL+OP	4.5 puis 4.4	LocalNet	Apprentissage puis groupe	Source & Atlas

4.4.2 Méthodes

En plus des paradigmes présentés ci-dessus, nous implémentons aussi la méthode traditionnelle **SyN**. Déjà présentée, il s'agit de l'une des méthodes conventionnelles les plus utilisées pour le recalage déformable en imagerie médicale. Cette méthode est implémentée dans ANTs et est exécutée avec les paramètres proposés par [250], utilisant la LNCC, trois résolutions pour une approche multi-échelle et un pas d'optimisation de 0,15. Les méthodes/paradigmes implémentés sont présentés en figure 4.1.

Il s'avère que la méthode **SyN** a beaucoup de difficultés à recalcr les images du jeu de données sur le template Allen. Ces difficultés proviennent surtout de la présence de singularités sur l'atlas, comme plusieurs troncatures aux extrémités du cerveau. En pratique, pour les études de groupes, l'algorithme *antsMultivariateTemplateConstruction2* de ANTs est préféré à l'utilisation basique de **SyN**. Cet algorithme crée, simultanément au recalage sur l'atlas Allen, un nouveau template provenant des images du jeu de données. Cependant, ce processus ne garantit pas forcément que le template généré soit aligné avec l'atlas fourni. L'utilisation de cet algorithme s'avère donc problématique. Pour remplacer l'atlas Allen, un nouveau template a été calculé *a posteriori* grâce aux images recalées obtenues avec **DL-IRIS**. Présentant un biais, cette approche nommée **SyN-DL** sert uniquement d'illustration aux performances de **SyN** lorsque le template utilisé n'est pas problématique.

En raison des différences entre les images et l'atlas, la fonction de coût employée est la variante locale de la NCC (LNCC). Pour tendre vers la régularité du champ [205], la fonction de régularisation suivante est choisie :

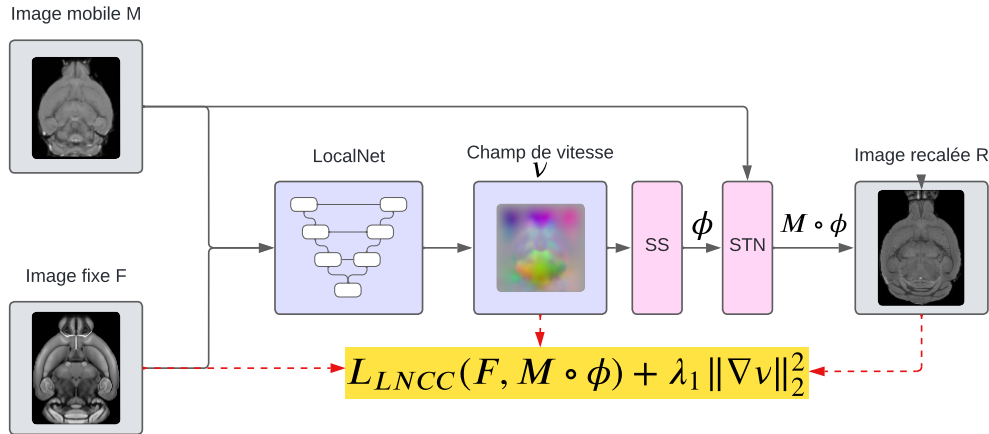


FIGURE 4.2 – Aperçu de la méthode de recalage direct proposée. Le bloc SS correspond à l’algorithme *Scaling and Squaring* [173]. Le bloc STN correspond au *Spatial Transformer Network*.

$$L_{reg} = \|\nabla v\|_2^2, \quad (4.10)$$

avec $\|\cdot\|_2$ la norme euclidienne et ∇ le gradient spatial.

De plus, pour se rapprocher des méthodes conventionnelles et obtenir des champs de déformation plus réguliers, l’approche difféomorphique des SVF est implémentée. Un champ de vitesse est obtenu en sortie du réseau, puis est intégré en champ de déformation grâce à l’algorithme de *Scaling and Squaring* [173]. La fonction de coût totale devient alors :

$$L_{LNCC}(F, M \circ \phi) + \lambda_1 \|\nabla v\|_2^2. \quad (4.11)$$

Le pipeline de recalage appliqué pour toutes les méthodes, sauf **SyN**, est présenté en figure 4.2.

Pour améliorer le recalage de toutes les méthodes, tous les calculs de fonction de coût ont été limités à la zone d’intérêt définie par l’extraction du cerveau, obtenue par les vérités terrain du jeu de données GIN et de l’atlas Allen.

4.4.3 Implémentation

L'ensemble du pipeline a été implémenté en utilisant PyTorch [156] et en exploitant les architectures de modèles et les méthodes d'augmentation de données fournies par MONAI [157]. Pour les méthodes **DL-IRIS**, **DL-3fold** et **OP**, les images du jeu de données et de l'atlas ont été augmentées à l'aide de petites transformations affines aléatoires afin d'exploiter pleinement les capacités du CNN. La stratégie d'augmentation repose sur l'application aléatoire de chaque opération avec une probabilité de 0,5. Elle comprend une rotation sur chaque axe entre -2° et $+2^\circ$, une translation entre -2 et $+2$ voxels, une mise à l'échelle par axe d'un facteur entre 0,95 et 1,05 ainsi qu'un décalage global d'intensité d'une valeur de 0,2.

L'ordinateur utilisé pour mener toutes ces expérimentations est équipé d'un processeur Intel Xeon W-2225 (4.10GHz) et d'une carte graphique NVIDIA RTX 3080 Ti (12 Go VRAM). Dans le but d'utiliser une architecture performante et validée, le réseau LocalNet [212, 213], de type U-Net [96], implémenté sur MONAI, est utilisé avec une profondeur encodeur-décodeur de 4 blocs. Ce réseau a été entraîné en utilisant l'optimiseur Adam [158] avec des paramètres par défaut et un taux d'apprentissage de 10^{-4} . Un planificateur de taux d'apprentissage a également été utilisé pour diviser par deux le taux d'apprentissage après avoir atteint un plateau pendant 20 époques. Chaque époque comprend 27 batchs d'une seule image pour pouvoir être contenus sur les 12 Go de VRAM du GPU.

L'hyper-paramètre λ_1 , contrôlant l'importance de la régularisation, a été testé pour différentes valeurs de λ_1 (0.001, 0.01, 0.1, 1, 10, 100, 1000) à la fois pour les modèles **OP** et les modèles **Pair**. Il s'avère que les meilleures performances de ces modèles sont atteintes simultanément lorsque λ_1 prend la valeur 1.0. Ainsi, λ_1 prendra pour valeur 1.0 dans le reste des expériences.

Pour chaque méthode, hormis **SyN**, le modèle final a été sélectionné en se basant sur l'itération ayant atteint la valeur de fonction de coût la plus basse. Pour les méthodes basées sur l'apprentissage, l'arrêt de l'entraînement s'est fait après convergence. Pour les méthodes basées sur l'optimisation, nous avons arbitrairement choisi de l'arrêter après 1000 époques, soit environ deux fois plus longtemps qu'un entraînement de deep learning classique.

Concernant la méthode **DL-3fold**, chaque cohorte respecte le même nombre d'individus par sous-division du jeu de données pour ne pas introduire de biais. Pour la méthode **DL-IRIS**, le modèle est entraîné sur le jeu de données IRIS et le modèle choisi est le modèle atteignant la valeur la plus basse pour la fonction de coût sur l'ensemble de validation.

Pour illustrer la rapidité du recalage obtenue grâce à l'optimisation d'un modèle pré-entraîné, le modèle **DL+OP** est optimisé sur seulement 10 époques d'optimisation.

4.4.4 Métriques d'évaluation

Étant donné que le problème est non supervisé, l'absence de vérité terrain pour l'image déformée ou pour le champ de déformation nous contraint à utiliser différentes métriques pour évaluer notre modèle.

Le jeu de données GIN étant segmenté en 10 structures, nous utilisons la métrique de Dice pour évaluer les performances du recalage. Cette segmentation est ensuite transportée sur l'atlas Allen grâce au recalage. L'atlas Allen est segmenté de manière beaucoup plus fine que GIN. Grâce à sa segmentation hiérarchique, il nous est possible de retrouver les 10 structures de GIN sur la segmentation de l'atlas en regroupant certaines sous-régions. Cependant, cela crée un biais car les deux segmentations n'ont pas été obtenues par les mêmes experts. Le score de Dice est ainsi calculé pour les 10 régions entre les segmentations de l'atlas et les segmentations des individus transportées sur l'atlas.

Les temps d'inférence et d'entraînement ont aussi été calculés pour toutes les méthodes considérées et sur le jeu de données entier pour pouvoir être comparés.

4.5 Résultats et discussions

L'ensemble des résultats sur les 10 régions est présenté dans le tableau 4.2 et le score moyen du coefficient de Dice sur 9 zones d'intérêt (Region Of Interest, ROI) pour les différentes méthodes est présenté dans le tableau 4.3. De manière surprenante, **SyN** affiche des performances médiocres par rapport à son utilisation générale en imagerie médicale. Cette différence s'explique principalement

TABLEAU 4.2 – Scores de Dice en pourcentage pour les 9 ROI de GIN. Une erreur de segmentation sur MAP6 a conduit à la suppression de la 10ème ROI sur l'ensemble de GIN. Les meilleurs scores par colonne sont indiqués en gras.

Méthodes	CB	CTX	GP
Affine	85.68 $\pm$ 3.36	88.70 $\pm$ 1.51	52.34 $\pm$ 15.4
SyN	85.55 $\pm$ 3.67	88.31 $\pm$ 1.02	56.60 $\pm$ 9.85
SyN-DL	90.43 $\pm$ 1.44	91.33 $\pm$ 0.62	66.18 $\pm$ 8.17
DL-IRIS	89.55 $\pm$ 1.55	89.81 $\pm$ 0.64	68.14 $\pm$ 6.53
DL-3fold	89.61 $\pm$ 1.55	88.68 $\pm$ 0.72	62.71 $\pm$ 11.0
Pair-Null	87.34 $\pm$ 3.00	89.70 $\pm$ 1.00	63.70 $\pm$ 8.79
Pair-Zero	90.31 $\pm$ 1.44	88.42 $\pm$ 1.32	69.24 $\pm$ 5.37
Pair	90.43 $\pm$ 1.51	89.28 $\pm$ 1.20	68.08 $\pm$ 5.47
OP	90.09 $\pm$ 1.46	88.87 $\pm$ 0.79	66.01 $\pm$ 6.41
DL+OP	90.15 $\pm$ 1.22	88.54 $\pm$ 0.66	66.24 $\pm$ 6.19
Méthodes	HPF	HY	MB
Affine	86.30 $\pm$ 1.72	80.67 $\pm$ 2.93	73.72 $\pm$ 5.20
SyN	80.85 $\pm$ 2.74	72.03 $\pm$ 5.25	54.51 $\pm$ 8.68
SyN-DL	89.23 $\pm$ 1.40	81.55 $\pm$ 2.28	76.62 $\pm$ 2.88
DL-IRIS	89.00 $\pm$ 1.54	82.00 $\pm$ 2.43	76.12 $\pm$ 3.21
DL-3fold	89.06 $\pm$ 1.73	81.69 $\pm$ 2.45	74.69 $\pm$ 3.42
Pair-Null	89.45 $\pm$ 1.56	81.83 $\pm$ 2.55	73.59 $\pm$ 4.10
Pair-Zero	88.56 $\pm$ 1.69	81.15 $\pm$ 2.52	75.21 $\pm$ 2.62
Pair	89.09 $\pm$ 1.62	82.12 $\pm$ 2.17	75.09 $\pm$ 2.62
OP	90.18 $\pm$ 1.16	82.53 $\pm$ 2.16	75.15 $\pm$ 2.79
DL+OP	89.87 $\pm$ 1.38	82.26 $\pm$ 2.07	74.86 $\pm$ 2.66
Méthodes	OLF	STR	TH
Affine	60.25 $\pm$ 4.50	83.94 $\pm$ 2.46	85.26 $\pm$ 1.51
SyN	66.92 $\pm$ 7.19	85.61 $\pm$ 2.83	77.58 $\pm$ 4.37
SyN-DL	67.93 $\pm$ 1.74	88.14 $\pm$ 2.02	87.17 $\pm$ 1.87
DL-IRIS	67.13 $\pm$ 4.41	88.55 $\pm$ 1.95	87.05 $\pm$ 1.66
DL-3fold	77.59 $\pm$ 2.52	88.16 $\pm$ 2.13	87.06 $\pm$ 1.91
Pair-Null	64.82 $\pm$ 4.75	87.88 $\pm$ 2.04	86.45 $\pm$ 1.40
Pair-Zero	73.01 $\pm$ 5.36	88.24 $\pm$ 2.26	86.77 $\pm$ 1.87
Pair	74.83 $\pm$ 3.91	88.74 $\pm$ 2.05	87.47 $\pm$ 1.70
OP	78.43 $\pm$ 2.00	88.68 $\pm$ 1.77	87.59 $\pm$ 1.58
DL+OP	76.90 $\pm$ 2.47	88.23 $\pm$ 1.85	87.35 $\pm$ 1.57

TABLEAU 4.3 – Score de Dice moyen et temps total de recalage des méthodes comparées sur le jeu de données GIN (9 ROI). Les meilleurs scores par colonne sont indiqués en gras.

Méthodes	Dice (%) $\uparrow$	Temps (min) $\downarrow$
Affine	77.43 ± 4.29	/
SyN	78.95 ± 3.51	142
SyN-DL	82.06 ± 2.49	/
DL-IRIS	81.93 ± 2.66	0
DL-3fold	82.14 ± 3.05	540
Pair-Null	80.53 ± 3.24	278
Pair-Zero	82.32 ± 2.72	452
Pair	82.79 ± 2.47	464
OP	83.06 ± 2.24	256
DL+OP	82.71 ± 2.23	8

par le fait que l’atlas contient uniquement des tissus cérébraux, tandis que les IRM incluent également des tissus extra-cérébraux, ce qui introduit des effets de bord dans la zone corticale. De plus, l’atlas choisi, sans aucune modification, est tronqué de manière nette sur un des bords du cerveau, ce qui accentue encore plus les effets de bord. Comme brièvement expliqué précédemment, pour pallier cette limitation, nous avons utilisé l’image recalée moyenne obtenue avec la méthode **DL-IRIS** comme template et avons relancé la méthode **SyN**. Ces résultats sont désignés sous le nom de **SyN-DL**. Cette approche obtient des résultats significativement meilleurs. Nous observons un comportement similaire en utilisant uniquement **Pair-Null**. L’ajout d’une architecture CNN, comme **Pair-Zero**, améliore le score moyen de Dice. En réintroduisant la paramétrisation des poids en fonction des entrées, la méthode **Pair** obtient les meilleurs résultats parmi les méthodes par paires. Notre approche **OP** surpasse toutes les autres, notamment **Pair**, en atteignant un score moyen de Dice de 83.06%.

Le temps requis pour entraîner ou optimiser et effectuer le recalage sur l’ensemble du jeu de données est indiqué dans le tableau 4.3. Comme nous effectuons un entraînement en 3-fold sur notre jeu de données, la méthode **DL-3fold** nécessite plus de temps que toutes les méthodes par paires ou globales. Au contraire,

la méthode **DL-IRIS** ne prend pas en compte le temps d'entraînement étant donné qu'on considère le modèle comme pré-entraîné. Pour recaler entièrement le jeu de données GIN, notre méthode **OP** est plus rapide que toutes les autres, sauf **SyN** et **DL-IRIS**. En utilisant **DL+OP**, le recalage du jeu de données est réalisé en moins de 10 minutes avec des performances proches de celles de **Pair** et nettement supérieures à celles de **DL-IRIS** dont la méthode est issue.

TABLEAU 4.4 – Score de Dice en pourcentage pour 3 ROI spécifiques : GP, HY et OLF représentent respectivement le globus pallidus, l'hypothalamus et les zones olfactives. Les meilleurs scores par colonne sont indiqués en gras.

Méthodes/Dice	GP ↑	HY ↑	OLF ↑
Affine	52.34 ± 15.4	80.67 ± 2.93	60.25 ± 4.50
SyN	60.12 ± 7.11	77.72 ± 4.48	73.73 ± 5.67
SyN-DL	66.18 ± 8.17	81.55 ± 2.28	67.93 ± 1.74
DL-IRIS	68.14 ± 6.53	82.00 ± 2.43	67.13 ± 4.41
DL-3fold	62.71 ± 11.0	81.69 ± 2.45	77.59 ± 2.52
Pair-Null	63.70 ± 8.79	81.83 ± 2.55	64.82 ± 4.75
Pair-Zero	69.24 ± 5.37	81.15 ± 2.52	73.01 ± 5.36
Pair	68.08 ± 5.47	82.12 ± 2.17	74.83 ± 3.91
OP	66.01 ± 6.41	82.53 ± 2.16	78.43 ± 2.00
DL+OP	66.24 ± 6.19	82.26 ± 2.07	76.90 ± 2.47

Nous avons décidé de porter une attention particulière à trois des dix structures anatomiques : le globus pallidus (GP), l'hypothalamus (HY) et les zones olfactives (OLF) (voir tableau 4.4). Le globus pallidus est une structure très petite et peu contrastée, tandis que l'hypothalamus et les zones olfactives sont des structures larges situées dans la partie inférieure du cerveau, caractérisées par un faible rapport signal/bruit provenant de la bobine de surface. Il est important de noter que les zones olfactives sont particulièrement sujettes aux problèmes de troncature de l'atlas, ce qui induit une variabilité significative. **OP** et **DL-3fold** surpassent les autres méthodes pour le recalage des zones olfactives grâce à la régularisation du jeu de données mentionnée précédemment, alors que les méthodes par paires donnent de mauvais résultats en raison de problèmes sur cette région chez certains individus du jeu de données. Cependant, **Pair** et **OP**

obtiennent de meilleurs résultats que les autres pour l'hypothalamus. Des conclusions relativement différentes sont tirées pour la structure du globus pallidus, qui est relativement petite (800 voxels) et particulièrement difficile à segmenter. Dans ce cas, les méthodes par paires obtiennent des résultats légèrement meilleurs que les autres.

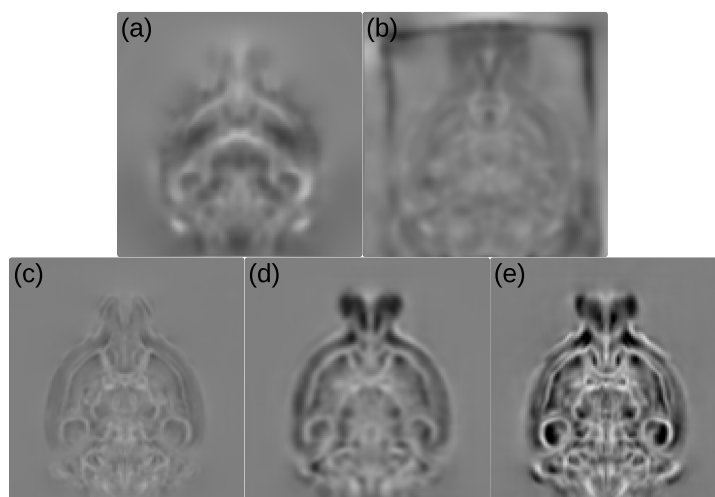


FIGURE 4.3 – Moyenne des log-Jacobiens des champs de déformation sur le jeu de données GIN. (a) **SyN**, (b) **SyN-DL**, (c) **DL-IRIS**, (d) **Pair**, (e) **OP**.

Les log-Jacobiens moyens des déformations sont présentés dans la figure 4.3. Les méthodes **OP** montrent des contrastes plus marqués dans les régions anatomiques que **Pair** et **DL-IRIS**, suggérant une meilleure qualité de recalage pour l'analyse VBM. Cependant, des valeurs de jacobien plus élevées, comparées à **SyN**, indiquent une expansion ou une compression plus importantes des structures. Cela implique que, malgré une meilleure performance du recalage, l'anatomie sous-jacente pourrait être trop modifiée. Afin de préserver la cohérence anatomique, une régularisation doit être mise en place pour maintenir les valeurs des jacobiens proches de celles des méthodes conventionnelles. Des conclusions numériques basées sur une étude de groupes restent à être menées pour confirmer ces observations.

4.6 Conclusion

Le paradigme d'entraînement **OP**, ainsi que sa variante combinée **DL+OP**, s'avère particulièrement adapté aux études de groupe. Il améliore les résultats du recalage au prix d'une optimisation de plusieurs heures. En étant non supervisé, il n'utilise pas de données annotées pour s'entraîner. En le considérant comme une méthode conventionnelle d'optimisation, il ne nécessite pas de grands jeux de données pour s'entraîner. En l'optimisant, non par paires d'images, mais sur le jeu de données entier, il s'adapte mieux au cadre du recalage de groupes.

À l'exception de certaines structures fines, cette approche produit de meilleurs résultats que les approches par paires ou par apprentissage profond conventionnelles. En l'optimisant à partir d'un modèle pré-entraîné sur la même tâche, il devient possible d'obtenir un recalage performant en moins d'une dizaine de minutes, comme le démontre le modèle **DL+OP**. Cependant, les jacobiens de la déformation présentent des différences notables avec les méthodes conventionnelles, s'expliquant par la représentation du champ de déformation par le CNN. Pour remédier à cela, nous envisageons de renforcer la régularisation de nos approches afin d'obtenir des jacobiens plus réguliers tout en conservant des recalages performants.

Toutefois, bien que cette approche réalise de meilleurs recalages de groupe, l'étude de groupes en elle-même nécessite plus d'informations que le simple recalage de chaque individu sur l'atlas. Dans certains cas, il est nécessaire d'obtenir le recalage de l'atlas sur les individus, le recalage inverse. Grâce à cela, il est possible de transférer les segmentations de l'atlas directement sur les individus, permettant par exemple de calculer les volumes de certaines structures anatomiques d'intérêt. Cette technique est connue sous le nom de segmentation par transfert d'atlas. Par ailleurs, les modifications nécessaires à l'obtention de ce recalage inverse régulariseront fortement nos champs de déformation. En obtenant le recalage inverse, nous pourrions ainsi comparer ces méthodes aux méthodes de segmentation spécifiques introduites dans le chapitre 2.

CHAPITRE 5

Améliorations pour le transfert d’atlas et l’inversibilité du champ

Sommaire

5.1 Introduction	96
5.2 Travaux relatifs	96
5.3 Méthode de recalage inverse	99
5.3.1 Métriques de validation	100
5.3.2 Expériences et validation	101
5.3.3 Résultats et discussions	101
5.4 Transfert d’atlas pour la segmentation des tissus cérébraux	106
5.5 Conclusion	107

5.1 Introduction

Afin d’obtenir des champs de déformation plus réguliers et de se rapprocher des jacobiens obtenus par les méthodes conventionnelles, nous souhaitons garantir l’inversibilité des recalages déformables produits par notre modèle. De plus, en générant simultanément les champs de déformation directs et inverses, nous nous affranchissons du besoin de segmenter les individus. L’obtention du recalage inverse permet le transfert de la segmentation fine de l’atlas Allen vers les cerveaux de souris à segmenter. Une solution aurait été de considérer l’approche du chapitre précédent et de l’appliquer pour recaler l’atlas sur chaque individu. Dans ce scénario, rien n’impose que le champ obtenu soit l’inverse de l’autre. En ajoutant une contrainte d’inversibilité, une certaine régularité est imposée aux champs de déformation obtenus, ce qui est pertinent pour l’analyse VBM. Pour ce faire, nous modifions l’architecture du modèle afin qu’il génère deux champs de déformation en sortie. La contrainte d’inversibilité évoquée est intégrée directement dans la fonction de coût. Cette amélioration, basée sur l’architecture et sur la fonction de coût, est appliquée aux trois méthodes **OP**, **Pair** et **DL-IRIS**.

Pour étudier l’effet des modifications apportées, les méthodes sont évaluées et comparées avec les méthodes initiales grâce au score de Dice et de ses variantes. En obtenant simultanément les recalages direct et inverse, ces approches permettent de transporter la segmentation de l’atlas sur chacun des sujets. Nous comparons ainsi pour l’ensemble des individus la qualité de la segmentation par transport d’atlas avec les méthodes de segmentation obtenues dans le chapitre 2.

5.2 Travaux relatifs

Obtenir un champ difféomorphique garantit des propriétés souhaitables comme la continuité, la différentiabilité ou la conservation de la topologie [172, 53, 54]. Parmi les méthodes conventionnelles les plus connues respectant ces propriétés, on retrouve les approches LDDMM, SyN et DARTEL. Adaptée au cadre du DL, cette propriété consiste à générer en sortie du modèle un champ de vitesse qui est ensuite intégré sur un certain pas de temps afin d’obtenir le

champ de déformation [251]. Elle permet aux méthodes DL de se rapprocher d'un recalage difféomorphique. Elle est aujourd'hui largement adoptée dans les travaux récents [209, 188, 252, 205, 223, 253], raison pour laquelle elle a déjà été implémentée dans le chapitre 4 grâce au *Scaling and Squaring*.

Pour obtenir des champs de déformation difféomorphiques, ils doivent respecter la propriété d'*inverse-consistency*. Cela consiste à recaler l'image source sur l'image cible, puis à recaler l'image cible sur l'image source pour obtenir l'inverse. Cette approche est implémentée avec un modèle capable de donner en sortie deux champs de déformation [254]. Un réseau est ensuite entraîné à inverser les champs de déformation en donnant en vérité terrain l'autre champ de déformation. Une fonction de coût symétrique basée sur la norme entre les quatre champs (deux pour le recalage direct et inverse et deux pour l'inverse des champs) est ajoutée pour entraîner ce réseau. Un décodeur supplémentaire est ajouté dans l'architecture du modèle pour obtenir un second champ de déformation [255]. La fonction de coût basée sur l'intensité est dupliquée pour les deux images recalées, puis une fonction de coût supplémentaire pénalise les différences d'intensités entre les images originales et les images recalées deux fois (par le recalage direct puis inverse). ICON [256] approxime la propriété d'*inverse-consistency* et ajoute une fonction de coût pénalisant la somme des différences au carré entre l'identité et la composition des champs de déformation. Si nous recalons M sur F via ϕ_1 et F sur M via ϕ_2 , la fonction de coût pour l'*inverse-consistency* est par exemple :

$$L_{inv}(\phi_1, \phi_2) = \|\phi_1 \circ \phi_2 - Id\|_2^2 + \|\phi_2 \circ \phi_1 - Id\|_2^2 \quad (5.1)$$

GradICON [217] propose plutôt de pénaliser la somme des différences au carré entre l'identité et le jacobien de la composition des champs de déformation. ConstrICON [257] étend ces formulations pour des recalages multi-étapes (composition en série de plusieurs réseaux de recalage) afin d'obtenir un recalage respectant l'*inverse-consistency* sur l'entièreté du recalage.

SYMNet [205] utilise la même stratégie que SyN appliquée aux approches DL. En recalant les deux images vers une image intermédiaire, deux champs de déformation sont obtenus. Le recalage direct est la composition des deux

champs, tandis que le recalage inverse est la composition inverse des deux champs. Cela impose la propriété d'*inverse-consistency*. La fonction de coût utilisée est simplement la LNCC sur les trois étapes des images. M et F sont recalées vers l'image intermédiaire avec ϕ_1 et ϕ_2 telle que $M \circ \phi_1 = F \circ \phi_2$ et le recalage entier est obtenu par composition $M \circ \phi_1 \circ \phi_2^{-1} = F$ et $F \circ \phi_2 \circ \phi_1^{-1} = M$, alors la fonction de coût symétrique est :

$$L_{sym}(F, M) = L_{LNCC}(M \circ \phi_1 \circ \phi_2^{-1}, F) + L_{LNCC}(F \circ \phi_2 \circ \phi_1^{-1}, M) + L_{LNCC}(M \circ \phi_1, F \circ \phi_2). \quad (5.2)$$

D'autres propriétés ont été étudiées comme par exemple la *cycle-consistency* ou l'ajout de régularisations supplémentaires. La *cycle-consistency* est proche de l'*inverse-consistency*. Elle consiste à recaler une image M sur F pour obtenir l'image recalée R . Elle recalc ensuite R sur M formant ainsi un cycle, assurant que le recalage direct est inversible. Cependant, cela ne recalc pas directement F sur M . Elle est similaire à l'*inverse-consistency* si le recalage parfait existe. Cette approche a notamment été implémentée par [258, 216, 250] en dupliquant la fonction de coût, puis par CycleMorph [210] en utilisant deux réseaux consécutifs et une fonction de coût spécifique. Ici, si on recalc M sur F avec ϕ_1 , puis qu'on retourne sur M avec ϕ_2 , un exemple de fonction de coût de *cycle-consistency* est :

$$L_{cycle} = L_{NCC}(M \circ \phi_1, F) + L_{NCC}(M \circ \phi_1 \circ \phi_2, M). \quad (5.3)$$

Plus directement, des fonctions de coût pénalisant le jacobien de la déformation ont aussi été étudiées pour obtenir des déformations presque difféomorphiques. Nous pouvons citer la fonction de coût anti-repliement, qui pénalise le gradient de la déformation pour chaque voxel où il existe un repliement [254, 215] ou la fonction de coût pénalisant les points négatifs du jacobien [259, 253]. En notant J_ϕ le jacobien associé à ϕ , N le cardinal de l'image Ω , un exemple de fonction de coût pénalisant la présence de jacobien négatif est :

$$L_{J_{det}} = \frac{1}{N} \sum_{p \in \Omega} \max(0, -J_\phi(p)). \quad (5.4)$$

Ce terme additionne sur l'image les termes négatifs du jacobien s'ils existent.

Dans la majorité des travaux présentés ci-dessus, l'effet régularisant des méthodes implémentées est toujours mesuré par la quantité de voxels où le jacobien est négatif et par la fonction de coût associée à la méthode introduite.

5.3 Méthode de recalage inverse

Le champ de déformation ϕ est ici dénommé ϕ_1 pour le recalage direct de M sur F . À l'inverse, le champ recalant F vers M est nommé ϕ_2 . Comme dans le chapitre précédent, la fonction de coût pour le recalage direct L_{dir} est toujours :

$$L_{dir}(F, M, \phi_1) = L_{LNCC}(F, M \circ \phi_1) + \lambda_1 \|\nabla v_1\|_2^2. \quad (5.5)$$

Pour que le modèle donne en sortie un champ de déformation supplémentaire, la tête de sortie du réseau est simplement dupliquée avec des paramètres qui lui sont propres (cf. figure 5.1). L'encodeur reste unique dans cette architecture. Cette modification peut être faite pour toute architecture. Le champ de déformation inverse doit tout d'abord recalculer l'image fixe sur l'image source. La fonction de coût est donc elle aussi dupliquée, mais de manière symétrique :

$$L_{sym}(F, M, \phi_2) = L_{LNCC}(M, F \circ \phi_2) + \lambda_1 \|\nabla v_2\|_2^2.$$

Afin que ϕ_2 soit autant régularisé que ϕ_1 , aucun coefficient supplémentaire n'est appliqué à cette partie symétrique. Pour imposer la propriété que $\phi_1 \approx \phi_2^{-1}$, une contrainte d'inversibilité similaire à [256] est ajoutée :

$$L_{inv}(F, M, \phi_1, \phi_2) = \frac{1}{2} (\|\phi_1 \circ \phi_2 - Id\|_2^2 + \|\phi_2 \circ \phi_1 - Id\|_2^2), \quad (5.6)$$

régulée par la suite par l'hyper-paramètre λ_2 . Ainsi, la fonction de coût totale est définie comme suit :

$$\begin{aligned} L = & L_{LNCC}(F, M \circ \phi_1) + \lambda_1 \|\nabla v_1\|_2^2 \\ & + L_{LNCC}(M, F \circ \phi_2) + \lambda_1 \|\nabla v_2\|_2^2 \\ & + \frac{\lambda_2}{2} (\|\phi_1 \circ \phi_2 - Id\|_2^2 + \|\phi_2 \circ \phi_1 - Id\|_2^2). \end{aligned} \quad (5.7)$$

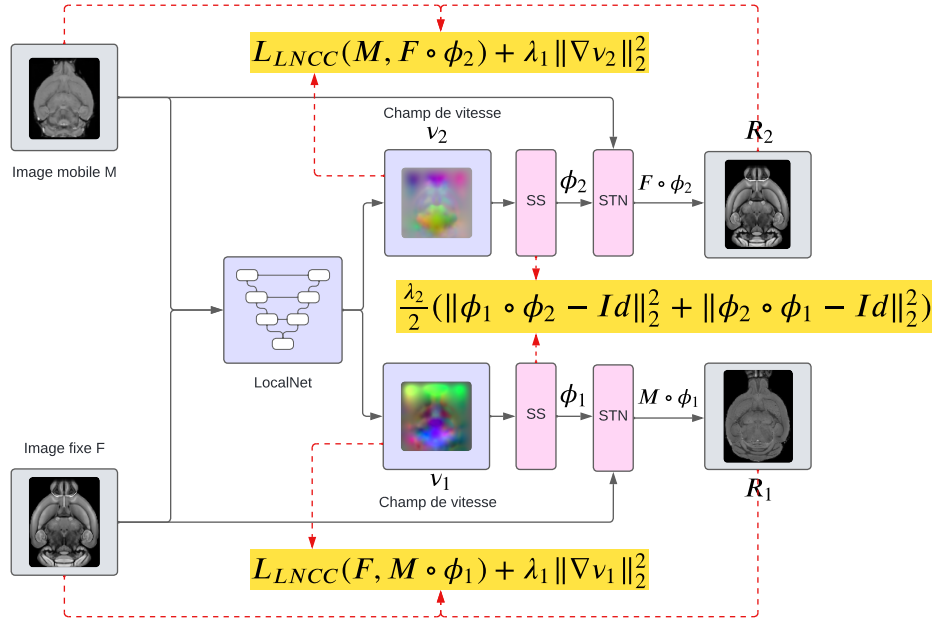


FIGURE 5.1 – Aperçu de la méthode proposée. Le bloc SS correspond à l'algorithme *Scaling and Squaring*. Le bloc STN correspond au *Spatial Transformer Network*.

Cette formulation garantit une régularisation adéquate du recalage et d'assurer la cohérence entre les transformations directes et inverses, favorisant ainsi des résultats plus précis et robustes dans le contexte de transfert d'atlas. L'obtention de champ inversible et d'une approximation correcte de son inverse permet de recalculer correctement l'individu sur l'atlas mais aussi l'atlas sur l'individu, autorisant ainsi la segmentation par transfert d'atlas.

5.3.1 Métriques de validation

Les scores de Dice des 10 régions annotées sur le jeu de données GIN sont comparés aux méthodes du chapitre précédent.

Concernant l'inversibilité, nous utilisons la fonction de coût introduite ci-dessus dans l'équation 5.6 pour mesurer l'inversibilité de la méthode et une formulation du Dice qu'on nomme Self-Dice, consistant à calculer les scores de Dice entre un individu et le même individu recalé sur l'atlas, puis sur lui-même.

5.3.2 Expériences et validation

S’agissant de la continuité du chapitre 4, les modèles sont à nouveau entraînés/optimisés sur le jeu de données GIN. L’architecture a été modifiée pour que le réseau puisse donner deux champs de déformation en sortie, comme indiqué plus haut sur la figure 5.1. Les fonctions de coût décrites dans les sections précédentes ont à nouveau été implémentées sur PyTorch et MONAI dans la continuité de nos travaux. Cette modification d’architecture intervient pour chaque modèle présenté dans le chapitre précédent. Par conséquent, les modèles entraînés ici se voient ajouter le suffixe *-sym* pour les distinguer des versions évaluées au chapitre précédent. Les modèles **OP**, **DL-IRIS** et **Pair** deviennent donc **OP-sym**, **DL-sym** et **Pair-sym**. Nous prendrons, pour tous les modèles entraînés, λ_1 égal à 1. De la même manière, λ_2 est choisi parmi (0.001, 0.01, 0.1, 1, 10, 100, 1000) pour les méthodes **Pair-sym** et **OP-sym**. Contrairement à λ_1 , le choix de λ_2 est plus complexe car nous nous attendons à ce que l’inversibilité et le recalage inverse pénalisent le recalage direct. L’influence de λ_2 est détaillée ci-dessous. Concernant **DL-IRIS-sym**, pour simuler l’obtention de ce modèle de manière externe, nous prenons arbitrairement λ_2 égal à 1. Nous ignorons ici l’effet de l’ajout des modifications sur le choix optimal de λ_1 .

5.3.3 Résultats et discussions

Pour choisir la valeur optimale de λ_2 des méthodes **OP-sym** et **Pair-sym**, le score de Dice moyen sur les 10 régions pour le recalage direct, le score de Dice moyen sur les 10 régions pour le recalage inverse, la métrique de Self-Dice ainsi que la fonction de coût pour l’inversibilité L_{inv} pour chacune des valeurs de λ_2 considérées sont présentés dans le tableau 5.1. Ces scores reflètent 3 critères différents : le recalage direct, le recalage inverse et l’inversibilité du champ. Il faut alors choisir le bon compromis. Alors que **OP** et **Pair** présentaient des comportements similaires pour λ_1 , cela n’est pas du tout le cas pour λ_2 .

Pour les méthodes **OP-sym**, le meilleur score de recalage direct est atteint pour $\lambda_2 = 0.01$ malgré un recalage inverse moyen. On remarque que l’inversibilité devient parfaite pour de grandes valeurs de λ_2 avec un bon compromis de performance à partir de $\lambda_2 = 100$. Dans tous les cas, le recalage direct obtient de

TABLEAU 5.1 – Comparatif des métriques pour différentes valeurs de λ_2 pour les méthodes **OP-sym** et **Pair-sym**. Le recalage affine et les méthodes **SyN** sont ajoutées en tant que références. Pour le recalage affine, le recalage déformable étant égal à l'identité, les performances de recalage direct et inverse sont égales et l'inversibilité est parfaite.

Méthode	λ_2	Dice moyen du recalage direct (%) $\uparrow$	Dice moyen du recalage inverse (%) $\uparrow$	Self-Dice (%) $\uparrow$	L_{inv} $\downarrow$
Affine		77.43 $\pm$ 4.29	77.43 $\pm$ 4.29	100.0 $\pm$ 0.00	0.0000 $\pm$ 0.0000
SyN		78.95 $\pm$ 3.51	65.29 $\pm$ 5.22	98.15 $\pm$ 0.29	0.0000 $\pm$ 0.0000
SyN-DL		82.06 $\pm$ 2.49	82.21 $\pm$ 2.80	98.34 $\pm$ 0.25	0.0000 $\pm$ 0.0000
DL-sym	(1)	82.09 $\pm$ 2.68	82.50 $\pm$ 2.97	97.09 $\pm$ 0.36	0.0031 $\pm$ 0.0004
OP-sym	0.001	82.70 $\pm$ 2.26	78.81 $\pm$ 2.88	92.97 $\pm$ 0.85	0.1369 $\pm$ 0.0161
	0.01	83.02 $\pm$ 2.19	79.40 $\pm$ 2.79	94.73 $\pm$ 0.55	0.0463 $\pm$ 0.0043
	0.1	82.57 $\pm$ 2.41	79.20 $\pm$ 2.82	96.21 $\pm$ 0.41	0.0048 $\pm$ 0.0005
	1	81.71 $\pm$ 3.06	78.17 $\pm$ 3.66	96.86 $\pm$ 0.35	0.0004 $\pm$ 0.0000
	10	81.65 $\pm$ 2.81	78.83 $\pm$ 3.34	97.13 $\pm$ 0.33	0.0001 $\pm$ 0.0000
	100	82.06 $\pm$ 2.47	79.59 $\pm$ 2.90	97.44 $\pm$ 0.31	0.0000 $\pm$ 0.0000
	1000	81.05 $\pm$ 2.91	82.11 $\pm$ 3.02	98.03 $\pm$ 0.26	0.0000 $\pm$ 0.0000
Pair-sym	0.001	82.91 $\pm$ 2.60	81.02 $\pm$ 2.84	94.05 $\pm$ 1.52	0.1002 $\pm$ 0.0451
	0.01	82.89 $\pm$ 2.61	80.83 $\pm$ 2.80	95.55 $\pm$ 0.85	0.0290 $\pm$ 0.0088
	0.1	82.41 $\pm$ 2.72	80.68 $\pm$ 2.85	97.02 $\pm$ 0.36	0.0033 $\pm$ 0.0005
	1	81.65 $\pm$ 2.74	81.11 $\pm$ 3.03	97.64 $\pm$ 0.32	0.0001 $\pm$ 0.0000
	10	79.47 $\pm$ 3.26	81.10 $\pm$ 3.20	98.11 $\pm$ 0.29	0.0000 $\pm$ 0.0000
	100	77.43 $\pm$ 4.29	78.94 $\pm$ 4.41	98.39 $\pm$ 0.23	0.0000 $\pm$ 0.0000
	1000	77.43 $\pm$ 4.29	78.94 $\pm$ 4.41	98.39 $\pm$ 0.23	0.0000 $\pm$ 0.0000

bons scores. Ces résultats sont cohérents avec la méthode d'optimisation utilisée. **OP-sym** étant optimisé sur l'ensemble du groupe, il produit un recalage direct performant peu importe les valeurs de λ_2 tandis que l'inversibilité et le recalage inverse s'améliorent progressivement avec λ_2 . De par la régularisation implicite offerte par l'optimisation par groupe, le recalage inverse nécessite des valeurs de λ_2 plus élevées, comprises entre 100 et 1000, pour obtenir de meilleurs résultats.

Pour les méthodes **Pair-sym**, le recalage inverse est très bon pour les faibles valeurs de λ_2 pour au final se dégrader avec de fortes valeurs de l'hyperparamètre. Contrairement aux méthodes **OP-sym**, les performances du recalage direct diminuent grandement pour les valeurs de λ_2 supérieures à 10. On observe aussi que les méthodes **Pair-sym** pour $\lambda_2 = 100$ et $\lambda_2 = 1000$ offrent exactement le même recalage, certainement trop régularisées. **Pair-sym** s'optimise pour chaque paire d'images et est ainsi beaucoup plus sensible à la régularisation, sa régularisation implicite étant plus faible. Pour de faibles valeurs de λ_2 , son recalage inverse est très performant mais son inversibilité ne peut atteindre des valeurs proches de celles de SyN. Dès que la régularisation devient plus importante, les performances en recalage direct ainsi qu'en recalage inverse sont dégradées. Pour **Pair-sym**, une exploration des valeurs de λ_2 comprises entre 0.1 et 1 aurait pu être plus judicieuse.

Contre toute attente, la méthode **DL-sym** est très compétitive, offrant un excellent compromis entre recalage direct, recalage inverse et inversibilité. L'ajout de *-sym* améliore même ses performances sur le recalage direct. Les performances de **DL-sym** s'expliquent par l'utilisation du jeu de données IRIS qui doit présenter de meilleures caractéristiques pour l'apprentissage du recalage inverse. Cependant, nous avons remarqué dans le chapitre 4 que **DL-sym** n'atteint pas d'aussi bonnes performances en recalage direct que les autres approches.

Dans l'ensemble, l'inversibilité permise par les modifications est toujours inférieure à l'inversibilité offerte par les méthodes **SyN**. Le choix optimal de λ_2 en fonction des trois critères étant complexe, nous faisons le choix de prioriser le recalage direct ainsi que l'inversibilité du champ au détriment du recalage inverse, en invoquant l'objectif de la régularisation du champ et de l'analyse VBM. Nous nous limitons d'abord aux modèles atteignant au moins des performances supérieures à **SyN-DL** en recalage direct, puis choisissons celui atteignant le

maximum de Self-Dice. Cela nous amène à considérer $\lambda_2 = 0.1$ pour la méthode **Pair-sym** et $\lambda_2 = 100$ pour **OP-sym** comme méthodes optimales. Comme indiqué ci-dessus, des valeurs mieux choisies de λ_2 grâce à des algorithmes de recherche auraient été préférables, mais cela aurait été très chronophage, tout comme l’entraînement d’autres modèles **OP-sym**. Ces choix offrent tout de même des recalages inverses corrects. Nous fixons ces valeurs de λ_2 pour la suite des résultats afin de ne pas surcharger les tableaux.

TABLEAU 5.2 – Comparaison du Dice de 3 régions d’intérêt pour chaque méthode avec et sans modifications pour le recalage inverse sur le jeu de données GIN. Les meilleurs scores pour les méthodes avec *-sym* pour chaque colonne sont en gras.

Méthodes	GP (%) ↑	HY (%) ↑	OLF (%) ↑
Affine	52.34 ± 15.4	80.67 ± 2.93	60.25 ± 4.50
SyN	60.12 ± 7.11	77.72 ± 4.48	73.73 ± 5.67
SyN-DL	66.18 ± 8.17	81.55 ± 2.28	67.93 ± 1.74
DL	68.14 ± 6.53	82.00 ± 2.43	67.13 ± 4.41
DL-sym	68.94 ± 6.04	81.70 ± 2.52	65.74 ± 4.56
Pair	68.08 ± 5.47	82.12 ± 2.17	74.83 ± 3.91
Pair-sym	68.07 ± 5.69	81.97 ± 2.12	69.95 ± 4.74
OP	66.01 ± 6.41	82.53 ± 2.16	78.43 ± 2.00
OP-sym	65.72 ± 6.30	81.53 ± 2.42	73.66 ± 2.75

Concernant les trois régions anatomiques mises en évidence dans le chapitre 4, les résultats sont présentés dans le tableau 5.2. Cependant, l’ajout des modifications n’entraîne aucun changement majeur dans les conclusions. Aucun effet n’est observé sur les régions GP et HY. En revanche, le recalage de la région OLF est fortement impacté par l’ajout du recalage inverse, bien que **OP-sym** conserve un avantage sur les autres méthodes. Cette région étant tronquée sur l’atlas et présentant de fortes variations inter-sujets, elle nécessite un recalage particulièrement important pour être correctement mise en correspondance. L’introduction d’une régularisation plus stricte explique ainsi la baisse de performance observée sur cette région.

Avant d’évaluer la segmentation par transfert d’atlas, nous présentons en

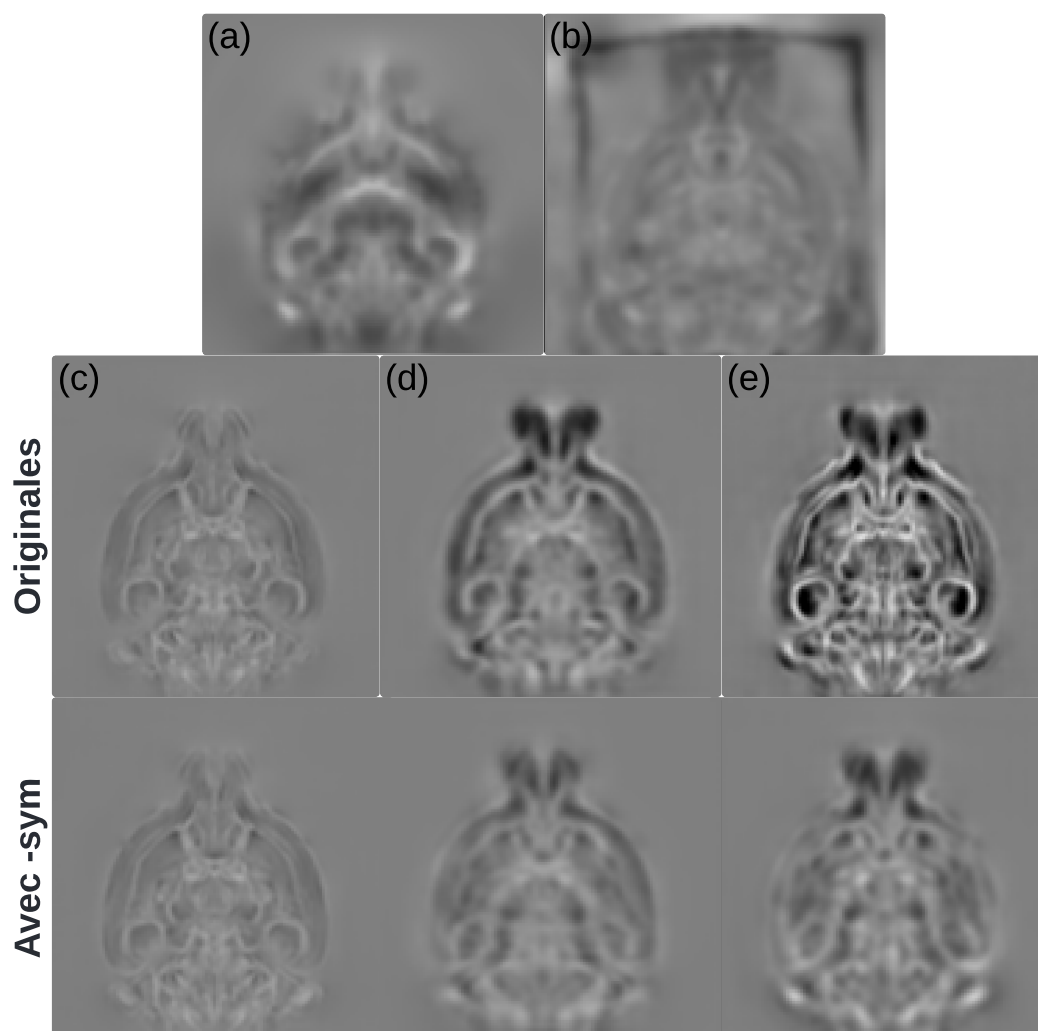


FIGURE 5.2 – Comparaison des moyennes des log-Jacobiens des déformations pour chaque méthode avec et sans modifications. La légende est la même que celle du chapitre précédent, (a) **SyN**, (b) **SyN-DL**, (c) **DL-IRIS**, (d) **Pair**, (e) **OP**. Les versions des modèles avec *-sym* sont situées en-dessous de l'image originale.

figure 5.2 les moyennes des jacobiens des déformations obtenues avec les méthodes modifiées comparées avec celles sans modifications. Nous remarquons l’effet très régularisateur des modifications apportées sur les méthodes **OP** et **Pair**, dont les contours sont beaucoup moins marqués, se rapprochant plus des méthodes **SyN**. L’ajout de *-sym* ne présente aucun effet notable sur le jacobien des déformations pour la méthode **DL-IRIS**. Nous obtenons donc des déformations beaucoup plus lisses et proches des méthodes conventionnelles pour un coût en performances léger.

5.4 Transfert d’atlas pour la segmentation des tissus cérébraux

Cette section met en rapport la segmentation par transfert d’atlas, obtenue grâce au recalage inverse présenté dans ce chapitre, avec les méthodes de segmentation des tissus cérébraux obtenues dans le chapitre 2. Pour ce travail, le jeu de données utilisé est *Feminad*, comme dans le chapitre portant sur la segmentation. Un unique modèle **OP-sym** a été optimisé sur *Feminad* afin d’obtenir les recalages. Toutefois, la segmentation fine de l’atlas Allen regroupée en tissus GM, WM et CSF ne correspond pas à la segmentation des vérités terrain obtenue par SPMmouse. En effet, l’atlas employé est celui pour la connectivité cérébrale en IRMf où le liquide cébrospinal ainsi qu’une partie de la substance blanche ne sont pas présents sur la segmentation.

Afin de garantir la cohérence avec les vérités terrain, nous utilisons directement les cartes de segmentation de SPMmouse plutôt que la segmentation de l’atlas Allen. Ces cartes de probabilités, déjà recalées sur l’atlas Allen, indiquent pour chaque voxel la probabilité d’appartenance à l’une des trois classes. Nous appliquons donc le recalage inverse à ces cartes pour les adapter aux individus, puis procédons à leur binarisation afin d’obtenir la segmentation finale. Les scores de Dice obtenus pour les trois tissus sont présentés dans le tableau 5.3, aux côtés des résultats des méthodes de segmentation du chapitre 2.

Le transfert d’atlas permet d’obtenir une meilleure segmentation de la matière blanche. Cependant, le liquide cébrospinal est à nouveau très mal segmenté.

TABLEAU 5.3 – Comparaison des scores de Dice par classe obtenus avec les méthodes de segmentation du chapitre 2 ainsi que la méthode de segmentation par transfert d’atlas. Les classes GM, WM et CSF correspondent respectivement à la matière grise, la matière blanche et le liquide cébrospinal. Les meilleurs scores pour chaque colonne sont en gras.

Type	Modèle	GM ↑	WM ↑	CSF ↑
Directe	UNet	0.929 ± 0.009	0.672 ± 0.017	0.292 ± 0.086
Directe	UNETR	0.940 ± 0.009	0.702 ± 0.021	0.353 ± 0.126
Recalage	SyN-DL	0.787 ± 0.138	0.647 ± 0.112	0.060 ± 0.023
Recalage	OP-sym	0.821 ± 0.143	0.701 ± 0.119	0.138 ± 0.039

Un exemple est illustré en figure 5.3. Nous remarquons que la segmentation par transfert d’atlas fournit des zones assez homogènes et lisses pour la matière blanche et grise, contrairement aux méthodes directes. En observant les artefacts de la vérité terrain, la conclusion reste la même : la mauvaise qualité des vérités terrains rend complexe la comparaison numérique des résultats obtenus. Toutefois, les résultats obtenus par **SyN** et **OP-sym** sont assez proches, indiquant à nouveau une mauvaise segmentation de départ avant le recalage, malgré l’utilisation de cartes de probabilité.

5.5 Conclusion

Dans ce chapitre, nous avons proposé des modifications visant à obtenir la segmentation par transfert d’atlas en calculant simultanément le champ de déformation direct et le champ de déformation inverse. Les modifications consistent en la modification de l’architecture pour ajouter une seconde sortie et l’intégration d’une fonction de coût spécifique à l’irréversibilité des champs de déformation.

Si ces ajustements entraînent une légère baisse des performances en termes de score de Dice, ils permettent néanmoins de réaliser le transfert d’atlas pour segmenter les individus. Associée à la méthode **OP** du chapitre précédent, notre approche **OP-sym** demeure performante pour le recalage direct pour un recalage inverse moins performant que dans d’autres approches. Nous faisons remarquer que la validation par Dice obtenue dans le chapitre précédent et ce chapitre pour

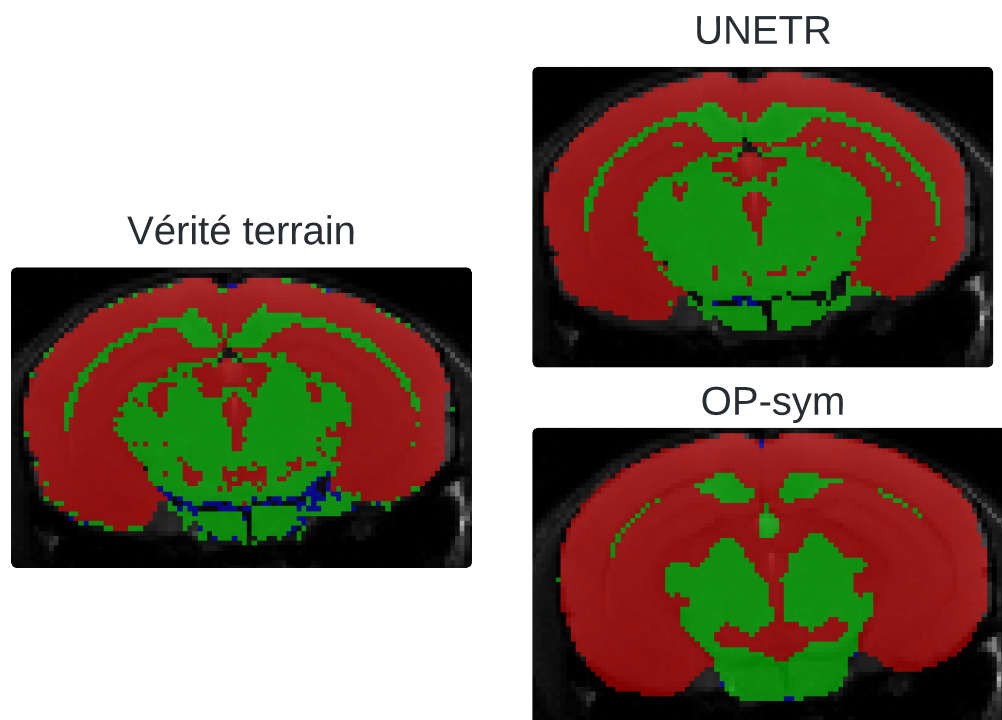


FIGURE 5.3 – Comparaison entre une vérité terrain, une segmentation directe avec UNETR et une segmentation par transfert d'atlas avec la méthode **OP-sym**.

l'introduction des modifications de *-sym* constitue en elle-même une segmentation par transfert d'atlas, non pas sur les tissus cérébraux mais sur 10 régions anatomiques, démontrant la bonne performance du recalage par apprentissage profond pour le transfert d'atlas.

Néanmoins, les jacobiens des déformations obtenus se rapprochent visuellement de ceux des méthodes conventionnelles, ce qui est attribuable à la régularisation induite par le recalage inverse. Cette convergence entre les approches basées sur l'apprentissage profond et les méthodes classiques devra être confirmée dans le chapitre suivant, lors de leur application à l'analyse VBM.

CHAPITRE 6

Application à l'analyse VBM

Sommaire

6.1 Introduction	110
6.2 Etat de l'art des pipelines VBM	111
6.2.1 Pipeline VBM	111
6.2.2 Intégration de méthodes par apprentissage profond	112
6.3 Application au jeu de données Feminad	113
6.3.1 Intérêt de l'étude	113
6.3.2 Pipelines comparés	116
6.3.3 Évaluation du recalage sur Feminad	116
6.3.4 Résultats	117
6.4 Conclusion	119

6.1 Introduction

Comme explicitée dans le premier chapitre, la morphométrie basée voxel constitue une méthode essentielle pour étudier les différences structurelles cérébrales en neurosciences. Qu'il s'agisse d'explorer les effets de pathologies, de traitements ou de différences biologiques naturelles entre groupes, elle est utilisée pour quantifier les variations anatomiques à partir d'IRM. Dans le contexte des études précliniques sur modèles murins, l'analyse VBM offre une opportunité unique d'examiner les changements induits sur les lignées génétiquement modifiées.

Les étapes de prétraitement indispensables à l'analyse VBM, notamment le recalage des images, influencent fortement la qualité des résultats obtenus [48]. Par ailleurs, ces étapes demeurent à la fois chronophages [68] et complexes à adapter aux modèles murins [66]. À l'inverse, les approches basées sur le deep learning, développées dans cette thèse, offrent une alternative prometteuse en alliant rapidité, robustesse et précision.

Ce chapitre propose un état de l'art des pipelines VBM actuellement utilisés pour l'homme puis pour les modèles murins. Nous mettons en exergue le manque de recours à des méthodes par DL dans ces pipelines. Les rares méthodes DL intégrées présentent les mêmes problèmes que certaines méthodes conventionnelles : elles ne sont pas directement adaptables à d'autres espèces.

Ce chapitre emploie ainsi les méthodes d'apprentissage profond développées tout au long de cette thèse pour former un pipeline VBM entier. Ce pipeline est appliqué à Feminad, un jeu de données portant sur l'étude du dimorphisme sexuel chez la souris atteinte d'Alzheimer. Nous comparons alors quantitativement les résultats obtenus grâce à ce pipeline avec ceux d'un pipeline utilisant ANTs et SPMmouse.

6.2 Etat de l'art des pipelines VBM

6.2.1 Pipeline VBM

Les pipelines conventionnels de VBM reposent généralement sur des méthodes bien établies. Chez l'homme, plusieurs pipelines automatiques pour l'analyse VBM ont déjà été proposés comme FSL [137], SPM [260] ou FreeSurfer [261]. Cependant, tous ces pipelines se basent sur des méthodes spécifiques à l'homme, rendant complexes leur adaptation aux souris dues aux différences de distribution des tissus cérébraux, de contrastes et d'échelles. Comme vu dans les chapitres précédents, chaque étape de l'analyse VBM a été étudiée et des méthodes ont été développées pour leur application aux études précliniques en neuro-imagerie.

Un pipeline d'analyse VBM semi-automatique pour la souris est proposé par [262] et est basé exclusivement sur ANTs. Sans entrer dans le détail, il utilise *N3BiasFieldCorrection* pour corriger le biais d'intensité, le recalage *Greedy-SyN* sur un template généré, *Atropos* pour la segmentation, *CreateJacobianDeterminantImage* pour les jacobiens et enfin FSL pour l'analyse statistique. Deux autres analyses [263, 264] utilisent plutôt *Unified Segmentation* pour la segmentation et DARTEL pour le recalage. Dans tous les cas, les résultats biologiques de l'analyse obtenus sont comparés avec d'autres résultats. Cependant, des résultats vérités terrain n'existent pas. Lorsque les résultats ne sont pas consistants, les jeux de données ou l'acquisition sont remis en question plutôt que les méthodes utilisées pour effectuer l'analyse VBM.

Plus récemment, avec le développement d'un pipeline VBM dédié au petit animal axé sur le temps d'exécution des algorithmes, un framework de validation des analyses VBM basé sur la simulation d'atrophie ou d'hypertrophie dans une structure choisie est proposé [66]. La validation est faite grâce à l'utilisation de métriques entre les données de contrôle et les données simulées comme le DSC. Le framework de validation est utilisé sur plusieurs pipelines VBM dont les seules différences résident dans les paramètres de la méthode SyN utilisée pour le recalage. Ce framework de validation est disponible en ligne, mais est malheureusement inutilisable en l'état, n'ayant pas été mis à jour depuis

2016. De plus, la simulation d'atrophie ou d'hypertrophie d'une unique région à compenser par le recalage nous paraît assez limitée pour évaluer l'impact global d'une méthode de recalage sur l'analyse VBM. Ce travail met ainsi en évidence la difficulté de valider différents pipelines VBM pour une même étude préclinique. À notre connaissance, ce sont les seules études portant sur les pipelines VBM pour les études précliniques.

6.2.2 Intégration de méthodes par apprentissage profond

L'utilisation de méthodes basées DL est encore très limitée pour les études VBM, même chez l'homme. FreeSurfer implémente SynthStrip [265] et SynthSeg [266] pour l'extraction et la segmentation, puis récemment EasyReg [267] et SynthMorph [268] pour le recalage affine et déformable afin de permettre leur utilisation dans des analyses VBM. Ces méthodes ne sont intégrées qu'en novembre 2024 dans une version bêta de FreeSurfer pour certains pipelines d'études. Elles sont toutes basées sur l'utilisation de données synthétiques pour le cerveau humain afin d'éviter les problèmes de contraste. Parallèlement, FSL a intégré TRuENET [269] en août 2024, un modèle DL de segmentation de la WM pour l'homme.

En octobre 2024, un pipeline entier (extraction, segmentation et recalage sur atlas) pour le prétraitement d'analyse VBM utilisant entièrement des méthodes basées DL a été diffusé pour le cerveau humain [68]. Ce pipeline met surtout l'accent sur la vitesse d'exécution des méthodes pour traiter des jeux de données très volumineux. Les images sont d'abord prétraitées avec CAT12 [270] puis le biais est corrigé avec *N4BiasFieldCorrection* d'ANTs. Deepbet [271] extrait le cerveau des images, torchreg [272] effectue le recalage affine, une multitude de réseaux basés sur nnU-Net [273] segmentent les tissus cérébraux en subdivisant l'image en 27 patches de petites tailles puis un modèle SYMNet [205], ici entraîné de manière supervisée sur les champs de déformations obtenus par CAT12, effectue le recalage déformable. Les deux pipelines VBM (CAT12 et deepmriprep) montrent de très grandes similarités au niveau des résultats des analyses VBM, atteignant des corrélations supérieures d'au moins 90%. Ces résultats étaient attendus dans la mesure où l'étape la plus impactante de l'analyse est apprise de

manière supervisée sur CAT12. Malgré l'approche entièrement supervisée des méthodes introduites, les modèles sont entraînés sur des jeux de données très volumineux et sont systématiquement testés sur des jeux de données externes pour démontrer la robustesse du pipeline.

Chez la souris, ou même chez les rongeurs, aucune analyse VBM n'a, à notre connaissance, été réalisée en utilisant ces nouvelles méthodes. De plus, les rares intégrations de méthodes DL dans des frameworks de neuro-imagerie utilisent toutes des modèles pré-entraînés sur des cerveaux humains, rendant impossible leur utilisation directe pour les études précliniques, si ce n'est en entraînant soi-même les modèles sur des jeux de données spécifiques.

6.3 Application au jeu de données Feminad

L'impact de chaque étape de l'analyse VBM a déjà fait l'objet de nombreuses études chez l'homme [59, 60, 61, 62, 48], comme nous l'avons déjà présenté dans les limites de l'analyse VBM dans le chapitre 1. Pour la souris, ou pour les modèles animaux en général, hormis les deux études citées précédemment [66, 65], l'impact des différents pipelines de prétraitement sur les résultats de l'analyse VBM n'a pas été étudié. C'est pourquoi l'objet de ce chapitre est d'effectuer deux analyses VBM sur un même jeu de données en choisissant soit la méthode de recalage **SyN** soit **OP-sym**. Nous étudions ici l'impact de notre méthode de recalage sur une analyse VBM avec une autre réalisée de façon conventionnelle.

6.3.1 Intérêt de l'étude

Nous mobilisons ici le jeu de données Feminad (cf. tableau 1.1 page 22 du chapitre 1). Pour rappel, celui-ci porte sur l'étude du dimorphisme sexuel chez la souris atteinte d'Alzheimer. Nous prenons en compte ici seulement les souris WT imagées à 2 mois pour ne pas étudier l'effet d'Alzheimer, ni celui de l'âge. L'étude de groupes correspond alors à étudier l'effet du sexe sur le cerveau entre 16 mâles et 17 femelles âgés de 2 mois. Les WT sont des souris C57BL/6, la souche de souris la plus utilisée pour les études précliniques [40]. Restreindre

l'étude à ces souris nous permet donc de comparer nos résultats directement avec ceux issus de la littérature, et de valider au moins qualitativement les analyses VBM obtenues. Les régions du cerveau soumises à un dimorphisme sexuel chez la souris âgée de 2 à 3 mois sont présentées dans le tableau 6.1 selon les résultats présentés par [274, 275, 263]. Nous notons le cas particulier du cortex cérébral où des sous-régions différentes selon le sexe sont détectées par une même étude [274].

Région	Abréviation	Mâle	Femelle
Cortex cérébral	CTX	[263, 274]	[274]
Noyau du lit de la strie terminale	BST	[263, 275]	—
Hypothalamus postérieur	vHY	[263, 275]	[274]
Hippocampe antérieur	dHIP	—	[263, 274]
Caudoputamen ventral	vCP	[263, 275]	—
Caudoputamen dorsal	dCP	—	[263, 275]
Cortex cérébelleux médial (+ lobules II-IX)	VERM	[263, 274]	—
Cortex cérébelleux latéral (+ paraflocculus)	HEM	—	[263, 274]
Amygdale basolatérale	BLA	—	[263, 274]
Amygdale médiale	MEA	[263, 275]	—

TABLEAU 6.1 – Régions cérébrales présentant un dimorphisme sexuel chez les souris. La présence d'un article dans une colonne mâle ou femelle indique que cet article a trouvé des volumes significativement plus volumineux sur cette région pour le sexe considéré. Deux colonnes peuvent donc être remplies pour la même région. Les abréviations des régions sont principalement tirées de la segmentation de l'atlas Allen. Tableau modifié et inspiré de [263].

Nous identifions ces zones sur l'atlas Allen et les présentons en figure 6.1 pour mieux les localiser par la suite.

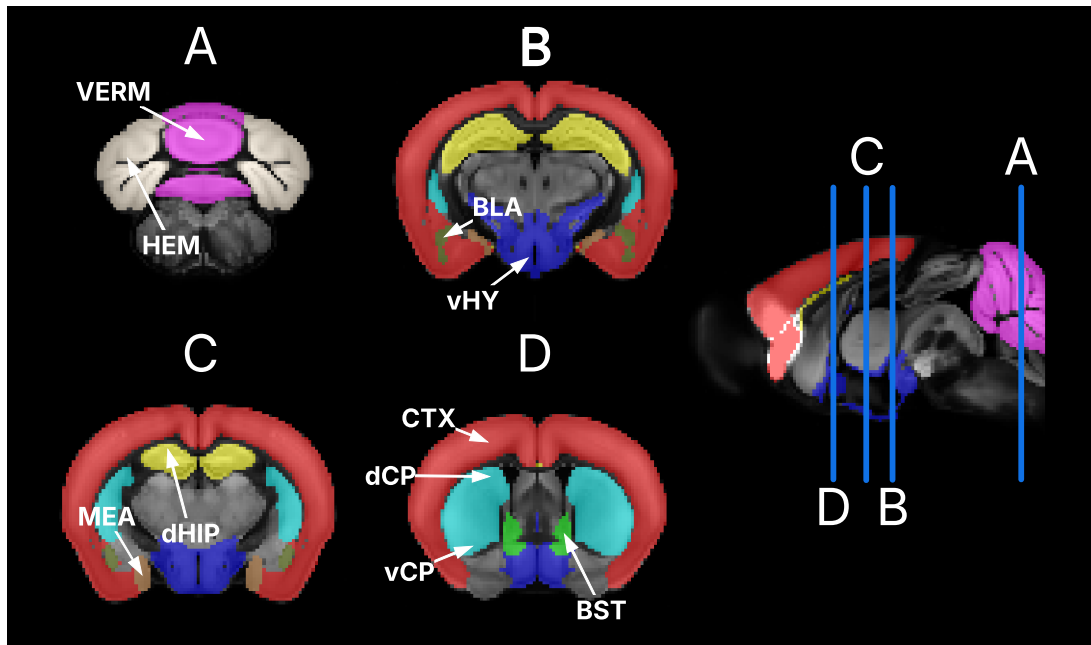


FIGURE 6.1 – Zones d'intérêt du cerveau pour l'étude Feminad identifiées sur l'atlas Allen. En rouge le cortex cérébral (CTX), en vert le noyau du lit de la strie terminale (BST), en bleu l'hypothalamus (HY), en jaune l'hippocampe (HY), en cyan le caudoputamen (CP), en rose le cortex cérébelleux médial (VERM), en beige le cortex cérébelleux latéral (HEM), en kaki l'amygdale basolatérale (BLA) et en brun l'amygdale médiale (MEA).

6.3.2 Pipelines comparés

Pour effectuer cette analyse, nous réalisons ici une étude TBM où nous ne segmentons pas le cerveau en tissus cérébraux. Seule la méthode de recalage est différente entre les analyses. L'étude des déformations des structures anatomiques du cerveau grâce au jacobien est donc réalisée sur l'intégralité du cerveau, et non uniquement sur la matière grise comme cela est généralement fait pour les études VBM.

Le premier pipeline est celui généralement utilisé avec ANTs et SPMmouse. Le biais est corrigé des images avec un algorithme N4 via *N4BiasFieldCorrection* d'ANTs. Les masques du cerveau sont obtenus grâce à SPMmouse. Le cerveau est alors extrait de l'image. Le recalage des individus vers l'atlas Allen est effectué avec *antsRegistration* en utilisant SyN avec les paramètres trouvés dans [250]. Le log-jacobien de la déformation est récupéré grâce à *CreateJacobianDeterminantImage*. Un noyau gaussien de 4 voxels est appliqué pour lisser les jacobiens. Les valeurs positives représentent des augmentations de volume, tandis que les valeurs négatives représentent des réductions de volume.

Le deuxième pipeline reprend exactement les mêmes points que le premier pipeline mais utilise la méthode **OP-sym** en lieu de SyN. La différence entre les pipelines réside donc dans la méthode de recalage, qui varie entre **SyN** et **OP-sym**.

Le dimorphisme sexuel est analysé grâce à un modèle linéaire général en prenant le sexe comme variable indépendante. Les différences significatives de volume local sont déterminées à l'aide de tests de Student. Les résultats sont ensuite corrigés en contrôlant le FDR avec un seuil à $p \leq 0.05$. Les valeurs restantes sont ensuite affichées sur l'atlas Allen pour leur visualisation et leur localisation.

6.3.3 Évaluation du recalage sur Feminad

Les méthodes **SyN** et **OP-sym** sont relancées avec exactement les mêmes paramètres que dans les chapitres précédents. Étant donné que l'étude Feminad porte sur l'effet du sexe à un âge donné, nous optimisons pour **OP-sym** un modèle pour chaque sous-groupe, à savoir un modèle optimisé sur les mâles

wild-type à 2 mois, et un modèle sur les femelles wild-type à 2 mois.

Nous n'avons pas de véritables métriques de performance sur ce jeu de données en raison de l'absence de segmentations de régions fiables. Toutefois, nous présentons en guise d'illustration en figure 6.2 les images moyennes recalées du jeu de données Feminad sur l'atlas, obtenues pour les deux méthodes considérées. Qualitativement, le recalage moyen obtenu par **OP-sym** semble meilleur, notamment dans le cortex cérébelleux (entouré en vert sur la figure) et dans les ventricules (entourés en jaune sur la figure).

6.3.4 Résultats

Les cartes sont superposées sur l'atlas Allen et sont présentées sur la figure 6.3 pour les deux pipelines étudiés.

Pour le pipeline **SyN**, peu de différences significatives sont réellement détectées. Seul le cortex cérébral est plus volumineux chez la femelle. Les mâles présentent un cortex cérébelleux médial ainsi que quelques zones de l'hypothalamus ventral et du caudoputamen plus volumineux que chez la femelle. Les régions significativement plus larges sont conformes avec les études citées précédemment. Cependant, aucune différence significative n'est trouvée dans les autres régions mentionnées dans le tableau 6.1.

Le pipeline **OP-sym** affiche au contraire beaucoup de différences significatives en comparaison avec **SyN**. Chez la femelle, le cortex cérébelleux latéral, le cortex cérébral, l'hippocampe antérieur ainsi que le caudoputamen ventral affichent des volumes significativement plus larges que chez le mâle. À l'inverse, le cortex cérébelleux médial, les amygdales basolatérales, l'hypothalamus postérieur, le noyau du lit de la strie terminale et le caudoputamen dorsal sont plus larges chez le mâle. Aucune différence significative n'est constatée pour l'amygdale médiale. **OP-sym** présente plus de différences significatives mais aussi des différences contraires aux études existantes. Nous trouvons en effet que les amygdales basolatérales sont plus larges chez le mâle à l'inverse de [263, 275]. Nous trouvons à nouveau l'inverse des résultats attendus sur le caudoputamen dorsal. Une asymétrie latérale est même indiquée pour le caudoputamen ventral. Les résultats pour l'hippocampe antérieur, l'hypothalamus postérieur, le cortex

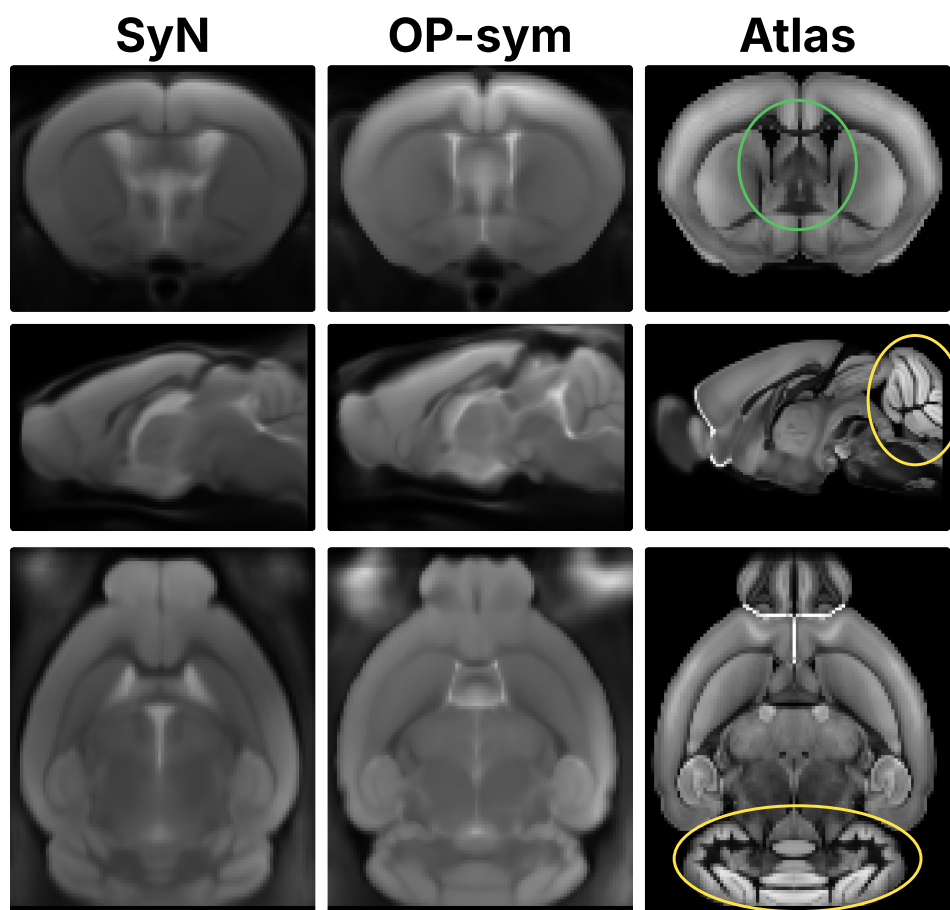


FIGURE 6.2 – Plusieurs coupes de l'image moyenne du jeu de données Feminad recalée sur l'atlas Allen pour les méthodes **SyN** et **OP-sym**. Les mêmes coupes pour l'atlas sont dans la colonne de droite. La région entourée en vert sur l'atlas est le cortex cérébelleux. La région entourée en jaune sur l'atlas représente les ventricules. La méthode **OP-sym** est qualitativement meilleure sur ces deux régions.

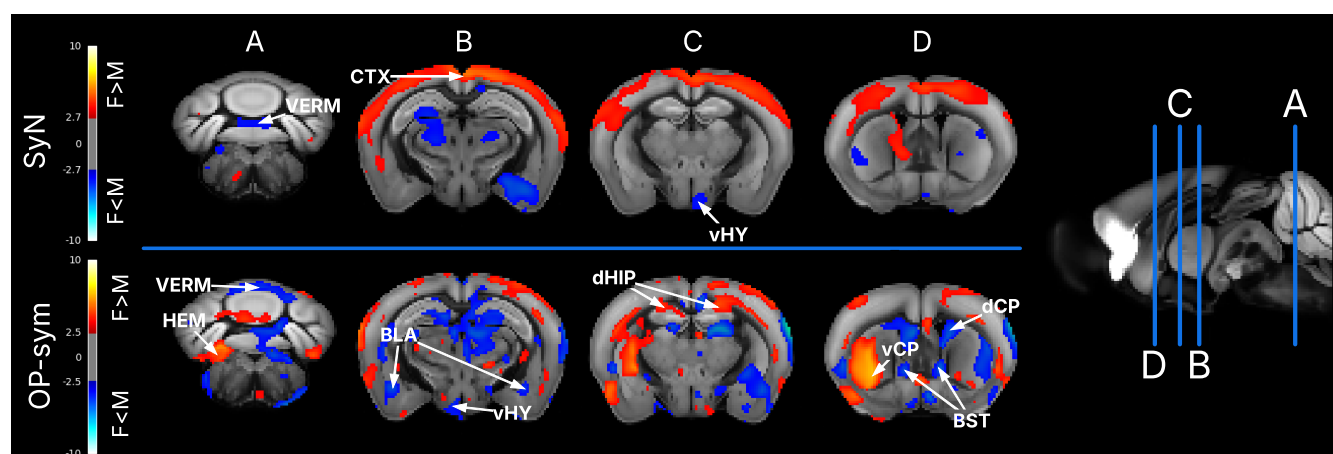


FIGURE 6.3 – Dimorphisme sexuel localisée dans le cerveau de souris sur Feminad pour les pipelines **SyN** et **OP-sym**. L'analyse VBM indique les régions significativement ($p \leq 0.05$ corrigé FDR) plus volumineuses chez la femelle (en rouge) ou chez le mâle (en bleu). Les régions sont superposées sur l'atlas Allen. Les quatre coupes coronales sont présentées par les lignes bleues sur la vue sagittale.

cérébral, le cortex cérébelleux ainsi que le noyau du lit de la strie terminale sont conformes aux trois études portant sur le dimorphisme sexuel.

Toutefois, les études citées présentent aussi entre elles des différences de résultats invoquant la résolution des images, le contraste ainsi que la différence des techniques d'analyses comme facteurs de ces différences. Dans notre étude, seule la méthode de recalage diffère entre les deux analyses et les résultats sont pourtant très différents. Les différences entre toutes ces études mettent en évidence le besoin d'un outil de validation pour l'analyse VBM.

6.4 Conclusion

Dans ce chapitre, nous avons comparé deux pipelines pour l'analyse VBM : l'un reposant sur des méthodes conventionnelles et l'autre exploitant les approches de deep learning développées dans cette thèse. Ces pipelines ont été appliqués au jeu de données Feminad, portant sur le dimorphisme sexuel du cerveau de souris.

En nous appuyant sur les études existantes, nous avons évalué la pertinence

des analyses issues des deux pipelines. Alors que le pipeline SyN identifie peu de régions présentant des différences significatives, le pipeline basé sur **OP-sym** détecte des variations subtiles, en grande partie cohérentes avec la littérature. Toutefois, l'absence d'une validation quantitative rigoureuse limite la confirmation des résultats obtenus.

Conclusion

Cette thèse a exploré l'analyse morphométrique du cerveau de souris à partir d'images obtenues par Imagerie par Résonance Magnétique (IRM). L'objectif principal était d'intégrer des approches d'Apprentissage Profond (Deep Learning, DL) pour résoudre trois problématiques majeures : l'extraction du cerveau, la segmentation des tissus cérébraux et le recalage d'images dans le contexte des études de groupes en Morphométrie Basée Voxel (Voxel-Based Morphometry, VBM). Un axe central de ce travail a été l'identification des limitations des méthodes existantes et leur adaptation dans le cadre particulier de l'analyse morphométrique appliquée aux modèles murins.

Concernant les tâches d'extraction et de segmentation, l'application des méthodes issues de l'état de l'art s'est révélée suffisante pour atteindre des performances satisfaisantes en matière de qualité et de rapidité d'exécution. En revanche, le recalage d'images a nécessité le développement d'une nouvelle approche, dénommée **OP**, présentée au chapitre 4. Cette méthode hybride, située entre le recalage par paires et l'apprentissage profond, a été conçue pour le recalage d'images en contexte de groupes. En optimisant simultanément les poids du réseau sur l'ensemble du groupe à analyser, **OP** exploite les déformations communes pour améliorer la précision du recalage. L'évaluation de **OP** sur un jeu de données annotées a montré qu'elle surpassait les méthodes classiques de recalage par paires ainsi que les approches d'apprentissage profond traditionnelles. De plus, lorsqu'elle est appliquée à un modèle pré-entraîné, **OP** offre des performances équivalentes aux méthodes par paires avec un coût de calcul

significativement réduit. Cette approche prolonge ainsi les travaux sur les *Deep Image Prior* en les adaptant au recalage de groupes.

Afin d'améliorer la régularisation des déformations, une version symétrique de **OP**, dénommée **OP-sym**, a été introduite. Elle intègre des contraintes de régularisation, développées au chapitre 5, permettant d'obtenir des transformations plus régulières, avec des jacobiens comparables à ceux des méthodes conventionnelles. L'évaluation qualitative a mis en évidence que cette optimisation globale du recalage améliore la stabilité et la cohérence des déformations.

Enfin, **OP-sym** a été appliquée pour réaliser une analyse VBM du dimorphisme sexuel du cerveau de souris et est comparée à SyN (Symmetric Image Normalization). Alors que SyN a identifié peu de différences significatives entre les groupes, **OP-sym** a démontré une sensibilité accrue aux variations subtiles des structures cérébrales, corroborant certaines études préexistantes tout en soulignant des différences inattendues sur certaines régions.

Malgré les avancées obtenues, plusieurs limitations demeurent. L'efficacité des méthodes de segmentation reste fortement dépendante de la qualité des vérités terrain, et l'absence d'annotations de référence complètes complexifie l'évaluation des performances. De plus, l'application du recalage pour la segmentation par transfert d'atlas a été limitée à un seul atlas, alors que les méthodes conventionnelles utilisent généralement plusieurs atlas pour améliorer la robustesse. Toutefois, bien que l'approche **OP-sym** permette d'obtenir des déformations plus régulières et proches des méthodes conventionnelles, son application à l'analyse VBM n'a été évaluée que qualitativement.

Ainsi, plusieurs axes de recherche futurs pourraient prolonger ces travaux. Il serait intéressant d'explorer des approches semi-supervisées ou auto-supervisées afin d'exploiter au mieux les possibilités offertes par les Réseaux Neuraux Convolutifs (Convolutional Neural Networks, CNN) pour le recalage. Par ailleurs, une étude plus approfondie des architectures neuronales plus adaptées à la modélisation des déformations pourrait permettre d'améliorer encore la précision du recalage. Une hypothèse serait d'utiliser une architecture de type Transformers avec un recalage hiérarchique (comme [186]), d'ajouter une semi-supervision grâce à des labels sur le cerveau ou des régions pour se démarquer des méthodes conventionnelles (comme préconisé par [185]), puis de l'optimiser

avec l'approche **OP**. Concernant l'analyse en elle-même, la mise en place d'une évaluation quantitative systématique des différents pipelines d'analyse VBM est essentielle afin d'éviter la variabilité des résultats selon les outils utilisés. Une étude plus approfondie des analyses VBM avec toutes les approches présentées dans les chapitres 4 et 5 doit être conduite. De plus, omis dans ces travaux et considéré comme prétraitement, le recalage affine aurait mérité d'être étudié plus en détail. Si nous avons ici fait l'hypothèse qu'un bon recalage déformable compense les petites erreurs du recalage affine, l'effet inverse est aussi vrai. Avec un meilleur recalage affine, le recalage déformable se contente de recaler l'anatomie interne au cerveau et non de compenser les erreurs de l'effet global. Il aurait été pertinent d'analyser l'effet de plusieurs méthodes de recalage affine sur les résultats du recalage déformable, ainsi que de déterminer si l'implémentation de méthodes basées sur l'apprentissage profond est bénéfique dans ce cadre.

Enfin, nous espérons que les approches d'apprentissage profond non supervisées seront prochainement intégrées dans des logiciels spécifiques à l'analyse morphométrique des cerveaux de souris, comme cela commence à être le cas pour l'IRM cérébrale humaine. Cette transition faciliterait l'adoption de ces méthodes dans les études précliniques et permettrait une analyse plus fine des structures cérébrales, contribuant ainsi à la recherche biomédicale.

Bibliographie

- [1] Serge O. DUMOULIN, Alessio FRACASSO, Wietske van der ZWAAG, Jeroen C. W. SIERO et Natalia PETRIDOU. « Ultra-high field MRI : Advancing systems neuroscience towards mesoscopic human brain function ». eng. In : *NeuroImage* 168 (mar. 2018), p. 345-357. DOI : 10.1016/j.neuroimage.2017.01.028.
- [2] Aleksandar DENIC, Slobodan I. MACURA, Prasanna MISHRA, Jeffrey D. GAMEZ, Moses RODRIGUEZ et Istvan PIRKO. « MRI in Rodent Models of Brain Disorders ». In : *Neurotherapeutics* 8.1 (jan. 2011), p. 3-18. DOI : 10.1007/s13311-010-0002-4.
- [3] Juan Eugenio IGLESIAS et Mert R. SABUNCU. « Multi-atlas segmentation of biomedical images : A survey ». In : *Medical Image Analysis* 24.1 (août 2015), p. 205-219. DOI : 10.1016/j.media.2015.06.012.
- [4] Aristeidis SOTIRAS, Christos DAVATZIKOS et Nikos PARAGIOS. « Deformable Medical Image Registration : A Survey ». In : *IEEE Transactions on Medical Imaging* 32.7 (juil. 2013). Conference Name : IEEE Transactions on Medical Imaging, p. 1153-1190. DOI : 10.1109/TMI.2013.2265603.
- [5] Fnu NEHA, Deepshikha BHATI, Deepak Kumar SHUKLA, Sonavi Makarand DALVI, Nikolaos MANTZOU et Safa SHUBBAR. *U-Net in Medical Image Segmentation : A Review of Its Applications Across Modalities*. arXiv :2412.02242 [eess]. Déc. 2024. DOI : 10.48550/arXiv.2412.02242.
- [6] Ahmad HAMMOUDEH et Stéphane DUPONT. *Deep learning in medical image registration : introduction and survey*. arXiv :2309.00727 [eess]. Jan. 2024. DOI : 10.48550/arXiv.2309.00727.
- [7] Asif T. CHINWALLA et al. « Initial sequencing and comparative analysis of the mouse genome ». en. In : *Nature* 420.6915 (déc. 2002). Publisher : Nature Publishing Group, p. 520-562. DOI : 10.1038/nature01262.

- [8] Simon FESTING et Robin WILKINSON. « The ethics of animal research. Talking Point on the use of animals in scientific research ». In : *EMBO Reports* 8.6 (juin 2007), p. 526-530. DOI : 10.1038/sj.embor.7400993.
- [9] P. MUKHERJEE, S. ROY, D. GHOSH et S. K. NANDI. « Role of animal models in biomedical research : a review ». en. In : *Laboratory Animal Research* 38.1 (juil. 2022), p. 18. DOI : 10.1186/s42826-022-00128-1.
- [10] Benjamin W. LINDSEY et Vincent TROPEPE. « A comparative framework for understanding the biological principles of adult neurogenesis ». In : *Progress in Neurobiology* 80.6 (déc. 2006), p. 281-307. DOI : 10.1016/j.pneurobio.2006.11.007.
- [11] P. C. LAUTERBUR. « Image Formation by Induced Local Interactions : Examples Employing Nuclear Magnetic Resonance ». en. In : *Nature* 242.5394 (mar. 1973). Publisher : Nature Publishing Group, p. 190-191. DOI : 10.1038/242190a0.
- [12] Nicholas A. BOCK, Norman B. KONYER et R. Mark HENKELMAN. « Multiple-mouse MRI ». en. In : *Magnetic Resonance in Medicine* 49.1 (2003). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.10326>, p. 158-167. DOI : 10.1002/mrm.10326.
- [13] Donald W. McROBBIE, Elizabeth A. MOORE, Martin J. GRAVES et Martin R. PRINCE. *MRI from Picture to Proton*. 2^e éd. Cambridge : Cambridge University Press, 2006. DOI : 10.1017/CB09780511545405.
- [14] Peter A. BANDETTINI. « Twenty years of functional MRI : the science and the stories ». eng. In : *NeuroImage* 62.2 (août 2012), p. 575-588. DOI : 10.1016/j.neuroimage.2012.04.026.
- [15] Robert W. BROWN, Y.-C. Norman CHENG, E. Mark HAACKE, Michael R. THOMPSON et Ramesh VENKATESAN. *Magnetic Resonance Imaging : Physical Principles and Sequence Design*. en. Google-Books-ID : Z45SAwAAQBAJ. John Wiley & Sons, juin 2014.
- [16] P. J. HORE. *Nuclear Magnetic Resonance*. en. Google-Books-ID : L9umCAAAQBAJ. Oxford University Press, 2015.
- [17] G. Allan JOHNSON, Alexandra BADEA, Jeffrey BRANDENBURG, Gary COFER, Boma FUBARA, Song LIU et Jonathan NISSANOV. « Waxholm Space : An image-based reference for coordinating mouse brain research ». In : *NeuroImage* 53.2 (nov. 2010), p. 365-372. DOI : 10.1016/j.neuroimage.2010.06.067.

- [18] Joseph V. HAJNAL, David J. BRYANT, Larry KASUBOSKI, Pradip M. PATTANY, Beatrice De COENE, Paul D. LEWIS, Jacqueline M. PENNOCK, Angela OATRIDGE, Ian R. YOUNG et Graeme M. BYDDER. « Use of Fluid Attenuated Inversion Recovery (FLAIR) Pulse Sequences in MRI of the Brain ». en-US. In : *Journal of Computer Assisted Tomography* 16.6 (déc. 1992), p. 841.
- [19] Katarzyna KRUPA et Monika BEKIESIŃSKA-FIGATOWSKA. « Artifacts in Magnetic Resonance Imaging ». In : *Polish Journal of Radiology* 80 (fév. 2015), p. 93-106. DOI : 10.12659/PJR.892628.
- [20] I. C. WRIGHT, P. K. MCGUIRE, J. -B. POLINE, J. M. TRAVERE, R. M. MURRAY, C. D. FRITH, R. S. J. FRACKOWIAK et K. J. FRISTON. « A Voxel-Based Method for the Statistical Analysis of Gray and White Matter Density Applied to Schizophrenia ». In : *NeuroImage* 2.4 (déc. 1995), p. 244-252. DOI : 10.1006/nimg.1995.1032.
- [21] Andrea MECHELLI, Cathy J. PRICE, Karl J. FRISTON et John ASHBURNER. « Voxel-Based Morphometry of the Human Brain : Methods and Applications ». In : *Current Medical Imaging Reviews* 1.2 (juin 2005), p. 105-113. DOI : 10.2174/1573405054038726.
- [22] Cristina SCARPAZZA et Maria DE SIMONE. « Voxel-based morphometry : current perspectives ». en. In : *Neuroscience and Neuroeconomics* Volume 5 (juil. 2016), p. 19-35. DOI : 10.2147/NAN.S66439.
- [23] Jonathan C. LAU, Jason P. LERCH, John G. SLED, R. Mark HENKELMAN, Alan C. EVANS et Barry J. BEDELL. « Longitudinal neuroanatomical changes determined by deformation-based morphometry in a mouse model of Alzheimer's disease ». In : *NeuroImage* 42.1 (août 2008), p. 19-27. DOI : 10.1016/j.neuroimage.2008.04.252.
- [24] S. J. SAWIAK, N. I. WOOD, G. B. WILLIAMS, A. J. MORTON et T. A. CARPENTER. « Voxel-based morphometry in the R6/2 transgenic mouse reveals differences between genotypes not seen with manual 2D morphometry ». In : *Neurobiology of Disease* 33.1 (jan. 2009), p. 20-27. DOI : 10.1016/j.nbd.2008.09.016.
- [25] P. Dwight TAPP, Kevin HEAD, Elizabeth HEAD, Norton W. MILGRAM, Bruce A. MUGGENBURG et Min-Ying SU. « Application of an automated voxel-based morphometry technique to assess regional gray and white matter brain atrophy in a canine model of aging ». In : *NeuroImage* 29.1 (jan. 2006), p. 234-244. DOI : 10.1016/j.neuroimage.2005.07.043.

- [26] Donald G. McLAREN, Kristopher J. Kosmatka, Erik K. Kastman, Barbara B. BENDLIN et Sterling C. JOHNSON. « Rhesus Macaque Brain Morphometry : A Methodological Comparison of Voxel-Wise Approaches ». In : *Methods (San Diego, Calif.)* 50.3 (mar. 2010), p. 157-165. DOI : 10.1016/j.ymeth.2009.10.003.
- [27] John ASHBURNER et Karl J. FRISTON. « Unified segmentation ». In : *NeuroImage* 26.3 (juil. 2005), p. 839-851. DOI : 10.1016/j.neuroimage.2005.02.018.
- [28] Ed S. LEIN et al. « Genome-wide atlas of gene expression in the adult mouse brain ». en. In : *Nature* 445.7124 (jan. 2007). Publisher : Nature Publishing Group, p. 168-176. DOI : 10.1038/nature05453.
- [29] K. FRISTON. « CHAPTER 2 - Statistical parametric mapping ». In : *Statistical Parametric Mapping*. Sous la dir. de KARL FRISTON, JOHN ASHBURNER, STEFAN KIEBEL, THOMAS NICHOLS et WILLIAM PENNY. London : Academic Press, jan. 2007, p. 10-31. DOI : 10.1016/B978-012372560-8/50002-4.
- [30] Luiz K. FERREIRA, Breno S. DINIZ, Orestes V. FORLENZA, Geraldo F. BUSATTO et Marcus V. ZANETTI. « Neurostructural predictors of Alzheimer's disease : A meta-analysis of VBM studies ». In : *Neurobiology of Aging* 32.10 (oct. 2011), p. 1733-1741. DOI : 10.1016/j.neurobiolaging.2009.11.008.
- [31] Ching-Hung LIN, Chun-Ming CHEN, Ming-Kuei LU, Chon-Haw TSAI, Jin-Chern CHIOU, Jan-Ray LIAO et Jeng-Ren DUANN. « VBM Reveals Brain Volume Differences between Parkinson's Disease and Essential Tremor Patients ». English. In : *Frontiers in Human Neuroscience* 7 (juin 2013). Publisher : Frontiers. DOI : 10.3389/fnhum.2013.00247.
- [32] G. DOUAUD, V. GAURA, M. -J. RIBEIRO, F. LETHIMONNIER, R. MAROY, C. VERNY, P. KRYSTKOWIAK, P. DAMIER, A. -C. BACHOUD-LEVI, P. HANTRAYE et P. REMY. « Distribution of grey matter atrophy in Huntington's disease patients : A combined ROI-based and voxel-based morphometric study ». In : *NeuroImage* 32.4 (oct. 2006), p. 1562-1575. DOI : 10.1016/j.neuroimage.2006.05.057.
- [33] Igor NENADIC, Raka MAITRA, Kerstin LANGBEIN, Maren DIETZEK, Carsten LORENZ, Stefan SMESNY, Jürgen R. REICHENBACH, Heinrich SAUER et Christian GASER. « Brain structure in schizophrenia vs. psychotic bipolar I disorder : A VBM study ». In : *Schizophrenia Research* 165.2 (juil. 2015), p. 212-219. DOI : 10.1016/j.schres.2015.04.007.

- [34] Meng XU, Jipeng LIU, Qingguo LIU, Yu GONG, Yinyin LI, Jing ZHANG, Shufeng SHI et Yuanyuan SHI. « Preliminary study on early diagnosis of Alzheimer's disease in APP/PS1 transgenic mice using multimodal magnetic resonance imaging ». English. In : *Frontiers in Aging Neuroscience* 16 (fév. 2024). Publisher : Frontiers. DOI : 10.3389/fnagi.2024.1326394.
- [35] Avinash CHANDRA, George DERVENOULAS, Marios POLITIS et FOR THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE. « Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment ». en. In : *Journal of Neurology* 266.6 (juin 2019), p. 1293-1302. DOI : 10.1007/s00415-018-9016-3.
- [36] Diego Alves Rodrigues de SOUZA, Hervé MATHIEU, Jean-Christophe DELOULME et Emmanuel L. BARBIER. « Evaluation of kernel low-rank compressed sensing in preclinical diffusion magnetic resonance imaging ». English. In : *Frontiers in Neuroscience* 17 (juin 2023). Publisher : Frontiers. DOI : 10.3389/fnins.2023.1172830.
- [37] Jean-Christophe DELOULME, Sylvie GORY-FAURÉ, Franck MAUCONDUIT, Sophie CHAUVET, Julie JONCKHEERE, Benoit BOULAN, Erik MIRE, Jing XUE, Marion JANY, Caroline MAUCLER, Agathe A. DEPARIS, Olivier MONTIGON, Alexia DAOUST, Emmanuel L. BARBIER, Christophe BOSC, Nicole DEGLON, Jacques BROCARD, Eric DENARIER, Isabelle LE BRUN, Karin PERNET-GALLAY, Isabelle VILGRAIN, Phillip J. ROBINSON, Hana LAHRECH, Fanny MANN et Annie ANDRIEUX. « Microtubule-associated protein 6 mediates neuronal connectivity through Semaphorin 3E-dependent signalling for axonal growth ». en. In : *Nature Communications* 6.1 (juin 2015). Publisher : Nature Publishing Group, p. 7246. DOI : 10.1038/ncomms8246.
- [38] Alistair T PAGNAMENTA, Pierre HEEMERYCK, Hilary C MARTIN, Christophe BOSC, Leticia PERIS, Ivy USZYNSKI, Sylvie GORY-FAURÉ, Simon COULY, Charu DESHPANDE, Ata SIDDIQUI, Alaa A ELMONAIRY, WGS500 CONSORTIUM, GENOMICS ENGLAND RESEARCH CONSORTIUM, Sandeep JAYAWANT, Sarada MURTHY, Ian WALKER, Lucy LOONG, Peter BAUER, Frédérique VOSSIER, Eric DENARIER, Tanguy MAURICE, Emmanuel L BARBIER, Jean-Christophe DELOULME, Jenny C TAYLOR, Edward M BLAIR, Annie ANDRIEUX et Marie-Jo MOUTIN. « Defective tubulin detyrosination causes structural brain abnormalities with cognitive deficiency in humans and mice ». In : *Human Molecular Genetics* 28.20 (oct. 2019), p. 3391-3405. DOI : 10.1093/hmg/ddz186.
- [39] S J SAWIAK, N I WOOD, G B WILLIAMS, A J MORTON et T A CARPENTER. « SPMouse : A new toolbox for SPM in the animal brain ». en. In : ().

- [40] Kazuyuki MEKADA, Kuniya ABE, Ayumi MURAKAMI, Satoe NAKAMURA, Hatsumi NAKATA, Kazuo MORIWAKI, Yuichi OBATA et Atsushi YOSHIKI. « Genetic differences among C57BL/6 substrains ». eng. In : *Experimental Animals* 58.2 (avr. 2009), p. 141-149. DOI : 10.1538/expanim.58.141.
- [41] Allan R. JONES, Caroline C. OVERLY et Susan M. SUNKIN. « The Allen Brain Atlas : 5 years and beyond ». en. In : *Nature Reviews Neuroscience* 10.11 (nov. 2009). Publisher : Nature Publishing Group, p. 821-828. DOI : 10.1038/nrn2722.
- [42] Nicholas J. TUSTISON, Brian B. AVANTS, Philip A. COOK, Yuanjie ZHENG, Alexander EGAN, Paul A. YUSHKEVICH et James C. GEE. « N4ITK : improved N3 bias correction ». eng. In : *IEEE transactions on medical imaging* 29.6 (juin 2010), p. 1310-1320. DOI : 10.1109/TMI.2010.2046908.
- [43] Stephen M. SMITH. « Fast robust automated brain extraction ». en. In : *Human Brain Mapping* 17.3 (2002). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.10062>. DOI : 10.1002/hbm.10062.
- [44] F. SÉGONNE, A. M. DALE, E. BUSA, M. GLESSNER, D. SALAT, H. K. HAHN et B. FISCHL. « A hybrid approach to the skull stripping problem in MRI ». eng. In : *NeuroImage* 22.3 (juil. 2004), p. 1060-1075. DOI : 10.1016/j.neuroimage.2004.03.032.
- [45] Simon F. ESKILDSEN, Pierrick COUPÉ, Vladimir FONOV, José V. MANJÓN, Kelvin K. LEUNG, Nicolas GUIZARD, Shafik N. WASSEF, Lasse Riis ØSTERGAARD et D. Louis COLLINS. « BEaST : Brain extraction based on nonlocal segmentation technique ». In : *NeuroImage* 59.3 (fév. 2012), p. 2362-2373. DOI : 10.1016/j.neuroimage.2011.09.012.
- [46] Kelvin K. LEUNG, Josephine BARNES, Marc MODAT, Gerard R. RIDGWAY, Jonathan W. BARTLETT, Nick C. FOX et Sébastien OURSELIN. « Brain MAPS : An automated, accurate and robust brain extraction technique using a template library ». In : *NeuroImage* 55.3 (avr. 2011), p. 1091-1108. DOI : 10.1016/j.neuroimage.2010.12.067.
- [47] Yikang LIU, Hayreddin Said UNSAL, Yi TAO et Nanyin ZHANG. « Automatic Brain Extraction for Rodent MRI Images ». In : *Neuroinformatics* 18.3 (juin 2020), p. 395-406. DOI : 10.1007/s12021-020-09453-z.
- [48] Georgios ANTONOPOULOS, Shammi MORE, Federico RAIMONDO, Simon B. EICKHOFF, Felix HOFFSTAEDTER et Kaustubh R. PATIL. « A systematic comparison of VBM pipelines and their application to age prediction ». In : *NeuroImage* 279 (oct. 2023), p. 120292. DOI : 10.1016/j.neuroimage.2023.120292.

- [49] Ipek OGUZ, Honghai ZHANG, Ashley RUMPLE et Milan SONKA. « RATS : Rapid Automatic Tissue Segmentation in rodent brain MRI ». In : *Journal of Neuroscience Methods* 221 (jan. 2014), p. 175-182. DOI : 10.1016/j.jneumeth.2013.09.021.
- [50] Nigel CHOU, Jiarong WU, Jordan BAI BINGREN, Anqi QIU et Kai-Hsiang CHUANG. « Robust Automatic Rodent Brain Extraction Using 3-D Pulse-Coupled Neural Networks (PCNN) ». In : *IEEE Transactions on Image Processing* 20.9 (sept. 2011). Conference Name : IEEE Transactions on Image Processing, p. 2554-2564. DOI : 10.1109/TIP.2011.2126587.
- [51] Ziqi YU, Xiaoyang HAN, Shengjie ZHANG, Jianfeng FENG, Tingying PENG et Xiao-Yong ZHANG. « MouseGAN++ : Unsupervised Disentanglement and Contrastive Representation for Multiple MRI Modalities Synthesis and Structural Segmentation of Mouse Brain ». eng. In : *IEEE transactions on medical imaging* 42.4 (avr. 2023), p. 1197-1209. DOI : 10.1109/TMI.2022.3225528.
- [52] Riccardo DE FEO, Artem SHATILLO, Alejandra SIERRA, Juan Miguel VALVERDE, Olli GRÖHN, Federico GIOVE et Jussi TOHKA. « Automated joint skull-stripping and segmentation with Multi-Task U-Net in large mouse brain MRI databases ». eng. In : *NeuroImage* 229 (avr. 2021), p. 117734. DOI : 10.1016/j.neuroimage.2021.117734.
- [53] B. B. AVANTS, C. L. EPSTEIN, M. GROSSMAN et J. C. GEE. « Symmetric diffeomorphic image registration with cross-correlation : evaluating automated labeling of elderly and neurodegenerative brain ». eng. In : *Medical Image Analysis* 12.1 (fév. 2008), p. 26-41. DOI : 10.1016/j.media.2007.06.004.
- [54] Tom VERCAUTEREN, Xavier PENNEC, Aymeric PERCHANT et Nicholas AYACHE. « Diffeomorphic demons : Efficient non-parametric image registration ». In : *NeuroImage. Mathematics in Brain Imaging* 45.1, Supplement 1 (mar. 2009), S61-S72. DOI : 10.1016/j.neuroimage.2008.10.040.
- [55] Arno KLEIN, Jesper ANDERSSON, Babak A. ARDEKANI, John ASHBURNER, Brian AVANTS, Ming-Chang CHIANG, Gary E. CHRISTENSEN, D. Louis COLLINS, James GEE, Pierre HELLIER, Joo Hyun SONG, Mark JENKINSON, Claude LEPAGE, Daniel RUECKERT, Paul THOMPSON, Tom VERCAUTEREN, Roger P. WOODS, J. John MANN et Ramin V. PARSEY. « Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration ». In : *NeuroImage* 46.3 (juil. 2009), p. 786-802. DOI : 10.1016/j.neuroimage.2008.12.037.

- [56] Matthijs C. van EEDE, Jan SCHOLZ, M. Mallar CHAKRAVARTY, R. Mark HENKELMAN et Jason P. LERCH. « Mapping registration sensitivity in MR mouse brain images ». eng. In : *NeuroImage* 82 (nov. 2013), p. 226-236. DOI : 10.1016/j.neuroimage.2013.06.004.
- [57] Patricia MARTÍNEZ-TAZO, Alexandra SANTOS, Mohamed Kotb SELIM, Elena ESPINÓS-SOLER et Silvia DE SANTIS. « Sex matters : The MouseX DW-ALLEN Atlas for mice diffusion-weighted MR imaging ». In : *NeuroImage* 292 (avr. 2024), p. 120573. DOI : 10.1016/j.neuroimage.2024.120573.
- [58] Brian B. AVANTS, Paul YUSHKEVICH, John PLUTA, David MINKOFF, Marc KORCZYKOWSKI, John DETRE et James C. GEE. « The optimal template effect in hippocampus studies of diseased populations ». eng. In : *NeuroImage* 49.3 (fév. 2010), p. 2457-2466. DOI : 10.1016/j.neuroimage.2009.09.062.
- [59] Dorothée V. CALLAERT, Annemie RIBBENS, Frederik MAES, Stephan P. SWINNEN et Nicole WENDEROTH. « Assessing age-related gray matter decline with voxel-based morphometry depends significantly on segmentation and normalization procedures ». English. In : *Frontiers in Aging Neuroscience* 6 (juin 2014). Publisher : Frontiers. DOI : 10.3389/fnagi.2014.00124.
- [60] Venkateswaran RAJAGOPALAN et Erik P. PIORO. « Disparate voxel based morphometry (VBM) results between SPM and FSL softwares in ALS patients with frontotemporal dementia : which VBM results to consider ? » In : *BMC Neurology* 15.1 (mar. 2015), p. 32. DOI : 10.1186/s12883-015-0274-8.
- [61] Logan HAYNES, Amanda IP, Ivy Y. K. CHO, Dennis DIMOND, Christiane S. ROHR, Mercedes BAGSHAW, Deborah DEWEY, Catherine LEBEL et Signe BRAY. « Grey and white matter volumes in early childhood : A comparison of voxel-based morphometry pipelines ». In : *Developmental Cognitive Neuroscience* 46 (déc. 2020), p. 100875. DOI : 10.1016/j.dcn.2020.100875.
- [62] Xinqi ZHOU, Renjing WU, Yixu ZENG, Ziyu QI, Stefania FERRARO, Lei XU, Xiaoxiao ZHENG, Jialin LI, Meina FU, Shuxia YAO, Keith M. KENDRICK et Benjamin BECKER. « Choice of Voxel-based Morphometry processing pipeline drives variability in the location of neuroanatomical brain markers ». en. In : *Communications Biology* 5.1 (sept. 2022). Publisher : Nature Publishing Group, p. 1-12. DOI : 10.1038/s42003-022-03880-1.
- [63] John ASHBURNER. « A fast diffeomorphic image registration algorithm ». eng. In : *NeuroImage* 38.1 (oct. 2007), p. 95-113. DOI : 10.1016/j.neuroimage.2007.07.007.

- [64] Delfina BRAGGIO, Jimena BARBEITO-ANDRÉS, Paula GONZALEZ et Ignacio LARRABIDE. « Fine tuning VBM for mouse brain analysis : model adjustment using atrophy simulation ». In : *14th International Symposium on Medical Information Processing and Analysis*. T. 10975. SPIE, déc. 2018, p. 160-169. DOI : 10.1117/12.2506634.
- [65] Delfina BRAGGIO, Jimena BARBEITO-ANDRÉS, Paula GONZALEZ, Benedikt HALLGRÍMSSON et Ignacio LARRABIDE. « VBM sensitivity to localization and extent of mouse brain lesions : A simulation approach ». In : *Computer Methods and Programs in Biomedicine* 196 (nov. 2020), p. 105636. DOI : 10.1016/j.cmpb.2020.105636.
- [66] Robert J. ANDERSON, James J. COOK, Natalie DELPRATT, John C. NOULS, Bin GU, James O. MCNAMARA, Brian B. AVANTS, G. Allan JOHNSON et Alexandra BADEA. « Small Animal Multivariate Brain Analysis (SAMBA) – a High Throughput Pipeline with a Validation Framework ». en. In : *Neuroinformatics* 17.3 (juil. 2019), p. 451-472. DOI : 10.1007/s12021-018-9410-0.
- [67] Gustav MÅRTENSSON, Daniel FERREIRA, Tobias GRANBERG, Lena CAVALLIN, Ketil OPPEDAL, Alessandro PADOVANI, Irena REKTOROVA, Laura BONANNI, Matteo PARDINI, Milica G. KRAMBERGER, John-Paul TAYLOR, Jakub HORT, Jón SNÆDAL, Jaime KULISEVSKY, Frederic BLANC, Angelo ANTONINI, Patrizia MECOCCHI, Bruno VELLAS, Magda TSOLAKI, Iwona KŁOSZEWSKA, Hilka SOININEN, Simon LOVESTONE, Andrew SIMMONS, Dag AARSLAND et Eric WESTMAN. « The reliability of a deep learning model in clinical out-of-distribution MRI data : A multicohort study ». eng. In : *Medical Image Analysis* 66 (déc. 2020), p. 101714. DOI : 10.1016/j.media.2020.101714.
- [68] Lukas FISCH, Nils R. WINTER, Janik GOLTERMANN, Carlotta BARKHAU, Daniel EMDEN, Jan ERNSTING, Maximilian KONOWSKI, Ramona LEENINGS, Tiana BORBERS, Kira FLINKENFLÜGEL, Dominik GROTEGERD, Anna KRAUS, Elisabeth J. LEEHR, Susanne MEINERT, Frederike STEIN, Lea TEUTENBERG, Florian THOMAS-ODENTHAL, Paula USEMANN, Marco HERMESDORF, Hamidreza JAMALABADI, Andreas JANSEN, Igor NENADIC, Benjamin STRAUBE, Tilo KIRCHER, Klaus BERGER, Benjamin RISSE, Udo DANNLOWSKI et Tim HAHN. *deepmrip : Voxel-based Morphometry (VBM) Preprocessing via Deep Neural Networks*. arXiv :2408.10656 [eess]. Oct. 2024. DOI : 10.48550/arXiv.2408.10656.
- [69] D. W. SHATTUCK, S. R. SANDOR-LEAHY, K. A. SCHAPER, D. A. ROTTENBERG et R. M. LEAHY. « Magnetic resonance image tissue classification using a partial volume model ». eng. In : *NeuroImage* 13.5 (mai 2001), p. 856-876. DOI : 10.1006/nimg.2000.0730.

- [70] Y. ZHANG, M. BRADY et S. SMITH. « Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm ». en. In : *IEEE Transactions on Medical Imaging* 20.1 (jan. 2001), p. 45-57. DOI : 10.1109/42.906424.
- [71] Bjoern H. MENZE, Andras JAKAB, Stefan BAUER, Jayashree KALPATHY-CRAMER, Keyvan FARAHANI, Justin KIRBY, Yuliya BURREN, Nicole PORZ, Johannes SLOTBOOM, Roland WIEST, Levente LANCZI, Elizabeth GERSTNER, Marc-André WEBER, Tal ARBEL, Brian B. AVANTS, Nicholas AYACHE, Patricia BUENDIA, D. Louis COLLINS, Nicolas CORDIER, Jason J. CORSO, Antonio CRIMINISI, Tilak DAS, Hervé DELINGETTE, Çağatay DEMIRALP, Christopher R. DURST, Michel DOJAT, Senan DOYLE, Joana FESTA, Florence FORBES, Ezequiel GEREMIA, Ben GLOCKER, Polina GOLLAND, Xiaotao GUO, Andac HAMAMCI, Khan M. IFTEKHARUDDIN, Raj JENA, Nigel M. JOHN, Ender KONUKOGLU, Danial LASHKARI, José António MARIZ, Raphael MEIER, Sérgio PEREIRA, Doina PRECUP, Stephen J. PRICE, Tammy Riklin RAVIV, Syed M. S. REZA, Michael RYAN, Duygu SARIKAYA, Lawrence SCHWARTZ, Hoo-Chang SHIN, Jamie SHOTTON, Carlos A. SILVA, Nuno SOUSA, Nagesh K. SUBBANNA, Gabor SZEKELY, Thomas J. TAYLOR, Owen M. THOMAS, Nicholas J. TUSTISON, Gozde UNAL, Flor VASSEUR, Max WINTERMARK, Dong Hye YE, Liang ZHAO, Binsheng ZHAO, Darko ZIKIC, Marcel PRASTAWA, Mauricio REYES et Koen VAN LEEMPUT. « The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) ». In : *IEEE Transactions on Medical Imaging* 34.10 (oct. 2015). Conference Name : IEEE Transactions on Medical Imaging, p. 1993-2024. DOI : 10.1109/TMI.2014.2377694.
- [72] Mohsen GHAFORIAN, Nico KARSSEMEIJER, Tom HESKES, Mayra BERGKAMP, Joost WISSINK, Jiri OBELS, Karlijn KEIZER, Frank-Erik de LEEUW, Bram van GINNEKEN, Elena MARCHIORI et Bram PLATEL. « Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin ». In : *NeuroImage : Clinical* 14 (jan. 2017), p. 391-399. DOI : 10.1016/j.nicl.2017.01.033.
- [73] Guotai WANG, Maria A. ZULUAGA, Wenqi LI, Rosalind PRATT, Premal A. PATEL, Michael AERTSEN, Tom DOEL, Anna L. DAVID, Jan DEPREST, Sebastien OURSELIN et Tom VERCAUTEREN. « DeepIGeoS : A Deep Interactive Geodesic Framework for Medical Image Segmentation ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.7 (juil. 2019), p. 1559-1572. DOI : 10.1109/TPAMI.2018.2840695.
- [74] Saleha MASOOD, Muhammad SHARIF, Afifa MASOOD, Mussarat YASMIN et Mudassar RAZA. « A Survey on Medical Image Segmentation ». In : *Current Medical Imaging Reviews* 11.1 (fév. 2015), p. 3-14.

- [75] Risheng WANG, Tao LEI, Ruixia CUI, Bingtao ZHANG, Hongying MENG et Asoke K. NANDI. *Medical Image Segmentation Using Deep Learning : A Survey*. arXiv :2009.13120 [eess]. Déc. 2021. DOI : 10.1049/ipr2.12419.
- [76] Yuncong FENG, Haiying ZHAO, Xiongfei LI, Xiaoli ZHANG et Hongpeng LI. « A multi-scale 3D Otsu thresholding algorithm for medical image segmentation ». In : *Digital Signal Processing* 60 (jan. 2017), p. 186-199. DOI : 10.1016/j.dsp.2016.08.003.
- [77] Chi-Hoon LEE, Mark SCHMIDT, Albert MURTHA, Aalo BISTRITZ, Jörg SANDER et Russell GREINER. « Segmenting Brain Tumors with Conditional Random Fields and Support Vector Machines ». en. In : *Computer Vision for Biomedical Image Applications*. Sous la dir. d'Yanxi LIU, Tianzi JIANG et Changshui ZHANG. Berlin, Heidelberg : Springer, 2005, p. 469-478. DOI : 10.1007/11569541_47.
- [78] Raouia AYACHI et Nahla BEN AMOR. « Brain Tumor Segmentation Using Support Vector Machines ». en. In : *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Sous la dir. de Claudio SOSSAI et Gaetano CHEMELLO. Berlin, Heidelberg : Springer, 2009, p. 736-747. DOI : 10.1007/978-3-642-02906-6_63.
- [79] László LEFKOVITS, Szidónia LEFKOVITS et László SZILÁGYI. « Brain Tumor Segmentation with Optimized Random Forest ». en. In : *Brainlesion : Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Sous la dir. d'Alessandro CRIMI, Bjoern MENZE, Oskar MAIER, Mauricio REYES, Stefan WINZECK et Heinz HANDELS. Cham : Springer International Publishing, 2016, p. 88-99. DOI : 10.1007/978-3-319-55524-9_9.
- [80] Michael KASS, Andrew WITKIN et Demetri TERZOPOULOS. « Snakes : Active contour models ». en. In : *International Journal of Computer Vision* 1.4 (jan. 1988), p. 321-331. DOI : 10.1007/BF00133570.
- [81] Xiaoli ZHANG, Xiongfei LI et Yuncong FENG. « A medical image segmentation algorithm based on bi-directional region growing ». In : *Optik* 126.20 (oct. 2015), p. 2398-2404. DOI : 10.1016/j.ijleo.2015.06.011.
- [82] Y. BOYKOV, O. VEKSLER et R. ZABIH. « Fast approximate energy minimization via graph cuts ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.11 (nov. 2001). Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1222-1239. DOI : 10.1109/34.969114.

- [83] Stuart GEMAN et Donald GEMAN. « Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6.6 (nov. 1984). Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 721-741. DOI : 10.1109/TPAMI.1984.4767596.
- [84] R. MALLADI, J.A. SETHIAN et B.C. VEMURI. « Shape modeling with front propagation : a level set approach ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17.2 (fév. 1995). Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 158-175. DOI : 10.1109/34.368173.
- [85] T.F. CHAN et L.A. VESE. « Active contours without edges ». In : *IEEE Transactions on Image Processing* 10.2 (fév. 2001). Conference Name : IEEE Transactions on Image Processing, p. 266-277. DOI : 10.1109/83.902291.
- [86] K. K. D. RAMESH, G. Kiran KUMAR, K. SWAPNA, Debabrata DATTA et S. Suman RAJEST. « A Review of Medical Image Segmentation Algorithms ». en. In : *EAI Endorsed Transactions on Pervasive Health and Technology* 7.27 (avr. 2021). Number : 27, e6-e6. DOI : 10.4108/eai.12-4-2021.169184.
- [87] Yan XU, Rixiang QUAN, Weiting XU, Yi HUANG, Xiaolong CHEN et Fengyuan LIU. « Advances in Medical Image Segmentation : A Comprehensive Review of Traditional, Deep Learning and Hybrid Approaches ». en. In : *Bioengineering* 11.10 (oct. 2024). Number : 10 Publisher : Multidisciplinary Digital Publishing Institute, p. 1034. DOI : 10.3390/bioengineering11101034.
- [88] Prachee TYAGI, Tripty SINGH, Ravi NAYAR et Shiv KUMAR. « Performance comparison and analysis of medical image segmentation techniques ». In : *2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*. Fév. 2018, p. 1-6. DOI : 10.1109/ICCTAC.2018.8370392.
- [89] Sandra JARDIM, João ANTÓNIO et Carlos MORA. « Image thresholding approaches for medical image segmentation - short literature review ». In : *Procedia Computer Science*. CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2022 219 (jan. 2023), p. 1485-1492. DOI : 10.1016/j.procs.2023.01.439.
- [90] Jyrki MP. LÖTJÖNEN, Robin WOLZ, Juha R. KOIKKALAINEN, Lennart THURFJELL, Gunhild WALDEMAR, Hilka SOININEN et Daniel RUECKERT. « Fast and robust multi-atlas segmentation of brain magnetic resonance images ». In :

- NeuroImage* 49.3 (fév. 2010), p. 2352-2365. DOI : 10.1016/j.neuroimage.2009.10.026.
- [91] Chen ZU, Zhengxia WANG, Daoqiang ZHANG, Peipeng LIANG, Yonghong SHI, Dinggang SHEN et Guorong WU. « Robust multi-atlas label propagation by deep sparse representation ». In : *Pattern Recognition* 63 (mar. 2017), p. 511-517. DOI : 10.1016/j.patcog.2016.09.028.
- [92] Arsen PLAKSYVYI, Maria SKUBLEWSKA-PASZKOWSKA et Pawel POWROZNIK. « A Comparative Analysis of Image Segmentation Using Classical and Deep Learning Approach ». en. In : *Advances in Science and Technology Research Journal* 17.6 (déc. 2023), p. 127-139. DOI : 10.12913/22998624/172771.
- [93] Wenlu ZHANG, Rongjian LI, Houtao DENG, Li WANG, Weili LIN, Shuiwang JI et Dinggang SHEN. « Deep convolutional neural networks for multi-modality isointense infant brain image segmentation ». In : *NeuroImage* 108 (mar. 2015), p. 214-224. DOI : 10.1016/j.neuroimage.2014.12.061.
- [94] Jose DOLZ, Christian DESROSIERS et Ismail BEN AYED. « 3D fully convolutional networks for subcortical segmentation in MRI : A large-scale study ». In : *NeuroImage. Segmenting the Brain* 170 (avr. 2018), p. 456-470. DOI : 10.1016/j.neuroimage.2017.04.039.
- [95] Ian GOODFELLOW, Yoshua BENGIO et Aaron COURVILLE. *Deep Learning*. MIT Press, 2016.
- [96] Olaf RONNEBERGER, Philipp FISCHER et Thomas BROX. *U-Net : Convolutional Networks for Biomedical Image Segmentation*. arXiv :1505.04597 [cs]. Mai 2015. DOI : 10.48550/arXiv.1505.04597.
- [97] Jonathan LONG, Evan SHELHAMER et Trevor DARRELL. « Fully convolutional networks for semantic segmentation ». English. In : ISSN : 1063-6919. IEEE Computer Society, juin 2015, p. 3431-3440. DOI : 10.1109/CVPR.2015.7298965.
- [98] Hyeonwoo NOH, Seunghoon HONG et Bohyung HAN. « Learning Deconvolution Network for Semantic Segmentation ». In : 2015, p. 1520-1528.
- [99] Özgün ÇİÇEK, Ahmed ABDULKADIR, Soeren S. LIENKAMP, Thomas BROX et Olaf RONNEBERGER. « 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation ». en. In : *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Sous la dir. de Sebastien OURSELIN, Leo JOSKOWICZ, Mert R. SABUNCU, Gozde UNAL et William WELLS. Cham : Springer International Publishing, 2016, p. 424-432. DOI : 10.1007/978-3-319-46723-8_49.

- [100] Eric KERFOOT, James CLOUGH, Ilkay OKSUZ, Jack LEE, Andrew P. KING et Julia A. SCHNABEL. « Left-Ventricle Quantification Using Residual U-Net : 9th International Workshop on Statistical Atlases and Computational Models of the Heart : Atrial Segmentation and LV Quantification Challenges, STACOM 2018, held in conjunction with Medical Image Computing and Computer-Assisted Intervention, MICCAI 2018 ». In : *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges - 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Revised Selected Papers*. Sous la dir. d'Alistair YOUNG, Kawal RHODE, Mihaela POP, Jichao ZHAO, Kristin McLEOD, Shuo LI, Maxime SERMESANT et Tommaso MANSI. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, jan. 2019, p. 371-380. DOI : 10.1007/978-3-030-12029-0_40.
- [101] Nabil IBTEHAZ et M. Sohel RAHMAN. « MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation ». In : *Neural Networks* 121 (jan. 2020), p. 74-87. DOI : 10.1016/j.neunet.2019.08.025.
- [102] Zongwei ZHOU, Md Mahfuzur RAHMAN SIDDIQUEE, Nima TAJBAKHSH et Jianming LIANG. « UNet++ : A Nested U-Net Architecture for Medical Image Segmentation ». en. In : *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Sous la dir. de Danail STOYANOV, Zeike TAYLOR, Gustavo CARNEIRO, Tanveer SYEDA-MAHMOOD, Anne MARTEL, Lena MAIER-HEIN, João Manuel R.S. TAVARES, Andrew BRADLEY, João Paulo PAPA, Vasileios BELAGIANNIS, Jacinto C. NASCIMENTO, Zhi LU, Sailesh CONJETI, Mehdi MORADI, Hayit GREENSPAN et Anant MADABHUSHI. Cham : Springer International Publishing, 2018, p. 3-11. DOI : 10.1007/978-3-030-00889-5_1.
- [103] Md Zahangir ALOM, Chris YAKOPCIC, Mahmudul HASAN, Tarek M. TAHA et Vijayan K. ASARI. « Recurrent residual U-Net for medical image segmentation ». In : *Journal of Medical Imaging* 6.1 (mar. 2019). Publisher : SPIE, p. 014006. DOI : 10.1117/1.JMI.6.1.014006.
- [104] Alexey DOSOVITSKIY, Lucas BEYER, Alexander KOLESNIKOV, Dirk WEISSENORN, Xiaohua ZHAI, Thomas UNTERTHINER, Mostafa DEGHANI, Matthias MINDERER, Georg HEIGOLD, Sylvain GELLY, Jakob USZKOREIT et Neil HOULSBY. *An Image is Worth 16x16 Words : Transformers for Image Recognition at Scale*. arXiv :2010.11929 [cs]. Juin 2021. DOI : 10.48550/arXiv.2010.11929.
- [105] Ashish VASWANI, Noam SHAZEER, Niki PARMAR, Jakob USZKOREIT, Llion JONES, Aidan N. GOMEZ, Lukasz KAISER et Illia POLOSUKHIN. *Attention Is*

- All You Need*. arXiv :1706.03762 [cs]. Août 2023. DOI : 10.48550/arXiv.1706.03762.
- [106] Hu CAO, Yueyue WANG, Joy CHEN, Dongsheng JIANG, Xiaopeng ZHANG, Qi TIAN et Manning WANG. « Swin-Unet : Unet-Like Pure Transformer for Medical Image Segmentation ». en. In : *Computer Vision – ECCV 2022 Workshops*. Sous la dir. de Leonid KARLINSKY, Tomer MICHAELI et Ko NISHINO. Cham : Springer Nature Switzerland, 2023, p. 205-218. DOI : 10.1007/978-3-031-25066-8_9.
- [107] Youyang SHA, Yonghong ZHANG, Xuquan Ji et Lei HU. *Transformer-Unet : Raw Image Processing with Unet*. arXiv :2109.08417 [eess]. Sept. 2021. DOI : 10.48550/arXiv.2109.08417.
- [108] Ali HATAMIZADEH, Yucheng TANG, Vishwesh NATH, Dong YANG, Andriy MYRONENKO, Bennett LANDMAN, Holger R. ROTH et Daguang XU. « UNETR : Transformers for 3D Medical Image Segmentation ». en. In : 2022, p. 574-584.
- [109] Reza AZAD, Ehsan Khodapanah AGHDAM, Amelie RAULAND, Yiwei JIA, Atlas Haddadi AVVAL, Afshin BOZORGPOUR, Sanaz KARIMIJAFAFBIGLOO, Joseph Paul COHEN, Ehsan ADELI et Dorit MERHOF. « Medical Image Segmentation Review : The Success of U-Net ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.12 (déc. 2024). Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 10076-10095. DOI : 10.1109/TPAMI.2024.3435571.
- [110] Nahian SIDDIQUE, Paheding SIDIKE, Colin ELKIN et Vijay DEVABHAKTUNI. « U-Net and its variants for medical image segmentation : theory and applications ». In : *IEEE Access* 9 (2021). arXiv :2011.01118 [eess], p. 82031-82057. DOI : 10.1109/ACCESS.2021.3086020.
- [111] Albert GU et Tri DAO. *Mamba : Linear-Time Sequence Modeling with Selective State Spaces*. arXiv :2312.00752 [cs]. Mai 2024. DOI : 10.48550/arXiv.2312.00752.
- [112] Lianghui ZHU, Bencheng LIAO, Qian ZHANG, Xinlong WANG, Wenyu LIU et Xinggang WANG. *Vision Mamba : Efficient Visual Representation Learning with Bidirectional State Space Model*. arXiv :2401.09417 [cs]. Nov. 2024. DOI : 10.48550/arXiv.2401.09417.
- [113] Ali HATAMIZADEH et Jan KAUTZ. *MambaVision : A Hybrid Mamba-Transformer Vision Backbone*. arXiv :2407.08083 [cs]. Juil. 2024. DOI : 10.48550/arXiv.2407.08083.

- [114] Yue LIU, Yunjie TIAN, Yuzhong ZHAO, Hongtian YU, Lingxi XIE, Yaowei WANG, Qixiang YE, Jianbin JIAO et Yunfan LIU. *VMamba : Visual State Space Model*. arXiv :2401.10166 [cs]. Déc. 2024. DOI : 10.48550/arXiv.2401.10166.
- [115] Jun MA, Feifei LI et Bo WANG. *U-Mamba : Enhancing Long-range Dependency for Biomedical Image Segmentation*. arXiv :2401.04722 [eess]. Jan. 2024. DOI : 10.48550/arXiv.2401.04722.
- [116] Ziyang WANG, Jian-Qing ZHENG, Yichi ZHANG, Ge CUI et Lei LI. *Mamba-UNet : UNet-Like Pure Visual Mamba for Medical Image Segmentation*. arXiv :2402.05079 [eess]. Mar. 2024. DOI : 10.48550/arXiv.2402.05079.
- [117] Jiacheng RUAN, Jincheng LI et Suncheng XIANG. *VM-UNet : Vision Mamba UNet for Medical Image Segmentation*. arXiv :2402.02491 [eess]. Nov. 2024. DOI : 10.48550/arXiv.2402.02491.
- [118] Jiarun LIU, Hao YANG, Hong-Yu ZHOU, Yan XI, Lequan YU, Yizhou YU, Yong LIANG, Guangming SHI, Shaoting ZHANG, Hairong ZHENG et Shanshan WANG. *Swin-UMamba : Mamba-based UNet with ImageNet-based pretraining*. arXiv :2402.03302 [eess]. Mar. 2024. DOI : 10.48550/arXiv.2402.03302.
- [119] Mingya ZHANG, Zhihao CHEN, Yiyuan GE et Xianping TAO. *HMT-UNet : A hybrid Mamba-Transformer Vision UNet for Medical Image Segmentation*. arXiv :2408.11289 [eess] version : 2. Sept. 2024. DOI : 10.48550/arXiv.2408.11289.
- [120] Yufeng JIANG, Zongxi LI, Xiangyan CHEN, Haoran XIE et Jing CAI. *MLLA-UNet : Mamba-like Linear Attention in an Efficient U-Shape Model for Medical Image Segmentation*. arXiv :2410.23738 [eess] version : 1. Oct. 2024. DOI : 10.48550/arXiv.2410.23738.
- [121] Shengbo TAN, Rundong XUE, Shipeng LUO, Zeyu ZHANG, Xinran WANG, Lei ZHANG, Daji ERGU, Zhang YI, Yang ZHAO et Ying CAI. *SegKAN : High-Resolution Medical Image Segmentation with Long-Distance Dependencies*. arXiv :2412.19990 [eess] version : 2. Jan. 2025. DOI : 10.48550/arXiv.2412.19990.
- [122] Yibo ZHANG. *KM-UNet KAN Mamba UNet for medical image segmentation*. arXiv :2501.02559 [eess] version : 1. Jan. 2025. DOI : 10.48550/arXiv.2501.02559.
- [123] Ziming LIU, Yixuan WANG, Sachin VAIDYA, Fabian RUEHLE, James HALVERSON, Marin SOLJAČIĆ, Thomas Y. HOU et Max TEGMARK. *KAN : Kolmogorov-Arnold Networks*. arXiv :2404.19756 [cs]. Juin 2024. DOI : 10.48550/arXiv.2404.19756.

- [124] Yangfan NI, Duo ZHANG, Gege MA, Fan RAO, Yuanfeng WU, Lijun LU, Zhongke HUANG et Wentao ZHU. « A Multiscale Spatial Transformer U-Net for Simultaneously Automatic Reorientation and Segmentation of 3-D Nuclear Cardiac Images ». In : *IEEE Transactions on Radiation and Plasma Medical Sciences* 8.6 (juil. 2024). Conference Name : IEEE Transactions on Radiation and Plasma Medical Sciences, p. 632-645. doi : 10.1109/TRPMS.2024.3382318.
- [125] Mohamed S. ELMAHDY, Laurens BELJAARDS, Sahar YOUSEFI, Hessam SOKOOTI, Fons VERBEEK, U. A. van der HEIDE et Marius STARING. *Joint Registration and Segmentation via Multi-Task Learning for Adaptive Radiotherapy of Prostate Cancer*. arXiv :2105.01844 [cs, eess]. Mai 2021. doi : 10.48550/arXiv.2105.01844.
- [126] Yao SU, Zhentian QIAN, Lei MA, Lifang HE et Xiangnan KONG. « One-shot Joint Extraction, Registration and Segmentation of Neuroimaging Data ». In : *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. arXiv :2307.15198 [cs]. Août 2023, p. 2049-2060. doi : 10.1145/3580305.3599452.
- [127] Carole H. SUDRE, Wenqi LI, Tom VERCAUTEREN, Sebastien OURSELIN et M. JORGE CARDOSO. « Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations ». en. In : *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Sous la dir. de M. Jorge CARDOSO, Tal ARBEL, Gustavo CARNEIRO, Tanveer SYEDA-MAHMOOD, João Manuel R.S. TAVARES, Mehdi MORADI, Andrew BRADLEY, Hayit GREENSPAN, João Paulo PAPA, Anant MADABHUSHI, Jacinto C. NASCIMENTO, Jaime S. CARDOSO, Vasileios BELAGIANNIS et Zhi LU. Cham : Springer International Publishing, 2017, p. 240-248. doi : 10.1007/978-3-319-67558-9_28.
- [128] Tsung-Yi LIN, Priya GOYAL, Ross GIRSHICK, Kaiming HE et Piotr DOLLÁR. *Focal Loss for Dense Object Detection*. arXiv :1708.02002 [cs]. Fév. 2018. doi : 10.48550/arXiv.1708.02002.
- [129] Fausto MILLETARI, Nassir NAVAB et Seyed-Ahmad AHMADI. *V-Net : Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation*. arXiv :1606.04797 [cs]. Juin 2016. doi : 10.48550/arXiv.1606.04797.
- [130] Seyed Sadegh Mohseni SALEHI, Deniz ERDOGMUS et Ali GHOLIPOUR. « Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks ». en. In : *Machine Learning in Medical Imaging*. Sous la dir. de Qian WANG, Yinghuan SHI, Heung-Il SUK et Kenji SUZUKI. Cham : Springer International Publishing, 2017, p. 379-387. doi : 10.1007/978-3-319-67389-9_44.

- [131] Lucas FIDON, Wenqi LI, Luis C. GARCIA-PERAZA-HERRERA, Jinendra EKANAYAKE, Neil KITCHEN, Sébastien OURSELIN et Tom VERCAUTEREN. « Generalised Wasserstein Dice Score for Imbalanced Multi-class Segmentation Using Holistic Convolutional Networks ». en. In : *Brainlesion : Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Sous la dir. d'Alessandro CRIMI, Spyridon BAKAS, Hugo KUIJF, Bjoern MENZE et Mauricio REYES. Cham : Springer International Publishing, 2018, p. 64-76. DOI : 10.1007/978-3-319-75238-9_6.
- [132] Ian GOODFELLOW, Jean POUGET-ABADIE, Mehdi MIRZA, Bing XU, David WARDE-FARLEY, Sherjil OZAIR, Aaron COURVILLE et Yoshua BENGIO. « Generative Adversarial Nets ». In : *Advances in Neural Information Processing Systems*. T. 27. Curran Associates, Inc., 2014.
- [133] Zhun ZHONG, Liang ZHENG, Guoliang KANG, Shaozi LI et Yi YANG. « Random Erasing Data Augmentation ». en. In : *Proceedings of the AAAI Conference on Artificial Intelligence* 34.07 (avr. 2020). Number : 07, p. 13001-13008. DOI : 10.1609/aaai.v34i07.7000.
- [134] Sangdoo YUN, Dongyoon HAN, Seong Joon OH, Sanghyuk CHUN, Junsuk CHOE et Youngjoon YOO. « CutMix : Regularization Strategy to Train Strong Classifiers With Localizable Features ». In : 2019, p. 6023-6032.
- [135] D. SUN, F. DORNAIKA et N. BARRENA. « HSMix : Hard and soft mixing data augmentation for medical image segmentation ». In : *Information Fusion* 115 (mar. 2025), p. 102741. DOI : 10.1016/j.inffus.2024.102741.
- [136] Tom EELBODE, Jeroen BERTELS, Maxim BERMAN, Dirk VANDERMEULEN, Frederik MAES, Raf BISSCHOPS et Matthew B. BLASCHKO. « Optimization for Medical Image Segmentation : Theory and Practice When Evaluating With Dice Score or Jaccard Index ». In : *IEEE Transactions on Medical Imaging* 39.11 (nov. 2020). Conference Name : IEEE Transactions on Medical Imaging, p. 3679-3690. DOI : 10.1109/TMI.2020.3002417.
- [137] Stephen M. SMITH, Mark JENKINSON, Mark W. WOOLRICH, Christian F. BECKMANN, Timothy E. J. BEHRENS, Heidi JOHANSEN-BERG, Peter R. BANNISTER, Marilena DE LUCA, Ivana DROBNJAK, David E. FLITNEY, Rami K. NIAZY, James SAUNDERS, John VICKERS, Yongyue ZHANG, Nicola DE STEFANO, J. Michael BRADY et Paul M. MATTHEWS. « Advances in functional and structural MR image analysis and implementation as FSL ». eng. In : *NeuroImage* 23 Suppl 1 (2004), S208-219. DOI : 10.1016/j.neuroimage.2004.07.051.
- [138] David W SHATTUCK et Richard M LEAHY. « BrainSuite : An automated cortical surface identification tool ». In : *Medical Image Analysis* 6.2 (juin 2002), p. 129-142. DOI : 10.1016/S1361-8415(02)00054-3.

- [139] V. POPESCU, M. BATTAGLINI, W. S. HOOGSTRATE, S. C. J. VERFAILLIE, I. C. SLUIMER, R. A. van SCHIJNDEL, B. W. van DIJK, K. S. COVER, D. L. KNOL, M. JENKINSON, F. BARKHOF, N. de STEFANO et H. VRENKEN. « Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis ». In : *NeuroImage* 61.4 (juil. 2012), p. 1484-1494. DOI : 10.1016/j.neuroimage.2012.03.074.
- [140] Nikolaus WEISKOPF, Antoine LUTTI, Gunther HELMS, Marianne NOVAK, John ASHBURNER et Chloe HUTTON. « Unified segmentation based correction of R1 brain maps for RF transmit field inhomogeneities (UNICORT) ». In : *NeuroImage* 54.3 (fév. 2011), p. 2116-2124. DOI : 10.1016/j.neuroimage.2010.10.023.
- [141] Brian PATENAUDE, Stephen M. SMITH, David N. KENNEDY et Mark JENKINSON. « A Bayesian model of shape and appearance for subcortical brain segmentation ». In : *NeuroImage* 56.3 (juin 2011), p. 907-922. DOI : 10.1016/j.neuroimage.2011.02.046.
- [142] K KAZEMI et N NOORIZADEH. « Quantitative Comparison of SPM, FSL, and Brainsuite for Brain MR Image Segmentation ». In : *Journal of Biomedical Physics & Engineering* 4.1 (mar. 2014), p. 13-26.
- [143] On TSANG, Ali GHOLIPOUR, Nasser KEHTARNAVAZ, Kaundinya GOPINATH, Richard BRIGGS et Issa PANAHI. « Comparison of tissue segmentation algorithms in neuroimage analysis software tools ». In : *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. ISSN : 1558-4615. Août 2008, p. 3924-3928. DOI : 10.1109/IEMBS.2008.4650068.
- [144] Frederick KLAUSCHEN, Aaron GOLDMAN, Vincent BARRA, Andreas MEYER-LINDENBERG et Arvid LUNDERVOLD. « Evaluation of automated brain MR image segmentation and volumetry methods ». In : *Human Brain Mapping* 30.4 (juin 2008), p. 1310-1327. DOI : 10.1002/hbm.20599.
- [145] Jordan BAI, Thi Lan Huong TRINH, Kai-Hsiang CHUANG et Anqi QIU. « Atlas-based automatic mouse brain image segmentation revisited : model complexity vs. image registration ». eng. In : *Magnetic Resonance Imaging* 30.6 (juil. 2012), p. 789-798. DOI : 10.1016/j.mri.2012.02.010.
- [146] Jiehua LI, Xiaofeng LIU, Jiachen ZHUO, Rao P. GULLAPALLI et Jason M. ZARA. « An automatic rat brain extraction method based on a deformable surface model ». In : *Journal of Neuroscience Methods* 218.1 (août 2013), p. 72-82. DOI : 10.1016/j.jneumeth.2013.04.011.

- [147] Zhenrong FU, Lan LIN, Miao TIAN, Jingxuan WANG, Baiwen ZHANG, Ping-ping CHU, Shaowu LI, Muhammad Mohsin PATHAN, Yulin DENG et Shuicai WU. « Evaluation of five diffeomorphic image registration algorithms for mouse brain magnetic resonance microscopy ». eng. In : *Journal of Microscopy* 268.2 (nov. 2017), p. 141-154. DOI : 10.1111/jmi.12594.
- [148] Min Hyeok BAE, Rong PAN, Teresa WU et Alexandra BADEA. « Automated Segmentation of Mouse Brain Images Using Extended MRF ». In : *NeuroImage* 46.3 (juil. 2009), p. 717-725. DOI : 10.1016/j.neuroimage.2009.02.012.
- [149] D. MA, M. CARDOSO, M. MODAT, N. POWELL, H. HOLMES, M. LYTHGOE et S. OURSELIN. *Multi Atlas Segmentation applied to in vivo mouse brain MRI*. eng. Proceedings paper. Pages : 134-143 Place : Nice Publisher : MICCAI 2012 Workshop on Multi-Atlas Labeling. Oct. 2012.
- [150] Da MA, Manuel J. CARDOSO, Marc MODAT, Nick POWELL, Jack WELLS, Holly HOLMES, Frances WISEMAN, Victor TYBULEWICZ, Elizabeth FISHER, Mark F. LYTHGOE et Sébastien OURSELIN. « Automatic Structural Parcellation of Mouse Brain MRI Using Multi-Atlas Label Fusion ». en. In : *PLOS ONE* 9.1 (jan. 2014). Publisher : Public Library of Science, e86576. DOI : 10.1371/journal.pone.0086576.
- [151] Riccardo FEO et Federico GIOVE. « Towards an efficient segmentation of small rodents brain : A short critical review ». In : *Journal of Neuroscience Methods* 323 (juil. 2019), p. 82-89. DOI : 10.1016/j.jneumeth.2019.05.003.
- [152] Shahinur ALAM, Tae-Yeon EOM, Jeffrey STEINBERG, David ACKERMAN, J. Eric SCHMITT, Walter J. AKERS, Stanislav S. ZAKHARENKO et Khaled KHAIRY. « An End-To-End Pipeline for Fully Automatic Morphological Quantification of Mouse Brain Structures From MRI Imagery ». eng. In : *Frontiers in Bioinformatics* 2 (2022), p. 865443. DOI : 10.3389/fbinf.2022.865443.
- [153] Guohui RUAN, Jiaming LIU, Ziqi AN, Kaiibin WU, Chuanjun TONG, Qiang LIU, Ping LIANG, Zhifeng LIANG, Wufan CHEN, Xinyuan ZHANG et Yan-qi FENG. « Automated Skull Stripping in Mouse Functional Magnetic Resonance Imaging Analysis Using 3D U-Net ». eng. In : *Frontiers in Neuroscience* 16 (2022), p. 801769. DOI : 10.3389/fnins.2022.801769.
- [154] Md. Eshmam RAYED, S. M. Sajibul ISLAM, Sadia Islam NIHA, Jamin Rahman JIM, Md Mohsin KABIR et M. F. MRIDHA. « Deep learning for medical image segmentation : State-of-the-art advancements and challenges ». In : *Informatics in Medicine Unlocked* 47 (jan. 2024), p. 101504. DOI : 10.1016/j.imu.2024.101504.

- [155] Trupti Chetan KHERDE et Trupti BARASKAR. « A Comprehensive Survey of Image Segmentation for Medical Images ». In : *2024 4th International Conference on Sustainable Expert Systems (ICSES)*. Oct. 2024, p. 1137-1144. DOI : 10.1109/ICSES63445.2024.10763045.
- [156] Adam PASZKE, Sam GROSS, Francisco MASSA, Adam LERER, James BRADBURY, Gregory CHANAN, Trevor KILLEEN, Zeming LIN, Natalia GIMELSHEIN, Luca ANTIGA, Alban DESMAISON, Andreas KÖPF, Edward YANG, Zach DeVITO, Martin RAISON, Alykhan TEJANI, Sasank CHILAMKURTHY, Benoit STEINER, Lu FANG, Junjie BAI et Soumith CHINTALA. *PyTorch : An Imperative Style, High-Performance Deep Learning Library*. arXiv :1912.01703 [cs, stat]. Déc. 2019. DOI : 10.48550/arXiv.1912.01703.
- [157] M. Jorge CARDOSO, Wenqi LI, Richard BROWN, Nic MA, Eric KERFOOT, Yiheng WANG, Benjamin MURREY, Andriy MYRONENKO, Can ZHAO, Dong YANG, Vishwesh NATH, Yufan HE, Ziyue XU, Ali HATAMIZADEH, Andriy MYRONENKO, Wentao ZHU, Yun LIU, Mingxin ZHENG, Yucheng TANG, Isaac YANG, Michael ZEPHYR, Behrooz HASHEMIAN, Sachidanand ALLE, Mohammad Zalbagi DARESTANI, Charlie BUDD, Marc MODAT, Tom VERCAUTEREN, Guotai WANG, Yiwen LI, Yipeng HU, Yunguan FU, Benjamin GORMAN, Hans JOHNSON, Brad GENEREAUX, Barbaros S. ERDAL, Vikash GUPTA, Andres DIAZ-PINTO, Andre DOURSON, Lena MAIER-HEIN, Paul F. JAEGER, Michael BAUMGARTNER, Jayashree KALPATHY-CRAMER, Mona FLORES, Justin KIRBY, Lee A. D. COOPER, Holger R. ROTH, Daguang XU, David BERICAT, Ralf FLOCA, S. Kevin ZHOU, Haris SHUAIB, Keyvan FARAHANI, Klaus H. MAIER-HEIN, Stephen AYLWARD, Prerna DOGRA, Sebastien OURSELIN et Andrew FENG. *MONAI : An open-source framework for deep learning in healthcare*. arXiv :2211.02701 [cs]. Nov. 2022. DOI : 10.48550/arXiv.2211.02701.
- [158] Diederik P. KINGMA et Jimmy BA. *Adam : A Method for Stochastic Optimization*. arXiv :1412.6980 [cs]. Jan. 2017. DOI : 10.48550/arXiv.1412.6980.
- [159] C. STUDHOLME, R.T. CONSTABLE et J.S. DUNCAN. « Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model ». In : *IEEE Transactions on Medical Imaging* 19.11 (nov. 2000). Conference Name : IEEE Transactions on Medical Imaging, p. 1115-1127. DOI : 10.1109/42.896788.
- [160] Lisa Gottesfeld BROWN. « A survey of image registration techniques ». In : *ACM Comput. Surv.* 24.4 (déc. 1992), p. 325-376. DOI : 10.1145/146370.146374.
- [161] Junyu CHEN, Yihao LIU, Shuwen WEI, Zhangxing BIAN, Shalini SUBRAMANIAN, Aaron CARASS, Jerry L. PRINCE et Yong DU. « A survey on deep learning

- in medical image registration : New technologies, uncertainty, evaluation metrics, and beyond ». In : *Medical Image Analysis* 100 (fév. 2025), p. 103385. DOI : 10.1016/j.media.2024.103385.
- [162] F.L. BOOKSTEIN. « Principal warps : thin-plate splines and the decomposition of deformations ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11.6 (juin 1989). Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 567-585. DOI : 10.1109/34.24792.
- [163] M. FORNEFETT, K. ROHR et H. S. STIEHL. « Radial basis functions with compact support for elastic registration of medical images ». In : *Image and Vision Computing* 19.1 (jan. 2001), p. 87-96. DOI : 10.1016/S0262-8856(00)00057-3.
- [164] D. RUECKERT, L.I. SONODA, C. HAYES, D.L.G. HILL, M.O. LEACH et D.J. HAWKES. « Nonrigid registration using free-form deformations : application to breast MR images ». In : *IEEE Transactions on Medical Imaging* 18.8 (août 1999). Conference Name : IEEE Transactions on Medical Imaging, p. 712-721. DOI : 10.1109/42.796284.
- [165] Ruzena BAJCSY et Stane KOVAČIČ. « Multiresolution elastic matching ». In : *Computer Vision, Graphics, and Image Processing* 46.1 (avr. 1989), p. 1-21. DOI : 10.1016/S0734-189X(89)80014-3.
- [166] C. DAVATZIKOS. « Spatial transformation and registration of brain images using elastically deformable models ». eng. In : *Computer vision and image understanding : CVIU* 66.2 (mai 1997), p. 207-222. DOI : 10.1006/cviu.1997.0605.
- [167] G.E. CHRISTENSEN, R.D. RABBITT et M.I. MILLER. « Deformable templates using large deformation kinematics ». In : *IEEE Transactions on Image Processing* 5.10 (oct. 1996). Conference Name : IEEE Transactions on Image Processing, p. 1435-1447. DOI : 10.1109/83.536892.
- [168] Berthold K. P. HORN et Brian G. SCHUNCK. « Determining optical flow ». In : *Artificial Intelligence* 17.1 (août 1981), p. 185-203. DOI : 10.1016/0004-3702(81)90024-2.
- [169] Pierre HELLIER et Christian BARILLOT. « Cooperation between Local and Global Approaches to Register Brain Images ». en. In : *Information Processing in Medical Imaging*. Sous la dir. de Michael F. INSANA et Richard M. LEAHY. Berlin, Heidelberg : Springer, 2001, p. 315-328. DOI : 10.1007/3-540-45729-1_32.

- [170] J. -P. THIRION. « Image matching as a diffusion process : an analogy with Maxwell's demons ». In : *Medical Image Analysis* 2.3 (sept. 1998), p. 243-260. DOI : 10.1016/S1361-8415(98)80022-4.
- [171] Tom VERCAUTEREN, Xavier PENNEC, Aymeric PERCHANT et Nicholas AYACHE. « Symmetric log-domain diffeomorphic Registration : a demons-based approach ». eng. In : *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 11.Pt 1 (2008), p. 754-761. DOI : 10.1007/978-3-540-85988-8_90.
- [172] M. Faisal BEG, Michael I. MILLER, Alain TROUVÉ et Laurent YOUNES. « Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms ». en. In : *International Journal of Computer Vision* 61.2 (fév. 2005), p. 139-157. DOI : 10.1023/B:VISI.0000043755.93987.aa.
- [173] Vincent ARSIGNY, Olivier COMMOWICK, Xavier PENNEC et Nicholas AYACHE. « A Log-Euclidean Framework for Statistics on Diffeomorphisms ». en. In : *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*. Sous la dir. de Rasmus LARSEN, Mads NIELSEN et Jon SPORRING. Berlin, Heidelberg : Springer, 2006, p. 924-931. DOI : 10.1007/11866565_113.
- [174] Mark HOLDEN. « A Review of Geometric Transformations for Nonrigid Body Registration ». In : *IEEE Transactions on Medical Imaging* 27.1 (jan. 2008). Conference Name : IEEE Transactions on Medical Imaging, p. 111-128. DOI : 10.1109/TMI.2007.904691.
- [175] Monan WANG et Pengcheng LI. « A Review of Deformation Models in Medical Image Registration ». en. In : *Journal of Medical and Biological Engineering* 39.1 (fév. 2019), p. 1-17. DOI : 10.1007/s40846-018-0390-1.
- [176] G.E. CHRISTENSEN et H.J. JOHNSON. « Consistent image registration ». In : *IEEE Transactions on Medical Imaging* 20.7 (juil. 2001). Conference Name : IEEE Transactions on Medical Imaging, p. 568-582. DOI : 10.1109/42.932742.
- [177] Daniel RUECKERT, Paul ALJABAR, Rolf A. HECKEMANN, Joseph V. HAJNAL et Alexander HAMMERS. « Diffeomorphic Registration Using B-Splines ». en. In : *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*. Sous la dir. de Rasmus LARSEN, Mads NIELSEN et Jon SPORRING. Berlin, Heidelberg : Springer, 2006, p. 702-709. DOI : 10.1007/11866763_86.

- [178] H. J. JOHNSON et G. E. CHRISTENSEN. « Consistent landmark and intensity-based image registration ». eng. In : *IEEE transactions on medical imaging* 21.5 (mai 2002), p. 450-461. DOI : 10.1109/TMI.2002.1009381.
- [179] W. M. WELLS, P. VIOLA, H. ATSUMI, S. NAKAJIMA et R. KIKINIS. « Multi-modal volume registration by maximization of mutual information ». eng. In : *Medical Image Analysis* 1.1 (mar. 1996), p. 35-51. DOI : 10.1016/s1361-8415(01)80004-9.
- [180] F. MAES, A. COLLIGNON, D. VANDERMEULEN, G. MARCHAL et P. SUETENS. « Multimodality image registration by maximization of mutual information ». In : *IEEE Transactions on Medical Imaging* 16.2 (avr. 1997). Conference Name : IEEE Transactions on Medical Imaging, p. 187-198. DOI : 10.1109/42.563664.
- [181] John DUCHI, Elad HAZAN et Yoram SINGER. « Adaptive Subgradient Methods for Online Learning and Stochastic Optimization ». In : *Journal of Machine Learning Research* 12.61 (2011), p. 2121-2159.
- [182] Rohit JENA, Pratik CHAUDHARI et James C. GEE. *FireANTs : Adaptive Riemannian Optimization for Multi-Scale Diffeomorphic Registration*. arXiv :2404.01249 [cs]. Avr. 2024. DOI : 10.48550/arXiv.2404.01249.
- [183] Hava LESTER et Simon R. ARRIDGE. « A survey of hierarchical non-linear medical image registration ». In : *Pattern Recognition* 32.1 (jan. 1999), p. 129-149. DOI : 10.1016/S0031-3203(98)00095-8.
- [184] Torsten ROHLFING. « Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy : Widely Used but Unreliable ». In : *IEEE transactions on medical imaging* 31.2 (fév. 2012), p. 153-163. DOI : 10.1109/TMI.2011.2163944.
- [185] Rohit JENA, Deeksha SETHI, Pratik CHAUDHARI et James C. GEE. *Deep Learning in Medical Image Registration : Magic or Mirage ?* arXiv :2408.05839 [eess] version : 2. Sept. 2024. DOI : 10.48550/arXiv.2408.05839.
- [186] Zhuoyuan WANG, Haiqiao WANG et Yi WANG. *Pyramid Attention Network for Medical Image Registration*. arXiv :2402.09016 [cs] version : 1. Fév. 2024. DOI : 10.48550/arXiv.2402.09016.
- [187] Alexander BIGALKE, Lasse HANSEN, Tony C. W. MOK et Mattias P. HEINRICH. *Unsupervised 3D registration through optimization-guided cyclical self-training*. arXiv :2306.16997 [cs]. Juil. 2023.
- [188] Tony C. W. Mok et Albert C. S. CHUNG. *Large Deformation Diffeomorphic Image Registration with Laplacian Pyramid Networks*. arXiv :2006.16148 [cs, eess]. Juin 2020.

- [189] Guha BALAKRISHNAN, Amy ZHAO, Mert R. SABUNCU, John GUTTAG et Adrian V. DALCA. « An Unsupervised Learning Model for Deformable Medical Image Registration ». In : *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. arXiv :1802.02604 [cs]. Juin 2018, p. 9252-9260. DOI : 10.1109/CVPR.2018.00964.
- [190] Brian AVANTS, Nicholas J. TUSTISON et Gang SONG. « Advanced Normalization Tools : V1.0 ». In : *The Insight Journal* (juil. 2009). DOI : 10.54294/uvnhin.
- [191] Paul A. YUSHKEVICH, John PLUTA, Hongzhi WANG, Laura E.M. WISSE, Sandhitsu DAS et David WOLK. « IC-P-174 : Fast Automatic Segmentation of Hippocampal Subfields and Medial Temporal Lobe Subregions In 3 Tesla and 7 Tesla T2-Weighted MRI ». en. In : *Alzheimer's & Dementia* 12.7S_Part_2 (2016). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.jalz.2016.06.P126-P127>. DOI : 10.1016/j.jalz.2016.06.205.
- [192] Paul YUSHKEVICH. *pyushkevich/greedy*. original-date : 2016-01-07T11:50:23Z. Jan. 2025.
- [193] S. JOSHI, Brad DAVIS, Matthieu JOMIER et Guido GERIG. « Unbiased diffeomorphic atlas construction for computational anatomy ». eng. In : *NeuroImage* 23 Suppl 1 (2004), S151-160. DOI : 10.1016/j.neuroimage.2004.07.068.
- [194] Marc MODAT, Gerard R. RIDGWAY, Zeike A. TAYLOR, Manja LEHMANN, Josephine BARNES, David J. HAWKES, Nick C. FOX et Sébastien OURSELIN. « Fast free-form deformation using graphics processing units ». In : *Computer Methods and Programs in Biomedicine*. HP-MICCAI 2008 98.3 (juin 2010), p. 278-284. DOI : 10.1016/j.cmpb.2009.09.002.
- [195] Florence DRU, Pierre FILLARD et Tom VERCAUTEREN. « An ITK Implementation of the Symmetric Log-Domain Diffeomorphic Demons Algorithm ». en. In : *The Insight Journal* (sept. 2010). DOI : 10.54294/8vm9t2.
- [196] Matthew Michael McCORMICK, Xiaoxiao LIU, Luis IBANEZ, Julien JOMIER et Charles MARION. « ITK : enabling reproducible research and open science ». English. In : *Frontiers in Neuroinformatics* 8 (fév. 2014). Publisher : Frontiers. DOI : 10.3389/fninf.2014.00013.
- [197] Mattias P. HEINRICH, Mark JENKINSON, Michael BRADY et Julia A. SCHNABEL. « MRF-Based Deformable Registration and Ventilation Estimation of Lung CT ». In : *IEEE Transactions on Medical Imaging* 32.7 (juil. 2013). Conference Name : IEEE Transactions on Medical Imaging, p. 1239-1248. DOI : 10.1109/TMI.2013.2246577.

- [198] Mattias P. HEINRICH, Oskar MAIER et Heinz HANDELS. « Multi-modal Multi-Atlas Segmentation using Discrete Optimisation and Self-Similarities ». English. In : *CEUR Workshop Proceedings*. Jan. 2015, p. 27-30.
- [199] Stefan KLEIN, Marius STARING, Keelin MURPHY, Max A. VIERGEVER et Josien P. W. PLUIM. « elastix : a toolbox for intensity-based medical image registration ». eng. In : *IEEE transactions on medical imaging* 29.1 (jan. 2010), p. 196-205. DOI : 10.1109/TMI.2009.2035616.
- [200] Alex KRIZHEVSKY, Ilya SUTSKEVER et Geoffrey E HINTON. « ImageNet Classification with Deep Convolutional Neural Networks ». In : *Advances in Neural Information Processing Systems*. T. 25. Curran Associates, Inc., 2012.
- [201] Xiaohuan CAO, Jianhua YANG, Jun ZHANG, Dong NIE, Min-Jeong KIM, Qian WANG et Dinggang SHEN. « Deformable Image Registration based on Similarity-Steered CNN Regression ». eng. In : *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 10433 (sept. 2017), p. 300-308. DOI : 10.1007/978-3-319-66182-7_35.
- [202] Koen A. J. EPPENHOF et Josien P. W. PLUIM. « Pulmonary CT Registration Through Supervised Learning With Convolutional Neural Networks ». In : *IEEE Transactions on Medical Imaging* 38.5 (mai 2019). Conference Name : IEEE Transactions on Medical Imaging, p. 1097-1105. DOI : 10.1109/TMI.2018.2878316.
- [203] Max JADERBERG, Karen SIMONYAN, Andrew ZISSERMAN et Koray KAVUKCUOGLU. *Spatial Transformer Networks*. arXiv :1506.02025 [cs]. Fév. 2016.
- [204] Zhenlin XU et Marc NIETHAMMER. « DeepAtlas : Joint Semi-supervised Learning of Image Registration and Segmentation ». en. In : *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. Sous la dir. de Dinggang SHEN, Tianming LIU, Terry M. PETERS, Lawrence H. STAIB, Caroline ESSERT, Sean ZHOU, Pew-Thian YAP et Ali KHAN. Cham : Springer International Publishing, 2019, p. 420-429. DOI : 10.1007/978-3-030-32245-8_47.
- [205] Tony C. W. Mok et Albert C. S. CHUNG. *Fast Symmetric Diffeomorphic Image Registration with Convolutional Neural Networks*. arXiv :2003.09514 [cs]. Fév. 2021. DOI : 10.48550/arXiv.2003.09514.
- [206] Xi JIA, Alexander THORLEY, Wei CHEN, Huaqi QIU, Linlin SHEN, Iain B. STYLES, Hyung Jin CHANG, Ales LEONARDIS, Antonio de MARVAO, Declan P. O'REGAN, Daniel RUECKERT et Jinming DUAN. *Learning a Model-Driven*

- Variational Network for Deformable Image Registration*. arXiv :2105.12227 [cs, eess]. Mai 2021. DOI : 10.48550/arXiv.2105.12227.
- [207] Bob D. de Vos, Bas H. M. van der VELDEN, Jörg SANDER, Kenneth G. A. GILHUIJS, Marius STARING et Ivana IŠGUM. « Mutual information for unsupervised deep learning image registration ». In : *Medical Imaging 2020 : Image Processing*. T. 11313. SPIE, mar. 2020, p. 155-161. DOI : 10.1117/12.2549729.
- [208] Junyu CHEN, Ye LI, Yong DU et Eric C. FREY. « Generating Anthropomorphic Phantoms Using Fully Unsupervised Deformable Image Registration with Convolutional Neural Networks ». In : *Medical Physics* 47.12 (déc. 2020). arXiv :1912.02942 [cs, eess], p. 6366-6380. DOI : 10.1002/mp.14545.
- [209] Guha BALAKRISHNAN, Amy ZHAO, Mert R. SABUNCU, John GUTTAG et Adrian V. DALCA. « VoxelMorph : A Learning Framework for Deformable Medical Image Registration ». In : *IEEE Transactions on Medical Imaging* 38.8 (août 2019). arXiv :1809.05231 [cs], p. 1788-1800. DOI : 10.1109/TMI.2019.2897538.
- [210] Boah KIM, Dong Hwan KIM, Seong Ho PARK, Jieun KIM, June-Goo LEE et Jong Chul YE. *CycleMorph : Cycle Consistent Unsupervised Deformable Image Registration*. arXiv :2008.05772 [cs, eess, stat]. Août 2020. DOI : 10.48550/arXiv.2008.05772.
- [211] Adrian V. DALCA, Guha BALAKRISHNAN, John GUTTAG et Mert R. SABUNCU. « Unsupervised Learning of Probabilistic Diffeomorphic Registration for Images and Surfaces ». In : *Medical Image Analysis* 57 (oct. 2019). arXiv :1903.03545 [cs], p. 226-236. DOI : 10.1016/j.media.2019.07.006.
- [212] Yipeng HU, Marc MODAT, Eli GIBSON, Nooshin GHAVAMI, Ester BONMATI, Caroline M. MOORE, Mark EMBERTON, J. Alison NOBLE, Dean C. BARRATT et Tom VERCAUTEREN. « Label-driven weakly-supervised learning for multimodal deformable image registration ». In : *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. arXiv :1711.01666 [cs]. Avr. 2018, p. 1070-1074. DOI : 10.1109/ISBI.2018.8363756.
- [213] Yipeng HU, Marc MODAT, Eli GIBSON, Wenqi LI, Nooshin GHAVAMI, Ester BONMATI, Guotai WANG, Steven BANDULA, Caroline M. MOORE, Mark EMBERTON, Sébastien OURSELIN, J. Alison NOBLE, Dean C. BARRATT et Tom VERCAUTEREN. « Weakly-supervised convolutional neural networks for multimodal image registration ». en. In : *Medical Image Analysis* 49 (oct. 2018), p. 1-13. DOI : 10.1016/j.media.2018.07.002.

- [214] Jingfan FAN, Xiaohuan CAO, Pew-Thian YAP et Dinggang SHEN. « BIR-Net : Brain image registration using dual-supervised fully convolutional networks ». In : *Medical Image Analysis* 54 (mai 2019), p. 193-206. DOI : 10.1016/j.media.2019.03.006.
- [215] Dongyang KUANG et Tanya SCHMAH. *FAIM – A ConvNet Method for Unsupervised 3D Medical Image Registration*. arXiv :1811.09243 [cs]. Juin 2019. DOI : 10.48550/arXiv.1811.09243.
- [216] Dongyang KUANG. *On Reducing Negative Jacobian Determinant of the Deformation Predicted by Deep Registration Networks*. arXiv :1907.00068 [cs, eess]. Juin 2019.
- [217] Lin TIAN, Hastings GREER, François-Xavier VIALARD, Roland KWITT, Raúl San José ESTÉPAR, Richard Jarrett RUSHMORE, Nikolaos MAKRIS, Sylvain BOUX et Marc NIETHAMMER. *GradICON : Approximate Diffeomorphisms via Gradient Inverse Consistency*. arXiv :2206.05897 [cs, eess]. Oct. 2023. DOI : 10.48550/arXiv.2206.05897.
- [218] Zhengyang SHEN, Xu HAN, Zhenlin XU et Marc NIETHAMMER. *Networks for Joint Affine and Non-parametric Image Registration*. arXiv :1903.08811 [cs]. Mar. 2019.
- [219] Mingyuan MENG, Lei BI, Michael FULHAM, Dagan FENG et Jinman KIM. *Non-iterative Coarse-to-fine Transformer Networks for Joint Affine and Deformable Image Registration*. arXiv :2307.03421 [cs, eess] version : 1. Juil. 2023.
- [220] Jinxin LV, Zhiwei WANG, Hongkuan SHI, Haobo ZHANG, Sheng WANG, Yilang WANG et Qiang LI. *Joint Progressive and Coarse-to-fine Registration of Brain MRI via Deformation Field Integration and Non-Rigid Feature Fusion*. arXiv :2109.12384 [cs, eess]. Avr. 2022.
- [221] Miao KANG, Xiaojun HU, Weilin HUANG, Matthew R. SCOTT et Mauricio REYES. *Dual-Stream Pyramid Registration Network*. arXiv :1909.11966 [cs]. Avr. 2023.
- [222] Jelmer M. WOLTERINK, Jesse C. ZWIENENBERG et Christoph BRUNE. « Implicit Neural Representations for Deformable Image Registration ». en. In : *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*. ISSN : 2640-3498. PMLR, déc. 2022, p. 1349-1359.
- [223] Ankita JOSHI et Yi HONG. « Diffeomorphic Image Registration Using Lipschitz Continuous Residual Networks ». en. In : *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*. ISSN : 2640-3498. PMLR, déc. 2022, p. 605-617.

- [224] Junhao HU, Weijie GAN, Zhixin SUN, Hongyu AN et Ulugbek S. KAMILOV. *A Plug-and-Play Image Registration Network*. arXiv :2310.04297 [eess]. Mar. 2024. DOI : 10.48550/arXiv.2310.04297.
- [225] Constance FOURCADE, Ludovic FERRER, Noemie MOREAU, Gianmarco SANTINI, Aishlinn BRENNAN, Caroline ROUSSEAU, Marie LACOMBE, Vincent FLEURY, Mathilde COLOMBIÉ, Pascal JÉZÉQUEL, Mario CAMPONE, Mathieu RUBEAUX et Diana MATEUS. « Deformable image registration with deep network priors : a study on longitudinal PET images ». In : *Physics in Medicine & Biology* 67.15 (août 2022). arXiv :2111.11873 [cs, eess], p. 155011. DOI : 10.1088/1361-6560/ac7e17.
- [226] Junyu CHEN, Eric C. FREY, Yufan HE, William P. SEGARS, Ye LI et Yong DU. « TransMorph : Transformer for unsupervised medical image registration ». In : *Medical Image Analysis* 82 (nov. 2022). arXiv :2111.10480 [cs, eess], p. 102615. DOI : 10.1016/j.media.2022.102615.
- [227] Tony C. W. MOK et Albert C. S. CHUNG. *Affine Medical Image Registration with Coarse-to-Fine Vision Transformer*. arXiv :2203.15216 [cs]. Mar. 2022.
- [228] Jiacheng SHI, Yuting HE, Youyong KONG, Jean-Louis COATRIEUX, Huazhong SHU, Guanyu YANG et Shuo LI. *XMorpher : Full Transformer for Deformable Medical Image Registration via Cross Attention*. arXiv :2206.07349 [cs] version : 1. Juin 2022.
- [229] Tao GUO, Yinuo WANG, Shihao SHU, Diansheng CHEN, Zhouping TANG, Cai MENG et Xiangzhi BAI. *MambaMorph : a Mamba-based Framework for Medical MR-CT Deformable Registration*. arXiv :2401.13934 [cs] version : 4. Mar. 2024. DOI : 10.48550/arXiv.2401.13934.
- [230] Ziyang WANG, Jian-Qing ZHENG, Chao MA et Tao GUO. *VMambaMorph : a Multi-Modality Deformable Image Registration Framework based on Visual State Space Model with Cross-Scan Module*. arXiv :2404.05105 [cs]. Avr. 2024. DOI : 10.48550/arXiv.2404.05105.
- [231] Pingkun YAN, Sheng XU, Ardeshir R. RASTINEHAD et Brad J. WOOD. *Adversarial Image Registration with Application for MR and TRUS Image Fusion*. arXiv :1804.11024 [cs]. Oct. 2018. DOI : 10.48550/arXiv.1804.11024.
- [232] Minh Q. TRAN, Tuong DO, Huy TRAN, Erman TJIPUTRA, Quang D. TRAN et Anh NGUYEN. *Light-weight Deformable Registration using Adversarial Learning with Distilling Knowledge*. arXiv :2110.01293 [cs, eess]. Oct. 2021. DOI : 10.48550/arXiv.2110.01293.

- [233] Junyu CHEN, Yihao LIU, Shuwen WEI, Zhangxing BIAN, Shalini SUBRAMANIAN, Aaron CARASS, Jerry L. PRINCE et Yong DU. *A survey on deep learning in medical image registration : new technologies, uncertainty, evaluation metrics, and beyond*. arXiv :2307.15615 [cs, eess]. Sept. 2023.
- [234] Grant HASKINS, Uwe KRUGER et Pingkun YAN. « Deep Learning in Medical Image Registration : A Survey ». In : *Machine Vision and Applications* 31.1-2 (fév. 2020). arXiv :1903.02026 [cs, eess, q-bio], p. 8. DOI : 10.1007/s00138-020-01060-x.
- [235] Hanna SIEBERT, Lasse HANSEN et Mattias P. HEINRICH. *Fast 3D registration with accurate optimisation and little learning for Learn2Reg 2021*. arXiv :2112.03053 [eess]. Déc. 2021. DOI : 10.48550/arXiv.2112.03053.
- [236] *Robust Non-rigid Registration Challenge for Expansion Microscopy - Grand Challenge*. en.
- [237] *JaneliaSciComp/bigstream*. original-date : 2020-10-05T14:53:21Z. Déc. 2024.
- [238] Dmitry ULYANOV, Andrea VEDALDI et Victor LEMPITSKY. « Deep Image Prior ». In : *International Journal of Computer Vision* 128.7 (juil. 2020). arXiv :1711.10925 [cs], p. 1867-1888. DOI : 10.1007/s11263-020-01303-4.
- [239] Max-Heinrich LAVES, Sontje IHLER et Tobias ORTMAIER. *Deformable Medical Image Registration Using a Randomly-Initialized CNN as Regularization Prior*. arXiv :1908.00788 [eess]. Août 2019. DOI : 10.48550/arXiv.1908.00788.
- [240] ENZO FERRANTE, Ozan OKTAY, Ben GLOCKER et Diego H. MILONE. « On the Adaptability of Unsupervised CNN-Based Deformable Image Registration to Unseen Image Domains ». en. In : *Machine Learning in Medical Imaging*. Sous la dir. d'Yinghuan SHI, Heung-Il SUK et Mingxia LIU. Cham : Springer International Publishing, 2018, p. 294-302. DOI : 10.1007/978-3-030-00919-9_34.
- [241] Tobias FECHTER et Dimos BALTAS. *One Shot Learning for Deformable Medical Image Registration and Periodic Motion Tracking*. arXiv :1907.04641 [cs]. Fév. 2020. DOI : 10.48550/arXiv.1907.04641.
- [242] Yunlu ZHANG, Xue WU, H Michael GACH, Harold LI et Deshan YANG. « GroupRegNet : a groupwise one-shot deep learning-based 4D image registration method ». In : *Physics in medicine and biology* 66.4 (fév. 2021), p. 045030. DOI : 10.1088/1361-6560/abd956.

- [243] Weicheng CHI, Zhiming XIANG et Fen GUO. « Few-shot learning for deformable image registration in 4DCT images ». In : *The British Journal of Radiology* 95.1129 (jan. 2022), p. 20210819. DOI : 10.1259/bjr.20210819.
- [244] Yu SUN, Xiaolong WANG, Zhuang LIU, John MILLER, Alexei A. EFROS et Moritz HARDT. *Test-Time Training with Self-Supervision for Generalization under Distribution Shifts*. arXiv :1909.13231 [cs, stat]. Juil. 2020. DOI : 10.48550/arXiv.1909.13231.
- [245] Yossi GANDELSMAN, Yu SUN, Xinlei CHEN et Alexei A. EFROS. *Test-Time Training with Masked Autoencoders*. arXiv :2209.07522 [cs]. Sept. 2022. DOI : 10.48550/arXiv.2209.07522.
- [246] Wentao ZHU, Yufang HUANG, Daguang XU, Zhen QIAN, Wei FAN et Xiaohui XIE. *Test-Time Training for Deformable Multi-Scale Image Registration*. en. arXiv :2103.13578 [cs, eess]. Mar. 2021.
- [247] Yicheng CHEN, Shengxiang JI, Yuelin XIN, Kun HAN et Xiaohui XIE. *Op-tron : Better Medical Image Registration via Training in the Loop*. arXiv :2308.15216 [cs] version : 1. Août 2023.
- [248] Yicheng CHEN, Shengxiang JI, Yuelin XIN, Kun HAN et Xiaohui XIE. *On-the-Fly Guidance Training for Medical Image Registration*. arXiv :2308.15216 [cs] version : 4. Déc. 2023. DOI : 10.48550/arXiv.2308.15216.
- [249] Tony C. W. MOK, Zi LI, Yingda XIA, Jiawen YAO, Ling ZHANG, Jingren ZHOU et Le LU. « Deformable Medical Image Registration Under Distribution Shifts with Neural Instance Optimization ». en. In : *Machine Learning in Medical Imaging*. Sous la dir. de Xiaohuan CAO, Xuanang XU, Islem REKIK, Zhiming CUI et Xi OUYANG. Cham : Springer Nature Switzerland, 2024, p. 126-136. DOI : 10.1007/978-3-031-45673-2_13.
- [250] Shengyu ZHAO, Tingfung LAU, Ji LUO, Eric I.-Chao CHANG et Yan XU. « Unsupervised 3D End-to-End Medical Image Registration with Volume Tweening Network ». In : *IEEE Journal of Biomedical and Health Informatics* 24.5 (mai 2020). arXiv :1902.05020 [cs], p. 1394-1404. DOI : 10.1109/JBHI.2019.2951024.
- [251] Adrian V. DALCA, Guha BALAKRISHNAN, John GUTTAG et Mert R. SABUNCU. « Unsupervised Learning for Fast Probabilistic Diffeomorphic Registration ». In : t. 11070. arXiv :1805.04605 [cs]. 2018, p. 729-738. DOI : 10.1007/978-3-030-00928-1_82.

- [252] Yungeng ZHANG, Yuru PEI et Hongbin ZHA. « Learning Dual Transformer Network for Diffeomorphic Registration ». en. In : *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Sous la dir. de Marleen de BRUIJNE, Philippe C. CATTIN, Stéphane COTIN, Nicolas PADOY, Stefanie SPEIDEL, Yefeng ZHENG et Caroline ESSERT. Cham : Springer International Publishing, 2021, p. 129-138. DOI : 10.1007/978-3-030-87202-1_13.
- [253] Kun HAN, Shanlin SUN, Xiangyi YAN, Chenyu YOU, Hao TANG, Junayed NAUSHAD, Haoyu MA, Deying KONG et Xiaohui XIE. « Diffeomorphic Image Registration with Neural Velocity Field ». en. In : *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Waikoloa, HI, USA : IEEE, jan. 2023, p. 1869-1879. DOI : 10.1109/WACV56688.2023.00191.
- [254] Jun ZHANG. *Inverse-Consistent Deep Networks for Unsupervised Deformable Image Registration*. arXiv :1809.03443 [cs]. Sept. 2018.
- [255] Abdullah NAZIB, Clinton FOOKES, Olivier SALVADO et Dimitri PERRIN. « A Multiple Decoder Cnn For Inverse Consistent 3d Image Registration ». In : *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. ISSN : 1945-8452. Avr. 2021, p. 904-907. DOI : 10.1109/ISBI48211.2021.9433911.
- [256] Hastings GREER, Roland KWITT, Francois-Xavier VIALARD et Marc NIETHAMMER. *ICON : Learning Regular Maps Through Inverse Consistency*. arXiv :2105.04459 [cs]. Juin 2021. DOI : 10.48550/arXiv.2105.04459.
- [257] Hastings GREER, Lin TIAN, Francois-Xavier VIALARD, Roland KWITT, Sylvain BOUX, Raul San Jose ESTEPAR, Richard RUSHMORE et Marc NIETHAMMER. *Inverse Consistency by Construction for Multistep Deep Registration*. arXiv :2305.00087 [cs] version : 1. Avr. 2023.
- [258] Dongyang KUANG. « Cycle-Consistent Training for Reducing Negative Jacobian Determinant in Deep Registration Networks ». en. In : *Simulation and Synthesis in Medical Imaging*. Sous la dir. de Ninon BURGOS, Ali GOOYA et David SVOBODA. Cham : Springer International Publishing, 2019, p. 120-129. DOI : 10.1007/978-3-030-32778-1_13.
- [259] Thao Thi Ho, Woo Jin KIM, Chang Hyun LEE, Gong Yong JIN, Kum Ju CHAE et Sanghun CHOI. « An unsupervised image registration method employing chest computed tomography images and deep neural networks ». en. In : *Computers in Biology and Medicine* 154 (mar. 2023), p. 106612. DOI : 10.1016/j.compbiomed.2023.106612.

- [260] K. J. FRISTON, A. P. HOLMES, K. J. WORSLEY, J.-P. POLINE, C. D. FRITH et R. S. J. FRACKOWIAK. « Statistical parametric maps in functional imaging : A general linear approach ». en. In : *Human Brain Mapping* 2.4 (1994). _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.460020402>, p. 189-210. DOI : 10.1002/hbm.460020402.
- [261] Bruce FISCHL. « FreeSurfer ». eng. In : *NeuroImage* 62.2 (août 2012), p. 774-781. DOI : 10.1016/j.neuroimage.2012.01.021.
- [262] Marco PAGANI, Mario DAMIANO, Alberto GALBUSERA, Sotirios A. TSAFTARIS et Alessandro GOZZI. « Semi-automated registration-based anatomical labelling, voxel based morphometry and cortical thickness mapping of the mouse brain ». In : *Journal of Neuroscience Methods* 267 (juil. 2016), p. 62-73. DOI : 10.1016/j.jneumeth.2016.04.007.
- [263] Cassandra E. MEYER, Florian KURTH, Stefano LEPORE, Josephine L. GAO, Hadley JOHNSONBAUGH, Mandavi R. OBEROI, Stephen J. SAWIAK et Allan MACKENZIE-GRAHAM. « In vivo magnetic resonance images reveal neuroanatomical sex differences through the application of voxel-based morphometry in C57BL/6 mice ». In : *NeuroImage* 163 (déc. 2017), p. 197-205. DOI : 10.1016/j.neuroimage.2017.09.027.
- [264] Keigo HIKISHIMA, Yuji KOMAKI, Fumiko SEKI, Yasuyuki OHNISHI, Hirotaka J. OKANO et Hideyuki OKANO. « In vivo microscopic voxel-based morphometry with a brain template to characterize strain-specific structures in the mouse brain ». en. In : *Scientific Reports* 7.1 (mar. 2017). Publisher : Nature Publishing Group, p. 85. DOI : 10.1038/s41598-017-00148-1.
- [265] Andrew HOOPES, Jocelyn S. MORA, Adrian V. DALCA, Bruce FISCHL et Malte HOFFMANN. « SynthStrip : skull-stripping for any brain image ». In : *NeuroImage* 260 (oct. 2022), p. 119474. DOI : 10.1016/j.neuroimage.2022.119474.
- [266] Benjamin BILLOT, Douglas N. GREVE, Oula PUONTI, Axel THIELSCHER, Koen VAN LEEMPUT, Bruce FISCHL, Adrian V. DALCA et Juan Eugenio IGLESIAS. « SynthSeg : Segmentation of brain MRI scans of any contrast and resolution without retraining ». In : *Medical Image Analysis* 86 (mai 2023), p. 102789. DOI : 10.1016/j.media.2023.102789.
- [267] Juan Eugenio IGLESIAS. « A ready-to-use machine learning tool for symmetric multi-modality registration of brain MRI ». en. In : *Scientific Reports* 13.1 (avr. 2023). Publisher : Nature Publishing Group, p. 6657. DOI : 10.1038/s41598-023-33781-0.

- [268] Malte HOFFMANN, Benjamin BILLOT, Douglas N. GREVE, Juan Eugenio IGLESIAS, Bruce FISCHL et Adrian V. DALCA. « SynthMorph : Learning Contrast-Invariant Registration Without Acquired Images ». In : *IEEE transactions on medical imaging* 41.3 (mar. 2022), p. 543-558. DOI : 10.1109/TMI.2021.3116879.
- [269] Vaanathi SUNDARESAN, Giovanna ZAMBONI, Peter M. ROTHWELL, Mark JENKINSON et Ludovica GRIFFANTI. « Triplanar ensemble U-Net model for white matter hyperintensities segmentation on MR images ». In : *Medical Image Analysis* 73 (oct. 2021), p. 102184. DOI : 10.1016/j.media.2021.102184.
- [270] Christian GASER, Robert DAHNKE, Paul M. THOMPSON, Florian KURTH, Eileen LUDERS et null THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE. « CAT : a computational anatomy toolbox for the analysis of structural MRI data ». eng. In : *GigaScience* 13 (jan. 2024), giae049. DOI : 10.1093/gigascience/giae049.
- [271] Lukas FISCH, Stefan ZUMDICK, Carlotta BARKHAU, Daniel EMDEN, Jan ERNSTING, Ramona LEENINGS, Kelvin SARINK, Nils R. WINTER, Benjamin RISSE, Udo DANNLOWSKI et Tim HAHN. « deepbet : Fast brain extraction of T1-weighted MRI using Convolutional Neural Networks ». In : *Computers in Biology and Medicine* 179 (sept. 2024), p. 108845. DOI : 10.1016/j.compbimed.2024.108845.
- [272] Lukas FISCH. *torchreg - Lightweight image registration library using PyTorch*. original-date : 2023-08-09T13:06:45Z. Août 2023.
- [273] Fabian ISENSEE, Paul F. JAEGER, Simon A. A. KOHL, Jens PETERSEN et Klaus H. MAIER-HEIN. « nnU-Net : a self-configuring method for deep learning-based biomedical image segmentation ». en. In : *Nature Methods* 18.2 (fév. 2021). Publisher : Nature Publishing Group, p. 203-211. DOI : 10.1038/s41592-020-01008-z.
- [274] Shoshana SPRING, Jason P. LERCH et R. Mark HENKELMAN. « Sexual dimorphism revealed in the structure of the mouse brain using three-dimensional magnetic resonance imaging ». en. In : *NeuroImage* 35.4 (mai 2007), p. 1424-1433. DOI : 10.1016/j.neuroimage.2007.02.023.
- [275] Christina CORRE, Miriam FRIEDEL, Dulcie A. VOUSDEN, Ariane METCALF, Shoshana SPRING, Lily R. QIU, Jason P. LERCH et Mark R. PALMERT. « Separate effects of sex hormones and sex chromosomes on brain structure and function revealed by high-resolution magnetic resonance imaging and spatial navigation assessment of the Four Core Genotype mouse model ». eng. In : *Brain Structure & Function* 221.2 (mar. 2016), p. 997-1016. DOI : 10.1007/s00429-014-0952-0.

