

Synthèse de diffusion sur variété riemannienne à partir d'imagerie structurelle en haute-résolution

par

Benoit ANCTIL-ROBITAILLE

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE LA MAÎTRISE
AVEC MÉMOIRE EN GÉNIE DES TECHNOLOGIES DE L'INFORMATION
M. Sc. A.

MONTRÉAL, LE 10 JANVIER 2022

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Benoit Anctil-Robitaille, 2022



Cette licence Creative Commons signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette oeuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'oeuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE:

M. Hervé Lombaert, Directeur de mémoire

Département de génie logiciel et des technologies de l'information, École de technologie supérieure

M. Christian Desrosiers, co-Directeur

Département de génie logiciel et des technologies de l'information, École de technologie supérieure

M. Jean-Marc Lina, Président du jury

Département de génie électrique, École de technologie supérieure

M. Maxime Descoteaux, Membre du jury

Faculté des sciences, Université de Sherbrooke

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 6 DÉCEMBRE 2021

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

AVANT-PROPOS

Ce mémoire présente les résultats de mes travaux de recherche effectués à l’École de technologie supérieure sous la supervision des professeurs Hervé Lombaert et Christian Desrosiers. Cette recherche avait pour but l’exploration et le développement de méthodes pour la génération d’images de diffusion mathématiquement valides à partir d’images structurelles en haute-résolution.

Mes travaux ont mené à la publication de deux articles pour lesquels je suis le premier auteur : **un article de workshop**¹ à l’International Workshop on Computational Diffusion MRI à MICCAI 2020 et **un article de journal**² qui, au moment d’écrire ces lignes, est en révision. Le présent mémoire détaille le contenu de ces travaux dans le Chapitre 2. En plus des publications scientifiques, mes travaux ont mené au développement de la librairie ouverte torch-vectorized³ qui permet l’apprentissage de tenseurs de diffusion sur GPU avec Pytorch.

¹ <https://arxiv.org/abs/2004.00173>

² <https://arxiv.org/abs/2108.04135>

³ <https://github.com/banctilrobitaille/torch-vectorized>

REMERCIEMENTS

Ce mémoire n'aurait pu voir le jour sans l'apport inestimable de nombreuses personnes. D'abord, j'aimerais remercier les membres du jury qui ont pris le temps de lire avec minutie et d'apprécier mon travail. J'aimerais aussi chaleureusement remercier mes directeur et codirecteur de recherche Hervé Lombaert et Christian Desrosiers pour la confiance qu'ils m'ont démontrée en me prenant sous leurs ailes, leur humanité et leur dévouement envers ma réussite. Vous m'avez permis d'exprimer mes idées, de développer ma pensée et ultimement vous m'avez donné les outils nécessaires pour faire de moi un meilleur ingénieur chercheur et j'en suis extrêmement reconnaissant. Merci aussi à mon collègue de longue date Pierre-Luc avec qui j'ai traversé les aléas des études supérieures et qui est maintenant un ami très cher ainsi qu'à Karthik, Sukesh, Matilde et Mélanie. Vous m'avez tous alimenté à votre manière par votre passion pour la recherche et votre engagement, malgré les défis sanitaires qui ont parsemé la dernière année. Merci à Ravichandran, Irene et tous les collègues du centre de recherche de Jülich de m'avoir initié à la recherche scientifique et m'avoir si bien accueilli lors de mon séjour en Allemagne, *alles gute!*

Sur une note plus personnelle, je ne peux passer sous silence le support inconditionnel de ma famille et amis. Merci à mes parents, Jeannine et Jacôt, qui m'ont merveilleusement bien outillé en me transmettant leur goût d'apprendre et leur curiosité très tôt dans ma vie. Merci à mes soeurs Marie-Hélène et Catherine qui sont pour moi des modèles et des sources d'inspiration au quotidien. Merci à mes amis, Francis-Gabriel, éternel partenaire avec qui je peux étancher ma soif de science et avec qui j'ai toujours l'impression d'être compris, Pascal pour ta simplicité, ton authenticité et tous ces moments de folie qui sont bien gravés dans ma mémoire et qui ont fait de mon parcours universitaire une période mémorable de ma vie, Hugo et tous les membres du club scientifique RockÉTS pour les expériences uniques. Finalement, merci à ma partenaire de vie, Évelyne. Ton énergie et ta positivité contagieuses m'ont donné le souffle nécessaire pour compléter ce marathon qu'est la maîtrise.

Synthèse de diffusion sur variété riemannienne à partir d'imagerie structurelle en haute-résolution

Benoit ANCTIL-ROBITAILLE

RÉSUMÉ

Les contraintes physiques et cliniques entourant l'imagerie pondérée en diffusion (DWI) limitent souvent la résolution spatiale des images produites à des voxels jusqu'à 8 fois plus grands que ceux des images T1w. Ainsi, les informations détaillées contenues dans les images T1w pourraient aider à la synthèse d'images de diffusion à plus haute résolution. Cependant, la nature non euclidienne de l'imagerie de diffusion empêche les modèles génératifs profonds actuels de synthétiser des images physiquement plausibles. Par conséquent, est-il possible de développer un réseau de neurones profonds exploitant la géométrie du cerveau, codée dans les images structurelles en haute résolution, pour générer des images de diffusion valides au même niveau de détail ? Pour y répondre, nous développons deux objectifs : 1) la synthèse d'images de diffusion à partir d'images structurelles et 2) la super-résolution d'images de diffusion. Nous proposons donc la première architecture de réseau de neurones riemannien pour la génération directe de tenseurs de diffusion (DT) et de fonctions de distribution d'orientation de diffusion (dODF) à partir d'images T1w haute-résolution. L'intégration de la métrique log-euclidienne dans un objectif d'apprentissage garantit, contrairement aux réseaux euclidiens standards, la synthèse mathématiquement valide de la diffusion. De plus, notre approche améliore l'erreur quadratique moyenne de l'anisotropie fractionnelle (FA MSE) entre la diffusion synthétisée et réelle de plus de 23% et la similarité cosinus entre les directions principales de près de 5% par rapport aux méthodes comparées. La diffusion générée est validée en comparant les tractogrammes résultants à des données réelles attendues. La méthode proposée produit des faisceaux de fibres similaires aux données réelles avec des *streamlines* ayant moins de 3% de différence de longueur, moins de 1% de différence de volume et une apparence visuellement similaire. Bien que l'approche présentée dans ce mémoire soit capable de générer des images de diffusion en haute-résolution à partir d'entrées structurelles en moins de 15 secondes, les limites de l'inférence de la diffusion en s'appuyant uniquement sur des images T1w sont discutées. Les résultats suggèrent néanmoins une relation entre la géométrie de haut niveau du cerveau et l'architecture globale de la matière blanche.

Mots-clés: synthèse de diffusion, apprentissage sur variétés, géométrie riemannienne

Manifold-aware Synthesis of High-resolution Diffusion from Structural Imaging

Benoit ANCTIL-ROBITAILLE

ABSTRACT

The physical and clinical constraints surrounding diffusion-weighted imaging (DWI) often limit the spatial resolution of the produced images to voxels up to 8 times larger than those of T1w images. Thus, the detailed information contained in T1w images could help in the synthesis of diffusion images in higher resolution. However, the non-Euclidean nature of diffusion imaging hinders current deep generative models from synthesizing physically plausible images. Therefore, is it possible to develop a deep neural network exploiting the geometry of the brain, encoded in high-resolution structural images, to generate valid diffusion images at the same level of detail ? To answer this question, we develop two objectives : 1) the synthesis of diffusion images from structural images and 2) the super-resolution of diffusion images. We therefore propose the first Riemannian network architecture for the direct generation of diffusion tensors (DT) and diffusion orientation distribution functions (dODF) from high-resolution T1w images. Our integration of the Log-Euclidean Metric into a learning objective guarantees, unlike standard Euclidean networks, the mathematically-valid synthesis of diffusion. Furthermore, our approach improves the fractional anisotropy mean squared error (FA MSE) between the synthesized diffusion and the ground-truth by more than 23% and the cosine similarity between principal directions by almost 5% when compared to our baselines. We validate our generated diffusion by comparing the resulting tractograms to our expected real data. We observe similar fiber bundles with streamlines having less than 3% difference in length, less than 1% difference in volume, and a visually close shape. While our method is able to generate high-resolution diffusion images from structural inputs in less than 15 seconds, we acknowledge and discuss the limits of diffusion inference solely relying on T1w images. Our results nonetheless suggest a relationship between the high-level geometry of the brain and the overall white matter architecture.

Keywords: diffusion synthesis, manifold-valued data learning, riemannian geometry

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTÉRATURE	5
1.1 Le cerveau et la neuroimagerie	5
1.1.1 Le cerveau	5
1.1.2 L'imagerie par résonance magnétique	6
1.1.3 Principe de l'IRM structurelle	8
1.1.4 Principes de l'IRM de diffusion	9
1.1.5 Les compromis de l'IRM de diffusion	10
1.2 Modèles mathématiques et estimation de la diffusion	12
1.2.1 Imagerie par tenseurs de diffusion (DTI)	12
1.2.2 Fonction de distribution d'orientations de diffusion (dODF)	14
1.2.3 Variétés riemanniennes et géométrie différentielle	17
1.2.4 Géométrie riemannienne pour le traitement de tenseurs de diffusion	18
1.2.5 Métrique Log-Euclidienne sur DT	18
1.2.6 Géométrie riemannienne pour le traitement d'ODF	19
1.2.7 Re-paramétrisation racine-carré des ODF	20
1.2.8 Métrique Log-Euclidienne sur ODF	21
1.3 Synthèse d'images médicales et méthodes d'apprentissage	21
1.3.1 Réseaux de neurones profonds	23
1.3.2 Fonctions d'activation	24
1.3.3 Optimisation	25
1.3.4 Normalisation	25
1.3.5 Réseau neuronal convolutif (CNN) et pleinement convolutif (FCNN)	26
1.3.6 Convolution	27
1.3.7 Sous-échantillonnage	28
1.3.8 Réseau antagoniste génératif (GAN)	29
1.3.9 Réseau antagoniste génératif cyclique (CycleGAN)	31
1.4 Travaux connexes	33
1.4.1 Synthèse d'images structurelles-à-diffusion	33
1.4.2 Apprentissage de données à valeurs sur variété riemannienne	34
1.5 Conclusion	35
CHAPITRE 2 MÉTHODOLOGIE	39
2.1 Motivations	39
2.2 Méthode proposée	43
2.3 Apprentissage antagoniste	44
2.4 Fonction de perte cyclique	45
2.5 Régularisation par information préalable	46

2.6	Pondération de la fonction de perte	48
2.7	Fonction objectif complète	49
CHAPITRE 3 EXPÉRIMENTATIONS ET RÉSULTATS		51
3.1	Données et pré-traitement	51
3.2	Détails d'implémentation	52
3.3	Paramètres d'entraînement	53
3.4	Approches de référence	54
3.4.1	Manifold-Aware WGAN	54
3.4.2	Manifold-Aware U-Net	55
3.4.3	U-Net	55
3.5	Synthèse de DT et ODF	56
3.6	Évaluation	56
3.7	Tractographie	57
3.7.1	Génération et regroupement des <i>streamlines</i>	57
3.7.2	Évaluation des <i>streamlines</i> et des faisceaux	58
3.8	Analyse de la diffusion synthétisée	59
3.9	Longueur et volume des <i>streamlines</i>	63
3.10	Forme des <i>streamlines</i>	64
CHAPITRE 4 DISCUSSION		67
4.1	Performance de la synthèse de diffusion	67
4.2	Performance de la tractographie	69
4.2.1	DT vs. ODF	69
4.2.2	Ensembles de fibres	71
CONCLUSION ET RECOMMANDATIONS		73
5.1	Résumé des contributions	73
5.2	Limitations et travaux futurs	75
ANNEXE I	RÉTROPROPAGATION POUR L'APPRENTISSAGE DE TENSEURS DE DIFFUSION	77
ANNEXE II	ÉVALUATION QUANTITATIVE SUPPLÉMENTAIRE DE LA TRACTOGRAPHIE	79
ANNEXE III	ÉVALUATION QUALITATIVE SUPPLÉMENTAIRE DES IMAGES GÉNÉRÉES	83
BIBLIOGRAPHIE		85

LISTE DES TABLEAUX

	Page	
Tableau 3.1	Détails d'implémentation de nos réseaux générateurs G_X et G_Y et réseaux discriminateurs D_X et D_Y	54
Tableau 3.2	La similarité cosinus entre les orientations principales réelles et générées par la méthode proposée avec un nombre croissant de sujets appariés utilisé comme information préalable dans eq(2.5). Comme les métriques sont plus pertinentes dans les régions qui encodent des informations de fibres cérébrales, elles sont rapportés avec un seuil de FA croissant de 0.2 et 0.5.	60
Tableau 3.3	L'erreur quadratique moyenne de FA (FA MSE), la géodésique et la similarité cosinus entre les orientations principales des images de diffusion réelles et prédites par les différentes méthodes. Comme les métriques sont plus pertinentes dans les régions qui encodent des informations de fibres cérébrales, elles sont rapportées avec un seuil de FA croissant de 0.2 et 0.5.	61
Tableau 3.4	La quantité générée de tenseurs qui ne sont pas SPD et d'ODF qui ne sont pas des PDF par les méthodes comparées pour cinq sujets de test sélectionnés aléatoirement. L'apprentissage sur variété riemannienne (i.e., <i>manifold-awareness</i>) assure que les données de diffusion soient exactement sur leur varitété riemannienne respective.	62
Tableau 3.5	Mesures sur les faisceaux de fibres segmentés depuis la tractographie sur les données réelles et générées.	64
Tableau 3.6	Métriques de correspondance entre les faisceaux de fibres segmentés depuis la tractographie sur les données réelles et générées.	65

LISTE DES FIGURES

	Page
Figure 0.1 Schéma des chapitres principaux de ce mémoire.	3
Figure 1.1 À gauche, les trois tissus principaux du cerveau soit la matière grise (GM), la matière blanche (WM) et le liquide cérébrospinal (CSF) d'une image IRM du cerveau. À droite, l'unité de base du cerveau, le neurone.	6
Figure 1.2 Exemple de tranches axiale, coronale et sagittale d'images de cerveau acquises à l'aide de séquences d'IRM structurelle T1w (en haut) et T2w (en bas).	8
Figure 1.3 À gauche, l'axone, prolongement tubulaire du neurone. À droite, les molécules d'eau se déplacent plus rapidement le long de l'axone que perpendiculairement à son orientation principale.	9
Figure 1.4 Acquisition d'un signal pondéré en diffusion. 1) Acquisition d'un signal de référence S_0 . 2) Application d'un premier gradient magnétique orienté dans une direction donnée. 3) Diffusion des molécules d'eau selon un mouvement Brownien. 4) Application d'un second gradient magnétique dans la direction opposée au premier. 5) Mesure du signal atténué par le processus de diffusion des molécules d'eau.	10
Figure 1.5 Les compromis de l'IRM de diffusion. À gauche, l'obtention d'images en haute-résolution nécessite un long temps d'acquisition ce qui favorise la présence d'artéfacts indésirables. À droite, l'atteinte d'un haut SNR nécessite un gradient magnétique de puissance modérée, ce qui diminue aussi la sensibilité à la diffusion et le contraste des images.	11
Figure 1.6 Première ligne, un exemple des cartes de l'anisotropie fractionnelle pour une coupe dans le plan sagittal, coronal et axial d'un sujet. Deuxième ligne, un exemple des cartes de l'anisotropie fractionnelle colorée par l'orientation pour une coupe dans le plan sagittal, coronal et axial d'un sujet.	14
Figure 1.7 Méthodes de reconstruction du signal de diffusion. En haut, la reconstruction du signal en tenseur de diffusion et en bas, la reconstruction du signal HARDI en dODF par l'imagerie <i>q-ball</i>	15

Figure 1.8	Base orthonormale des harmoniques sphériques. Illustration des premières harmoniques sphériques d'ordre ℓ paire. Adaptée de (Descoteaux, 2015).	16
Figure 1.9	À gauche, la projection au plan tangent $T_{\mathbf{p}_0}\mathcal{M}$ du point \mathbf{p}_1 appartenant à la variété \mathcal{M} à l'aide de l'application log de la métrique g . À droite, l'interpolation euclidienne entre deux points appartenant à une variété riemannienne peut se retrouver hors variété contrairement à l'interpolation suivant la géodésique.	17
Figure 1.10	Un réseau de neurones profond de type <i>feed-forward</i> et son unité de base, le neurone.	23
Figure 1.11	À gauche, la fonction d'activation sigmoïde et à droite, la fonction d'activation ReLU.	24
Figure 1.12	En a), les activations de la fonction ReLU lorsque les données ne sont pas normalisées. En b), les activations de la fonction ReLU lorsque les données sont normalisées.	26
Figure 1.13	À gauche, un CNN prend une entrée \mathbf{x} et applique successivement des opérations de convolution, normalisation et sous-échantillonnage pour générer une sortie \mathbf{y} de dimension spatiale moindre. À droite, un réseau FCNN de type U-Net (Ronneberger, Fischer & Brox, 2015) prend une entrée \mathbf{x} et apprend une sortie \mathbf{y} de même dimension spatiale.	27
Figure 1.14	En a), l'opération de convolution entre un signal d'entrée et un filtre 3×3 à valeurs réelles et en b), l'opération de convolution entre un signal d'entrée à valeur sur variété Riemannienne \mathcal{S}_{++}^3 et un filtre 3×3 . à valeurs dans \mathbb{R}^9	28
Figure 1.15	En a), l'échantillonnage maximum (<i>max pooling</i>) et en b), l'échantillonnage moyen (<i>average pooling</i>)	29
Figure 1.16	Architecture d'un réseau antagoniste génératif (GAN).	30
Figure 1.17	Cycle avant de l'architecture CycleGAN (Zhu, Park, Isola & Efros, 2017). Le générateur G_{AB} prend en entrée une image réelle du domaine source A et génère son homologue dans le domaine cible B. G_{BA} prend en entrée l'image générée par G_{AB} et tente de restaurer l'image initiale. D_B évalue la plausibilité de l'image générée par rapport à la distribution réelle des données du domaine B.	32

Figure 1.18	L'architecture <i>manifold-aware</i> Wasserstein GAN (MA-WGAN) proposée par Huang, Wu & Van Gool (2019). Les applications log et exp sont utilisées pour générer des images avec des valeurs dans le plan tangent à la variété riemannienne sous-jacente aux données à partir d'un vecteur de bruit en entrée. Tirée de Huang <i>et al.</i> (2019).	35
Figure 2.1	Exemple d'entrées et de sorties de l'architecture proposée. Une image T1w en HR (première colonne), FA colorée par la direction de la vraie diffusion en LR (deuxième colonne), FA colorée par la direction de la diffusion générée en HR (troisième colonne) et l'image T1w recouverte en HR (dernière colonne) d'un sujet de test.	42
Figure 2.2	Le cycle avant de l'architecture proposée. G_Y génère un volume DT/ODF en haute-résolution sur la variété riemannienne appropriée (e.g. S_{++}^3 pour DT et \mathbb{S}^n pour ODF) en utilisant les applications log et exp. D_Y évalue la qualité des images produites et fournit un retour à G_Y . G_X reconstruit l'image T1w d'entrée originale à partir de la diffusion HR générée $\log(\exp(G_Y(\mathbf{x})))$	44
Figure 2.3	À gauche, le cycle avant proposée : Une image T1w est traduite en image de diffusion DT/ODF sur la variété S_{++}^3/\mathbb{S}^n en haute-résolution avant d'être restaurée dans le domaine des images T1w où la perte cyclique est calculée. À droite, le cycle arrière proposée : Une image de diffusion DT/ODF sur-échantillonnée dans le plan tangent $\mathcal{T}_{\text{Id}}/\mathcal{T}_{\mathbf{u}}$ est traduite dans le domaine des images T1w avant d'être restaurée en image de diffusion sur $\mathcal{T}_{\text{Id}}/\mathcal{T}_{\mathbf{u}}$ où la perte cyclique est calculée.	47
Figure 2.4	À gauche, l'histogramme de la FA typique pour un sujet de test. L'histogramme est débalancé vers des valeurs faible de FA avec une FA moyenne inférieure à 0.2. À droite, la densité des valeurs de FA de la diffusion générée avec et sans pondération de la fonction de perte par la FA.	48
Figure 3.1	L'architecture de la fonction résiduelle de sous-échantillonnage \mathcal{F} qui est utilisée pour réduire la résolution des images de diffusion synthétisées HR avant l'évaluation de D_Y	52
Figure 3.2	Visualisation de trois faisceaux segmentés à partir de tractogrammes réels et synthétisés. (À gauche), tractus corticospinal gauche (CST), fascicule arqué (AF) et fascicule longitudinal inférieur (ILF) segmentés à partir de données DT réelles et générées. (À	

droite), CST gauche, AF et ILF segmentés à partir de données ODF réelles et synthétisées.	59
Figure 3.3 Comparaison de la FA colorée par l'orientation principale des DT générés par les méthodes comparées pour les tranches sagittale, coronale et axiale d'un sujet de test.	60
Figure 3.4 Métriques sur les tranches sagittales, axiales et coronales entre le volume de diffusion haute-résolution généré pour un sujet de test aléatoire et le volume réel suréchantillonné. (Première ligne) FA MAE, (ligne du centre) la distance géodésique et (ligne du bas) la similarité cosinus.	62
Figure 3.5 Longueur des <i>streamlines</i> moyenne (en mm) pour les faisceaux segmentés de la tractographie DT (ligne du haut) et de la tractographie ODF (ligne du bas).	63
Figure 4.1 GFA de la diffusion basse-résolution réelle (première colonne), diffusion réelle suréchantillonnée (deuxième colonne) et la diffusion haute-résolution généré par les méthodes comparées. On peut remarquer la capacité de notre réseau à récupérer des détails anatomiques fins qui sont bien visibles dans l'image T1W haute-résolution, mais pas dans la diffusion originale.	68
Figure 4.2 Évaluation qualitative des ODF générés par les méthodes comparées. Première ligne, GFA colorée par l'orientation principale d'une tranche coronale, d'un sujet de test. Deuxième ligne, les ODF générées et la GFA d'une région avec croisement de fibres. Troisième ligne, une vue rapprochée des ODF générées et de la GFA. Dernière ligne, une vue rapprochée des pics des ODF générées.	70

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

2D	2 dimensions
3D	3 dimensions
AF	Arcuate Fascicle - Fascicule arqué
AR	Acoustic Radiation - Rayonnement acoustique
C	Cingulum
CB	Cerebellum - Cervelet
CC	Corpus Callosum - Corps calleux
CGAN	Conditional GAN - Réseau antagoniste génératif conditionnel
CNII	Visual Nerve - Nerf visuel
CNIII	Oculomotor Nerve - Nerf oculomoteur
CNIV	Trochlear Nerve - Nerf trochléaire
CNV	Trigeminal Nerve - Nerf trijumeau
CNVII	Facial Nerve - Nerf facial
CNVIII	Auditory Nerve - Nerf auditif
CNX	Vagus Nerve - Nerf vague
CNN	Convolutional Neural Network - Réseau neuronal convolutif
CS	Corticostriatal Pathway - Voie corticostriatale
CSF	Cerebrospinal Fluid - Liquide céphalo-rachidien
CST	Corticospinal Tract - Faisceau corticospinal
CT	Corticothalamic Pathway - Voie corticothalamique
CT	Computed Tomography - Tomodensitométrie
CTT	Central Tegmental Tract - Faisceau central de la calotte
DLF	Dorsal Longitudinal Fasciculus - Fascicule longitudinal dorsal

DNN	Deep Neural Network - Réseau de neurones profond
dODF	Diffusion Orientation Distribution Function - Fonction de distribution
DT	Diffusion tensor - Tenseur de diffusion
DTI	Diffusion tensor imaging - Imagerie par tenseur de diffusion
DWI	Diffusion Weighted Imaging - Imagerie pondérée en diffusion
EMC	Extreme Capsule - Capsule extrême
EPI	Echo-planar Imaging - Imagerie echo-planaire
F	Fornix
FA	Fractional Anisotropy - Anisotropie fractionnelle
FAT	Frontal Aslant Tract - Faisceau oblique frontal
FCNN	Fully Convolutional Neural Network - Réseau neuronal pleinement convolutif
FPT	Fronto-pontine tract - Voie fronto-pontique
GAN	Generative Adversarial Network - Réseau antagoniste génératif
GFA	Generalized Fractional Anisotropy - Anisotropie fractionnelle généralisée
GM	Grey Matter - Matière grise
HARDI	High Angular Resolution Diffusion Imaging - Imagerie de diffusion à haute résolution angulaire
HR	High-resolution - Haute résolution
ICP	Inferior Cerebellar Peduncle - Pédoncule cérébelleux inférieur
IFOF	Inferior Fronto-occipital Fasciculus - Fascicule fronto-occipital inférieur
ILF	Inferior Longitudinal Fascicle - Fascicule longitudinal inférieur
IQT	Image Quality Transfer - Transfert de qualité d'image
IRM	Imagerie par résonance magnétique
JSD	Jensen-Shannon Divergence - Divergence de Jensen-Shannon
LL	Lateral Lemniscus - Lemnisque latéral

LR	Low-resolution - Basse résolution
LSGAN	Least Squares GAN
MA-WGAN	Manifold-Aware WGAN - WGAN conscient de la variété
MCP	Middle Cerebellar Peduncle - Pédoncule cérébelleux moyen
MD	Mean Diffusion - Diffusion moyenne
MdLF	Middle Longitudinal Fasciculus
ML	Medial Lemniscus - Lemnisque médial
MLF	Middle Longitudinal Fascicle - Fascicule longitudinal moyen
MLP	Multilayer Perceptron - Perceptron multicouche
MSE	Mean Squared Error - Erreur quadratique moyenne
ODF	Orientation Distribution Function - Fonction de distribution d'orientations
OL	Overlap - Chevauchement
OPT	Occipitopontine Tract - Faisceau occipitopontin
OR	Overreach - Portée excessive
OR	Optic Radiation - Rayonnement optique
PC	Posterior Commissure - Commissure postérieure
PDF	Probability Density Function - Fonction de densité de probabilité
PPT	Parietopontine Tract - Faisceau pariétopontin
PVE	Partial Volume Effect - Effet de volume partiel
QBI	Q-Ball Imaging - Imagerie <i>q-ball</i>
ReLU	Rectified Linear Unit - Fonction unité linéaire rectifiée
RST	Rubrospinal Tract - Faisceau rubrospinal
SC-GAN	Self-Attention Conditional GAN - GAN conditionnel avec auto-attention
SCP	Superior Cerebellar Peduncle - Faisceau pédonclérospinal cérébelleux supérieur

SLF	Superior Longitudinal Fascicle - Fascicule longitudinal supérieur
SNR	Signal-to-noise Ratio - Rapport signal sur bruit d'orientations de diffusion
SPD	Symmetric Positive Definite - Symétrique positive définie
SRCNN	Super-resolution CNN - CNN super-résolution
STT	Spinothalamic Tract - Faisceau spinothalamique
T1w	T1-weighted - Pondéré en T1
tanh	Hyperbolic Tangent - Fonction tangente hyperbolique
TPT	Temporopontine Tract - Faisceau temporopontine
U	U-fiber - Fibre U
UF	Uncinate Fascicle - Fascicule unciné
WGAN	Wasserstein GAN
WM	White Matter - Matière blanche
VOF	Vertical Occipital Fasciculus - Fascicule occipital vertical

LISTE DES SYMBOLES ET UNITÉS DE MESURE

\mathcal{S}_{++}^*	Variété des matrices symétriques positives définies
\mathcal{S}_{++}^3	Variété des matrices symétriques positives définies 3×3
\mathbb{S}^n	n-Sphère
mm	Millimètre

INTRODUCTION

L'Homme est constamment en quête d'une meilleure compréhension des mécanismes biologiques et des structures composant son propre corps. Historiquement, le développement des connaissances médicales s'est fait par l'entremise d'analyses histologiques *post-mortem* et d'observations externes du corps humain. Cependant, depuis la découverte des rayons X par Wilhelm Conrad Roentgen en 1895, les techniques d'imagerie médicale *in-vivo* non invasives n'ont cessé d'évoluer. Aujourd'hui, l'imagerie par résonnance magnétique (IRM) est un outil indispensable permettant d'obtenir des images tridimensionnelles précises et détaillées d'une vaste majorité d'organes internes au corps humain. Parmi ces organes, le cerveau demeure un des plus complexes et des moins bien compris. Au cœur de celui-ci, la matière blanche, composée de milliards de fibres cérébrales, s'occupe d'acheminer l'information d'une région à l'autre du cerveau. L'étude de ces fines structures tubulaires est un élément central dans l'avancement de la compréhension du fonctionnement du cerveau, mais aussi dans le diagnostic et le traitement de diverses maladies neurodégénératives (Huang, Friedland & Auchus, 2007; Gattellaro, Minati, Grisoli, Mariani, Carella, Osio, Ciceri, Albanese & Bruzzone, 2009b) et cognitives (Kantarci, Murray, Schwarz, Reid, Przybelski, Lesnick, Zuk, Raman, Senjem, Gunter, Boeve, Knopman, Parisi, Petersen, Jack & Dickson, 2017; Neuner, Kupriyanova, Stöcker, Huang, Posnansky, Schneider, Tittgemeyer & Shah, 2010) ainsi qu'à la planification neurochirurgicale (Costabile, Alaswad, D'Souza, Thompson & Ormond, 2019). Des séquences d'IRM spécifiques à l'imagerie de la matière blanche et divers modèles mathématiques ont donc été développés dans les dernières décennies afin de permettre une caractérisation toujours plus juste des fibres cérébrales. Établies sur le principe du mouvement brownien des molécules d'eau, ces séquences, dites pondérées en diffusion, souffrent cependant toujours de diverses limitations physiques et cliniques en comparaison aux séquences d'acquisition d'IRM structurelles conventionnelles (p. ex. T1w). En effet, la séquence IRM utilisée dans l'imagerie de diffusion, l'imagerie écho-planaire (EPI), induit un compromis entre le temps d'acquisition, le rapport signal sur bruit (SNR) et la

résolution des images produites. De ce fait, les voxels, l'équivalent du pixel en trois dimensions, des images acquises par IRM de diffusion peuvent atteindre jusqu'à huit fois la taille des voxels d'une image structurelle d'un même cerveau (Shi, Cheng, Wang, Yap & Shen, 2016). Il a été démontré que cette imprécision a des répercussions sur les analyses subséquentes des images de diffusion (Oouchi, Yamada, Sakai, Kizu, Kubota, Ito & Nishimura, 2007; Alexander, Hasan, Lazar, Tsuruda & Parker, 2001). Le développement de méthodes permettant d'améliorer la résolution de telles images dans un contexte clinique réaliste est donc une étape essentielle pour une meilleure évaluation de la matière blanche.

L'intelligence artificielle, de par sa capacité à modéliser des relations non linéaires complexes, s'est frayée un chemin parmi les technologies les plus prometteuses dans de multiples domaines, dont l'analyse d'images médicales. Particulièrement, l'apprentissage profond s'est démarqué dans des tâches telles que la segmentation d'images médicales, mais aussi dans la synthèse et la super-résolution d'images. Le développement de méthodes d'apprentissage automatique appliquées à la synthèse et à la super-résolution d'images de diffusion est donc une avenue intéressante, mais toujours peu exploitée. Pour cause, les réseaux de neurones artificiels conventionnels établis sur des opérations mathématiques euclidiennes peinent à comprendre la nature riemannienne des modèles mathématiques de diffusion. Cette incapacité se traduit par des estimations de diffusion qui sont physiquement invalides ou imprécises.

Ce mémoire est donc dédié au développement d'un modèle d'apprentissage profond permettant la synthèse d'images de diffusion en haute-résolution tout en assurant la validité mathématique de celles-ci. Dans ce mémoire, nous nous attaquons aux limites des réseaux de neurones traditionnels en intégrant des éléments de la géométrie riemannienne à même leur architecture. Le but étant d'élargir l'application de ces réseaux aux images dites à valeurs sur variétés riemanniennes comme les images à tenseurs de diffusion (DTI) ou à fonctions de distribution d'orientations de diffusion (dODF). Comme application, nous testons la faisabilité d'exploiter

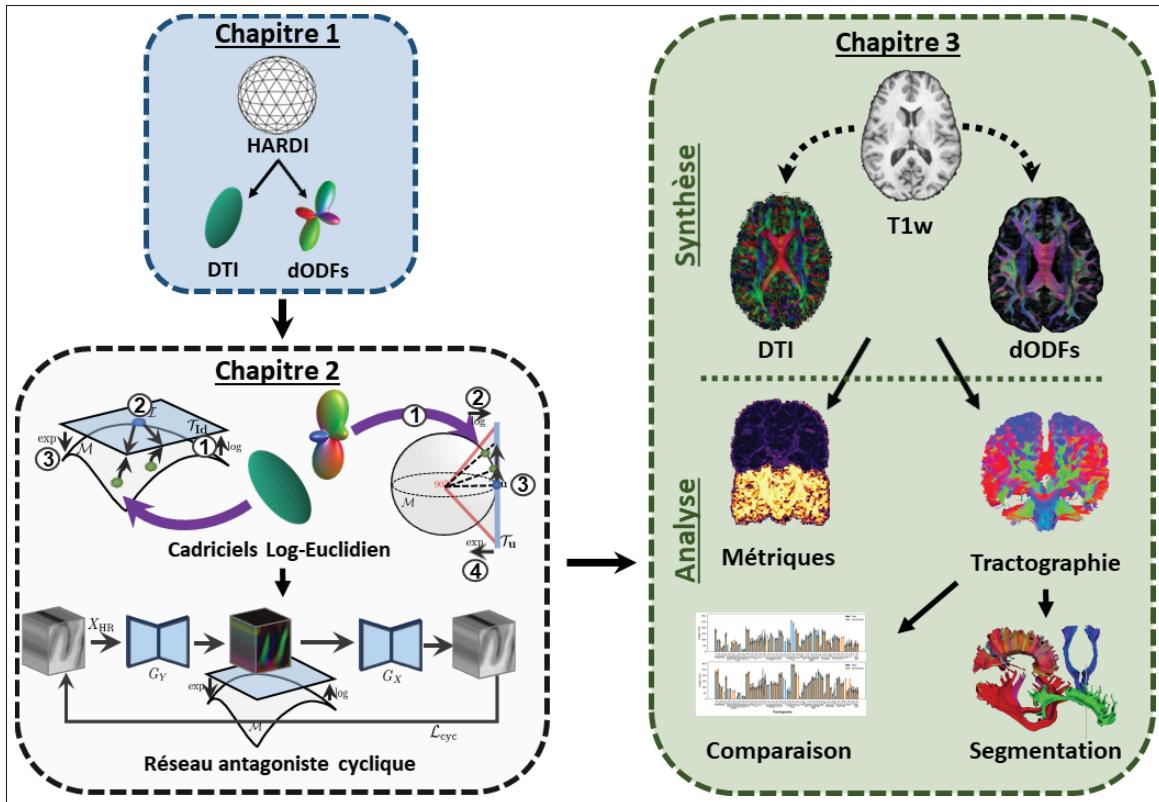


Figure 0.1 Schéma des chapitres principaux de ce mémoire.

l’information contenue dans les images T1w en haute résolution afin d’améliorer la qualité des images de diffusion en basse résolution. En utilisant comme guide des images IRM structurelles afin d’améliorer la résolution des images de diffusion, ce mémoire traite par le fait même de la **question de recherche suivante** : *Est-il possible d’exploiter la géométrie structurelle à haut-niveau du cerveau afin de prédire ses populations de fibres cérébrales sous-jacentes ?*

Pour explorer cette question de recherche, les objectifs spécifiques suivants sont développés :

1. **Objectif 1** - La synthèse d’images de diffusion à partir d’images structurelles par apprentissage de données sur variétés riemanniennes ;
2. **Objectif 2** - La génération faiblement supervisée de super-résolution d’imagerie de diffusion à partir d’images structurelles.

Le développement de ces objectifs amène concrètement les contributions scientifiques suivantes :

1. Le premier réseau d'apprentissage profond riemannien de type CycleGAN pour le mappage entre des images à valeurs réelles et des images à valeurs se trouvant sur les variétées riemanniennes \mathcal{S}_{++}^3 et \mathbb{S}^n ;
2. Le premier modèle d'apprentissage profond pour la super-résolution de DTI et dODF guidée par des images IRM structurelles non appariées et avec un minimum de supervision;
3. Une analyse détaillée des images de diffusion synthétisées à partir d'images structurelles incluant l'évaluation des cartes scalaires de diffusion, de la tractographie et des faisceaux segmentés.

En plus des contributions notées ci-haut, les travaux détaillés dans ce mémoire ont mené à la publication d'**un article de workshop**⁴ et d'une présentation orale à l'International Workshop on Computational Diffusion MRI à MICCAI 2020 ainsi qu'à **un article de journal**⁵ qui, au moment d'écrire ces lignes, est en processus de soumission.

Ce mémoire est composé de cinq chapitres. D'abord, les concepts théoriques et la littérature nécessaires à la compréhension du mémoire sont présentés dans le premier chapitre *Revue de littérature*. Ces concepts incluent les principes de l'IRM de diffusion, les modèles de tenseurs de diffusion et de dODF et les réseaux de neurones artificiels. Dans le deuxième chapitre, la méthode proposée est détaillée. Le chapitre *Expérimentations et résultats* présente les résultats quantitatifs et qualitatifs obtenues. Une discussion des résultats est faite dans le quatrième chapitre et finalement, les limitations de la méthodes et des recommandations sont détaillées dans la conclusion.

⁴ <https://arxiv.org/abs/2004.00173>

⁵ <https://arxiv.org/abs/2108.04135>

CHAPITRE 1

REVUE DE LITTÉRATURE

1.1 Le cerveau et la neuroimagerie

Le cerveau demeure aujourd’hui l’organe le moins bien compris du corps humain. Bien qu’il soit au cœur de processus aussi fondamentaux que le ressenti, la perception ou la motricité, son architecture complexe et microscopique en complique son étude. Malgré tout, grâce à des méthodes d’imagerie médicale non invasives toujours plus sophistiquées, il est aujourd’hui possible de l’imager *in-vivo* en quelques minutes. Les sections suivantes présentent brièvement l’anatomie du cerveau humain et les techniques d’imagerie médicale nécessaires à la compréhension du mémoire.

1.1.1 Le cerveau

Le cerveau est composé de trois types de tissus soit la matière grise, la matière blanche et le liquide cérébrospinal tel qu’illustre à la Figure 1.1.

La matière grise est principalement composée des corps cellulaires des neurones, présent par milliards (Williams & Herrup, 1988) et responsable de recevoir, traiter et de retransmettre les influx nerveux du cerveau. La matière blanche, quant à elle, est principalement composée de faisceaux d’axones qui servent à interconnecter les neurones des différentes régions de la matière grise. Finalement, le CSF est le liquide dans lequel baigne le cerveau.

La caractérisation de chacun de ces tissus et l’étude de leurs fonctions permettent de mieux comprendre le cerveau et sont essentielles pour le traitement et le diagnostic de différentes maladies neurologiques. Par exemple, l’épaisseur de la matière grise peut être utilisée pour diagnostiquer de façon précoce l’Alzheimer (Querbes, Aubry, Pariente, Lotterie, Démonet, Duret, Puel, Berry, Fort, Celsis et al., 2009). La matière blanche elle, est utile pour l’étude de maladies comme le Parkinson (Gattellaro, Minati, Grisolí, Mariani, Carella, Osio, Ciceri, Albanese & Bruzzone,

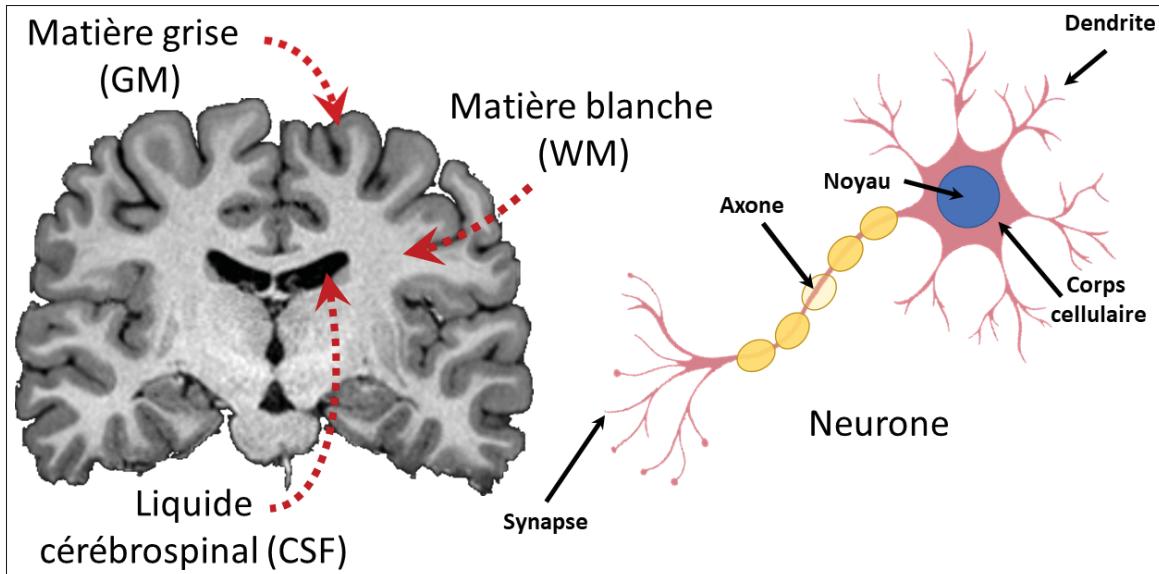


Figure 1.1 À gauche, les trois tissus principaux du cerveau soit la matière grise (GM), la matière blanche (WM) et le liquide cérébrospinal (CSF) d'une image IRM du cerveau. À droite, l'unité de base du cerveau, le neurone.

2009a; Melzer, Watts, MacAskill, Pitcher, Livingston, Keenan, Dalrymple-Alford & Anderson, 2013), l'Alzheimer (Agosta, Pievani, Sala, Geroldi, Galluzzi, Frisoni & Filippi, 2011) ou la schizophrénie (Davis, Stewart, Friedman, Buchsbaum, Harvey, Hof, Buxbaum & Haroutunian, 2003). Finalement, l'analyse du CSF est utile pour identifier diverses conditions en lien avec le cerveau et la moelle épinière telles que la méningite bactérienne et la neurosyphilis (Seehusen, Reeves & Fomin, 2003).

Les sections suivantes présentent une des technologies les plus communes pour l'imagerie des tissus cérébraux, l'imagerie par résonance magnétique.

1.1.2 L'imagerie par résonance magnétique

L'imagerie par résonance magnétique permet d'imager une multitude d'organes *in-vivo*, de façon non invasive et sans exposition à des rayons ionisants. Fondée sur les principes de la résonance magnétique nucléaire (Bloch, 1946; Purcell, Torrey & Pound, 1946), la technologie derrière l'IRM exploite les propriétés physiques des noyaux atomiques combinées à un champ magnétique

puissant et à une énergie radiofréquence afin de générer des images (Mansfield & Grannell, 1973; Lauterbur, 1973).

La physique quantique montre qu'un noyau atomique a un moment magnétique. Dans un champ magnétique B_0 , sa précession peut être déterminée par la fréquence de Larmor ω donnée par l'équation suivante :

$$\omega = \gamma \cdot B_0 \quad (1.1)$$

Chaque isotope a un rapport gyromagnétique γ différent donc une fréquence de Larmor différente. Du point de vue de l'IRM, le noyau d'hydrogène est le plus pertinent, car il est présent dans l'eau qui compose plus de 60% d'un corps humain adulte moyen (Jéquier & Constant, 2010). Il est donc intéressant d'étudier le comportement des noyaux d'hydrogène afin de caractériser des lésions, de localiser des tumeurs ou de comprendre les microstructures sous-jacentes à une variété d'organes, comme le cerveau.

Lorsqu'un patient est placé le scanner IRM, la plupart des noyaux d'hydrogène de la partie du corps qui est imagée s'alignent avec le fort champ magnétique homogène B_0 du scanner. Des impulsions radiofréquences (RF) ayant exactement la même fréquence que la fréquence de Larmor des noyaux d'hydrogène (42,58 MHz/Tesla) est alors transmise aux noyaux. Ces signaux RF externes fournissent l'énergie nécessaire aux noyaux pour les faire passer dans un état de haute énergie et synchroniser leur *spins* électromagnétiques. L'alignement et la synchronisation des *spins* crée un signal radio pouvant être détecté par une bobine RF. Différentes atténuations du signal sont par la suite mesurées à l'aide de séquences IRM spécifiques aux tissus à imager. Ce sont ces atténuations du signal qui donne le contraste aux tissus dans les images produites.

1.1.3 Principe de l'IRM structurelle

L'imagerie par résonance magnétique structurelle permet d'imager et de visualiser les différentes structures et anomalies physiques du cerveau. Les séquences dites pondérées en T1 (T1w) et pondérées en T2 (T2w) sont les séquences d'IRM structurelle les plus communes. Suite à l'absorption de l'énergie des pulsations RF externes par les protons d'hydrogène et à la synchronisation des *spins* électromagnétiques, les protons retrouvent un état d'équilibre énergétique. Cette période de retour à l'équilibre est appelée temps de relaxation et est spécifique à chaque tissu. Pour les images T1w et T2w, les temps de relaxation T1 et T2 sont attendus avant de lire le signal RF résiduel émis par les protons des tissus imaginés. Selon le taux de relaxation des tissus au temps T1 et T2, le signal RF lu par la bobine du scanner est plus ou moins intense. C'est cette variation dans le signal RF qui permet de donner le contraste aux images T1w et T2w.

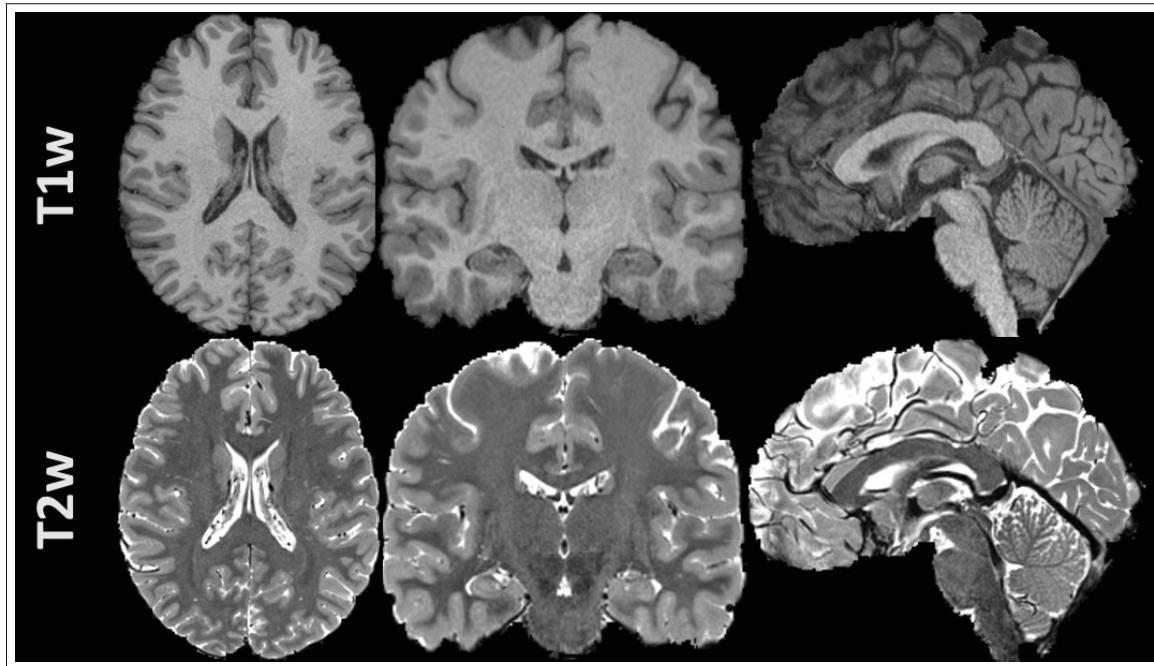


Figure 1.2 Exemple de tranches axiale, coronale et sagittale d'images de cerveau acquises à l'aide de séquences d'IRM structurelle T1w (**en haut**) et T2w (**en bas**).

Les images T1w et T2w ont généralement une bonne résolution spatiale avec des voxels pouvant atteindre une taille inférieure au millimètre cube et un bon contraste des tissus mous tel qu'illusté à la Figure 1.2.

1.1.4 Principes de l'IRM de diffusion

Sensible au mouvement des molécules d'eau, l'IRM de diffusion permet de caractériser les microstructures sous-jacentes aux tissus imaginés. Dans la matière blanche du cerveau, la diffusion est majoritairement contrainte par les parois des ensembles de fibres cérébrales tel qu'illustré dans la Figure 1.3. Dans ces structures tubulaires nommées axones, l'eau se déplace plus rapidement le long des parois que dans la direction perpendiculaire à celles-ci.

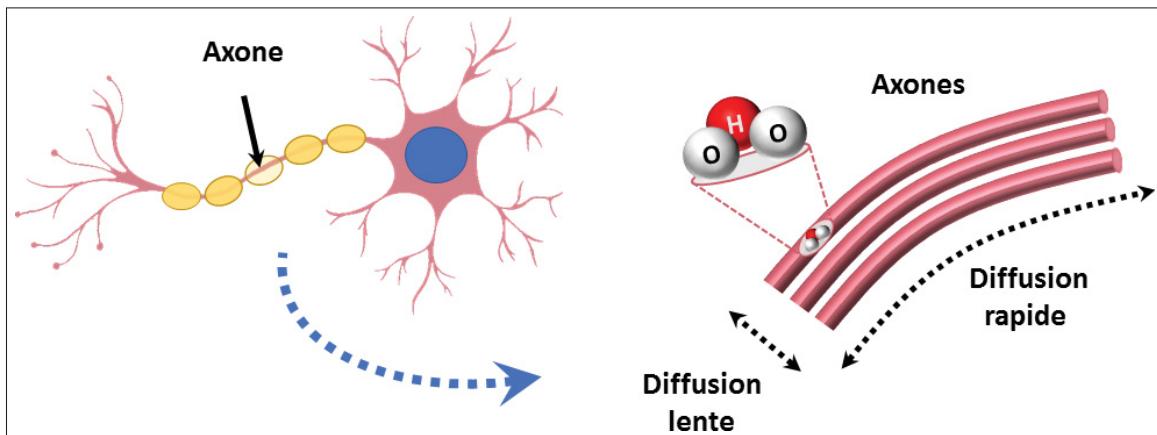


Figure 1.3 À gauche, l'axone, prolongement tubulaire du neurone. À droite, les molécules d'eau se déplacent plus rapidement le long de l'axone que perpendiculairement à son orientation principale.

Cette anisotropie de diffusion se reflète dans les images pondérées en diffusion (DWI) du cerveau par une intensité plus élevée là où il y a présence de fibres cérébrales que dans les régions où l'eau se déplace plus librement comme dans les ventricules. En prenant des images DWI sous plusieurs angles, il est alors possible d'estimer, avec l'aide de modèles mathématiques, la population de fibres cérébrales présente dans chaque voxel de l'image.

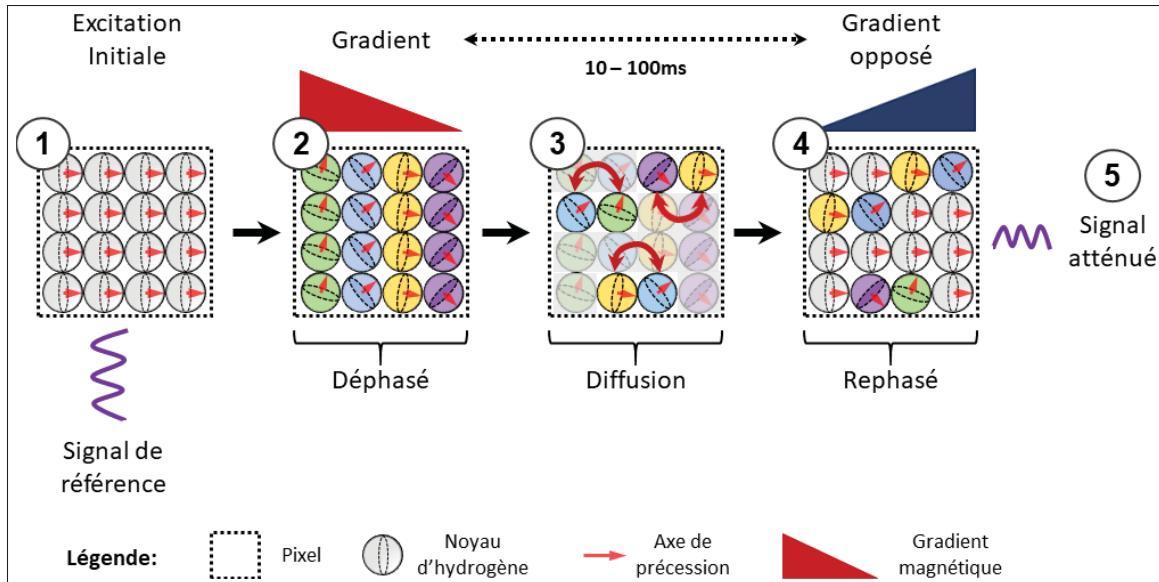


Figure 1.4 Acquisition d'un signal pondéré en diffusion. **1)** Acquisition d'un signal de référence S_0 . **2)** Application d'un premier gradient magnétique orienté dans une direction donnée. **3)** Diffusion des molécules d'eau selon un mouvement Brownien. **4)** Application d'un second gradient magnétique dans la direction opposée au premier. **5)** Mesure du signal atténué par le processus de diffusion des molécules d'eau.

Comme pour l'IRM structurelle, les noyaux d'hydrogène sont d'abord alignés par un fort champ magnétique B_0 et excités à l'aide de pulsations RF. Une image est acquise à ce moment et est utilisée comme référence pré-diffusion nommée image S_0 . Ultérieurement, deux gradients magnétiques opposés sont appliqués successivement dans un intervalle entre 10 et 100 millisecondes comme illustré dans la Figure 1.4. Plus les molécules d'eau se diffusent pendant l'intervalle de temps entre l'application des deux gradients magnétiques opposés, plus la différence est grande entre le signal S_0 et le signal mesuré post-diffusion. Pour finir, ce processus est répété dans de multiples orientations nommées *b-vectors* et parfois à plusieurs intensités de gradient magnétique *b-values*.

1.1.5 Les compromis de l'IRM de diffusion

L'IRM de diffusion souffre principalement de deux compromis illustrés à la Figure 1.5. Premièrement, il y a un compromis entre la résolution des images produites, la robustesse aux

artefacts de mouvement et le temps d'acquisition. Un temps d'acquisition prolongé permet d'acquérir des images avec une meilleure résolution, mais favorise du même coup la présence d'artefacts de mouvement dans les images produites. Un long temps d'acquisition n'est pas favorable non plus d'un point de vue clinique, car généralement peu agréable pour le patient et plus coûteux.

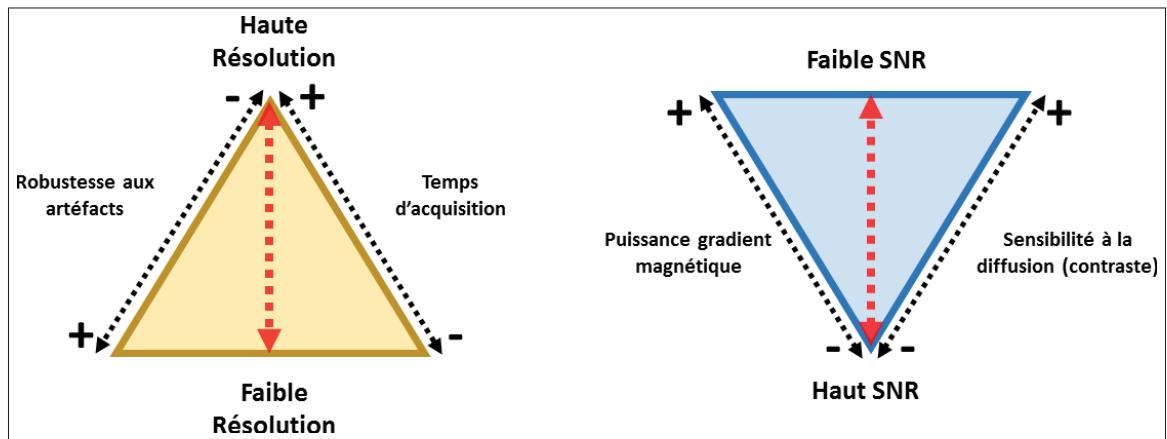


Figure 1.5 Les compromis de l'IRM de diffusion. À gauche, l'obtention d'images en haute-résolution nécessite un long temps d'acquisition ce qui favorise la présence d'artefacts indésirables. À droite, l'atteinte d'un haut SNR nécessite un gradient magnétique de puissance modérée, ce qui diminue aussi la sensibilité à la diffusion et le contraste des images.

Le deuxième compromis important vise le rapport signal sur bruit, le contraste dans les images produites et la puissance des gradients magnétiques. En effet, l'utilisation de forts gradients magnétiques accentue la différence entre le signal de référence pré-diffusion S_0 et le signal mesuré post-diffusion. Cette différence est favorable, car elle augmente le contraste entre les régions à forte diffusion et celles à faible diffusion. Cependant, cette sensibilité accrue à la diffusion augmente la présence de bruit dans les images produites et donc, réduit leur SNR.

En raison de ces compromis, les images de diffusion souffrent généralement d'un plus long temps d'acquisition, d'une moins bonne résolution spatiale et d'un plus faible SNR que les images structurelles comme les images T1w.

1.2 Modèles mathématiques et estimation de la diffusion

Dans cette section, deux des méthodes les plus utilisées pour l'estimation de la diffusion sont détaillées soient, l'imagerie par tenseurs de diffusion et par fonctions de distribution d'orientations de diffusion. Ces modèles font partie intégrante de ce mémoire et il est important de comprendre leurs propriétés mathématiques et leurs limitations afin d'apprécier les contributions du présent travail.

1.2.1 Imagerie par tenseurs de diffusion (DTI)

Le tenseur de diffusion, introduit en 1992 par Basser et al., modélise la fonction de densité de probabilité (PDF) de la diffusion des molécules d'eau à l'aide d'une distribution gaussienne en trois dimensions. Concrètement, la fonction gaussienne en 3D, visualisée par un ellipsoïde, estime la probabilité P qu'une molécule d'eau initialement à une position x_0 se trouve à une position x_τ après une diffusion de durée τ . Cette probabilité est notée :

$$P(x_\tau|x_0, \tau) = \frac{1}{\sqrt{(4\pi\tau)^3|\mathbf{D}|}} \exp\left(-\frac{1}{4\tau}(x_\tau - x_0)^T \mathbf{D}^{-1} (x_\tau - x_0)\right), \quad (1.2)$$

où $|\mathbf{D}|$ est le déterminant du tenseur de diffusion \mathbf{D} . Le tenseur est donc proportionnelle à la matrice de covariance de dimension 3×3 qui paramètre la fonction gaussienne 3D estimant la PDF de la diffusion et est noté comme suit :

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} \quad (1.3)$$

Comme \mathbf{D} est symétrique, six de ses neufs valeurs sont uniques. Afin d'estimer ces six valeurs, l'équation de Stejskal-Tanner suivant est résolue :

$$S(b, \mathbf{g}) = S_0 e^{b\mathbf{g}^T \mathbf{D}\mathbf{g}} \quad (1.4)$$

Ici, S_0 est l'image de référence acquise pré-diffusion, \mathbf{g} est le vecteur unité représentant l'orientation du gradient magnétique (i.e., *b-vector*), b est la puissance du gradient magnétique (i.e., *b-value*) et \mathbf{D} est le tenseur de diffusion à estimer. Comme on tente de résoudre une équation à six inconnus, un minimum de six images DWI acquises avec une orientation \mathbf{g} différente sont nécessaires pour solutionner eq(1.4).

Les tenseurs de diffusion ont des propriétés mathématiques particulières ; ils sont symmétriques, positifs et définis (SPD). Par conséquent, un tenseur \mathbf{D} est décomposable en trois valeurs propres strictement réelles et positives $\lambda_1 \geq \lambda_2 \geq \lambda_3$ et trois vecteurs propres $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ de la façon suivante :

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)^T. \quad (1.5)$$

Le vecteur propre \mathbf{u}_1 , associé à la valeur propre la plus grande λ_1 , indique la direction vers laquelle la vitesse de diffusion des molécules d'eau est la plus grande. Par conséquent, l'orientation principal du tenseur est alignée avec la population de fibres sous-jacentes comme illustré à la Figure 1.7. À partir des valeurs et vecteurs propres de \mathbf{D} , plusieurs mesures peuvent être extraites. Parmi les plus utilisées, on y retrouve l'anisotropie fractionnelle (FA) et la carte de FA colorée par l'orientation principale de la diffusion. La FA est une valeur scalaire dans l'intervalle $[0, 1]$ mesurant l'anisotropie de la diffusion qui est calculée à partir des valeurs propres du tenseur de diffusion comme suit :

$$FA(\lambda_1, \lambda_2, \lambda_3) = \sqrt{\frac{1}{2} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2}}{\sqrt{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}} \quad (1.6)$$

Plus la valeur de FA d'un voxel est grande, plus la diffusion est directionnelle à cet endroit et donc plus la probabilité que ce voxel exprime une population de fibres cérébrales est importante. Un seuil de $FA \geq 0.2$ est communément utilisé pour distinguer les voxels qui encodent l'information

de fibres des autres voxels. La carte de FA colorée par la direction, ou *color FA*, quant à elle, est une image RGB où la couleur est fonction de l'orientation principale de la diffusion et la saturation est fonction de la FA. Un exemple de carte de ces deux mesures est représenté dans la Figure 1.6.

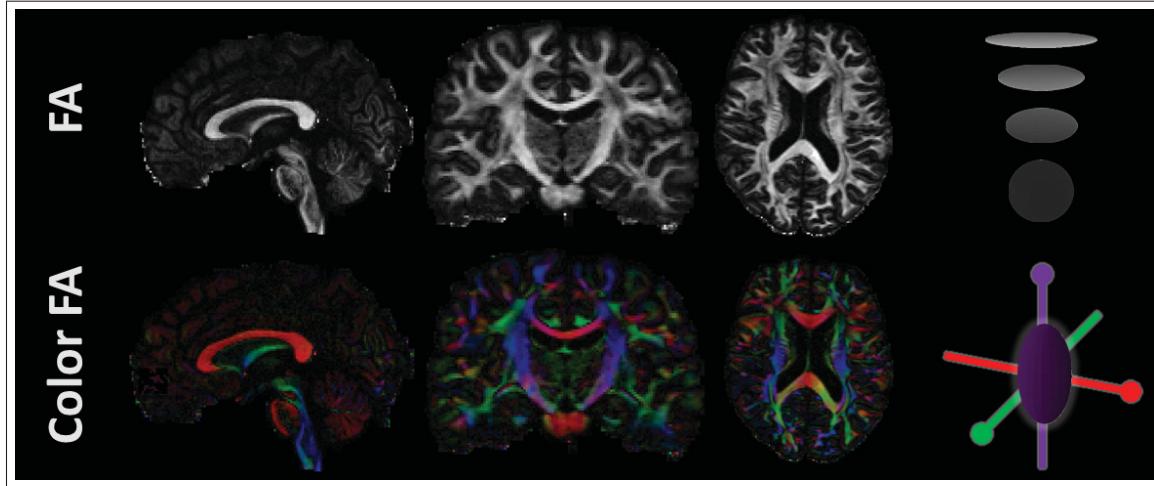


Figure 1.6 **Première ligne**, un exemple des cartes de l'anisotropie fractionnelle pour une coupe dans le plan sagittal, coronal et axial d'un sujet. **Deuxième ligne**, un exemple des cartes de l'anisotropie fractionnelle colorée par l'orientation pour une coupe dans le plan sagittal, coronal et axial d'un sujet.

1.2.2 Fonction de distribution d'orientations de diffusion (dODF)

Le modèle de tenseur de diffusion, qui estime la PDF de la diffusion par une fonction gaussienne en 3D, souffre de nombreuses limitations. Entre autres, la réduction du phénomène de diffusion en un processus gaussien ne permet pas au DT de correctement modéliser le croisement de plusieurs fibres dans un même voxel. Par conséquent, les méthodes d'estimation de la diffusion dites sans-modèle, comme l'estimation par fonctions de distribution d'orientations (ODF), sont de plus en plus utilisées. En estimant directement la vraie PDF de la diffusion, et ce sans modèle, les ODF permettent d'approximer avec une plus grande fidélité que le DTI les configurations complexes de fibres.

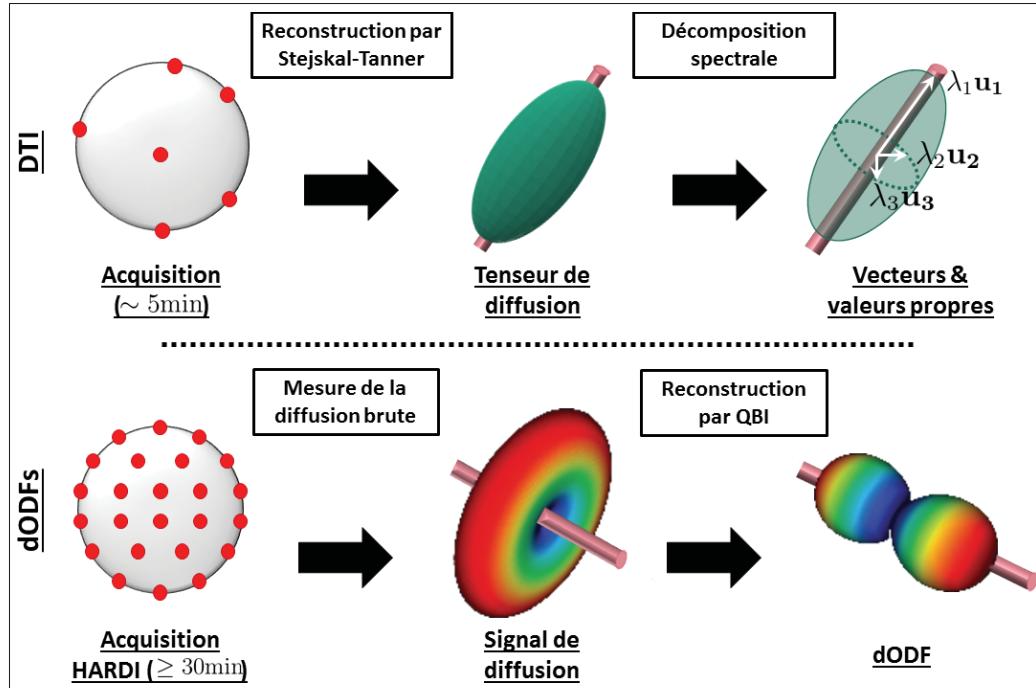


Figure 1.7 Méthodes de reconstruction du signal de diffusion. **En haut**, la reconstruction du signal en tenseur de diffusion et **en bas**, la reconstruction du signal HARDI en dODF par l'imagerie *q-ball*.

Pour ce faire, l'imagerie de diffusion à haute résolution angulaire (HARDI) est utilisée. Bien qu'il existe plus d'une méthode pour recueillir des données HARDI et pour reconstruire les dODF, ce mémoire se concentre sur l'échantillonnage sphérique monocoque (i.e., *single shell HARDI*) et la reconstruction de dODF par l'imagerie *q-ball* (QBI). L'échantillonnage sphérique monocoque consiste à mesurer l'atténuation du signal de diffusion sur une sphère uniformément discrétisée en N parcelles où chaque parcelle $n \in \{1, \dots, N\}$ définit l'orientation du gradient magnétique d'une acquisition. Plus N est grand, plus l'estimation de la direction de la diffusion est précise, mais plus l'acquisition est longue et le traitement des données demandant en ressources. Typiquement, un minimum de $N = 60$ acquisitions est nécessaire pour une résolution angulaire acceptable. À partir des N mesures acquises sur la sphère, aussi nommée coque ou *shell*, les dODF sont reconstruites par QBI. Comme on peut le voir dans la Figure 1.7, le signal de diffusion est maximal dans le plan perpendiculaire à la population de fibres sous-jacentes, alors que la dODF est alignée à celle-ci. L'imagerie *q-ball*, fondée sur la transformation de

Funk-Radon, transforme donc le signal de diffusion de sorte que le maximum de la fonction sur la sphère soit aligné avec la population de fibres sous-jacentes. Cette représentation, que l'on nomme fonction de distribution d'orientations de diffusion, est plus intuitive que le signal brut HARDI en plus d'être une estimation lisse et sans modèle de la vraie PDF de diffusion.

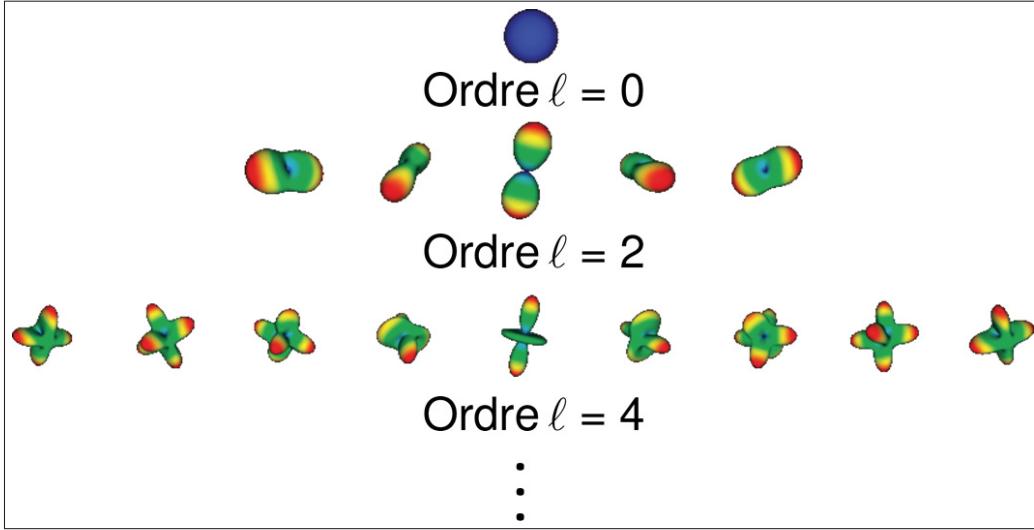


Figure 1.8 Base orthonormale des harmoniques sphériques. Illustration des premières harmoniques sphériques d'ordre ℓ paire. Adaptée de (Descoteaux, 2015).

Les ODF sont donc des fonctions PDF sur la sphère unitaire \mathbb{S}^2 . En pratique, les ODF sont représentées par un histogramme à N valeurs où chaque valeur est une estimation de l'ODF pour une orientation $s \in \mathbb{S}^2$. Généralement, un minimum de $N = 362$ valeurs uniformément distribuées sur \mathbb{S}^2 sont utilisées. La représentation et le traitement des ODF sous cette forme nécessite donc beaucoup de mémoire et de temps de calcul. Pour cette raison, les ODF sont la plupart du temps représentées à l'aide d'un nombre K d'harmoniques sphériques orthonormées formant une base pour les fonctions sur \mathbb{S}^2 . En utilisant un nombre de fonctions de base K suffisamment grand, toute fonction sur \mathbb{S}^2 , incluant les ODF, peut être approximée avec une bonne fidélité. Les harmoniques sphériques d'ordre 4, qui comprennent un ensemble de 15 fonctions de base illustrées à la Figure 1.8, représentent généralement un bon compromis entre la précision de l'estimation des ODF et les ressources nécessaires à leur stockage et à leur traitement.

1.2.3 Variétés riemanniennes et géométrie différentielle

Une variété riemannienne (\mathcal{M}, g) est un espace courbe et différentiable \mathcal{M} qui est équipé d'une métrique riemannienne g permettant entre autres de définir l'opération de produit intérieur $\langle \cdot, \cdot \rangle$ et une distance dans cet espace (Pennec, Sommer & Fletcher, 2019).

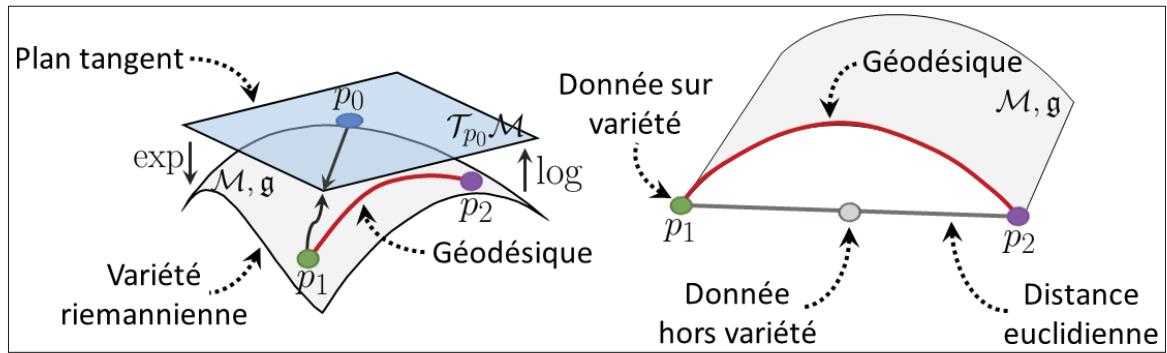


Figure 1.9 **À gauche**, la projection au plan tangent $T_{p_0}\mathcal{M}$ du point p_1 appartenant à la variété \mathcal{M} grâce à l'application de la carte log de la métrique g . **À droite**, l'interpolation euclidienne entre deux points appartenant à une variété riemannienne peut se retrouver hors variété contrairement à l'interpolation suivant la géodésique.

Grâce aux applications exponentielle (\exp) et logarithme (\log) définies par la métrique g , il est possible d'établir une relation entre les points $\mathbf{p} \in \mathcal{M}$ et leur homologue dans un espace tangent à \mathcal{M} (Hopf & Rinow, 1931), noté $T_{\mathbf{p}}\mathcal{M}$, de sorte que :

$$\forall \mathbf{p} \in \mathcal{M}, \log_{\mathbf{p}}(\mathbf{q}) \in T_{\mathbf{p}}\mathcal{M} \quad (1.7)$$

$$\forall \mathbf{v} \in T_{\mathbf{p}}\mathcal{M}, \exp_{\mathbf{p}}(\mathbf{v}) \in \mathcal{M} \quad (1.8)$$

Dans l'espace tangent à un point de référence \mathbf{p} , la distance géodésique entre \mathbf{p} et un autre point $\mathbf{q} \in \mathcal{M}$ est donnée par la norme du vecteur $\vec{\mathbf{pq}} \in T_{\mathbf{p}}\mathcal{M} = \log_{\mathbf{p}}(\mathbf{q})$. Contrairement à la distance euclidienne, la distance géodésique considère la topologie de \mathcal{M} tel qu'illustré à la Figure 1.9. L'interpolation le long de la géodésique assure donc que la donnée interpolée reste à la surface de \mathcal{M} , ce qui n'est pas le cas suivant la distance euclidienne.

1.2.4 Géométrie riemannienne pour le traitement de tenseurs de diffusion

Telle que mentionné dans la Section 1.2.1, les tenseurs de diffusion sont des matrices 3×3 , ici notées \mathbf{M} , qui sont symétriques positives et définies (SPD) et qui peuvent être décomposées, à l'aide de la décomposition spectrale, en trois valeurs propres réelles et positives Σ et trois vecteurs propres correspondants \mathbf{U} tel que $\mathbf{M} = \mathbf{U}\Sigma\mathbf{U}^\top$.

En raison de leurs propriétés SPD, les tenseurs reposent sur une variété non linéaire notée

$$\mathcal{S}_{++}^3 = \{\mathbf{M} \in \mathbb{R}^{3 \times 3}, \mathbf{M} = \mathbf{M}^\top, \mathbf{x}^\top \mathbf{M} \mathbf{x} > 0, \forall \mathbf{x} \in \mathbb{R}^3, \|\mathbf{x}\|_2 > 0\} \quad (1.9)$$

\mathcal{S}_{++}^3 n'est pas un espace vectoriel (c'est-à-dire que la combinaison linéaire de deux éléments dans \mathcal{S}_{++}^3 peut se trouver en dehors de cet espace), donc l'utilisation d'opérations euclidiennes standard pour le calcul de statistiques sur les tenseurs de diffusion peut conduire à des effets indésirables comme l'effet de gonflement (*swelling effect*) documenté dans (Arsigny, Fillard, Pennec & Ayache, 2006). De plus, omettre la géométrie de la variété \mathcal{S}_{++}^3 lors de l'interpolation ou la synthèse de DT en exploitant des outils ne reposant que sur des opérations euclidiennes peut résulter en des tenseurs qui ne sont pas SPD. De tels tenseurs sont physiquement incorrects et doivent être évités. Pour traiter avec précision les DT, (Arsigny *et al.*, 2006) a proposé d'équiper \mathcal{S}_{++}^3 d'une métrique riemannienne dite log-euclidienne décrite dans la section suivante.

1.2.5 Métrique Log-Euclidienne sur DT

Les matrices de tenseurs de diffusion sont bien définies dans la métrique log-euclidienne, où un logarithme matriciel et une exponentielle peuvent être facilement traités dans une métrique et peuvent toujours être mappés vers des tenseurs SPD valides (Arsigny *et al.*, 2006). Soit $\mathbf{M} = \mathbf{U}\Sigma\mathbf{U}^\top$ la décomposition en éléments propres d'une matrice symétrique \mathbf{M} . Le calcul des

applications logarithme et exponentielle d'un tenseur, noté \log_{Id} et \exp_{Id} , est défini comme suit :

$$\forall \mathbf{P} \in \mathcal{S}_{++}^3, \log_{\text{Id}}(\mathbf{P}) = \mathbf{U} \log(\boldsymbol{\Sigma}) \mathbf{U}^\top \in \mathcal{T}_{\text{Id}} \quad (1.10)$$

$$\forall \mathbf{S} \in \mathcal{T}_{\text{Id}}, \exp_{\text{Id}}(\mathbf{S}) = \mathbf{U} \exp(\boldsymbol{\Sigma}) \mathbf{U}^\top \in \mathcal{S}_{++}^3 \quad (1.11)$$

En utilisant ces définitions, la distance géodésique entre deux points sur la variété \mathcal{S}_{++}^3 , \mathbf{P}_1 et \mathbf{P}_2 , peut alors être exprimée comme

$$dist(\mathbf{P}_1, \mathbf{P}_2) = \|\log_{\text{Id}}(\mathbf{P}_1) - \log_{\text{Id}}(\mathbf{P}_2)\|_2. \quad (1.12)$$

1.2.6 Géométrie riemannienne pour le traitement d'ODF

Les fonctions de distribution d'orientation, représentées ici par $\mathbf{p}(\mathbf{s})$, sont des fonctions de densité de probabilité modélisant la diffusion des molécules d'eau à n'importe quel angle \mathbf{s} sur la 2-sphère \mathbb{S}^2 . L'espace \mathcal{P} de ces PDF forme l'ensemble

$$\mathcal{P} = \{\mathbf{p} : \mathbb{S}^2 \rightarrow \mathbb{R}^+ \mid \forall \mathbf{s} \in \mathbb{S}^2, \mathbf{p}(\mathbf{s}) \geq 0; \int_{\mathbf{s} \in \mathbb{S}^2} \mathbf{p}(\mathbf{s}) d\mathbf{s} = 1\} \quad (1.13)$$

L'ensemble \mathcal{P} n'est pas un espace vectoriel, mais une variété différentielle non linéaire qui, tout comme la variété \mathcal{S}_{++}^3 susmentionnée, doit être équipée d'une métrique riemannienne efficace pour calculer avec précision les statistiques sur celle-ci (Srivastava, Jermyn & Joshi, 2007). Heureusement, les PDF peuvent être reparamétrés de plusieurs manières, ce qui conduit à des variétés connues avec des opérations riemanniennes de forme fermée et efficaces en termes de complexité de calcul. Le reparamétrage par racine carrée des PDF est particulièrement pratique car il aboutit à une variété sphérique de Hilbert unitaire avec une métrique \mathbb{L}^2 (Srivastava *et al.*, 2007).

1.2.7 Re-paramétrisation racine-carré des ODF

A l'aide du reparamétrage racine carrée d'ODF $\psi(\mathbf{s}) = \sqrt{\mathbf{p}(\mathbf{s})}$, $\forall \mathbf{s} \in \mathbb{S}^2$, l'espace ψ peut être vu comme l'orthant positif d'une sphère de Hilbert unitaire :

$$\psi = \left\{ \psi : \mathbb{S}^2 \rightarrow \mathbb{R}^+ \mid \forall \mathbf{s} \in \mathbb{S}^2, \psi(\mathbf{s}) \geq 0; \int_{\mathbf{s} \in \mathbb{S}^2} \psi^2(\mathbf{s}) d\mathbf{s} = 1 \right\} \quad (1.14)$$

où les applications exponentielles et logarithmes ainsi que la géodésique sont définies sous une forme fermée. En pratique, les ODF sont représentés sous forme d'histogrammes de N groupes (*bins*) où chaque groupe représente une orientation sur une 2-sphère discrétisée. Dans le cadre du traitement d'ODF spatiallement organisés en 3D au sein d'une image, travailler directement avec les ODF reparamétrés $\psi(\mathbf{s})$ nécessiterait des ressources considérables. En effet, N est typiquement de l'ordre de quelques centaines. Pour alléger ce fardeau computationnel, $\psi(\mathbf{s})$ peut être représenté sous une forme plus compacte en utilisant une base de fonctions harmoniques sphérique (Descoteaux, Angelino, Fitzgibbons & Deriche, 2007; Cheng, Ghosh, Jiang & Deriche, 2009) comme suit :

$$\psi(\mathbf{s}) = \sum_{i=1}^K c_i \mathbf{B}_i(\mathbf{s}) \quad (1.15)$$

Ici, $K \ll N$ est le nombre de fonctions de base orthonormées utilisées pour représenter $\psi(\mathbf{s})$ et $\{\mathbf{B}_i\}_{i \in K}$ est l'ensemble de fonctions harmoniques sphériques formant la base comme dans (Descoteaux *et al.*, 2007). À partir de cette représentation paramétrique, un cadre log-euclidien efficace a été proposé dans (Cheng *et al.*, 2009) et est présenté dans la Section 1.2.6 suivante. De la même manière que la FA du modèle tenseur de diffusion, l'anisotropie fractionnelle généralisée (GFA) (Tuch, 2004a) de l'ODF peut être calculée à partir des coefficients des harmoniques sphériques comme dans l'éq(1.16) ci-dessous :

$$\text{GFA} = \sqrt{1 - \frac{(c_0^0)}{\sum_{k=0}^L \sum_{m=-k}^k (c_k^m)^2}}. \quad (1.16)$$

Dans cette formulation, la GFA mesure la distance entre l'ODF et la distribution uniforme.

1.2.8 Métrique Log-Euclidienne sur ODF

Étant donné la représentation paramétrique de $\psi(\mathbf{s})$ dans l'eq(1.15), la racine carrée de tout ODF peut être exprimée par ses coordonnées riemanniennes $\mathbf{c} = (c_1, c_2, \dots, c_K)^\top$ et donne la famille de probabilités PF_K :

$$PF_K = \left\{ p(\mathbf{s} | \mathbf{c}) = \left(\sum_{i=1}^K c_i \mathbf{B}_i(\mathbf{s}) \right)^2 : \int_{\mathbb{S}} p(\mathbf{s} | \mathbf{c}) d\mathbf{s} = \sum_{i=1}^K c_i^2 = 1, \sum_{i=1}^K c_i \mathbf{B}_i(\mathbf{s}) \geq 0, \forall \mathbf{s} \in \mathbb{S}^2 \right\}. \quad (1.17)$$

Ensuite, l'espace des paramètres PS_K peut être défini comme

$$PS_K = \left\{ \mathbf{c} \mid \|\mathbf{c}\| = \sqrt{\sum_{i=1}^K c_i^2} = 1, \sum_{i=1}^K c_i \mathbf{B}_i(\mathbf{s}) \geq 0, \forall \mathbf{s} \in \mathbb{S}^2 \right\} \quad (1.18)$$

qui est aussi un sous-ensemble de la variété sphérique \mathbb{S}^{K-1} . La sphère, étant une variété simple et bien étudiée, rend le cadre log-euclidien pour le calcul des ODF simple et efficace, comme on le voit dans l'eq(1.19) et l'eq(1.20) ci-dessous :

$$\forall \mathbf{c} \in PS_K \subset S^{k-1}, \log_{\mathbf{u}}(\mathbf{c}) = \frac{\mathbf{c} - \mathbf{u} \cos \Psi}{\|\mathbf{c} - \mathbf{u} \cos \Psi\|_2} \Psi, \text{ où } \Psi = \arccos(\langle \mathbf{u} | \mathbf{c} \rangle) \quad (1.19)$$

$$\forall \mathbf{v}_c \in \mathcal{T}_{\mathbf{u}}, \exp_{\mathbf{u}}(\mathbf{v}_c) = \mathbf{u} \cos \Psi + \frac{\mathbf{v}_c}{\|\mathbf{v}_c\|_2} \sin \Psi, \text{ où } \Psi = \|\mathbf{v}_c\|_2. \quad (1.20)$$

Ici, \mathbf{u} est la fonction de distribution d'orientation uniforme définie comme $\mathbf{u} = (1, 0, \dots, 0)$. Les applications riemanniennes définies dans l'eq(1.19) et l'eq(1.20) permettent de calculer des statistiques sur les ODF avec précision et d'assurer la validité des ODF résultants.. En plus, le cadre log-euclidien propose une géodésique simple entre deux ODF paramétrés $p(\cdot | \mathbf{c})$ et $p(\cdot | \mathbf{c}')$ comme suit :

$$dist(p(\cdot | \mathbf{c}), p(\cdot | \mathbf{c}')) = \|\log_{\mathbf{u}}(\mathbf{c}) - \log_{\mathbf{u}}(\mathbf{c}')\|_2. \quad (1.21)$$

1.3 Synthèse d'images médicales et méthodes d'apprentissage

La synthèse d'images médicales, processus par lequel une image médicale est générée numériquement, a connu un intérêt croissant dans les dernières décennies. Pour cause, la synthèse

d'images médicales promet des applications cliniques intéressantes telles que le remplacement de procédures d'acquisition longues et coûteuses, la réduction de l'exposition des patients à des rayons ionisants ou l'amélioration de la qualité des images acquises (Wang, Lei, Fu, Wynne, Curran, Liu & Yang, 2021). Cet intérêt a d'abord mené au développement de méthodes permettant de synthétiser des images en utilisant des règles explicitement définies par des experts et spécifiques à chaque modalité (Wang *et al.*, 2021). Par exemple, Ramsey & Oliver (1998); Ramsey, Arwood, Scaperoth & Oliver (1999); Stanescu, Hans-Sonke, Stavrev & Fallone (2006) ont proposé des méthodes permettant la génération d'images similaires à celles acquises par tomodensitométrie (CT) à partir d'images structurelles IRM. Bien que ces méthodes aient démontré un certain succès pour la génération d'images structurelles (e.g. T1w à CT), l'élaboration manuelle de fonctions de transfert entre deux modalités exprimant des informations de nature différente (e.g. T1w à diffusion) reste fastidieuse (Wang *et al.*, 2021).

Récemment, les méthodes d'apprentissage automatique, et plus particulièrement les réseaux de neurones profonds, ont surpassé les algorithmes de synthèse d'images à base de règles (Wang *et al.*, 2021). Les algorithmes d'apprentissage automatique ne nécessitent pas la création manuelle de fonctions de transfert entre modalités, mais plutôt une fonction de perte qui mesure la différence entre la sortie de l'algorithme et le résultat attendu. En minimisant cette fonction de perte à l'aide de données, les réseaux de neurones peuvent apprendre la synthèse d'images médicales avec plus de flexibilité et à partir d'entrées ayant un lien complexe avec la sortie du réseau. Les réseaux de neurones profonds ont donc permis le développement de nouvelles applications de la synthèse d'images médicales, telles que la super-résolution (Kim, Do & Park, 2018; Sánchez & Vilaplana, 2018; Chen, Shi, Christodoulou, Xie, Zhou & Li, 2018) et la génération de modalités non-structurelles (Son, Yong Park, Byeon & Park, 2019; Schilling, Blaber, Huo, Newton, Hansen, Nath, Shafer, Williams, Resnick, Rogers, Anderson & Landman, 2019; Sanaat, Arabi, Mainta, Garibotto & Zaidi, 2020). Une revue des concepts entourant les réseaux de neurones profonds ainsi que les architectures de synthèse d'images les plus importantes sera faite dans les sections suivantes.

1.3.1 Réseaux de neurones profonds

Un réseau de neurones profond est une fonction $\mathbf{y} = \mathbf{f}(\mathbf{x}; \theta)$ qui estime le mappage $\mathbf{f}' : \mathbf{x} \rightarrow \mathbf{y}$ entre des données en entrée \mathbf{x} et une valeur de sortie \mathbf{y} à l'aide de paramètres appris θ . Selon la nature du problème, les valeurs d'entrées \mathbf{x} et de sorties \mathbf{y} du réseau peuvent prendre plusieurs formes et être de dimension variable. Le perceptron multicouche (*multilayer perceptron MLP*), illustré à la Figure 1.10, est une extension du perceptron proposé par Rosenblatt (1958). Ce modèle représente la forme la plus simple des réseaux de neurones profonds à propagation avant (*feed-forward*).

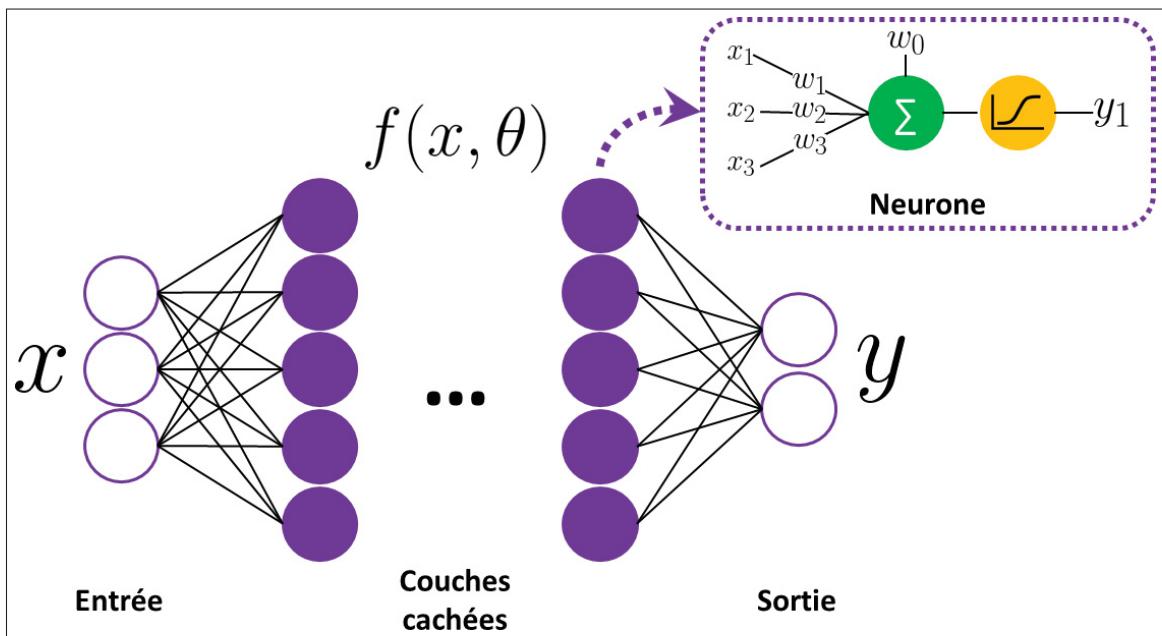


Figure 1.10 Un réseau de neurones profond de type *feed-forward* et son unité de base, le neurone.

Dans ce type de réseau, les informations circulent de la couche d'entrée à la couche de sortie en se voyant appliquer une succession de transformations linéaires et de fonctions d'activation. Un MLP de l couches est donc défini par une composition de fonctions affines f et de fonctions d'activation σ tel que $\mathbf{f}(\mathbf{x}; \theta) = (\sigma_l(f_l) \circ \dots \circ \sigma_2(f_2) \circ \sigma_1(f_1))(\mathbf{x})$.

1.3.2 Fonctions d'activation

Il a été démontré par Cybenko (1989) et par le théorème d'approximation universelle (Hornik, Stinchcombe & White, 1989) qu'un MLP pour lequel les activations σ sont sigmoïdales peut approximer toute fonction réelle et continue avec une précision raisonnable. Outre la fonction sigmoïde, la fonction tangente hyperbolique (tanh) et la fonction unité linéaire rectifiée (ReLU) illustrées à la Figure 1.11 sont aujourd'hui parmi les fonctions d'activation les plus utilisées. Ces fonctions jouent un rôle important dans la puissance expressive des réseaux de neurones profonds en leur prodiguant une non-linéarité nécessaire à l'approximation de fonctions complexes.

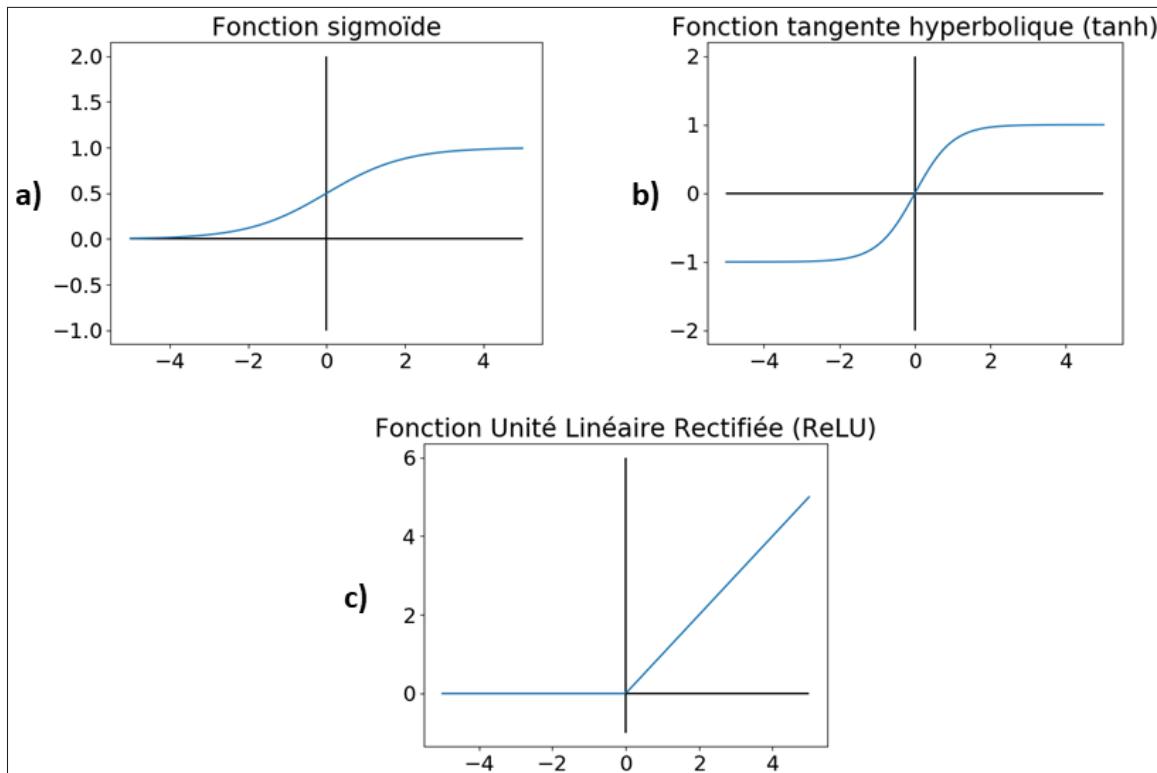


Figure 1.11 À gauche, la fonction d'activation sigmoïde et à droite, la fonction d'activation ReLU.

Les fonctions d'activations utilisées dans les réseaux de neurones profonds sont continues, différentiables sur l'ensemble de leur domaine et généralement monotones. Ces propriétés permettent d'entraîner les réseaux de neurones à l'aide de la rétro-propagation du gradient présentée à la Section 1.3.3 suivante.

1.3.3 Optimisation

L'optimisation des réseaux de neurones à propagation avant se fait à l'aide de l'algorithme de rétro-propagation du gradient (*backpropagation*) (Rumelhart, Hinton & Williams, 1986). Cet algorithme permet de calculer les gradients d'une fonction de perte \mathcal{L} par rapport aux paramètres $w_{ij} \in \theta$ du réseau entraîné, et ce grâce aux règles de dérivation en chaîne. Suivant cet algorithme, une quantité Δw_{ij} peut être calculée et ajoutée à chacun des paramètres du réseau afin de minimiser \mathcal{L} de sorte que

$$\Delta w_{ij} = -\eta \frac{\partial \mathcal{L}}{\partial w_{ij}}, \quad (1.22)$$

où η est le taux d'apprentissage (*learning rate*). La méthode de mise à jour des paramètres dépend de l'optimiseur utilisé. La descente de gradient stochastique (SGD) (Kiefer & Wolfowitz, 1952), qui est à la base des méthodes d'optimisation itératives récentes comme Adagrad (Duchi, Hazan & Singer, 2011) ou Adam (Kingma & Ba, 2015), met à jour les paramètres appris θ pour chaque exemple d'entraînement de la façon suivante :

$$\theta_{t+1} = \theta_t - \eta \frac{\partial \mathcal{L}}{\partial \theta_t}, \quad (1.23)$$

Dans l'éq(1.23) ci-dessus, θ_t représente les paramètres du réseau de neurones avant le pas d'optimisation et θ_{t+1} représente les paramètres du réseau mis à jour.

1.3.4 Normalisation

La mise à jour des paramètres des réseaux de neurones par descente de gradient lors de l'entraînement introduit des changements dans les distributions des activations du réseau (*internal covariate shift*). Les méthodes de normalisation, telle que la normalisation par lot (*batch normalization*) (Ioffe & Szegedy, 2015), permettent de décaler et de mettre à l'échelle ces distributions. En plus d'accélérer la convergence du réseau, ceci a pour effet de stabiliser l'entraînement en diminuant la variabilité dans les distributions des activations au cours de l'entraînement.

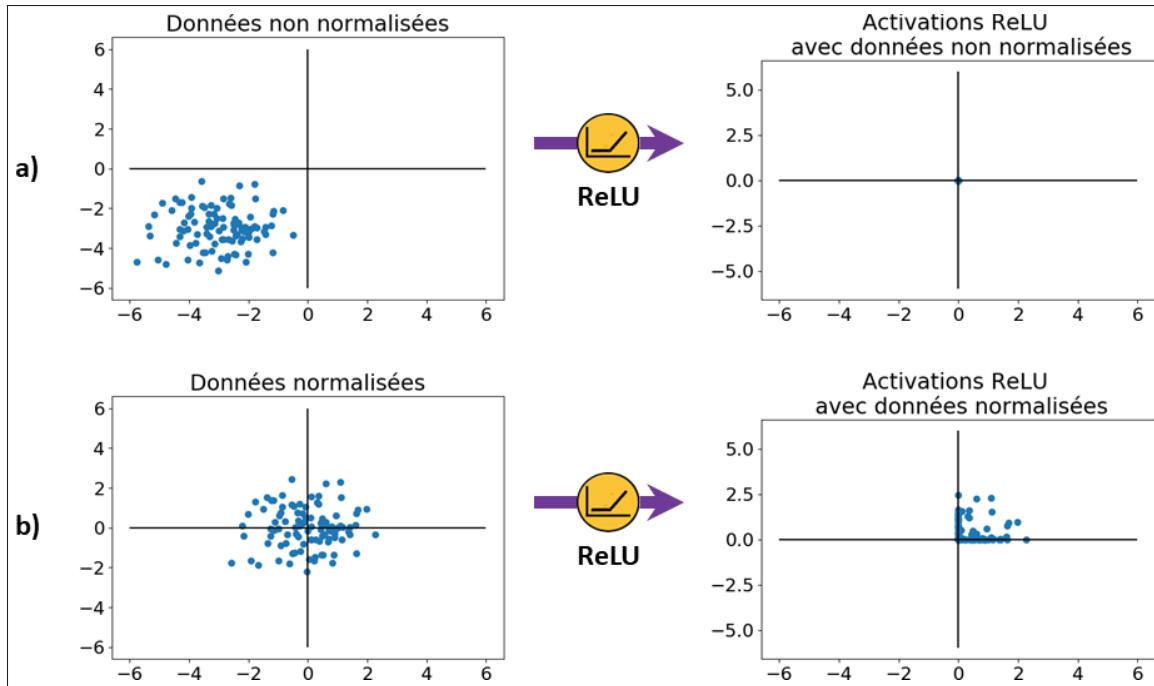


Figure 1.12 En a), les activations de la fonction ReLU lorsque les données ne sont pas normalisées. En b), les activations de la fonction ReLU lorsque les données sont normalisées.

La Figure 1.12 démontre l’effet de la normalisation qui permet de propager un signal négatif, malgré l’utilisation d’une fonction d’activation ReLU, en centrant les activations (en bleu) à zéro.

1.3.5 Réseau neuronal convolutif (CNN) et pleinement convolutif (FCNN)

Le perceptron multicouche décrit à la Section 1.3.1 permet d’extraire des caractéristiques à partir de données qui sont sous forme de vecteurs. Cependant, lorsque l’entrée ou la sortie d’un réseau est une image (e.g., pour la synthèse ou la classification d’images), les réseaux convolutifs (CNN) (LeCun, Haffner, Bottou & Bengio, 1999) et pleinement convolutifs (FCNN) (Long, Shelhamer & Darrell, 2015) sont plus appropriés. En effet, les opérations de convolution et de sous-échantillonnage, qui sont à la base des CNN, sont particulièrement efficaces pour l’extraction hiérarchique de caractéristiques dans les images. Pour cause, les opérations de

convolution et de sous-échantillonnage permettent d'exploiter l'organisation spatiale et la corrélation locale dans les images naturelles tout en favorisant le partage de paramètres.

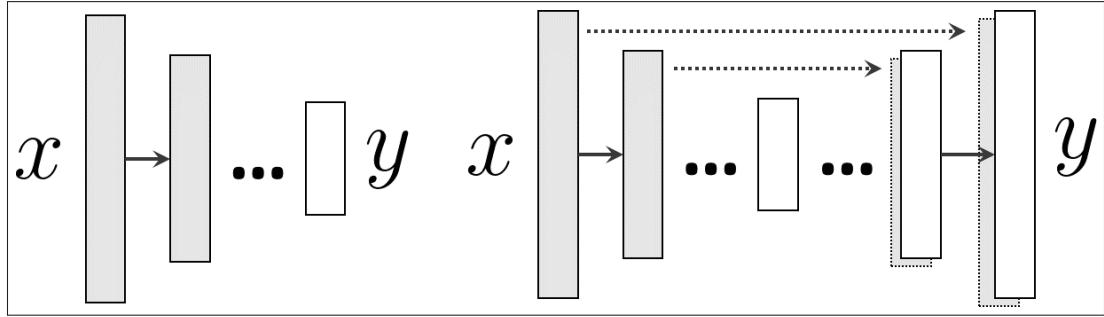


Figure 1.13 **À gauche**, un CNN prend une entrée x et applique successivement des opérations de convolution, normalisation et sous-échantillonnage pour générer une sortie y de dimension spatiale moindre. **À droite**, un réseau FCNN de type U-Net (Ronneberger *et al.*, 2015) prend une entrée x et apprend une sortie y de même dimension spatiale.

1.3.6 Convolution

La convolution permet d'extraire des caractéristiques locales en calculant la réponse d'un signal, ici une image, à un filtre, tel que démontré à la Figure 1.14. Ce filtre, généralement de taille 3×3 , 5×5 ou 7×7 est appliqué pas-à-pas sur l'ensemble de l'image indépendamment de sa localisation spatiale.

La réponse est calculée à chaque déplacement du filtre à l'aide du produit scalaire entre les valeurs de l'images dans le champs récepteur du filtre et le filtre lui-même. L'ensemble des réponses forment une nouvelle image nommée carte de caractéristiques (*feature map*) où l'intensité des pixels varie selon la similarité locale entre l'image et le filtre. En utilisant un ensemble de filtres différents, il est possible de détecter plusieurs caractéristiques distinctes dans une même image et de regrouper les cartes de caractéristiques produites en canaux dans un même volume. Chaque canal du volume résultant indique la similarité, définie par le produit scalaire, entre l'image et le filtre associé au canal. L'opération de convolution traditionnelle nécessite donc que les données d'entrée respectent deux propriétés essentielles : 1) être organisées sous forme de grille et 2) exister dans un espace équipé d'un produit intérieur, par exemple un espace vectoriel. Ces deux

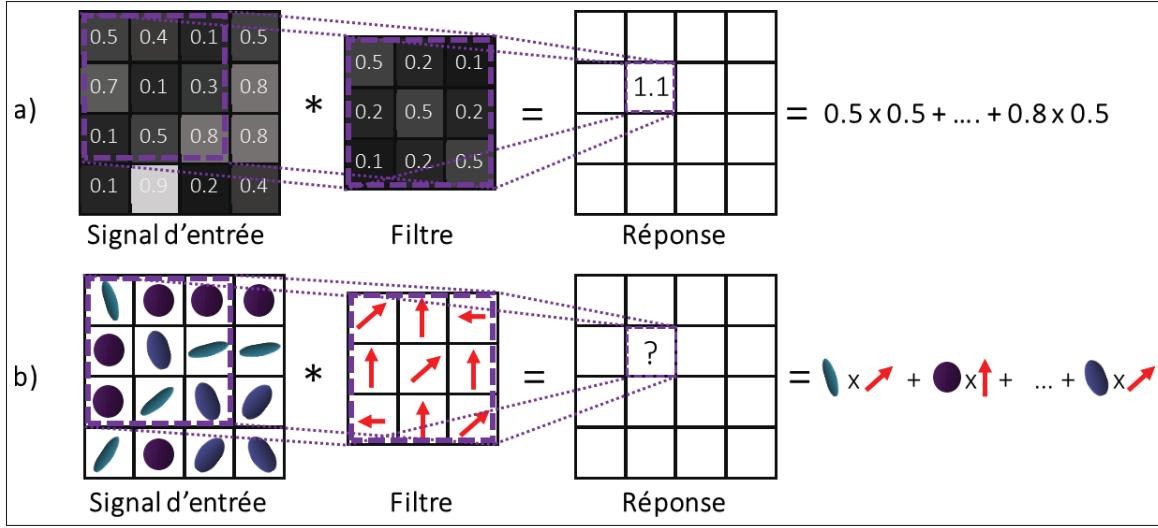


Figure 1.14 En a), l'opération de convolution entre un signal d'entrée et un filtre 3×3 à valeurs réelles et en b), l'opération de convolution entre un signal d'entrée à valeurs sur varitété Riemannienne S^3_{++} et un filtre 3×3 . à valeurs dans \mathbb{R}^9 .

propriétés permettent à la fois de définir la façon dont le filtre parcourt les données d'entrée, mais aussi de calculer la réponse entre les valeurs de l'image et le filtre. Lorsque les données ne respectent pas les deux conditions nommées ci-dessus, comme c'est le cas avec les images DTI ou ODF, la convolution est indéfinie (voir Figure 1.14 b).

1.3.7 Sous-échantillonnage

Les méthodes de sous-échantillonnage les plus utilisées dans les réseaux de neurones profonds sont l'échantillonnage maximum (*max pooling*) et l'échantillonnage moyen (*average pooling*). Ces fonctions, qui sont invariantes à la permutation, permettent de réduire la résolution spatiale des cartes de caractéristique à différents niveaux des réseaux de neurones. Concrètement, l'opération se fait *pas-à-pas* sur toute l'image sous l'action d'un noyau couvrant généralement 2×2 pixels, tel qu'illustré à la Figure 1.15.

En prenant soit la valeur maximale ou la valeur moyenne sous le champs récepteur du noyau et en déplaçant celui-ci sans chevauchement, la résolution spatiale de la sortie est amoindrie. Comme les images naturelles ont généralement des pixels voisins corrélés entre eux, l'échantillonnage a

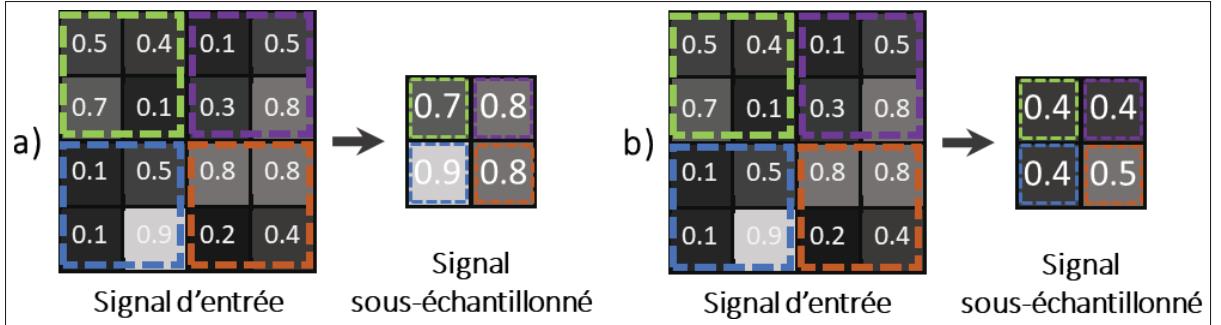


Figure 1.15 En a), l'échantillonnage maximum (*max pooling*) et en b), l'échantillonnage moyen (*average pooling*)

pour effet de réduire les informations redondantes tout en induisant une certaine robustesse à la translation locale aux réseaux.

1.3.8 Réseau antagoniste génératif (GAN)

Les réseaux antagonistes génératifs (GAN) (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville & Bengio, 2014) permettent l'apprentissage non supervisé de la distribution d'un ensemble de données réelles par l'entremise d'un jeu minimax à deux joueurs. Tel qu'illustré à la Figure 1.16, un réseau générateur \mathcal{G} prend en entrée un vecteur de bruit \mathbf{z} et tente de générer des échantillons $\mathcal{G}(\mathbf{z})$ suffisamment similaires à ceux d'une distribution cible qu'un deuxième réseau \mathcal{D} ne puisse dire si les échantillons qui lui sont présentés sont réels ou pas. Le discriminateur \mathcal{D} , quant à lui, est entraîné simultanément afin de pouvoir reconnaître si les échantillons proviennent de la distribution réelle \mathbb{P}_{data} ou de la distribution estimée. L'objectif minimax d'un tel réseau est noté

$$\min_{\mathcal{G}} \max_{\mathcal{D}} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{data}} [\log(\mathcal{D}(\mathbf{x}))] + \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_z} [1 - \log(\mathcal{D}(\mathcal{G}(\mathbf{z})))] \quad (1.24)$$

où \mathbf{x} est issue de la distribution réelle et \mathbf{z} provient d'une distribution a priori. À terme, il a été démontré que eq(1.24) minimise la divergence de Jensen-Shannon (JSD) entre la distribution réelle et la distribution apprise (Goodfellow *et al.*, 2014). Dans un contexte de synthèse d'images, les échantillons générés $\mathcal{G}(\mathbf{z})$ et réels \mathbf{x} sont des images. Par conséquent, l'architecture des réseaux

\mathcal{G} et \mathcal{D} sont des réseaux de neurones convolutifs (LeCun *et al.*, 1999; Radford, Metz & Chintala, 2016).

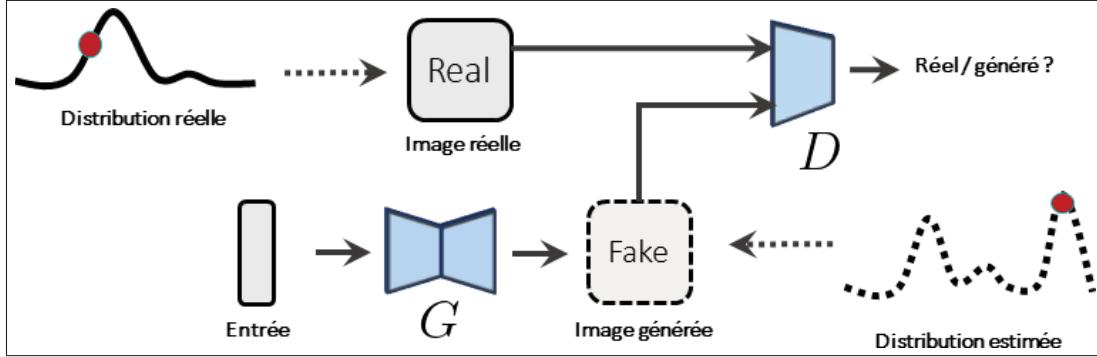


Figure 1.16 Architecture d'un réseau antagoniste génératif (GAN).

L'architecture GAN souffre néanmoins de plusieurs problèmes d'optimisation comme l'effondrement des modes (*mode collapse*) ou la disparition du gradient (*gradient vanishing*) (Salimans, Goodfellow, Zaremba, Cheung, Radford & Chen, 2016; Goodfellow, 2017). Pour palier à ces faiblesses, des alternatives ont été proposées. Par exemple, Least Squares GAN (LSGAN) (Mao, Li, Xie, Lau, Wang & Paul Smolley, 2017) et Wasserstein GAN (WGAN) (Arjovsky, Chintala & Bottou, 2017) proposent de remplacer respectivement l'entropie croisée, utilisée comme fonction de perte dans (Goodfellow *et al.*, 2014), par une norme ℓ_2 et par la distance de Wasserstein. Les fonctions objectifs utilisées par LSGAN sont décrites comme suit :

$$\min_{\mathcal{D}} \mathcal{L}_{\text{LSGAN}}(\mathcal{D}) = \frac{1}{2} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{\text{data}}(\mathbf{x})} [(\mathcal{D}(\mathbf{x}) - 1)^2] + \frac{1}{2} \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_{\mathbf{z}}(\mathbf{z})} [(\mathcal{D}(\mathcal{G}(\mathbf{z})))^2] \quad (1.25)$$

$$\min_{\mathcal{G}} \mathcal{L}_{\text{LSGAN}}(\mathcal{G}) = \frac{1}{2} \mathbb{E}_{\mathbf{z} \sim \mathbb{P}_{\mathbf{z}}(\mathbf{z})} [(\mathcal{D}(\mathcal{G}(\mathbf{z})))^2] \quad (1.26)$$

où les échantillons réels $\mathbf{x} \sim \mathbb{P}_{\text{data}}(\mathbf{x})$ et générés $\mathcal{G}(\mathbf{z})$ se voient attribuer des étiquettes binaires. En mesurant la distance des échantillons par rapport à la frontière de décision réel/génééré avec une norme ℓ_2 , la fonction de perte de LSGAN fournit au générateur des gradients plus informatifs que ceux de l'entropie croisée, pour les échantillons irréalistes mais correctement classifiés. Ce changement par rapport à (Goodfellow *et al.*, 2014) apporte plus de stabilité lors de l'entraînement tout en adressant le problème de disparition du gradient (Mao *et al.*, 2017).

1.3.9 Réseau antagoniste génératif cyclique (CycleGAN)

Les réseaux génératifs adversaires tels que présentés à la Section 1.3.8 ne permettent pas de conditionner les valeurs de sortie du générateur \mathcal{G} . Cette limitation pose problème lorsque la tâche à accomplir nécessite un contrôle sur les caractéristiques des échantillons générés comme pour la synthèse d’images médicales. Les réseaux génératifs adversaires conditionnels (CGAN) introduits dans (Mirza & Osindero, 2014) ont d’abord adressé cette limitation en ajoutant une variable de conditionnement en entrée au générateur et au discriminateur. Suivant la même idée, (Isola, Zhu, Zhou & Efros, 2017) a proposé l’architecture *pix2pix*, fondée sur CGAN, pour la tâche de translation image-à-image (*image-to-image translation*). Dans cette approche, le générateur prend en entrée une image $\mathbf{x} \sim \mathbb{P}_{\mathbf{X}}(\mathbf{x})$ ainsi qu’un vecteur issu d’une distribution a priori $\mathbf{z} \sim \mathbb{P}_{\mathbf{z}}(\mathbf{z})$ et génère une image $\hat{\mathbf{y}}$ partageant la même géométrie que \mathbf{x} , tout en ayant une apparence différente (e.g., la conversion d’images en noir et blanc en couleur). Le discriminateur dans (Isola *et al.*, 2017) est alors entraîné avec des paires d’images réelles (\mathbf{x}, \mathbf{y}) et des paires d’images $(\mathbf{x}, \hat{\mathbf{y}})$ où l’image de sortie est générée.

L’approche proposée par Isola *et al.* (2017) nécessite d’avoir des paires d’images (\mathbf{x}, \mathbf{y}) où \mathbf{x} est une image d’entrée provenant d’un domaine source \mathbf{X} et \mathbf{y} est sa représentation homologue réelle dans un domaine cible \mathbf{Y} . Or, pour de nombreuses tâches, l’accès à des images d’entraînement pairées est limité, voir même impossible. L’architecture CycleGAN, présenté dans (Zhu *et al.*, 2017), résout ce problème en permettant la translation image-à-image sans données pairées à l’aide de deux générateurs, G_{AB} et G_{BA} , et deux discriminateurs, D_A et D_B , tel qu’illustré à la Figure 1.17.

Les générateurs G_{AB} et G_{BA} ont pour tâche la translation d’images provenant du domaine source A vers leur représentation homologue dans le domaine cible B, et vice versa. La fonction objectif proposée dans (Zhu *et al.*, 2017) est composée de pertes adversaires, assurant la correspondance entre les distributions des données générées et des données réelles, et cycliques (eq(1.28)), assurant la cohérence entre les représentations d’une même image dans les domaines source et cible.

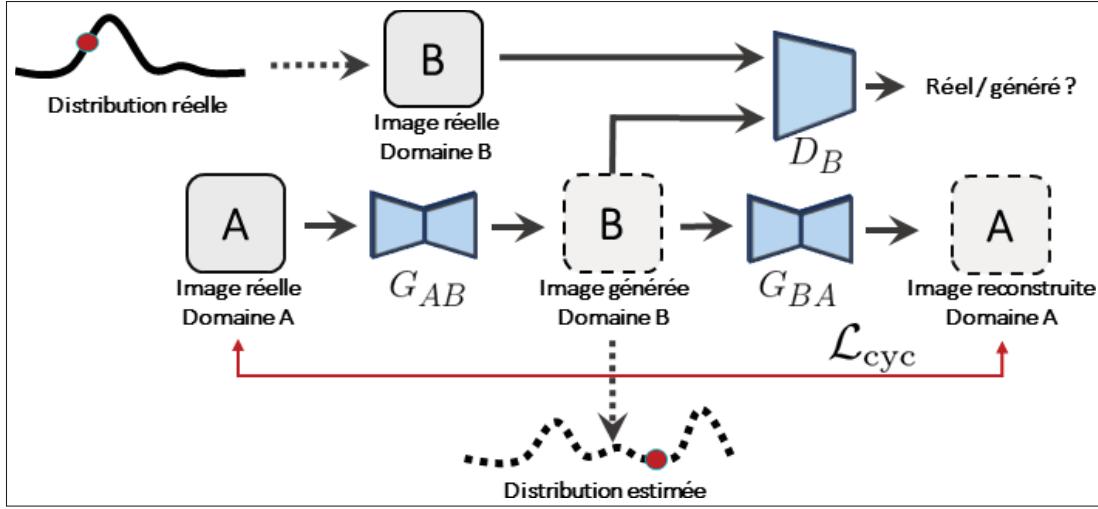


Figure 1.17 Cycle avant de l’architecture CycleGAN (Zhu *et al.*, 2017). Le générateur G_{AB} prend en entrée une image réelle du domaine source A et génère son homologue dans le domaine cible B. G_{BA} prend en entrée l’image générée par G_{AB} et tente de restaurer l’image initiale. D_B évalue la plausibilité de l’image générée par rapport à la distribution réelle des données du domaine B.

L’objectif complet d’un tel réseau est noté

$$\begin{aligned} \mathcal{L}(G_{AB}, G_{BA}, D_A, D_B) &= \mathcal{L}_{\text{GAN}}(G_{AB}, D_B, A, B) \\ &\quad + \mathcal{L}_{\text{GAN}}(G_{BA}, D_A, B, A) \\ &\quad + \lambda \mathcal{L}_{\text{cyc}}(G_{AB}, G_{BA}) \end{aligned} \quad (1.27)$$

où \mathcal{L}_{cyc} est la perte cyclique décrite par l’eq(1.28) ci-dessous :

$$\begin{aligned} \mathcal{L}_{\text{cyc}}(G_{AB}, G_{BA}) &= \\ \mathbb{E}_{\mathbf{a} \sim \mathbb{P}_A} [\| (G_{BA}(G_{AB}(\mathbf{a})) - \mathbf{a} \|_1] &+ \mathbb{E}_{\mathbf{b} \sim \mathbb{P}_B} [\| (G_{AB}(G_{BA}(\mathbf{b})) - \mathbf{b} \|_1]. \end{aligned} \quad (1.28)$$

Les sections suivantes présentent certaines applications des réseaux génératifs pour la génération d’images de diffusion à partir d’images structurelles, ainsi que des travaux connexes récents sur l’apprentissage de données à valeurs sur variété riemannienne.

1.4 Travaux connexes

1.4.1 Synthèse d'images structurelles-à-diffusion

Parmi la littérature, (Gu, Knutsson, Nilsson & Eklund, 2019) étudie la génération d'images d'anisotropie fractionnelle et de diffusion moyenne (MD) à partir d'images structurelles sous-échantillonnées. Pour ce faire, les auteurs utilisent un CycleGAN pour apprendre les relations intermodales entre les images T1w et les cartes FA/MD, et traduire avec succès les unes vers les autres. Ces travaux démontrent que les images structurelles partagent suffisamment d'informations avec l'anisotropie de diffusion des tissus pour synthétiser des tranches FA et MD 2D plausibles. De même, dans (Lan, Toga & Sepehrband, 2020), un GAN conditionnel avec auto-attention *self-attention* (SC-GAN) est employé pour générer des images FA et MD à partir de différentes modalités d'entrée, y compris des images T1w structurelles. Leurs résultats indiquent que les informations contextuelles 3D et l'optimisation antagoniste sont des éléments importants pour la synthèse des données de diffusion. Dans (Zhong, Wang, Li, Xue, Liu, Wang, Gao, Wang, Yang & Li, 2020), un duo de GAN *dual GAN* avec un discriminateur markovien (Li & Wand, 2016) sont utilisés pour l'harmonisation des métriques FA et MD inter-sites dérivées d'images DT. Enfin, dans (Son *et al.*, 2019), l'IRM fonctionnelle en combinaison avec des entrées T1w structurelles est envoyée à un réseau de type CNN pour générer des tenseurs de diffusion.

Cet ensemble de travaux démontre le potentiel des modèles génératifs pour la traduction de données d'imagerie structurelle vers des données de diffusion. Néanmoins, ils restent limités en n'exploitant pas pleinement les informations en haute-résolution contenues dans les images structurelles, en considérant soit une version sous-échantillonnée des entrées T1w, soit des tranches 2D avec un contexte limité (Gu *et al.*, 2019; Lan *et al.*, 2020; Son *et al.*, 2019). De plus, même si les cartes scalaires de diffusion FA et MD sont cliniquement utiles, elles ignorent la plupart du temps les orientations des fibres et, conséquemment, sont d'un intérêt limité pour des tâches telles que la tractographie ou la visualisation du connectome. En ce qui concerne les tenseurs générés dans (Son *et al.*, 2019), les auteurs ne fournissent aucune garantie sur leur

validité mathématique, telle que les propriétés SPD, ni sur leur utilité dans une tâche en aval comme la tractographie.

1.4.2 Apprentissage de données à valeurs sur variété riemannienne

Les modèles d'apprentissage profond sont bien adaptés pour modéliser des données situées dans un espace vectoriel euclidien. Cependant, les opérations euclidiennes à partir desquelles elles sont construites, par exemple la convolution ou l'échantillonnage (*pooling*), ne sont pas bien définies sur les variétés courbes. De plus, l'application de la géométrie euclidienne à des données à valeurs sur variété riemannienne, telles que les images DT et ODF, a des effets secondaires bien documentés (Arsigny *et al.*, 2006). Par conséquent, des études utilisant des réseaux de neurones pour le traitement précis des données sur les variétés riemanniennes ont commencé à émerger (Chakraborty, Bouza, Manton & Vemuri, 2019; Brooks, Schwander, Barbaresco, Schneider & Cord, 2019). Cependant, ces travaux nécessitent des modifications substantielles des modèles d'apprentissage profond connus et appellent à des recherches plus approfondies dans un ensemble plus large de scénarios.

Une autre voie pour le traitement de données à valeurs sur variété riemannienne réside dans la conception de métriques riemanniennes efficaces d'un point de vue de la complexité de calcul. Suivant cette voie, (Arsigny *et al.*, 2006) a proposé une métrique log-euclidienne pour traiter des données se trouvant sur la variété S_{++}^* avec des applications aux tenseurs de diffusion. Grâce aux applications log et exp définies dans (Arsigny *et al.*, 2006), on peut traiter des tenseurs à l'aide d'opérations euclidiennes et garantir que les tenseurs traités conservent leurs propriétés SPD. De même, un cadre log-euclidien a été proposé dans (Cheng *et al.*, 2009) pour le calcul de statistiques sur des fonctions de distribution d'orientation et appliqué à l'ODF de diffusion. Ces deux cadres, combinés à la rétropropagation matricielle des couches spectrales présentée dans (Ionescu, Vantzos & Sminchisescu, 2015), constituent les fondements des approches récentes d'apprentissage de données à valeurs sur variété riemannienne. Par exemple, dans (Huang & Van Gool, 2017), les auteurs ont intégré la métrique log-euclidienne dans leur modèle d'apprentissage profond appelé SPDNet pour apprendre des matrices SPD compactes et discriminantes. Bien

que SPDNet offre un moyen d'apprendre des données sur \mathcal{S}_{++}^* , ce modèle n'a pas été conçu pour l'apprentissage de matrices SPD spatialement organisées et volumétriques comme dans DT.

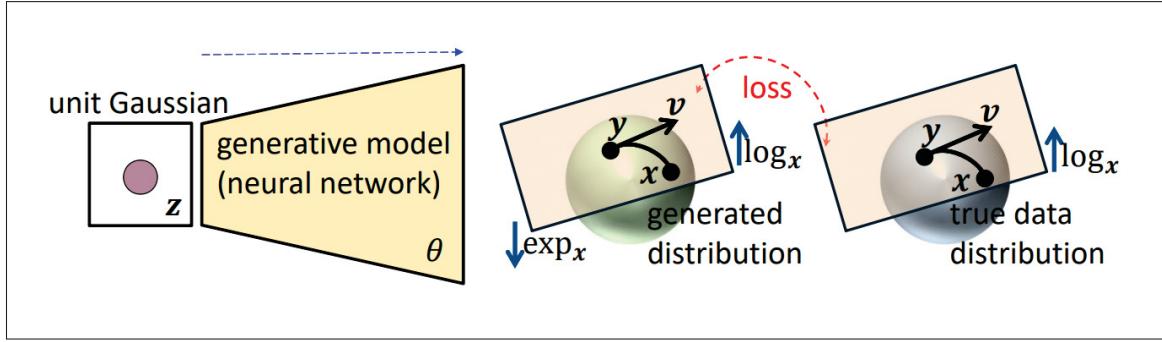


Figure 1.18 L'architecture *manifold-aware* Wasserstein GAN (MA-WGAN) proposée par Huang *et al.* (2019). Les applications log et exp sont utilisées pour générer des images avec des valeurs dans le plan tangent à la variété riemannienne sous-jacente aux données à partir d'un vecteur de bruit en entrée. Image tirée de Huang *et al.* (2019).

Plus récemment, (Huang *et al.*, 2019) a proposé un Wasserstein GAN (WGAN) (Arjovsky *et al.*, 2017) exploitant la métrique log-euclidienne pour synthétiser des DT plausibles, parmi d'autres types de données à valeurs sur variété. Pour assurer la validité des données générées, les auteurs projettent la sortie de leur réseau générateur dans un espace vectoriel en utilisant l'application log de (Arsigny *et al.*, 2006) avant l'évaluation du discriminateur. L'opération exp est ensuite utilisée sur la sortie synthétisée pour récupérer des DT valides. Malgré sa capacité à générer une diffusion mathématiquement valide, le modèle de (Huang *et al.*, 2019) produit des images qui ne sont conditionnées par aucune information spécifique au sujet réel (par exemple, une image T1w) et, par conséquent, ont moins de valeur clinique. De plus, ce travail ne porte que sur la génération de DT en 2D, ce qui limite encore une fois la valeur des données générées.

1.5 Conclusion

Dans ce chapitre, les concepts nécessaires à la compréhension de la méthode proposée dans ce mémoire ont été détaillés. D'abord, un survol du cerveau et ses principaux tissus a été fait. Par la suite, les techniques d'imagerie par résonance magnétique permettant d'imager le cerveau,

particulièrement l'IRM structurelle et l'IRM de diffusion, ont été présentées et certaines de leur limitations ont été discutées.

La section Section 1.2 a détaillé les deux modèles mathématiques pour l'estimation de la diffusion qui sont étudiés dans ce travail soient les tenseurs de diffusion et les fonctions de distribution d'orientations de diffusion. Un lien a été fait entre ces modèles et la théorie entourant les variétés riemanniennes et la géométrie différentielle.

Les techniques récentes de synthèse d'images par apprentissage automatique en lien avec les sections suivantes du mémoire ont été présentées dans la Section 1.3 et une présentation des travaux connexes au mémoire a été faite. Comme décrit dans la Section 1.4, les approches pour la génération de diffusion à partir d'images structurelles ont des limitations importantes. D'une part, ces méthodes n'exploitent pas l'information en haute-résolution des images structurelles et d'autre part, elles peinent à garantir la validité mathématique de la diffusion générée ou sont limitées à la génération de cartes scalaires de diffusion.

Ce mémoire propose une nouvelle architecture d'apprentissage profond qui exploite les informations détaillées des images structurelles en haute-résolution pour guider la synthèse d'images DT et ODF dans le même espace en haute-résolution. Fondée sur l'architecture CycleGAN (Zhu *et al.*, 2017), la solution proposée exploite les représentations intermodales inhérentes aux images structurelles et de diffusion pour apprendre de manière faiblement supervisée à les traduire les unes vers les autres. Cette recherche aborde les limites actuelles des modèles d'apprentissage profond construits sur des opérations euclidiennes en intégrant un cadre riemannien, le cadre log-euclidien, pour les calculs statistiques sur DT et ODF directement dans le modèle. Un tel cadre riemannien dans la procédure d'apprentissage du réseau impose une synthèse valide des données de diffusion qui reposent sur une variété riemannienne déterminée. En contraignant la diffusion générée à reposer sur la variété riemannienne des matrices symétriques positives définies 3×3 (\mathcal{S}_{++}^3) pour DT et sur la variété n-Sphere \mathbb{S}^n pour ODF, nous garantissons la cohérence mathématique de la solution. Contrairement aux modèles génératifs profonds existants qui se sont concentrés sur la génération de cartes scalaires dérivées de la diffusion (Li & Wand,

2016; Gu *et al.*, 2019; Lan *et al.*, 2020), le modèle proposé produit des schémas de diffusion complets, ici les matrices 3×3 des DT et les coefficients des harmoniques sphériques des ODF. Cette différence importante permet d'effectuer la tractographie, la visualisation du tractogramme et la segmentation des faisceaux de fibres, en plus du calcul des cartes scalaires directement à partir de la sortie du réseau tout en ne nécessitant qu'une image T1w en entrée.

CHAPITRE 2

MÉTHODOLOGIE

Le contenu des prochains chapitres représente une adaptation des articles de *workshop* intitulé *Manifold-Aware CycleGAN for High-Resolution Structural-to-Diffusion Synthesis* et de journal intitulé *Manifold-Aware Synthesis of High-Resolution Diffusion from Structural Imaging*. Le premier article a fait l'objet d'une présentation orale à l'*MICCAI 2020 International Workshop on Computational Diffusion MRI (CDMRI)* tandis que le deuxième est actuellement en processus de soumission.

2.1 Motivations

L'IRM de diffusion est d'une importance cruciale dans de nombreuses tâches, y compris le diagnostic de troubles cognitifs (Kelly, Jahanshad, Zalesky, Kochunov, Agartz, Alloza, Andreassen, Arango, Banaj, Bouix, Bousman, Brouwer, Bruggemann, Bustillo, Cahn, Calhoun, Cannon, Carr, Catts, Chen, Chen, Chiapponi, Cho, Ciullo, Corvin, Crespo-Facorro, Cropley, De Rossi, Diaz-Caneja, Dickie, Ehrlich, Fan, Faskowitz, Fatouros-Bergman, Flyckt, Ford, Fouche, Fukunaga, Gill, Glahn, Gollub, Goudzwaard, Guo, Gur, Gur, Gurholt, Hashimoto, Hatton, Henskens, Hibar, Hickie, Hong, Horacek, Howells, Hulshoff Pol, Hyde, Isaev, Jablensky, Jansen, Janssen, Jönsson, Jung, Kahn, Kikinis, Liu, Klauser, Knöchel, Kubicki, Lagopoulos, Langen, Lawrie, Lenroot, Lim, Lopez-Jaramillo, Lyall, Magnotta, Mandl, Mathalon, McCarley, McCarthy-Jones, McDonald, McEwen, McIntosh, Melicher, Mesholam-Gately, Michie, Mowry, Mueller, Newell, O'Donnell, Oertel-Knöchel, Oestreich, Paciga, Pantelis, Pasternak, Pearlson, Pellicano, Pereira, Pineda Zapata, Piras, Potkin, Preda, Rasser, Roalf, Roiz, Roos, Rotenberg, Satterthwaite, Savadjiev, Schall, Scott, Seal, Seidman, Shannon Weickert, Whelan, Shenton, Kwon, Spalletta, Spaniel, Sprooten, Stäblein, Stein, Sundram, Tan, Tan, Tang, Temmingh, Westlye, Tønnesen, Tordesillas-Gutierrez, Doan, Vaidya, Van Haren, Vargas, Vecchio, Velakoulis, Voineskos, Voyvodic, Wang, Wan, Wei, Weickert, Whalley, White, Whitford, Wojcik, Xiang, Xie, Yamamori, Yang, Yao, Zhang, Zhao, Van Erp, Turner, Thompson & Donohoe, 2018; Kantarci

et al., 2017; Neuner *et al.*, 2010), l'étude de maladies neurodégénératives (Huang *et al.*, 2007; Gattellaro *et al.*, 2009b) et la planification neurochirurgicale (Costabile *et al.*, 2019). Néanmoins, l'imagerie pondérée en diffusion (DWI) souffre d'un faible rapport signal sur bruit (SNR) et d'une mauvaise résolution spatiale résultant de limitations physiques et cliniques telles que l'utilisation de l'imagerie écho-planaire (EPI) et le temps d'acquisition limité du patient. En effet, le compromis induit entre la résolution de l'image, le SNR et le temps d'imagerie (Poot, Jeurissen, Bastiaensen, Veraart, Van Hecke, Parizel & Sijbers, 2013) dans l'acquisition des images DWI se traduit souvent par des images avec une taille de voxel jusqu'à 8 fois plus grande que celle d'autres modalités courantes telles que les images T1w structurelles, e.g., $2 \times 2 \times 2$ mm . pour DWI vs $1 \times 1 \times 1$ mm pour T1w (Shi *et al.*, 2016). Par conséquent, il a été montré que la présence accrue d'effet de volume partiel (PVE) dans les images DW, en raison de leur faible résolution spatiale, peut altérer leur analyse ultérieure (Oouchi *et al.*, 2007; Alexander *et al.*, 2001).

Ces limitations ont déclenché le développement de méthodes de post-traitement visant à améliorer la résolution spatiale des volumes de diffusion en basse résolution. L'estimation de la diffusion en haute résolution (HR) à partir d'images en basse résolution (LR) a d'abord été explorée avec des méthodes basées sur l'interpolation (Arsigny *et al.*, 2006; Yap, An, Chen & Shen, 2014; Dyrby, Lundell, Burke, Reislev, Paulson, Ptito & Siebner, 2014). Ces méthodes ré-échantillonnent les images existantes sur une grille à plus haute résolution et constituent aujourd'hui une étape par défaut dans les outils de traitement d'IRM de diffusion tels que TractoFlow (Theaud, Houde, Boré, Rheault, Morency & Descoteaux, 2020) et MRtrix3 (Tournier, Smith, Raffelt, Tabbara, Dhollander, Pietsch, Christiaens, Jeurissen, Yeh & Connelly, 2019). Bien que certains détails anatomiques fins puissent être améliorés avec une telle approche, l'interpolation sera toujours limitée par la grossièreté inhérente des données de diffusion d'origine, car elles reposent exclusivement sur des informations intra-image.

L'apprentissage automatique offre un moyen efficace d'exploiter la richesse des informations contenues dans les images HR pour la synthèse de l'imagerie de diffusion dans la même résolution, surpassant ainsi l'interpolation. Dans (Alexander, Zikic, Ghosh, Tanno, Wottschel,

Zhang, Kaden, Dyrby, Sotiroopoulos, Zhang & Criminisi, 2017), un cadre de transfert de qualité d'image (IQT) entièrement supervisé utilisant des forêts aléatoires (*random forests*) est proposé pour apprendre une correspondance non linéaire entre des données de diffusion appariées de faible qualité et de haute qualité. De même, dans (Elsaid & Wu, 2019), un SRCNN 2D supervisé est utilisé pour le même objectif. Les auteurs démontrent que l'apprentissage d'une application reliant une entrée LR vers sa version HR permet non seulement de récupérer des détails anatomiques mieux que l'interpolation, mais peut également aider dans des tâches en aval telles que la tractographie. Cependant, ces approches nécessitent des données de diffusion en haute résolution qui sont coûteuses et difficiles à acquérir. De plus, les méthodes de Alexander *et al.* (2017); Elsaid & Wu (2019) n'ont été testées que sur de petits ensembles de données comprenant respectivement un maximum de 23 sujets et 3 sujets.

En parallèle, les réseaux de neurones profonds offrent des techniques d'apprentissage non supervisées qui ne nécessitent que quelques échantillons d'entraînement appariés pour apprendre des tâches de synthèse spécifiques. Plus particulièrement, les Generative Adversarial Networks (GAN) (Goodfellow *et al.*, 2014) ont été utilisés avec succès pour la synthèse de modalités manquantes (Dar, Yurt, Karacan, Erdem, Erdem & Cukur, 2019), la traduction d'image à image (Zhu *et al.*, 2017; Lei, Harms, Wang, Liu, Shu, Jani, Curran, Mao, Liu & Yang, 2019) et la super-résolution d'images (Sánchez & Vilaplana, 2018), pour n'en nommer que quelques-unes. Par conséquent, les modèles génératifs profonds pourraient être une solution clé pour la synthèse de la diffusion en haute résolution à partir d'images non appariées, exprimant un niveau plus élevé de détails structurels tels que les images T1w.

Alors que la synthèse de signaux DWI bruts avec une résolution angulaire appropriée est une tâche extrêmement gourmande en ressources, les architectures d'apprentissage profond actuelles ont du mal à générer des schémas de reconstruction de diffusion plausibles, notamment le tenseur de diffusion (DT) ou les fonctions de distribution d'orientation (ODF) en raison de leur nature non-euclidienne (Huang *et al.*, 2019). En effet, chaque voxel d'une image DT repose sur une variété riemannienne de matrices symétriques définies positives (SPD) 3×3 (Arsigny *et al.*, 2006), et les ODF peuvent être représentés comme des points sur une variété n-Sphère \mathbb{S}^n (Cheng *et al.*,

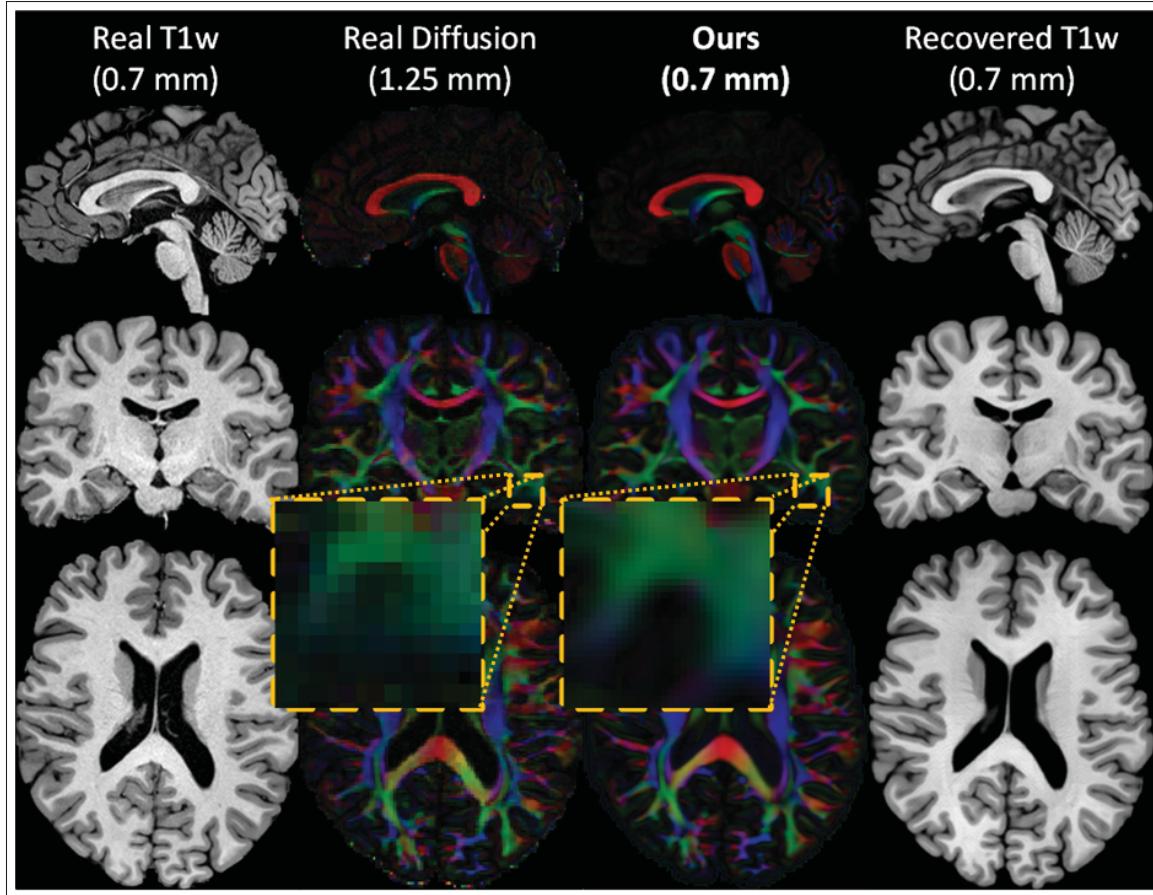


Figure 2.1 Exemple d’entrées et de sorties de l’architecture proposée. Une image T1w en HR (**première colonne**), FA colorée par la direction de la vraie diffusion en LR (**deuxième colonne**), FA colorée par la direction de la diffusion générée en HR (**troisième colonne**) et l’image T1w recouverte en HR (**dernière colonne**) d’un sujet de test.

2009). Dans le contexte de la synthèse d’images, l’incapacité des réseaux à capturer la géométrie de la variété riemannienne non linéaire sous-jacente aux données entraîne la génération d’images invraisemblables qui manquent les propriétés mathématiques importantes de l’imagerie de diffusion (Huang *et al.*, 2019). Par conséquent, les limites des réseaux de neurones profonds (DNN) actuels ont entravé le développement de modèles génératifs en imagerie de diffusion, qui se sont principalement limités à la synthèse de cartes scalaires DT telles que l’anisotropie fractionnelle (FA) et la diffusion moyenne (MD).

2.2 Méthode proposée

Soit X le domaine à valeurs réelles des images structurelles (par exemple, T1w) et Y le domaine à valeurs sur la variété des images de diffusion (par exemple, DT ou ODF). Nous visons à apprendre les fonctions d'application (*map*) $G_Y : X_{\text{HR}} \mapsto Y_{\text{HR}}$ et $G_X : Y_{\text{HR}} \mapsto X_{\text{HR}}$ qui traduisent des images T1w haute résolution en images de diffusion haute résolution et inversement. Néanmoins, nous avons principalement accès à des images HR T1w non appariées et à des données de diffusion LR qui diffèrent en termes de sujet, de résolution et de nature. Pour résoudre ce problème, nous proposons une architecture Manifold-Aware CycleGAN (MA-CycleGAN) (voir Figure 2.2) qui gère intrinsèquement à la fois la traduction de domaine et la super-résolution de diffusion, tout en tenant compte de la géométrie riemannienne des données en intégrant les métriques log-euclidienne présentées à la Section 1.2.

Nous entraînons notre réseau avec des échantillons d'apprentissage non appariés $\{\mathbf{x}_i\}_{i=1}^N$ où $\mathbf{x}_i \in X_{\text{HR}}$ est une image structurelle 3D, et $\{\mathbf{y}_j\}_{j=1}^M$ où $\mathbf{y}_j \in Y_{\text{LR}}$ est une image de diffusion dans une résolution spatiale inférieure. La validité mathématique des DT et ODF générés est assurée en projetant chaque voxel synthétisé sur le plan tangent à leur contrepartie isotrope (c'est-à-dire la matrice d'identité 3×3 pour DT et la distribution uniforme pour ODF). Pour ce faire, les applications exponentielles et logarithmes empruntées à leur variété riemannienne sous-jacente sont utilisées.

Deux discriminateurs D_X et D_Y évaluent respectivement les images HR T1w synthétisées $G_X(\mathbf{y})$ et les images de diffusion LR produites par le réseau. Ces dernières sont obtenues avec une fonction

$$\mathcal{F}(G_Y(\mathbf{x})) = \downarrow G_Y(\mathbf{x}) + \mathcal{F}(\downarrow G_X(\mathbf{x})), \quad (2.1)$$

où \mathcal{F} est une fonction résiduelle apprise et \downarrow est l'opération de sous-échantillonnage trilinéaire pour réduire la résolution. En combinant les pertes de reconstruction cyclique voxel-à-voxel et antagoniste dans un seul objectif, le MA-CycleGAN proposé est capable d'exploiter les informations tissulaires locales et la géométrie globale des images structurelles en haute-résolution pour produire des données de diffusion nettes, réalistes et utilisables.

Dans les sections suivantes, nous détaillons nos objectifs cycliques et antagonistes incorporant à la fois les applications exp et log des cadres riemanniens susmentionnés et une stratégie de sous-échantillonnage apprise. Ensuite, nous présentons notre mécanisme d’attention basé sur l’anisotropie fractionnelle qui aide le réseau à reconstruire les régions exprimant des informations de faisceaux de fibres avec une plus grande fidélité.

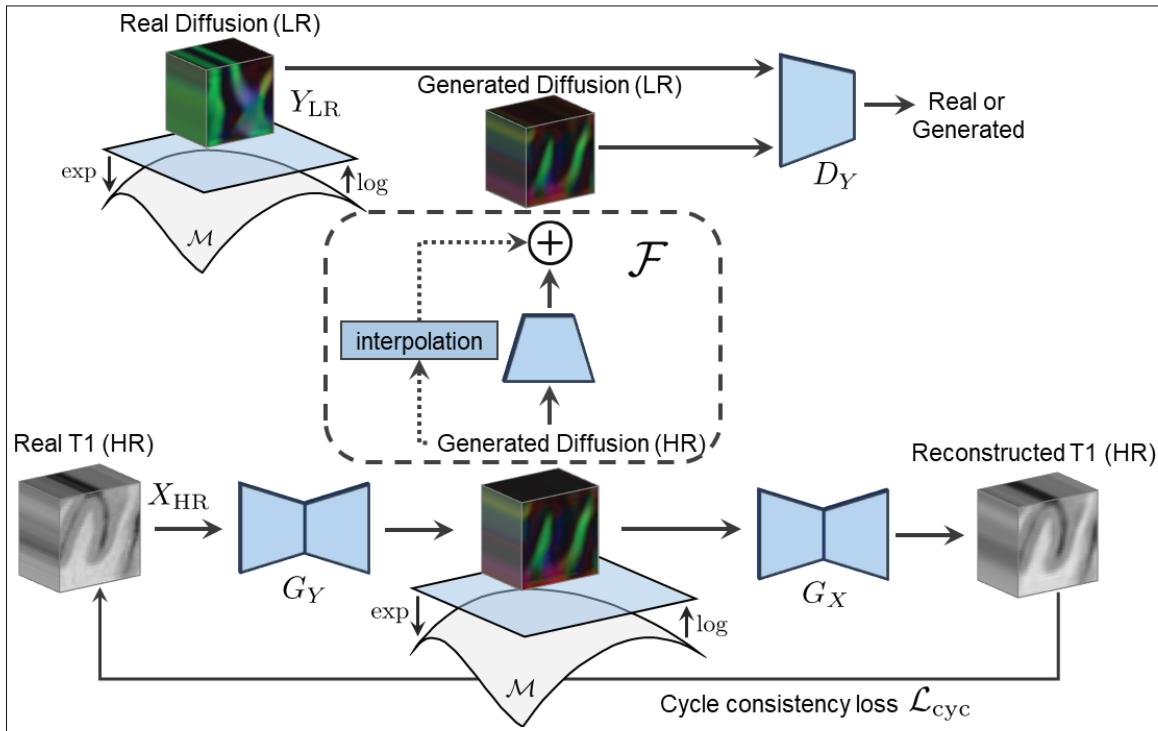


Figure 2.2 Le cycle avant de l’architecture proposée. G_Y génère un volume DT/ODF en haute-résolution sur la variété riemannienne appropriée (e.g. S^3_{++} pour DT et \mathbb{S}^n pour ODF) en utilisant les applications log et exp. D_Y évalue la qualité des images produites et fournit un retour à G_Y . G_X reconstruit l’image T1w d’entrée originale à partir de la diffusion HR générée notée $\log(\exp(G_Y(\mathbf{x})))$.

2.3 Apprentissage antagoniste

Dans une configuration GAN standard (Goodfellow *et al.*, 2014), un réseau générateur G essaie de générer des échantillons proches de la vraie distribution de données afin qu’un réseau discriminateur D soit incapable de faire la distinction entre les exemples réels et les exemples générés. Suivant l’architecture CycleGAN, notre méthode utilise deux générateurs et deux

discriminateurs notés G_X, G_Y, D_X et D_Y . Ici, G_X prend en entrée un lot de volumes de diffusion sur-échantillonnés et essaie de tromper D_X en générant des volumes T1w haute résolution réalistes. De même, G_Y prend les volumes HR T1w en entrée et essaie de générer des volumes de diffusion plausibles dans le même espace HR.

Comme nous n'avons accès qu'aux données de diffusion LR, la diffusion HR synthétisée est sous-échantillonnée à l'aide d'une fonction résiduelle apprise \mathcal{F} avant l'évaluation de D_Y comme indiqué dans Figure 2.2.

Nos objectifs antagonistes suivent la formulation LSGAN telle que décrite par Mao *et al.* (2017) et sont exprimés comme suit :

$$\begin{aligned} \min_{D_Y} \mathcal{L}_{\text{LSGAN}}(G_Y, D_Y, X_{\text{HR}}, Y_{\text{LR}}) = \\ \frac{1}{2} \mathbb{E}_{\mathbf{y} \sim \mathbb{P}_{Y_{\text{LR}}}} [(D_Y(\log(\mathbf{y})) - 1)^2] + \frac{1}{2} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{X_{\text{HR}}}} [(D_Y(\mathcal{F}(G_Y(\mathbf{x}))))^2] \end{aligned} \quad (2.2)$$

$$\begin{aligned} \min_{D_X} \mathcal{L}_{\text{LSGAN}}(G_X, D_X, Y_{\text{LR}}, X_{\text{HR}}) = \\ \frac{1}{2} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{X_{\text{HR}}}} [(D_X(\mathbf{x}) - 1)^2] + \frac{1}{2} \mathbb{E}_{\mathbf{y} \sim \mathbb{P}_{Y_{\text{HR}}}} [(D_X(G_X(\uparrow \log(\mathbf{y}))))^2] \end{aligned} \quad (2.3)$$

où ‘ \uparrow ’ représente le sur-échantillonnage trilinéaire et \log est l'application logarithme définie dans l'eq(1.10) ou l'eq(1.19) selon la méthode de reconstruction de la diffusion utilisée, DT ou ODF. Il est important de noter que le sur-échantillonnage des données de diffusion réelles et l'évaluation du discriminateur des volumes de diffusion générés sont effectués dans le domaine log-euclidien pour tenir compte de la géométrie non linéaire de la variété sous-jacente aux données de diffusion.

2.4 Fonction de perte cyclique

Les pertes antagonistes à elles seules ne sont pas suffisantes pour générer des données de diffusion HR. En effet, D_Y n'évalue que les données sous-échantillonées et, par conséquent, ne peut pas aider G_Y à s'améliorer au-delà d'un certain niveau de précision. Par conséquent, la perte de cohérence cyclique notée dans l'eq(2.4) permet non seulement d'assurer la cohérence

structurelle des images synthétisées à travers les modalités, mais fournit également des gradients en haute résolution importants pour entraîner G_Y .

Notre perte de cohérence cyclique est composée de trois termes : 1) l'erreur entre le volume structurel HR original \mathbf{x} et le volume reconstruit $G_X(G_Y(\mathbf{x}))$, 2) l'erreur entre la diffusion sur-échantillonnée $\uparrow \log(\mathbf{y})$ et sa reconstruction HR $G_Y(G_X(\uparrow \log(\mathbf{y})))$, et 3) l'erreur entre la diffusion LR originale \mathbf{y} et le volume récupéré sous-échantillonné $\mathcal{F}(G_Y(G_X(\uparrow \log(\mathbf{y}))))$.

Les combiner en une seule fonction donne

$$\begin{aligned} \mathcal{L}_{\text{cyc}}(G_Y, G_X) = & \lambda_{\text{cyc}_X} \underbrace{\mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{X_{\text{HR}}}} [\|G_X(G_Y(\mathbf{x})) - \mathbf{x}\|_1]}_{\text{Cycle avant HR}} \\ & + \frac{1}{2} \lambda_{\text{cyc}_Y} \underbrace{\mathbb{E}_{\mathbf{y} \sim \mathbb{P}_{Y_{\text{LR}}}} [\|G_Y(G_X(\uparrow \log(\mathbf{y}))) - \uparrow \log(\mathbf{y})\|_1]}_{\text{Cycle arrière HR}} \\ & + \frac{1}{2} \lambda_{\text{cyc}_Y} \underbrace{\mathbb{E}_{\mathbf{y} \sim \mathbb{P}_{Y_{\text{LR}}}} [\|\mathcal{F}(G_Y(G_X(\uparrow \log(\mathbf{y})))) - \log(\mathbf{y})\|_1]}_{\text{Cycle arrière LR}} \end{aligned} \quad (2.4)$$

La norme ℓ_1 est utilisée dans l'éq(2.4) pour mesurer les erreurs de reconstruction de cycle avant et arrière, car elle est moins sensible aux valeurs extrêmes que la norme ℓ_2 et permet typiquement de synthétiser des images plus réalistes et avec moins d'artefacts (Zhao, Gallo, Frosio & Kautz, 2016). Encore une fois, l'application logarithme de la métrique log-euclidienne est utilisée pour projeter les données générées et les données réelles à valeurs sur variété sur un plan tangent avant de calculer la perte de cohérence de cycle. De plus, deux paramètres λ_{cyc_X} et λ_{cyc_Y} contrôlent la contribution des deux cycles et ont été ajustés empiriquement. Les cycles complets, avant et arrière, y compris les applications riemannienne et le calcul des pertes dans le plan tangent aux variétés, sont illustrés dans la Figure 2.3.

2.5 Régularisation par information préalable

En utilisant une perte de cohérence de cycle dans les deux sens, le modèle CycleGAN est capable d'apprendre une application bijective entre deux domaines à l'aide d'exemples non appariés

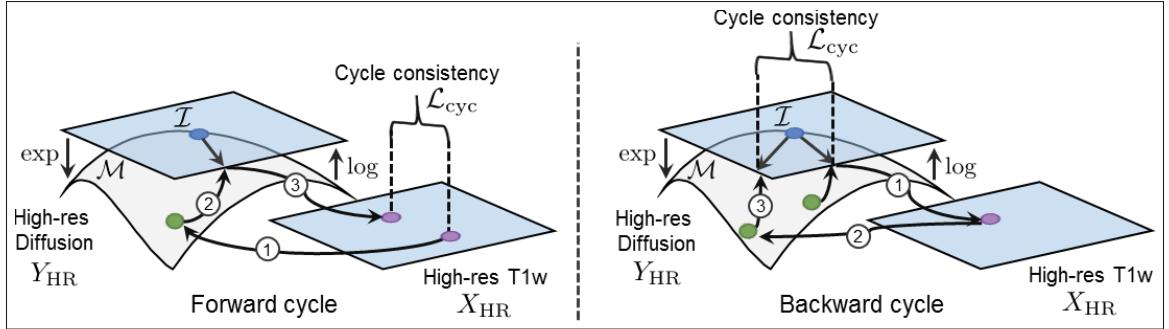


Figure 2.3 **À gauche**, le cycle avant proposée : Une image T1w est traduite en image de diffusion DT/ODF sur la variété S^3_{++}/\mathbb{S}^n en haute-résolution avant d'être restaurée dans le domaine des images T1w où la perte cyclique est calculée. **À droite**, le cycle arrière proposée : Une image de diffusion DT/ODF sur-échantillonnée dans le plan tangent $\mathcal{T}_{\text{Id}}/\mathcal{T}_{\text{u}}$ est traduite dans le domaine des images T1w avant d'être restaurée en image de diffusion sur $\mathcal{T}_{\text{Id}}/\mathcal{T}_{\text{u}}$ où la perte cyclique est calculée.

(Zhu *et al.*, 2017). Cependant, pour de nombreux problèmes de traduction inter-domaines, l'espace de solution est grand et le modèle ne converge pas nécessairement vers une solution qui satisfait d'importantes propriétés spécifiques au domaine (Lu, Zhou, Song, Ren & Yu, 2019). Ceci est problématique, notamment dans le cas de la synthèse d'images médicales où les images générées doivent non seulement être réalistes du point de vue du discriminateur, mais aussi être fidèles aux résultats attendus des tâches en aval et aux propriétés anatomiques connues. Ainsi, pour assurer la convergence du modèle vers des solutions plausibles, nous introduisons une perte *a priori* comme suit :

$$\begin{aligned} \mathcal{L}_{\text{prior}}(G_Y, G_X) = & \lambda_{\text{prior}_X} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{X_{\text{HR}}}, \mathbf{y} \sim \mathbb{P}_{Y_{\text{LR}}}} [\|G_Y(\mathbf{x}_i) - \uparrow \log(\mathbf{y}_i)\|_1] \\ & + \lambda_{\text{prior}_Y} \mathbb{E}_{\mathbf{x} \sim \mathbb{P}_{X_{\text{HR}}}, \mathbf{y} \sim \mathbb{P}_{Y_{\text{LR}}}} [\|G_X(\uparrow \log(\mathbf{y}_i)) - \mathbf{x}_i\|_1] \end{aligned} \quad (2.5)$$

où \mathbf{x}_i et \mathbf{y}_i sont des volumes appariés provenant d'un nombre limité de sujets. Avec cette perte, la diffusion HR synthétisée reste proche de la diffusion sur-échantillonnée réelle tout en intégrant des haute-fréquences contenues dans les images structurelles HR d'entrée.

2.6 Pondération de la fonction de perte

Les voxels exprimant des informations sur les faisceaux de fibres ont généralement des valeurs de FA plus élevées que ceux représentant d'autres tissus comme la matière grise (GM) ou le liquide céphalo-rachidien (CSF). Donc, la méthode de synthèse de diffusion proposée se doit d'être particulièrement précise dans les régions avec une anisotropie fractionnelle plus élevée. Cependant, comme on le voit dans la Figure 2.4, les voxels avec une FA élevée sont sous-représentés par rapport à ceux avec des valeurs plus faibles. Par conséquent, ce problème de déséquilibre conduit la génération du réseau vers une diffusion ayant une FA plus proche de la moyenne. Pour pallier ce problème, nous pondérons l'erreur de diffusion dans \mathcal{L}_{cyc} et $\mathcal{L}_{\text{prior}}$ par le FA/GFA du volume cible à chaque voxel.

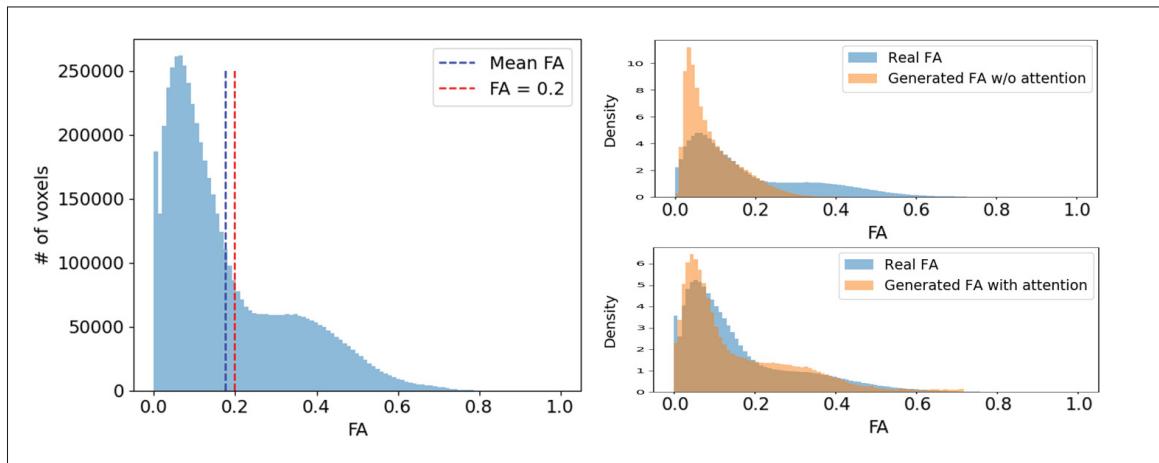


Figure 2.4 **À gauche**, l'histogramme de la FA typique pour un sujet de test. L'histogramme est débalancé vers des valeurs faible de FA avec une FA moyenne inférieure à 0.2. **À droite**, la densité des valeurs de FA de la diffusion générée avec et sans pondération de la fonction de perte par la FA.

L'avantage d'un tel mécanisme peut être observé sur les graphiques de densité du côté droit de la Figure 2.4 qui présentent une distribution FA plus fidèle lors de l'utilisation de la pondération des fonctions de perte par l'anisotropie fractionnelle proposée.

2.7 Fonction objectif complète

En combinant tous les termes de perte, notre fonction objectif complète est donnée par

$$\begin{aligned}\mathcal{L} = & -\mathcal{L}_{\text{LSGAN}}(G_X, D_X, Y_{\text{LR}}, X_{\text{HR}}) - \mathcal{L}_{\text{LSGAN}}(G_Y, D_Y, X_{\text{HR}}, Y_{\text{LR}}) \\ & + \mathcal{L}_{\text{cyc}}(G_Y, G_X) + \mathcal{L}_{\text{prior}}(G_Y, G_X)\end{aligned}\quad (2.6)$$

Comme dans les approches classiques d'apprentissage antagoniste, nous entraînons les générateurs et les discriminateurs simultanément en résolvant un problème mini-max :

$$G_X^*, G_Y^* = \arg \min_{G_X, G_Y} \max_{D_X, D_Y} \mathcal{L}(G_X, G_Y, D_X, D_Y) \quad (2.7)$$

En résumé, l'eq(2.6) combine les erreurs des images structurelles HR et de la diffusion LR de manière antagoniste et *voxel-par-voxel*.

CHAPITRE 3

EXPÉRIMENTATIONS ET RÉSULTATS

3.1 Données et pré-traitement

Les données d'IRM T1w et de diffusion de 1 065 sujets de la version HCP1200 du *Human Connectome Project* (HCP) (Van Essen, Smith, Barch, Behrens, Yacoub & Ugurbil, 2013) sont utilisées pour évaluer le modèle de synthèse de diffusion proposé. Les images T1w (0.7mm^3 voxels, $\text{FOV}=224\text{mm}$, matrice=320, 256 coupes sagittales) et la diffusion (1.25mm^3 voxels, sequence=Spin-echo EPI, temps de répétition (TR)=5520 ms, temps d'écho (TE)=89.5ms) ont été acquises avec un scanner Siemens Skyra 3T (Sotiroopoulos, Jbabdi, Xu, Andersson, Moeller, Auerbach, Glasser, Hernandez, Sapiro, Jenkinson, Feinberg, Yacoub, Lenglet, Van Essen, Ugurbil & Behrens, 2013) et traitées de manière minimale suivant (Glasser, Sotiroopoulos, Wilson, Coalson, Fischl, Andersson, Xu, Jbabdi, Webster, Polimeni, Van Essen & Jenkinson, 2013).

Les tenseurs de diffusion ont été calculés à l'aide du logiciel DSI Studio toolbox (Jiang, Van Zijl, Kim, Pearlson & Mori, 2006) et les dODF estimés à l'aide de la méthode *Constant Solid Angle* (CSA) (Aganj, Lenglet, Sapiro, Yacoub, Ugurbil & Harel, 2010) de la librairie DIPY (Garyfallidis, Brett, Amirbekian, Rokem, Van Der Walt, Descoteaux & Nimmo-Smith, 2014). Les ODF de diffusion ont ensuite été reparamétrisés suivant l'approche décrite dans la Section 1.2.7 et estimés en utilisant une base modifiée formée de fonctions harmoniques sphériques de 4ème ordre (Descoteaux *et al.*, 2007). Les volumes DT et ODF ont tous deux été transformés dans le domaine log-euclidien en utilisant leur cadre riemannien respectif décrit dans la Section ???. Une fois dans le domaine log-euclidien, ces volumes ont été sur-échantillonnés à la même résolution spatiale que les images T1w, grâce à une interpolation trilinéaire, et alignés à leur image HR T1w correspondante. Dans les expériences, ces volumes de diffusion sur-échantillonnés et alignés sont considérés comme la “vérité terrain” (*ground truth*). Les images T1w ont été mises à l'échelle dans l'intervalle [0,1] par normalisation min-max. Enfin, les volumes structurels et

de diffusion, en haute et basse résolution, ont été décomposés en *patches* superposés de 32^3 et 18^3 voxels respectivement. Les volumes sont traités par *patch* par le modèle pour deux raisons importantes : 1) limiter la mémoire requise par le modèle pour calculer les activations et les sorties du réseau, et 2) augmenter le nombre d'exemples d'apprentissage.

3.2 Détails d'implémentation

Les deux réseaux générateurs, G_X et G_Y , suivent le modèle U-Net proposé par Çiçek, Abdulkadir, Lienkamp, Brox & Ronneberger (2016). Les dernières couches d'activation de G_X et G_Y ont été remplacées pour s'adapter aux différentes plages de valeurs des données de sortie. Dans toutes les expériences, une activation de type sigmoïde est utilisée dans G_X pour générer des images T1w dans l'intervalle [0,1]. Pour la génération de la diffusion, la dernière fonction d'activation de G_Y varie en fonction du modèle de diffusion généré. Pour la génération de DT, une fonction tangente hyperbolique dure (*hard tanh*) est utilisée et, pour la génération d'ODF, une activation tangente hyperbolique (*tanh*) est utilisée.

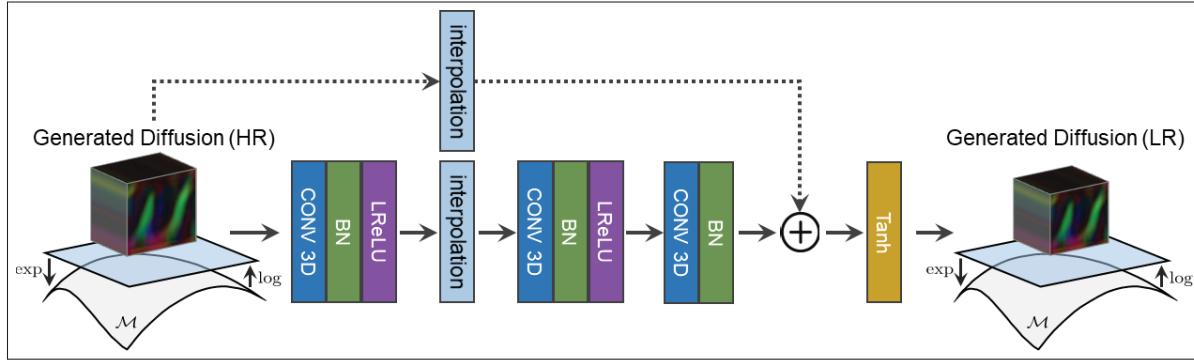


Figure 3.1 L'architecture de la fonction résiduelle de sous-échantillonnage \mathcal{F} qui est utilisée pour réduire la résolution des images de diffusion synthétisées HR avant l'évaluation de D_Y .

De plus, la quantité de canaux d'entrée et de sortie des réseaux est définie en fonction de la dimension des données d'entrée et de sortie, comme décrit dans la Table 3.1. Les entrées DT sont de forme $9 \times 32 \times 32 \times 32$, où les 9 canaux représentent les matrices 3×3 des tenseurs de diffusion sous forme vectorielle à chaque voxel. Les entrées ODF sont de forme $C \times 32 \times 32 \times 32$

où C dépend de l'ordre de la base de fonctions harmoniques sphériques qui est utilisée pour les représenter. En raison de limitations des ressources de calcul, les ODF de diffusion sont estimés à l'aide des fonctions harmoniques sphériques de 4ème ordre, donnant un total de coefficients $C = 15$.

Les réseaux discriminateurs D_X et D_Y suivent l'architecture du discriminateur *Super-Resolution GAN* (SRGAN) proposée par Ledig, Theis, Huszár, Caballero, Cunningham, Acosta, Aitken, Tejani, Totz, Wang & Shi (2017) où les convolutions 2D sont remplacées par des convolutions 3D. De plus, le nombre de cartes de caractéristiques (*feature maps*) a été réduit dans les couches de convolution du modèle à 32, 64, 128 et 256, comme suggéré dans (Sánchez & Vilaplana, 2018) pour les données volumétriques. Dans tous les scénarios, D_X évalue les volumes T1w qui sont de dimension $1 \times 32 \times 32 \times 32$ tandis que D_Y prend en entrée des volumes de dimension $9 \times 18 \times 18 \times 18$ pour DT et $15 \times 18 \times 18 \times 18$ pour les ODF.

3.3 Paramètres d'entraînement

Pour entraîner et évaluer le modèle proposé, 70% (746 sujets) des 1065 sujets ont été sélectionnés pour l'apprentissage du modèle, 20% (213 sujets) pour la validation et 10% (106 sujets) pour les tests. Sur les 746 sujets d'apprentissage, entre 0 et 75 sujets ont été utilisés comme information *a priori* appariées dans l'éq(2.5). Pour former les ensembles d'entraînement, de validation et de test, 50 000 *patchs* de diffusion LR et T1w HR non appariés ont été sélectionnés de façon aléatoire parmi les sujets d'entraînement restants, 10 000 *patchs* parmi les sujets de validation et 5 000 *patchs* parmi les sujets de test. Le même nombre de *patchs* (i.e., 50 000) a été sélectionné parmi les sujets conservés comme information *a priori*, c'est-à-dire des paires d'images T1w HR et de diffusion sur-échantillonnée et alignées, pour former un ensemble d'entraînement apparié. Bien que le nombre de *patchs* ait été maintenu constant dans l'ensemble d'entraînement apparié, le nombre de sujets choisis aléatoirement pour extraire ces *patchs* a été ajusté expérimentalement. Pour ce faire, le réseau a été entraîné avec un nombre croissant de sujets appariés, comme indiqué dans la Table 3.2, et la meilleure configuration a été conservée pour les expériences subséquentes.

Tableau 3.1 Détails d’implémentation de nos réseaux générateurs G_X et G_Y et réseaux discriminateurs D_X et D_Y .

Réseau	Forme d’entrée	Forme de sortie	Dernière activation
G_Y (T1w \mapsto DT)	$1 \times 32 \times 32 \times 32$	$9 \times 32 \times 32 \times 32$	Hardtanh
G_X (DT \mapsto T1w)	$9 \times 32 \times 32 \times 32$	$1 \times 32 \times 32 \times 32$	Sigmoïde
G_Y (T1w \mapsto ODF)	$1 \times 32 \times 32 \times 32$	$15 \times 32 \times 32 \times 32$	Tanh
G_X (ODF \mapsto T1w)	$15 \times 32 \times 32 \times 32$	$1 \times 32 \times 32 \times 32$	Sigmoïde
D_X (T1w)	$1 \times 32 \times 32 \times 32$	1	Linéaire
D_Y (DT)	$9 \times 18 \times 18 \times 18$	1	Linéaire
D_Y (ODF)	$15 \times 18 \times 18 \times 18$	1	Linéaire

Les réseaux ont été entraînés à l’aide d’un optimiseur Adam (Kingma & Ba, 2015) avec un taux d’apprentissage de 10^{-4} et des valeurs beta1, beta2 de 0,5 et 0,999. Les hyper-paramètres λ_{prior_X} , λ_{prior_Y} , λ_{cyc_X} et λ_{cyc_Y} ont été fixés expérimentalement à 10, 0,5, 5 et 0,25. De plus, une stratégie a été utilisée pour ajuster dynamiquement le taux d’apprentissage, réduisant celui-ci par un facteur de 10 à chaque plateau de 10 époques. Des lots (*batches*) de 8 *patches* ont été utilisés et les modèles ont été entraînés pour 35 époques ($\sim 220k$ pas) sur une carte graphique NVIDIA TITAN XP avec 12 Go de VRAM. Toutes les expériences ont été répétées trois fois avec une graine aléatoire (*random seed*) différente pour l’initialisation.

3.4 Approches de référence

Comme mentionné précédemment, des modèles d’apprentissage profond pour la synthèse de données à valeurs sur variété commencent tout juste à émerger. Donc, le nombre d’approches de référence, nécessitant une adaptation minimale pour l’évaluation du modèle proposé, est limité. Néanmoins, le modèle est comparé à trois approches décrites ci-dessous.

3.4.1 Manifold-Aware WGAN

La première approche de référence est une adaptation du réseau *Manifold-Aware WGAN* présenté par Huang *et al.* (2019) pour la génération conditionnelle de diffusion à partir

d’images T1w structurelles. Cette méthode, désignée comme “MA-WGAN” dans les résultats présentés ultérieurement, peut générer des images plausibles à valeurs variétés riemannniennes en incorporant les applications log-euclidiennes dans le réseau et fournit donc un point de comparaison naturel pour la méthode présentée dans ce mémoire. Pour cette approche de référence, le même générateur G_Y et le même discriminateur D_Y que dans le modèle proposé sont utilisés. Les applications log et exp utilisées dans le réseau sont modifiées en fonction du modèle de diffusion généré suivant la Section 1.2.

3.4.2 Manifold-Aware U-Net

La proposition faite dans ce mémoire est également comparée à un modèle U-Net supervisé. Pour cette référence, le même générateur que dans l’architecture proposée est utilisé, c’est-à-dire G_Y (Çiçek *et al.*, 2016), mais contrairement à la contribution de ce mémoire, l’entraînement est fait de manière supervisée à l’aide des images T1w HR appariés avec des volumes de diffusion sur-échantillonnés dans le domaine log-euclidien. Semblablement à l’approche de référence MA-WGAN, les applications exp et log du réseau sont changées en fonction du modèle de reconstruction de diffusion généré (DT ou ODF). Cette référence, désignée par “MA-U-Net” dans les résultats, permet de mesurer l’effet des fonctions de pertes antagoniste et de cohérence cyclique.

3.4.3 U-Net

La dernière approche de référence est un U-Net supervisé standard (Çiçek *et al.*, 2016) entraîné à l’aide d’images T1w HR appariées avec des images de diffusion LR sur-échantillonnée sans considérer la variété riemannienne. Cette méthode de référence, permet de mesurer le gain de performance de la méthode proposée, induit à la fois par la modélisation des variétés riemannniennes et l’utilisation d’échantillons supplémentaires non appariés. De plus, cette approche permet de valider que l’apprentissage sur variété riemannienne est nécessaire pour synthétiser des échantillons de diffusion mathématiquement réalistes et valides.

3.5 Synthèse de DT et ODF

D'abord le réseau et les approches de référence sont testés pour la tâche de synthèse de tenseurs de diffusion et d'ODF de diffusion. Dans cette configuration, le réseau présenté est entraîné avec des *patches* T1w HR et des *patches* de diffusion LR non appariés dans le domaine log-euclidien. De plus, un ensemble apparié de 50 000 *patches* provenant de 50 sujets préalablement choisi et extrait aléatoirement de l'ensemble d'entraînement est construit. L'ensemble d'entraînement apparié est utilisé comme information *a priori* pour la méthode suggérée et comme ensemble d'entraînement pour les approches de référence. De ce fait, la méthode suggérée et les approches de référence sont entraînées avec la même quantité de données appariées.

3.6 Évaluation

Trois métriques ont été considérées pour évaluer quantitativement la diffusion générée. Tout d'abord, la similarité cosinus est utilisée pour comparer l'orientation principale des fibres de chaque tenseur et ODF synthétisés à leur orientation réelle attendue :

$$\text{similarity}(\mathbf{a}, \mathbf{b}) = \left| \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right| \quad (3.1)$$

Pour récupérer l'orientation principale des ODF, les coordonnées sphériques auxquelles la valeur de l'ODF est maximale sont calculées en utilisant une sphère discrétisée de 724 sommets. Ces coordonnées sphériques sont ensuite converties en coordonnées euclidiennes pour obtenir leur vecteur d'orientation principal. Pour les DT, l'orientation principale est donnée par le vecteur propre associé à la plus grande valeur propre du DT comme décrit dans la Section 1.2.

De plus, la diffusion générée est comparée avec la diffusion réelle à l'aide de l'erreur quadratique moyenne (MSE) entre leur FA/GFA. La FA et le GFA sont calculées suivant l'éq(??) et l'éq(1.16). Cela permet d'évaluer la forme des DT/ODF synthétisés indépendamment de leur orientation.

Enfin, la distance géodésique moyenne entre les données générées et la diffusion réelle est mesurée en utilisant l'éq(1.12) pour les tenseurs de diffusion et l'éq(1.21) pour les ODF.

Cette dernière métrique mesure à la fois l'erreur d'orientation et l'erreur de l'anisotropie dans une seule valeur.

3.7 Tractographie

Pour évaluer l'intégrité des volumes de diffusion synthétisés par la méthode proposée, une tractographie du cerveau entier sur les données réelles et générées a été effectuée. Le produit de cette opération est un tractogramme composé de courbes 3D discrètes appelées *streamlines*. Les tractogrammes résultants sont ensuite segmenté en faisceaux (*bundles*) regroupant des fibres anatomiquement reliées. Les tractogrammes générés ont ensuite été comparés avec les données réelles, considérées comme vérité terrain, et des mesures quantitatives ont été calculées sur les *streamlines*. De même, les faisceaux segmentés sont comparés afin de mesurer à quel point la tractographie est impactée par le processus de synthèse de diffusion. Dans les sous-sections suivantes, chaque étape de l'analyse est présentée.

3.7.1 Génération et regroupement des *streamlines*

La tractographie a été effectuée à l'aide de l'algorithme EuDX (Garyfallidis, 2013) avec un pas (*step*) de 0,5 mm. Un angle maximum de 60 degrés a été utilisé entre les pas, en utilisant la direction principale des tenseurs de diffusion et les maxima des ODF. Les maxima ont été extraits des ODF à l'aide de scilpy¹. La sélection aléatoire des points de départ (*seeding*) a été faite avec 2 initialisations par voxel sur l'ensemble du masque de matière blanche, qui a été calculé à partir de l'image de vérité terrain T1w en utilisant Dipy (Garyfallidis *et al.*, 2014). Les *streamlines* d'une longueur inférieure à 10 mm ou supérieure à 300 mm ont été rejetés. Les tractogrammes du cerveau entier ont ensuite été segmentés à l'aide de RecoBundleX (Garyfallidis, Côté, Rheault, Sidhu, Hau, Petit, Fortin, Cunanne & Descoteaux, 2018), en utilisant 80 faisceaux (Yeh, Panesar, Fernandes, Meola, Yoshino, Fernandez-Miranda, Vettel & Verstynen, 2018) comme référence. Pour permettre une comparaison plus robuste, et parce que le placement des points initiaux des *streamline* est aléatoire et peut avoir un impact sur les streamlines reconstruites, la tractographie

¹ <https://github.com/scilus/scilpy>

a été effectuée cinq fois avec différentes graines aléatoires sur chaque volume de diffusion analysé.

3.7.2 Évaluation des *streamlines* et des faisceaux

La similitude entre les *streamlines* et les faisceaux reconstruits à partir des volumes de diffusion réels et générés est évaluée à l'aide de trois mesures : la longueur des *streamlines*, le volume des faisceaux et la correspondance de la forme des faisceaux. Tout d'abord, la longueur des streamlines (en mm) entre les tractogrammes réels et synthétisés est mesurée pour chaque faisceau segmenté. Le volume occupé par le tractogramme du cerveau entier et pour chaque faisceau segmenté est mesuré voxel par voxel et comparé. De plus, la similarité entre la forme des tractogrammes reconstruits et réels est mesurée à l'aide du coefficient Dice, de leur chevauchement (*overlap* – OL) et de leur dépassement (*overreach* – OR). L'OL, défini comme

$$\text{OL} = \frac{|B \cap A|}{|A|}, \quad (3.2)$$

où A , B sont des masques de faisceau binaires, quantifie combien le volume du faisceau A est reconstruit par le faisceau B . L'OR, exprimé comme

$$\text{OU} = \frac{|B \cup A| - |B \cap A|}{|A|}, \quad (3.3)$$

évalue la quantité de faisceau B qui dépasse le faisceau A . Les faisceaux segmentés ont été fusionnés pour permettre une comparaison par paires entre les données réelles et générées. Les faisceaux fusionnés avec moins de 100 *streamlines* ont été écartés de l'analyse. Pour évaluer la qualité globale de la reconstruction, le Dice, OL et OR entre les tractogrammes.

La figure 3.2 présente quelques-uns des faisceaux reconstruits à partir de données réelles et synthétisées.

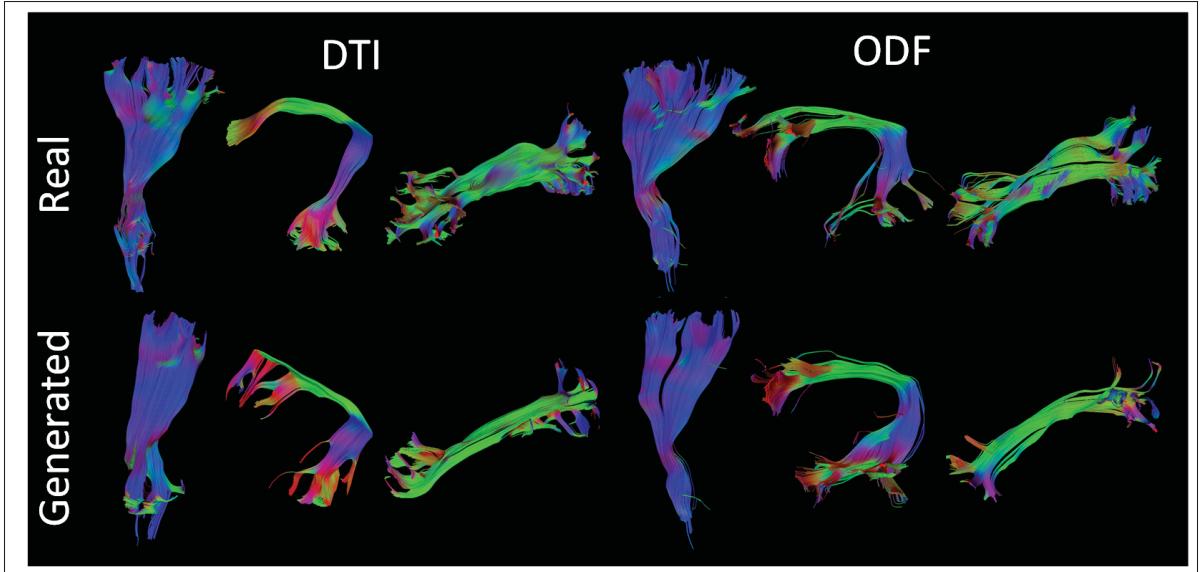


Figure 3.2 Visualisation de trois faisceaux segmentés à partir de tractogrammes réels et synthétisés. (**À gauche**), tractus corticospinal gauche (CST), fascicule arqué (AF) et fascicule longitudinal inférieur (ILF) segmentés à partir de données DT réelles et générées. (**À droite**), CST gauche, AF et ILF segmentés à partir de données ODF réelles et synthétisées.

3.8 Analyse de la diffusion synthétisée

Les performances du modèle sont rapportées dans la Table 3.2 avec un nombre croissant de sujets appariés utilisés comme information a priori dans l'éq(2.5). Pour les DT et les ODF, la similarité cosinus moyenne augmente avec le nombre de sujets appariés et atteint son maximum lorsque l'extraction des *patches* d'entraînement appariés proviennent de 50 sujets distincts. Comme le nombre de *patches* extraits demeure constant (50 000 *patches*) peu importe la taille du jeu de donnée apparié, 50 sujets représentent une moyenne de 1 000 *patches* par sujet. Cet échantillonnage de 1 000 *patches* par sujet semble un bon compromis entre la diversité et la couverture de l'échantillonnage. Dans cette configuration, l'approche proposée donne une similarité cosinus moyenne de $0.8648 \pm 5.5\text{e-}3$ et $0.8846 \pm 4.4\text{e-}3$ pour les voxels avec une FA ≥ 0.2 pour les DT et les ODF respectivement. Dans les voxels avec une FA ≥ 0.5 , une similarité cosinus moyenne de $0.9167 \pm 3.2\text{e-}3$ est atteinte pour les DT et $0.9425 \pm 1.4\text{e-}4$ pour les ODF. Comme indiqué

Tableau 3.2 La similarité cosinus entre les orientations principales réelles et générées par la méthode proposée avec un nombre croissant de sujets appariés utilisé comme information préalable dans eq(2.5). Comme les métriques sont plus pertinentes dans les régions qui encodent des informations de fibres cérébrales, elles sont rapportées avec un seuil de FA croissant de 0.2 et 0.5.

Sujets Appariés	Similarité Cosinus (DT)		Similarité Cosinus (ODF)	
	FA ≥ 0.2	FA ≥ 0.5	GFA ≥ 0.2	GFA ≥ 0.5
0	0.7315	0.7955	0.7029	0.7380
10	0.7795	0.8599	0.8664	0.9058
25	0.8152	0.8863	0.8745	0.9081
50	0.8648	0.9167	0.8846	0.9425
75	0.8317	0.8993	0.8674	0.9051

dans Table 3.3, cela correspond à des FA MSE de $0,0089 \pm 1.5\text{e-}3$ et $0,0159 \pm 1.4\text{e-}4$ pour les DT et à des GFA MSE de $0,0229 \pm 2.8\text{e-}4$ et $0,0614 \pm 1.3\text{e-}3$ pour les ODF.

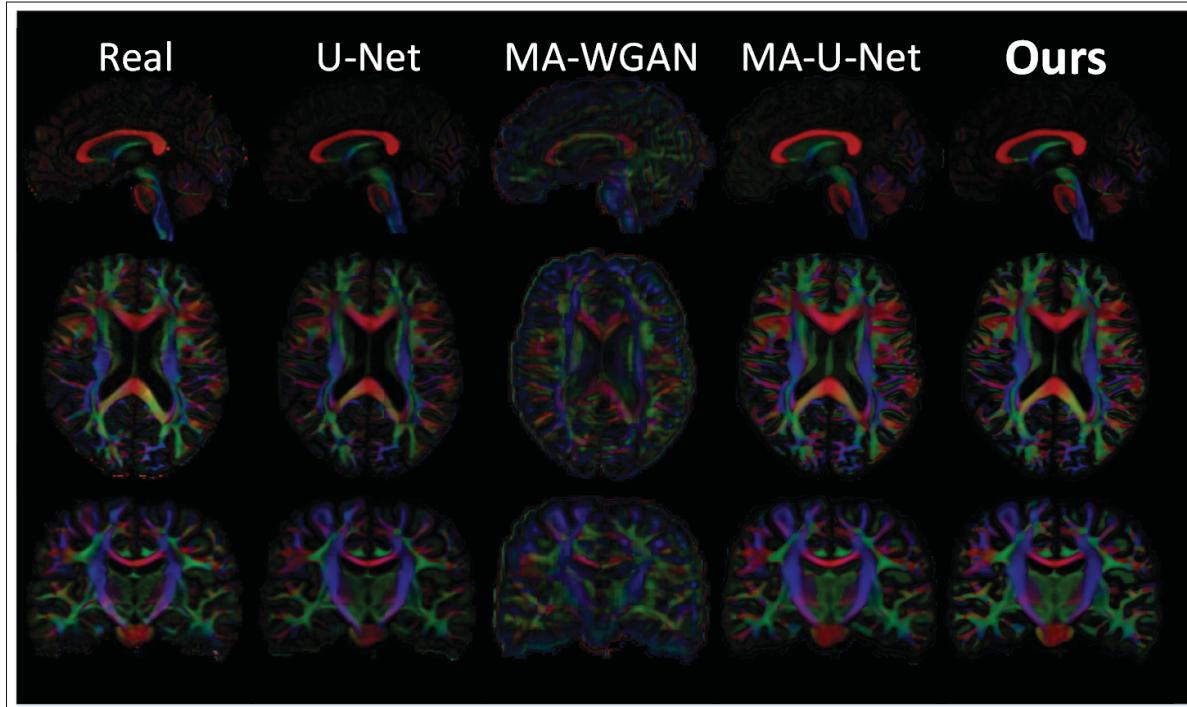


Figure 3.3 Comparaison de la FA colorée par l'orientation principale des DT générés par les méthodes comparées pour des tranches dans le plan sagittal, coronal et axial d'un sujet de test.

Tableau 3.3 L'erreur quadratique moyenne de FA (FA MSE), la géodésique et la similarité cosinus entre les orientations principales des images de diffusion réelles et prédites par les différentes méthodes. Comme les métriques sont plus pertinentes dans les régions qui encodent des informations de fibres cérébrales, elles sont rapportées avec un seuil de FA croissant de 0.2 et 0.5.

	Méthode	FA* MSE		Géodésique		Sim. Cosinus	
		FA* ≥ 0.2	FA* ≥ 0.5	FA* ≥ 0.2	FA* ≥ 0.5	FA* ≥ 0.2	FA* ≥ 0.5
DT	U-Net	0.0100	0.0207	2.2559	2.5457	0.8266	0.8795
	MA-WGAN	0.0294	0.0781	0.7791	1.046	0.5724	0.6246
	MA-U-Net	0.0088	0.0184	0.3947	0.5415	0.8288	0.8863
	Ours	0.0089	0.0159	0.3531	0.4809	0.8648	0.9167
ODF	U-Net	0.0220	0.0559	0.3639	0.6299	0.8702	0.9045
	MA-WGAN	0.0443	0.1173	0.5351	0.9121	0.6823	0.7278
	MA-U-Net	0.0219	0.0563	0.3611	0.6249	0.8710	0.9047
	Ours	0.0229	0.06145	0.3467	0.6238	0.8846	0.9425

* : FA est utilisée pour les DT et GFA pour les ODF.

A titre de comparaison, les performances des approches de référence lorsqu'elles sont entraînées avec les mêmes 50 sujets appariés sont rapportées dans la Table 3.3. Comme on peut le voir, le modèle proposé obtient les meilleures performances pour toutes les métriques lorsqu'il est entraîné sur des images DT, ainsi qu'une meilleure estimation de l'orientation des fibres et géodésique que les approches de référence lorsqu'il est entraîné avec les données ODF.

Par rapport au modèle U-Net entraîné de manière entièrement supervisé, qui n'impose pas de contrainte de variété sur la sortie, notre modèle améliore la FA MSE de 23,14%, la distance géodésique moyenne de 81,11% et la similarité cosinus moyenne de 4,23% dans les régions avec $FA \geq 0.5$ pour les DT. Cela montre l'intérêt d'imposer des contraintes de variété sur la sortie du réseau. Sans ces contraintes, U-Net génère en moyenne 3 843,5 tenseurs non SPD et 1 559,7 ODF non PDF comme indiqué dans le Table 3.4. L'approche proposée obtient également des performances statistiquement meilleures (t-test pairé, $p < 0,05$), pour les données DT et ODF, sur la similarité cosinus moyenne et la distance géodésique moyenne par rapport à MA-U-Net, qui n'inclut pas la perte cyclique et n'est entraîné qu'avec des données appariées.

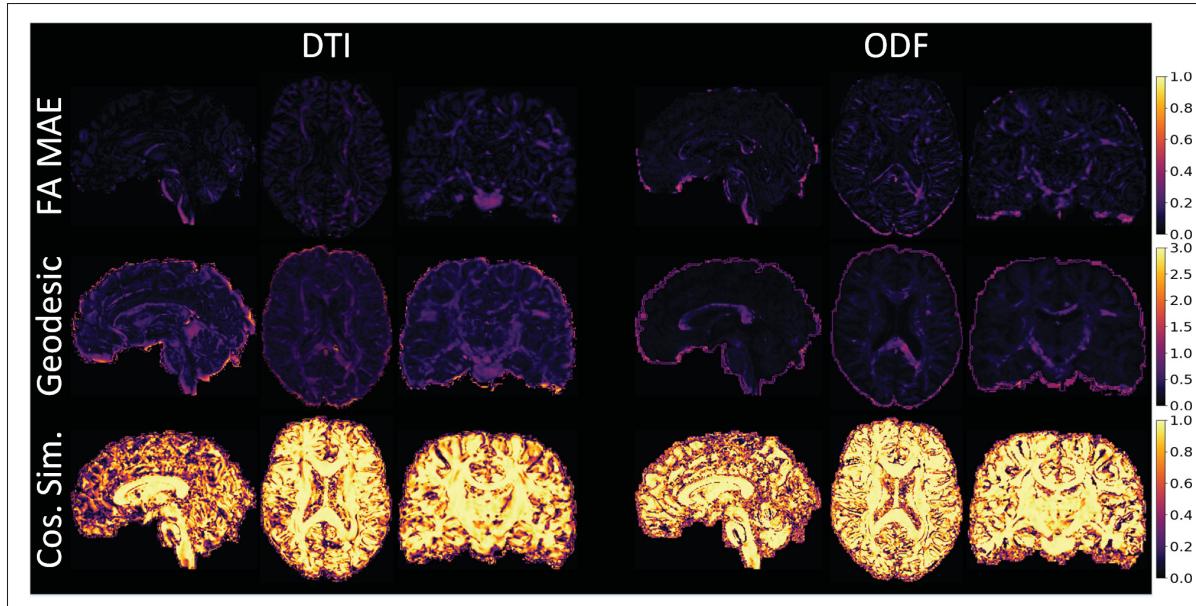


Figure 3.4 Métriques sur les tranches sagittales, axiales et coronales entre le volume de diffusion haute-résolution généré pour un sujet de test aléatoire et le volume réel suréchantillonné. (**Première ligne**) FA MAE, (**ligne du centre**) la distance géodésique et (**ligne du bas**) la similarité cosinus.

Tableau 3.4 La quantité générée de tenseurs qui ne sont pas SPD et d'ODF qui ne sont pas des PDF par les méthodes comparées pour cinq sujets de test sélectionnés aléatoirement. L'apprentissage sur variété riemannienne (i.e., *manifold-awareness*) assure que les données de diffusion soient exactement sur leur variété riemannienne respective.

Méthode	Apprentissage sur variété riemannienne	Tenseurs non SPD	ODF non PDF
U-Net	✗	3843.5 ± 481.22	1559.7 ± 122.3
MA-U-Net	✓	0	0
MA-WGAN	✓	0	0
Ours	✓	0	0

Les améliorations sont particulièrement importantes pour les voxels avec une $FA \geq 0.5$, où la méthode proposée obtient une similarité cosinus moyenne plus élevée de 3,43%, une distance géodésique réduite de 11,19% et un FA MSE amélioré de 13,58% pour les DT, ainsi qu'une

similarité cosinus moyenne plus élevée de 4,18% pour les ODF. Cela démontre l'impact de la perte pondérée par l'anisotropie décrite dans la Section 2.6, qui donne plus d'importance aux voxels avec des valeurs d'anisotropie plus grandes.

3.9 Longueur et volume des *streamlines*

La Figure II-1 rapporte la longueur moyenne des *streamlines* (en mm) et le volume (en voxels) pour les tractogrammes du cerveau entier reconstruits à partir de la diffusion réelle et générée. La Figure 3.5 illustre la longueur moyenne des *streamlines* et la Figure II-2 le volume moyen pour chacun des faisceaux segmentés.

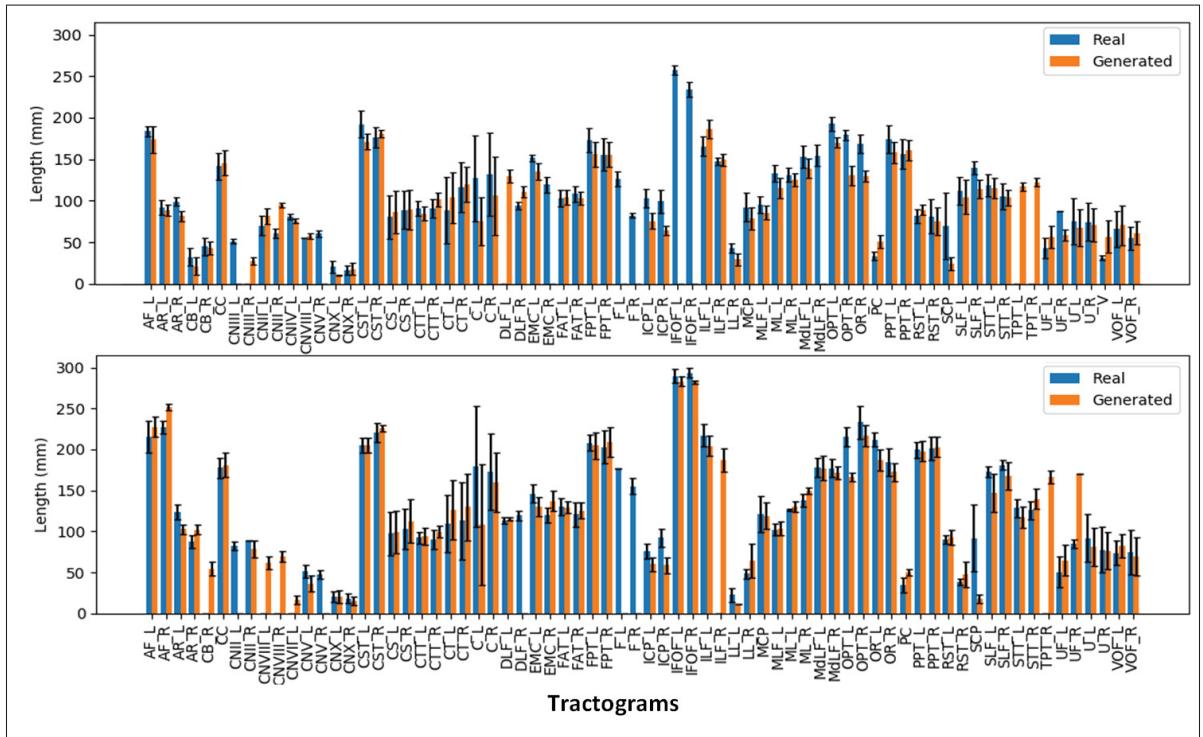


Figure 3.5 Longueur des *streamlines* moyenne (en mm) pour les faisceaux segmentés de la tractographie DT (**ligne du haut**) et de la tractographie ODF (**ligne du bas**).

Les résultats démontrent que les tractogrammes reconstruits sont de taille très similaire, mais que les faisceaux individuels segmentés à partir des données synthétisées ont tendance à être plus courts. En effet, en regardant les longueurs et volumes moyens des faisceaux rapportés

Tableau 3.5 Mesures sur les faisceaux de fibres segmentés depuis la tractographie sur les données réelles et générées.

		Réel (moyenne \pm ds)	Généré (moyenne \pm ds)
DT	Longueur (mm)	99.99 ± 54.37	85.01 ± 50.02
	Volume (voxels)	15963.62 ± 24199.55	14119.52 ± 24274.11
ODF	Longueur (mm)	120.94 ± 72.29	118.13 ± 71.57
	Volume (voxels)	24210.36 ± 42082.48	24408.78 ± 44255.32

dans le Table 3.5, on peut voir que les faisceaux segmentés à partir des tractogrammes DT/ODF générés sont respectivement 14,98% et 2,32% plus courts que ceux provenant des données réelles. Néanmoins, les ODF générées ont tendance à produire des faisceaux légèrement plus volumineux avec un volume moyen de ~24.408 voxels contre ~24.210 pour les ODF réelles. Alors que certains faisceaux n'ont été segmentés que sur les données réelles ou sur les données générées (IFOF vs. TPT par exemple), on peut observer que pour les faisceaux reconstruits dans les données réelles et générées les statistiques sont similaires similaires.

3.10 Forme des *streamlines*

La Figure II-1 montre les mesures de Dice, OL et OR entre les tractogrammes provenant de diffusion réelle et générée. Les mêmes mesures sont ensuite détaillées dans la Figure II-3 pour tous les faisceaux segmentés. On voit sur ces deux figures que, malgré leur similitude en longueur et en volume, l'espace occupé par les faisceaux segmentés à partir de données réelles et générées peut varier. Cette disparité peut être observée dans la Figure 3.2 où trois faisceaux segmentés réels et générés sont comparés. De plus, on observe une grande variabilité dans les métriques inter-faisceau. Par exemple, des faisceaux tels que le *Frontal Aslant Tract* (FAT), le *Frontopontine* (FPT) ou le *Vertical Occipital Fasciculus* (VOF) atteignent une correspondance Dice réel-généré élevée (coefficient de Dice près de 1) alors que le *Superior Cerebellar Peduncle* (SCP) et l'*Occipitopontine Tract* (OPT) ont une faible correspondance réel-généré. Cette grande variabilité inter-faisceau peut être d'avantage appréciée dans le Table 3.6 où la moyenne du coefficient Dice, OL et OR est notée avec leur écart type considérant tous les faisceaux segmentés.

Tableau 3.6 Métriques de correspondance entre les faisceaux de fibres segmentés depuis la tractographie sur les données réelles et générées.

	Dice (moyenne \pm ds)	OL (moyenne \pm ds)	OR (moyenne \pm ds)
DT	0.46 ± 0.24	0.38 ± 0.21	1.36 ± 0.86
ODF	0.54 ± 0.24	0.42 ± 0.19	1.23 ± 0.58

A partir de Table 3.6, on peut également voir que les faisceaux segmentés de l'ODF généré sont plus fidèles à leur contrepartie réelle que DT avec un Dice moyen de $0,54 \pm 0,24$, un OL moyen de $0,42 \pm 0,19$ et un OR moyen de $1,23 \pm 0,58$.

CHAPITRE 4

DISCUSSION

Dans ce mémoire, une nouvelle architecture d'apprentissage profond riemannienne pour la synthèse de données 3D à valeurs sur variétés riemanniennes est proposée et testée sur deux tâches : 1) la génération de tenseurs de diffusion et 2) la génération de fonctions de diffusion d'orientation de diffusion. Plus précisément, la faisabilité de générer des images DT et ODF en haute-résolution à partir d'images T1w structurelles HR et de diffusion LR non appariée a été explorée. La Table 3.4 montre qu'un modèle standard reposant uniquement sur des opérations euclidiennes ne parvient pas à capturer la géométrie des variétés riemanniennes des données de diffusion, ce qui mène à l'estimation d'une diffusion physiquement incorrecte. Pour atténuer ce problème, un cadriel. bâtit sur des avancées récentes dans le traitement des données à valeurs sur variétés riemanniennes (Arsigny *et al.*, 2006; Cheng *et al.*, 2009; Huang *et al.*, 2019), est proposé pour assurer la validité de la diffusion générée. Les propriétés des volumes générés ont été évaluées à l'aide des erreurs quadratiques moyennes des cartes FA/GFA, des distances géodésiques et des similitudes de cosinus entre les orientations principales réelles et prédictes des fibres. Pour évaluer l'intégrité de la diffusion synthétisée dans une application de diffusion typique, une tractographie a été effectuée sur les données générées et réelles et évaluée en comparant les longueurs, les volumes et les formes des tractogrammes et faisceaux segmentés résultants.

4.1 Performance de la synthèse de diffusion

La génération de DT/ODF reposant uniquement sur une image T1w est un problème mal posé pour lequel une même intensité T1w peut correspondre à plusieurs arrangements de fibres. Cependant, en fournissant les informations contextuelles requises pour que le réseau localise l'entrée structurelle, on observe que des arrangements de fibres fortement structurés peuvent être synthétisés avec succès par la méthode proposée et les approches de référence. Par conséquent, la géométrie de haut niveau du cerveau a un impact global sur l'organisation sous-jacente des

fibres. Cela semble être particulièrement vrai dans les régions d'anisotropie plus élevée où les faisceaux de fibres sont fortement organisés. En conséquence, une meilleure estimation de l'orientation des fibres principales dans les régions à FA/GFA plus élevé est observée contre une estimation généralement médiocre des orientations principales dans les régions à forte variance inter-sujets comme les ventricules.

L'utilisation d'images structurelles HR pour piloter la synthèse de la diffusion permet de récupérer des détails anatomiques fins et des contours plus nets mieux que les méthodes basées sur l'interpolation. En tirant parti des informations détaillées contenues dans les images structurelles HR, le réseau n'est pas limité par la grossièreté du signal de diffusion LR d'entrée, comme c'est le cas pour l'interpolation.

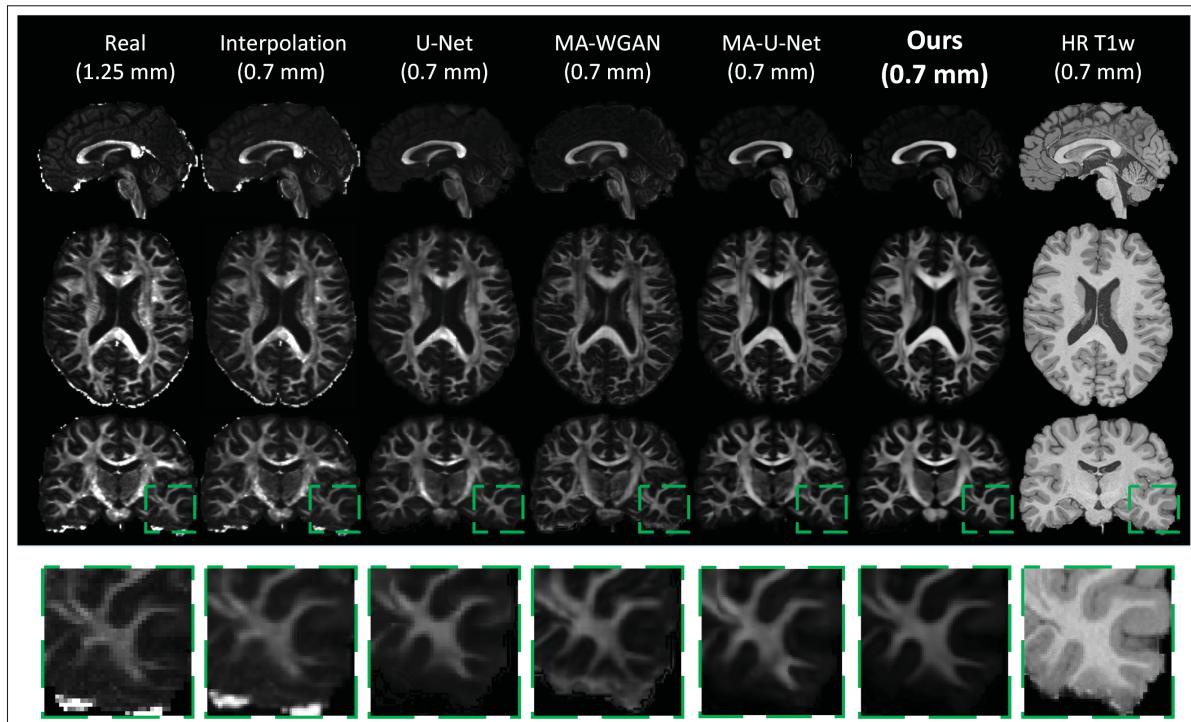


Figure 4.1 GFA de la diffusion basse-résolution réelle (**première colonne**), diffusion réelle suréchantillonnée (**deuxième colonne**) et la diffusion haute-résolution générée par les méthodes comparées. On peut remarquer la capacité de notre réseau à récupérer des détails anatomiques fins qui sont bien visibles dans l'image T1W haute-résolution, mais pas dans la diffusion originale.

Ce transfert d'informations des images HR aux images LR est renforcé par la perte cyclique qui préserve une bonne cohérence structurelle entre les entrées T1w HR et la diffusion générée. De plus, les pertes antagonistes et voxel-par-voxel sur les images appariées a priori aident à récupérer des modèles de fibres plausibles mieux que les approches de référence en tirant profit d'une diversité d'exemples non appariés de diffusion réelle. Par ailleurs, la conscience (*awareness*) de la variété riemannienne par le réseau lors de l'entraînement donne des métriques comparables ou meilleures par rapport aux architectures euclidiennes équivalentes. Cependant, l'ajout de contraintes de synthèse sur variétés riemanniennes a l'avantage de garantir les propriétés mathématiques de la diffusion synthétisée, indépendamment de la progression de l'entraînement du réseau.

4.2 Performance de la tractographie

4.2.1 DT vs. ODF

La Table 3.5 permet d'observer que les faisceaux segmentés à partir de la tractographie sur les données ODF ont tendance à être plus longs et plus volumineux que ceux segmentés à partir de la tractographie de données DT, ce qui est un comportement attendu. Si une comparaison approfondie entre la tractographie DT et ODF est hors de portée pour ce travail (cf (Farquharson, Tournier, Calamante, Fabinyi, Schneider-Kolsky, Jackson & Connelly, 2013; Thomas, Frank, Irfanoglu, Modi, Saleem, Leopold & Pierpaoli, 2014; Jeurissen, Descoteaux, Mori & Leemans, 2019)), on peut néanmoins apprécier que les DT et ODF synthétisés se comportent de manière similaire à leurs homologues réels dans le cadre de la tractographie.

De plus, on observe que les faisceaux segmentés à partir de la tractographie ODF ont un Dice moyen plus élevé, un chevauchement plus élevé et un débordement plus faible que les faisceaux segmentés à partir de la tractographie DT. Étant donné que le même algorithme a été utilisé pour effectuer la tractographie sur les deux ensembles de données, la principale différence est que la tractographie DT utilise une seule direction tandis que la tractographie ODF peut utiliser plusieurs maxima locaux pour progresser. En tant que tels, les écarts plus faibles dans les formes

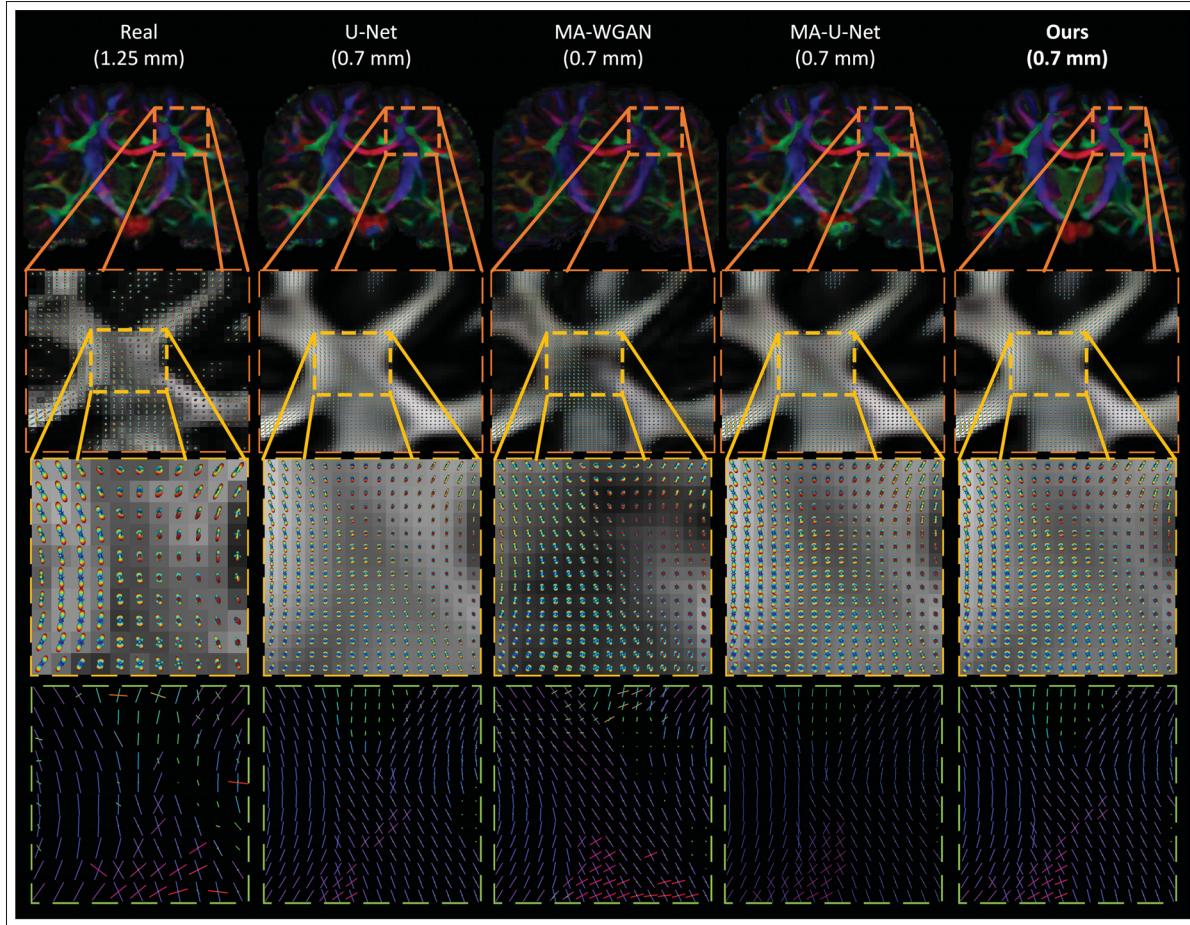


Figure 4.2 Évaluation qualitative des ODF générés par les méthodes comparées.

Première ligne, GFA colorée par l'orientation principale d'une tranche coronale, d'un sujet de test. **Deuxième ligne**, les ODF générées et la GFA d'une région avec croisement de fibres. **Troisième ligne**, une vue rapprochée des ODF générées et de la GFA. **Dernière ligne**, une vue rapprochée des pics des ODF générées.

des faisceaux issus de la tractographie ODF peuvent s'expliquer par le fait que les multiples directions réduisent la sensibilité de l'algorithme aux erreurs locales d'approximation de la direction. À l'inverse, la tractographie DT est connue pour être sensible à l'erreur d'estimation locale (Huang, Zhang, Van Zijl & Mori, 2004). Cette sensibilité, qui conduit souvent à l'arrêt prématûr ou au basculement vers un mauvais tract adjacent de l'algorithme de tractographie (Jeurissen *et al.*, 2019), peut grandement affecter la forme finale des tractogrammes.

4.2.2 Ensembles de fibres

Les études portant sur la génération de données DT et ODF étant encore peu nombreuses, il est difficile de tirer une conclusion définitive sur la qualité des *streamlines* générées sur les données synthétisées. Cependant, on observe à partir de la Figure II-1 que les tractogrammes du cerveau entier ont une longueur de *streamlines* similaire, occupent le même nombre de voxels et ont la même forme. Les faisceaux segmentés, s'ils sont extraits à la fois de données réelles et générées, ont également tendance à présenter la même longueur et le même volume moyens.

Bien que les métriques de forme de faisceaux rapportées puissent sembler indiquer une mauvaise correspondance entre les faisceaux générés à partir de données réelles et synthétisées, il convient de noter que la segmentation de faisceaux est une opération très variable. Par exemple, Rheault, De Benedictis, Daducci, Maffei, Tax, Romascano, Caverzasi, Morency, Corrivetti, Pestilli, Girard, Theaud, Zemmoura, Hau, Glavin, Jordan, Pomiecko, Chamberland, Barakovic, Goyette, Poulin, Chenot, Panesar, Sarubbo, Petit & Descoteaux (2020), qui a analysé la reproductibilité de la segmentation d'un seul faisceau entre experts humains et non-experts, rapporte une différence significative entre le volume de faisceaux segmentés entre experts et non-experts ainsi que les scores médians de Dice autour de 0,77 pour la reproductibilité intra-évaluateur, 0,65 pour la reproductibilité inter-évaluateur et 0,8 pour la reproductibilité avec un *gold standard*. De manière similaire, Schilling, Rheault, Petit, Hansen, Nath, Yeh, Girard, Barakovic, Rafael-Patino, Yu, Fischl-Gomez, Pizzolato, Ocampo-Pineda, Schiavi, Canales-Rodríguez, Daducci, Granziera, Innocenti, Thiran, Mancini, Wastling, Cocozza, Petracca, Pontillo, Mancini, Vos, Vakharia, Duncan, Melero, Manzanedo, Sanz-Morales, Peña-Melián, Calamante, Attyé, Cabeen, Korobova, Toga, Vijayakumari, Parker, Verma, Radwan, Sunaert, Emsell, De Luca, Leemans, Bajada, Haroon, Azadbakht, Chamberland, Genc, Tax, Yeh, Srikanchana, McKnight, Yang, Chen, Kelly, Yeh, Cochereau, Maller, Welton, Almairac, Seunarine, Clark, Zhang, Makris, Golby, Rathi, O'Donnell, Xia, Aydogan, Shi, Fernandes, Raemaekers, Warrington, Michielse, Ramírez-Manzanares, Concha, Aranda, Meraz, Lerma-Usabiaga, Roitman, Fekonja, Calarco, Joseph, Nakua, Voineskos, Karan, Grenier, Legarreta, Adluru, Nair, Prabhakaran, Alexander, Kamagata, Saito, Uchida, Andica, Masahiro, Bayrak, Gandini Wheeler-Kingshott, D'Angelo, Palesi, Savini,

Roland, Guevara, Houenou, López-López, Mangin, Poupon, Román, Vázquez, Maffei, Arantes, Andrade, Silva, Raja, Calhoun, Caverzasi, Sacco, Lauricella, Pestilli, Bullock, Zhan, Brignoni-Perez, Lebel, Reynolds, Nestrasil, Labounek, Lenglet, Paulson, Aulicka, Heilbronner, Heuer, Anderson, Landman & Descoteaux (2020) a analysé la variabilité de la segmentation de 14 faisceaux entre 42 groupes en utilisant à la fois la segmentation manuelle et automatique. Bien que peu de mesures soient rapportées, les chiffres indiquent un score Dice généralement faible, ainsi qu'une grande variabilité dans le volume des faisceaux et la longueur des *streamlines* selon le protocole utilisé. L'analyse de faisceaux spécifiques rapporte des scores Dice entre 0,4 et 0,6 pour la reproductibilité inter-protocole et inter-sujet et des scores Dice entre 0,6 et 0,8 pour la reproductibilité intra-protocole. Il est donc possible de théoriser que les scores Dice rapportés dans le présent travail peuvent être affectés par la variabilité inhérente au processus de segmentation. Ceci dit, l'entraînement par *patches* et le manque d'une métrique globale sur l'ensemble du connectome lors de l'entraînement du réseau peut aussi expliquer les différences de performance entre les métriques locales (FA MSE, sim. cosinue, géodésique) qui sont généralement bonnes et globale (Dice, OR, OL) qui sont plus faibles.

CONCLUSION ET RECOMMANDATIONS

5.1 Résumé des contributions

Ce mémoire présente une nouvelle architecture de réseau riemannienne pour la synthèse de tenseurs de diffusion et d'ODF de diffusion dans l'espace des images T1w en haute-résolution. Les résultats démontrent que l'architecture riemannienne proposée peut synthétiser des images de diffusion valides avec une amélioration de 5% sur l'estimation de l'orientation principales des fibres et une amélioration de 23% de FA MSE par rapport aux approches de référence. Cette amélioration des performances par rapport aux méthodes comparées montre l'avantage d'utiliser à la fois des échantillons appariés et non appariés dans un seul objectif en plus de l'intégration d'une métrique riemannienne à même l'architecture du réseau. De plus, contrairement aux modèles d'apprentissage profond euclidiens standard, qui génèrent en moyenne 3 844 tenseurs invalides et 1 560 ODF invalides par volume, le logiciel présenté dans ce mémoire garantit la cohérence mathématique des schémas de diffusion synthétisés, exempts de tenseurs et d'ODF invalides.

La question de recherche principale de ce mémoire soit, *Est-il possible d'exploiter la géométrie structurelle à haut-niveau du cerveau afin de prédire ses populations de fibres cérébrales sous-jacentes ?* a donc été abordé en développant avec succès les deux objectifs spécifiques suivant :

1. **Objectif 1** - La synthèse d'images de diffusion à partir d'images structurelles par apprentissage de données sur variétés riemanniennes ;
2. **Objectif 2** - La génération faiblement supervisée de super-résolution d'imagerie de diffusion à partir d'images structurelles.

L'atteinte de ces objectifs a davantage été validé en comparant non seulement les scalaires de diffusion, mais aussi les tractogrammes des volumes de diffusion générés à partir d'images T1w HR à leurs homologues réels. Il a été observé que la diffusion générée, à partir d'images T1w HR uniquement, partage des similitudes avec la diffusion réelle en termes de longueur de *streamlines* ainsi que de volume et de forme des faisceaux de fibres. La capacité du réseau à transférer des détails anatomiques fins des images T1w haute résolution aux images de diffusion a aussi été démontré qualitativement. Ce transfert d'informations permet la génération d'images avec des contours plus nets et un niveau de détails plus élevé qui ne pourrait pas être atteint avec des méthodes d'interpolation d'images traditionnelles.

Ainsi, les résultats présentés dans ce travail suggèrent que la géométrie de haut niveau du cerveau, exprimée dans les images structurelles T1w, peut être utilisée pour estimer l'organisation globale des faisceaux de fibres sous-jacente à celle-ci. Tirant parti de ce principe, notre méthode pourrait permettre la synthèse rapide de tenseurs de diffusion et d'ODF dans des situations où l'acquisition d'imagerie de diffusion n'est pas disponible. Plus généralement, ce travail présente la base d'un logiciel applicable à toutes les tâches de traduction d'images à valeurs réelles vers des données à valeurs sur variétés riemannienne.

En somme, l'atteinte des objectifs de recherche a mené aux contributions scientifiques originales suivantes :

1. Le premier réseau d'apprentissage profond riemannien de type CycleGAN pour le mappage entre des images à valeurs réelles et des images à valeurs se trouvant sur les variétés riemanniennes \mathcal{S}_{++}^3 et \mathbb{S}^n ;
2. Le premier modèle d'apprentissage profond pour la super-résolution de DTI et d'ODF guidée par des images IRM structurelles non appariées et avec un minimum de supervision ;

3. Une analyse détaillée des images de diffusion synthétisées à partir d'images structurelles incluant l'évaluation des cartes scalaires de diffusion, de la tractographie et des faisceaux segmentés.

5.2 Limitations et travaux futurs

L'approche présentée dans ce mémoire souffre de différentes limitations méthodologiques qui se doivent d'être détaillées. D'abord, il est important de souligner le fait que les métriques riemannniennes utilisées dans ce travail, soient les métriques log-euclidienne pour tenseurs de diffusion et pour ODF, sont des métriques ayant un champs d'action local seulement. Autrement dit, l'optimisation de telle métrique dans un objectif permet de réduire l'erreur de synthèse de diffusion au sein d'une région (i.e., une *patch*), mais ne tient pas compte d'informations à plus haut niveau comme la connectivité inter-régions du cerveau. De ce fait, même si le réseau présenté dans ce mémoire permet une approximation locale plus juste de la diffusion, il n'offre aucun support pour l'optimisation de métrique sur l'ensemble du connectome. Il serait intéressant de développer une telle métrique globale qui permettrait à la fois la réduction de l'erreur d'estimation locale, tout en assurant la convergence vers un connectome répondant aux caractéristiques anatomiques et de connectivité globale attendues.

Une autre limitation de la méthodologie présentée dans ce mémoire réside dans les données utilisées. En effet, le réseau a été entraîné sur des données de patients sains uniquement. On peut donc se demander quelle est la robustesse du réseau face à des images structurelles exprimant des déformations géométriques associées à des pathologies comme la présence d'une tumeur ou l'atrophie de la matière grise chez les patients atteints d'Alzheimer ? Dans de travaux futurs, il serait intéressant de tester l'algorithme avec comme entrée des images de patients ayant diverses maladies neurologiques et d'en évaluer la diffusion générée. Dans un même ordre d'idée, un réseau génératif tel que présenté dans ce travail pourrait aider à générer les fibres cérébrales

saines d'un patient atteint d'une maladie neurologique particulière. On peut penser à un réseau permettant de générer des portions de diffusions manquantes ou bruitées et qui permettrait d'aider les algorithmes de tractographie à poursuivre à travers ces régions.

Une autre avenue intéressante pour l'extension de ce projet serait d'exploiter la théorie entourant les fonctions réversibles produite par flux normalisés (*normalizing flows*) pour la super-résolution d'images de diffusion. En effet, lorsqu'on observe attentivement le réseau présenté, on remarque que la fonction de sous-échantillonnage non-linéaire apprise \mathcal{F} gagnerait à être réversible. Apprendre une fonction de sous-échantillonnage réversible grâce à un objectif cyclique et antagoniste permettrait à la fois de générer des images de diffusion à partir d'images structurelles, mais aussi de sur-échantillonner de façon intelligente les images de diffusion.

ANNEXE I

RÉTROPROPAGATION POUR L'APPRENTISSAGE DE TENSEURS DE DIFFUSION

Les applications $\log_{\text{Id}}(\cdot)$ et $\exp_{\text{Id}}(\cdot)$ garantissent que notre réseau de générateurs synthétise des tenseurs de diffusion valides. Les deux applications impliquent l'opérateur EIG sur des matrices symétriques 3×3 dont les gradients doivent être définis pour entraîner le réseau avec l'algorithme standard de rétropropagation. En utilisant la généralisation matricielle de la rétropropagation (Ionescu *et al.*, 2015; Huang & Van Gool, 2017), nous définissons les dérivées partielles de la fonction objectif à la couche k^{th} par rapport à un tenseur de diffusion généré \mathbf{M} comme suit :

$$\forall \mathbf{M}_{k-1} \in \mathbf{X}_{k-1}, \frac{\partial L^{(k)}}{\partial \mathbf{M}_{k-1}} = 2\mathbf{U} \left(\mathbf{K}^\top \circ \left(\mathbf{U}^\top \frac{\partial L^{(k')}}{\partial \mathbf{U}} \right)_{sym} \right) \mathbf{U}^\top + \mathbf{U} \left(\frac{\partial L^{(k')}}{\partial \Sigma} \right)_{diag} \mathbf{U}^\top, \quad (\text{A I-1})$$

with

$$\mathbf{M}_{k-1} = \mathbf{U} \Sigma \mathbf{U}^\top, \quad (\text{A I-2})$$

and

$$\mathbf{K}(i, j) = \begin{cases} \frac{1}{\sigma_i - \sigma_j} & i \neq j \\ 0 & i = j, \end{cases} \quad (\text{A I-3})$$

où $\mathbf{A}_{sym} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$ et \mathbf{A}_{diag} est \mathbf{A} avec des éléments non nuls uniquement dans sa diagonale. De plus, $\mathbf{X}_{k-1} \in \mathbb{R}^{B \times 9 \times D \times H \times W}$ est un lot de patchs 3D DT de taille $D \times H \times W$ et \mathbf{K} est construit sur les valeurs propres σ dans Σ de \mathbf{M} .

Les variations $\frac{\partial L^{(k')}}{\partial \mathbf{U}}$ et $\frac{\partial L^{(k')}}{\partial \Sigma}$ dans eq(A I-1) sont définies pour chaque tenseur \mathbf{M} pour les deux applications $\log_{\text{Id}}(\cdot)$ et $\exp_{\text{Id}}(\cdot)$ respectivement comme :

$$\frac{\partial L^{(k')}}{\partial \mathbf{U}} = 2 \left(\frac{\partial L^{(k+1)}}{\partial \mathbf{M}_k} \right)_{sym} \mathbf{U} \log(\Sigma), \quad (\text{A I-4})$$

$$\frac{\partial L^{(k')}}{\partial \Sigma} = \Sigma^{-1} \mathbf{U}^\top \left(\frac{\partial L^{(k+1)}}{\partial \mathbf{M}_k} \right)_{sym} \mathbf{U}, \quad (\text{A I-5})$$

et

$$\frac{\partial L^{(k')}}{\partial \mathbf{U}} = 2 \left(\frac{\partial L^{(k+1)}}{\partial \mathbf{M}_k} \right)_{sym} \mathbf{U} \exp(\Sigma), \quad (\text{A I-6})$$

$$\frac{\partial L^{(k')}}{\partial \Sigma} = \exp(\Sigma) \mathbf{U}^\top \left(\frac{\partial L^{(k+1)}}{\partial \mathbf{M}_k} \right)_{sym} \mathbf{U}. \quad (\text{A I-7})$$

A l'aide de eq(A I-1) et de l'application successive des applications $\exp_{\text{Id}}(\cdot)$ et $\log_{\text{Id}}(\cdot)$, on s'assure que les tenseurs générés sont et restent SPD tout au long de l'apprentissage . Une implémentation open source de ces fonctions différentiables est disponible sur notre Github^{1 2}.

¹ <https://github.com/banctilrobitaille/torch-vectorized>

² <https://torch-vectorized.readthedocs.io/en/latest/>

ANNEXE II

ÉVALUATION QUANTITATIVE SUPPLÉMENTAIRE DE LA TRACTOGRAPHIE

Cette annexe présente des figures supplémentaires, référencées dans le Chapitre 3 et dans le Chapitre 4, qui comparent les tractogrammes réels générés et attendus. Tout d'abord, nous comparons dans Figure II-1 le volume moyen, la longueur de streamline, les dés, l'OL et l'OR en considérant l'ensemble des tractogrammes du cerveau.

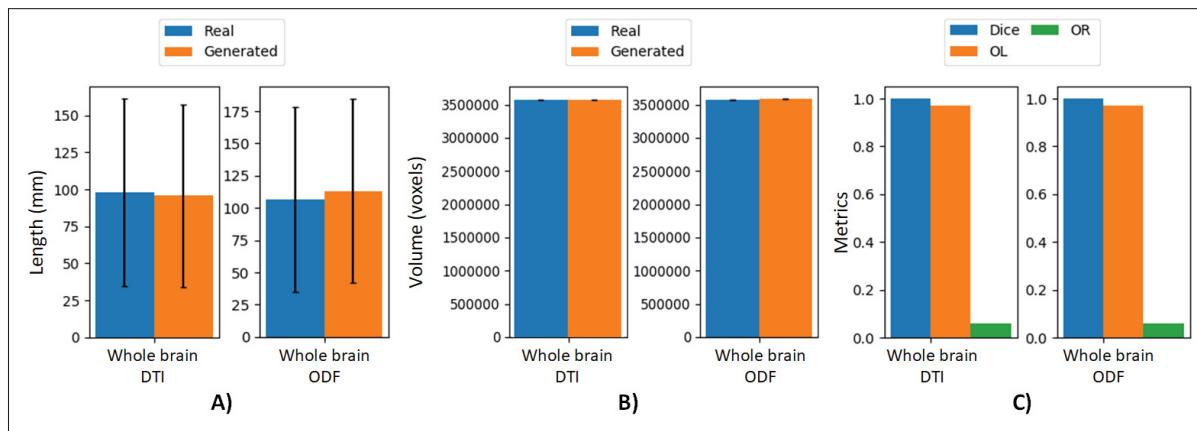


Figure-A II-1 En **A**), la longueur moyenne des *streamlines* (en mm). En **B**, le volume moyen des *streamlines* (en voxels) et en **C**), le Dice moyen, OL et OR.

Nous comparons ensuite le volume moyen de chaque paquet récupéré dans Figure II-2 ci-dessous :

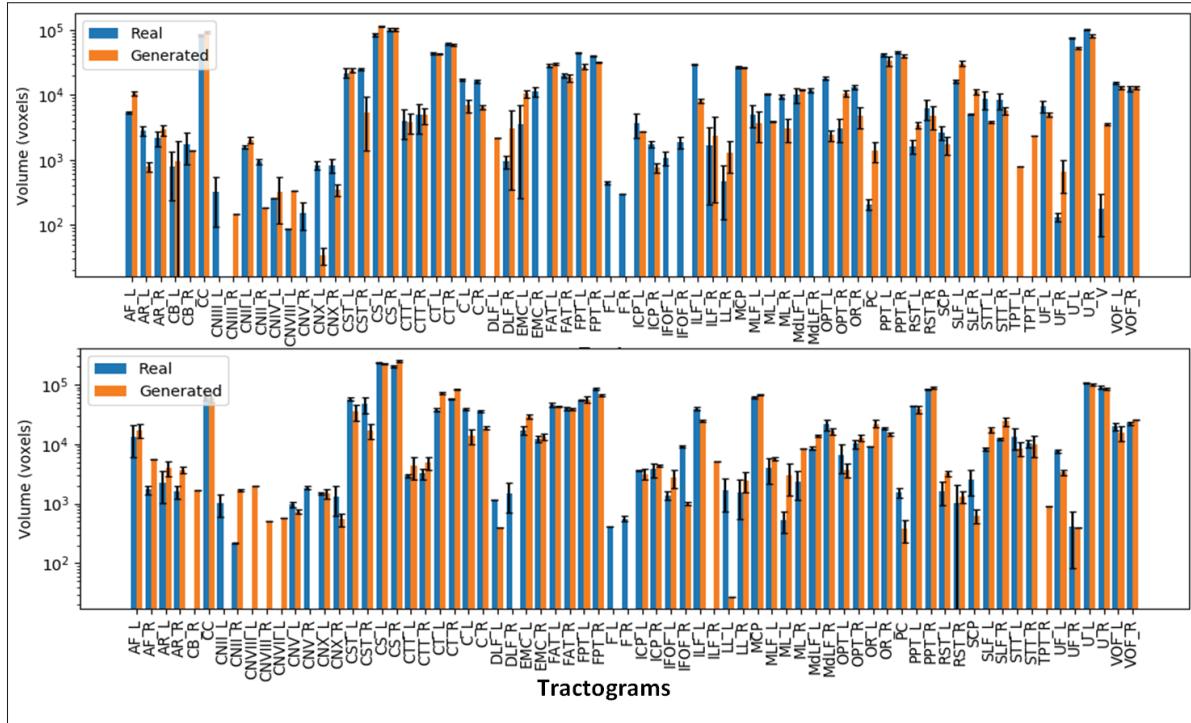


Figure-A II-2 **Ligne du haut :** Volume moyen (en voxels) des faisceaux réels et générés à partir de tenseurs de diffusion. **Ligne du bas :** Volume moyen (en voxels) des faisceaux réels et générés à partir d'ODF

Enfin, Figure II-3 présente le Dice, OL et OR produits par chaque ensemble de fibres.

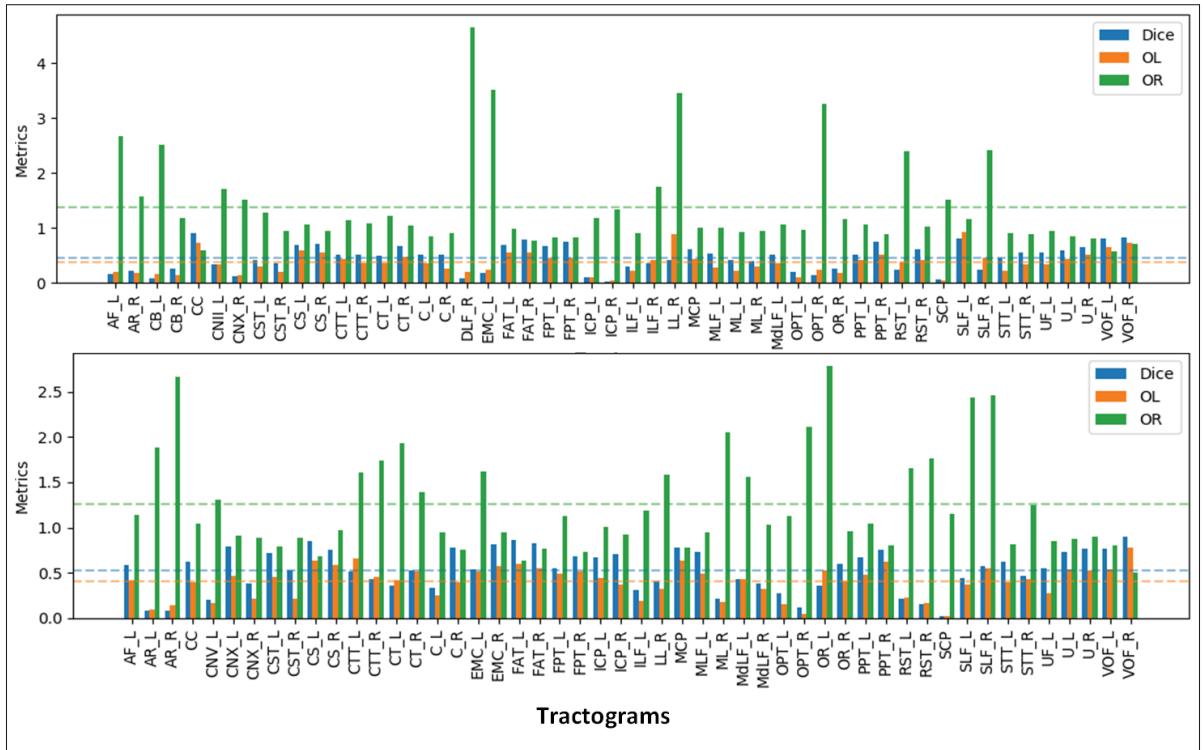


Figure-A II-3 Ligne du haut : Dice, OL et OR entre les faisceaux des tractogrammes réels et générés à partir de tenseurs de diffusion. **Ligne du bas :** Dice, OL et OR entre les faisceaux des tractogrammes réels et générés à partir d'ODF. Les lignes pointillées indique la valeur moyenne pour chaque métrique pour l'ensemble des faisceaux.

ANNEXE III

ÉVALUATION QUALITATIVE SUPPLÉMENTAIRE DES IMAGES GÉNÉRÉES

L'annexe suivante contient les cartes FA et FA colorée par l'orientation principale de cinq sujets de test supplémentaires. On compare d'abord dans Figure III-1 les FA générées et réelles d'une coupe sagittale, frontale et transversale. On compare ensuite dans Figure III-2 les FA colorées par l'orientation principale générées et réelles des cinq mêmes sujets. On peut apprécier comment notre méthode s'adapte à différentes géométries cérébrales et génère une FA et des orientations principales plausibles malgré les différences dans l'anatomie des sujets.

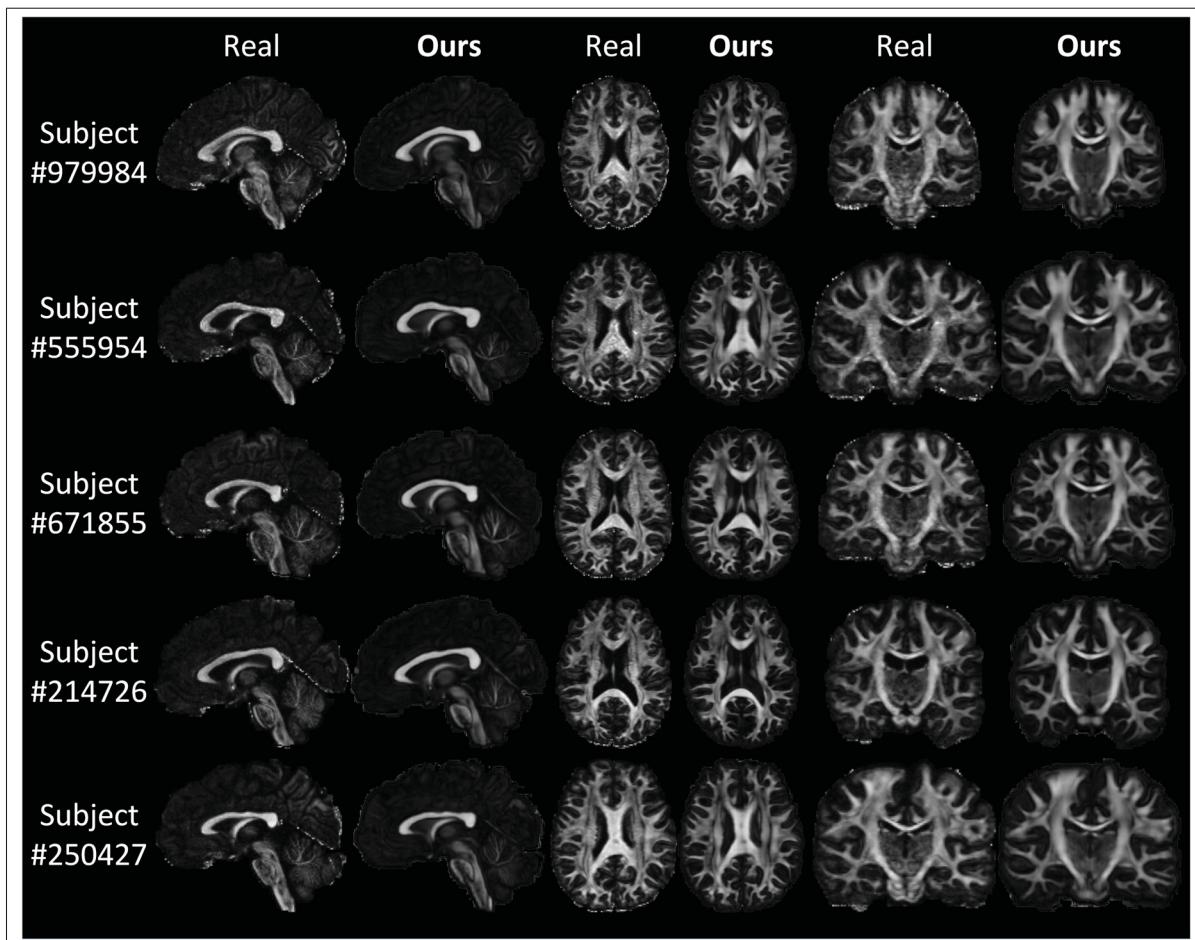


Figure-A III-1 Les cartes FA réelles et synthétisées de cinq sujets de test dans les plans sagittal, frontal et transversal.

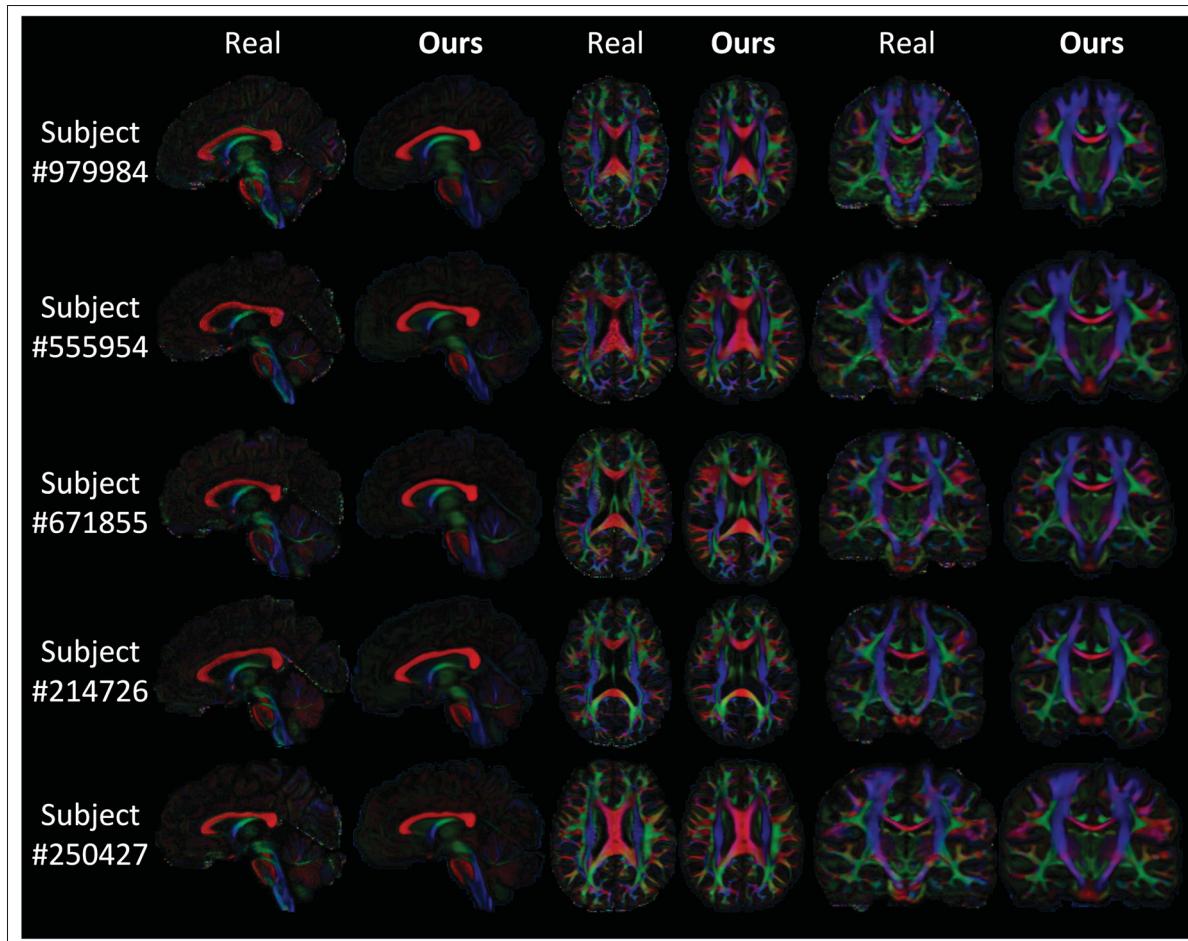


Figure-A III-2 Les cartes FA colorées par la direction principale réelles et synthétisées de cinq sujets de test dans les plans sagittal, frontal et transversal.

BIBLIOGRAPHIE

- Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K. & Harel, N. (2010). Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2), 554-566.
- Agosta, F., Pievani, M., Sala, S., Geroldi, C., Galluzzi, S., Frisoni, G. B. & Filippi, M. (2011). White matter damage in Alzheimer disease and its relationship to gray matter atrophy. *Radiology*, 853–863.
- Alexander, A. L., Hasan, K. M., Lazar, M., Tsuruda, J. S. & Parker, D. L. (2001). Analysis of partial volume effects in diffusion-tensor MRI. *Magnetic Resonance in Medicine : An Official Journal of the International Society for Magnetic Resonance in Medicine*, 45(5), 770-780.
- Alexander, D. C., Zikic, D., Ghosh, A., Tanno, R., Wottschel, V., Zhang, J., Kaden, E., Dyrby, T. B., Sotiropoulos, S. N., Zhang, H. & Criminisi, A. (2017). Image quality transfer and applications in diffusion MRI. *NeuroImage*, 283-298.
- Arjovsky, M., Chintala, S. & Bottou, L. (2017). Wasserstein generative adversarial networks. *International conference on machine learning*, pp. 214-223.
- Arsigny, V., Fillard, P., Pennec, X. & Ayache, N. (2006). Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 411-421.
- Basser, P. J., Mattiello, J. & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 259-267.
- Bloch, F. (1946). Nuclear Induction. *Phys. Rev.*, 460–474.
- Brooks, D., Schwander, O., Barbaresco, F., Schneider, J.-Y. & Cord, M. (2019). Riemannian batch normalization for SPD neural networks. *Advances in Neural Information Processing Systems*, pp. 15489-15500.
- Chakraborty, R., Bouza, J., Manton, J. & Vemuri, B. C. (2019). A Deep Neural Network for Manifold-Valued Data with Applications to Neuroimaging. *Information Processing in Medical Imaging*.
- Chen, Y., Shi, F., Christodoulou, A. G., Xie, Y., Zhou, Z. & Li, D. (2018). Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network.
- Cheng, J., Ghosh, A., Jiang, T. & Deriche, R. (2009). A Riemannian framework for orientation distribution function computing. *Lecture Notes in Computer Science (including subseries*

- Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).*
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. (2016). 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 424-432.
- Costabile, J. D., Alaswad, E., D'Souza, S., Thompson, J. A. & Ormond, D. R. (2019). Current Applications of Diffusion Tensor Imaging and Tractography in Intracranial Tumor Resection. *Frontiers in Oncology*, 426.
- Coupé, P., Manjón, J. V., Chamberland, M., Descoteaux, M. & Hiba, B. (2013). Collaborative patch-based super-resolution for diffusion-weighted images. *NeuroImage*, 246-261.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, 2(4), 303-314.
- Dar, S. U., Yurt, M., Karacan, L., Erdem, A., Erdem, E. & Cukur, T. (2019). Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks. *IEEE Transactions on Medical Imaging*, 38(10), 2375-2388.
- Davis, K. L., Stewart, D. G., Friedman, J. I., Buchsbaum, M., Harvey, P. D., Hof, P. R., Buxbaum, J. & Haroutunian, V. (2003). White matter changes in schizophrenia : evidence for myelin-related dysfunction. *Archives of general psychiatry*, 443–456.
- de Lutio, R., D'Aronco, S., Wegner, J. D. & Schindler, K. (2019). Guided Super-Resolution as Pixel-to-Pixel Transformation. *Proceedings of the IEEE International Conference on Computer Vision*, 8828-8836.
- Descoteaux, M. (2015). High Angular Resolution Diffusion Imaging (HARDI). *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1-25.
- Descoteaux, M., Angelino, E., Fitzgibbons, S. & Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 497-510.
- Du, J., Goh, A. & Qiu, A. (2012). Diffeomorphic metric mapping of high angular resolution diffusion imaging based on riemannian structure of orientation distribution functions. *IEEE Transactions on Medical Imaging*, 1021-1033.
- Du, J., Goh, A., Kushnarev, S. & Qiu, A. (2014). Geodesic regression on orientation distribution functions with its application to an aging study. *NeuroImage*, 416-426.

- Duchi, J., Hazan, E. & Singer, Y. (2011). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12(61), 2121-2159.
- Dyrby, T. B., Lundell, H., Burke, M. W., Reislev, N. L., Paulson, O. B., Ptito, M. & Siebner, H. R. (2014). Interpolation of diffusion weighted imaging datasets. *NeuroImage*, 202-213.
- Elsaid, N. M. & Wu, Y. C. (2019, jul). Super-Resolution Diffusion Tensor Imaging using SRCNN : A Feasibility Study. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2830-2834.
- Farquharson, S., Tournier, J.-D., Calamante, F., Fabinyi, G., Schneider-Kolsky, M., Jackson, G. D. & Connelly, A. (2013). White matter fiber tractography : why we need to move beyond DTI. *Journal of neurosurgery*, 118(6), 1367-1377.
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B. & Dale, A. M. (2004). Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, 14(1), 11-22. doi : 10.1093/cercor/bhg087.
- Garyfallidis, E. (2013). *Towards an accurate brain tractography*. (Thèse de doctorat, University of Cambridge).
- Garyfallidis, E., Brett, M., Amirkhanian, B., Rokem, A., Van Der Walt, S., Descoteaux, M. & Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, 8.
- Garyfallidis, E., Côté, M.-A., Rheault, F., Sidhu, J., Hau, J., Petit, L., Fortin, D., Cunanne, S. & Descoteaux, M. (2018). Recognition of white matter bundles using local and global streamline-based registration and clustering. *NeuroImage*, 170, 283-295. doi : <https://doi.org/10.1016/j.neuroimage.2017.07.015>. Segmenting the Brain.
- Gattellaro, G., Minati, L., Grisoli, M., Mariani, C., Carella, F., Osio, M., Ciceri, E., Albanese, A. & Bruzzone, M. (2009a). White matter involvement in idiopathic Parkinson disease : a diffusion tensor imaging study. *American Journal of Neuroradiology*, 1222–1226.
- Gattellaro, G., Minati, L., Grisoli, M., Mariani, C., Carella, F., Osio, M., Ciceri, E., Albanese, A. & Bruzzone, M. (2009b). White matter involvement in idiopathic Parkinson disease : a diffusion tensor imaging study. *American Journal of Neuroradiology*, 30(6), 1222-1226.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C. & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 105-124.

- Goodfellow, I. (2017). NIPS 2016 Tutorial : Generative Adversarial Networks.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*.
- Gu, X., Knutsson, H., Nilsson, M. & Eklund, A. (2019). *Generating Diffusion MRI Scalar Maps from T1 Weighted Images Using Generative Adversarial Networks*.
- He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Holdsworth, S. J., O'Halloran, R. & Setsompop, K. (2019). *The quest for high spatial resolution diffusion-weighted imaging of the human brain in vivo*. NMR in Biomedicine. John Wiley and Sons Ltd.
- Hopf, H. & Rinow, W. (1931). Ueber den Begriff der vollständigen differentialgeometrischen Fläche. *Commentarii Mathematici Helvetici*, 3(1), 209-225.
- Hornik, K., Stinchcombe, M. & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366.
- Huang, H., Zhang, J., Van Zijl, P. C. & Mori, S. (2004). Analysis of noise effects on DTI-based tractography using the brute-force and multi-ROI approach. *Magnetic Resonance in Medicine*, 52(3), 559-565.
- Huang, J., Friedland, R. & Auchus, A. (2007). Diffusion tensor imaging of normal-appearing white matter in mild cognitive impairment and early Alzheimer disease : preliminary evidence of axonal degeneration in the temporal lobe. *American journal of neuroradiology*, 28(10), 1943-1948.
- Huang, Z. & Van Gool, L. (2017). A riemannian network for SPD matrix learning. *31st AAAI Conference on Artificial Intelligence, AAAI 2017*, 2036-2042.
- Huang, Z., Wu, J. & Van Gool, L. (2019). Manifold-Valued Image Generation with Wasserstein Generative Adversarial Nets. *Proceedings of the AAAI Conference on Artificial Intelligence*, 3886-3893.
- Ioffe, S. & Szegedy, C. (2015). Batch normalization : Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015*, 1, 448-456.

- Ionescu, C., Vantzos, O. & Sminchisescu, C. (2015). *Matrix backpropagation for deep networks with structured layers.*
- Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125-1134.
- Jenkinson, M. & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 143-156.
- Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, 825-841.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. (2012). Review FSL. *NeuroImage*, 782-790.
- Jéquier, E. & Constant, F. (2010). Water as an essential nutrient : the physiological basis of hydration. *European journal of clinical nutrition*, 115–123.
- Jeurissen, B., Descoteaux, M., Mori, S. & Leemans, A. (2019). Diffusion MRI fiber tractography of the brain. *NMR in Biomedicine*, 32(4), e3785.
- Jiang, H., Van Zijl, P. C., Kim, J., Pearlson, G. D. & Mori, S. (2006). DtStudio : Resource program for diffusion tensor computation and fiber bundle tracking. *Computer Methods and Programs in Biomedicine*, 106-116.
- Kantarci, K., Murray, M. E., Schwarz, C. G., Reid, R. I., Przybelski, S. A., Lesnick, T., Zuk, S. M., Raman, M. R., Senjem, M. L., Gunter, J. L., Boeve, B. F., Knopman, D. S., Parisi, J. E., Petersen, R. C., Jack, C. R. & Dickson, D. W. (2017). White-matter integrity on DTI and the pathologic staging of Alzheimer's disease. *Neurobiology of Aging*, 172-179.
- Kelly, S., Jahanshad, N., Zalesky, A., Kochunov, P., Agartz, I., Alloza, C., Andreassen, O. A., Arango, C., Banaj, N., Bouix, S., Bousman, C. A., Brouwer, R. M., Bruggemann, J., Bustillo, J., Cahn, W., Calhoun, V., Cannon, D., Carr, V., Catts, S., Chen, J., Chen, J. X., Chen, X., Chiapponi, C., Cho, K. K., Ciullo, V., Corvin, A. S., Crespo-Facorro, B., Cropley, V., De Rossi, P., Diaz-Caneja, C. M., Dickie, E. W., Ehrlich, S., Fan, F. M., Faskowitz, J., Fatouros-Bergman, H., Flyckt, L., Ford, J. M., Fouche, J. P., Fukunaga, M., Gill, M., Glahn, D. C., Gollub, R., Goudzwaard, E. D., Guo, H., Gur, R. E., Gur, R. C., Gurholt, T. P., Hashimoto, R., Hatton, S. N., Henskens, F. A., Hibar, D. P., Hickie, I. B., Hong, L. E., Horacek, J., Howells, F. M., Hulshoff Pol, H. E., Hyde, C. L., Isaev, D., Jablensky, A., Jansen, P. R., Janssen, J., Jönsson, E. G., Jung, L. A., Kahn, R. S., Kikinis, Z., Liu, K., Klauser, P., Knöchel, C., Kubicki, M., Lagopoulos, J., Langen, C., Lawrie, S., Lenroot, R. K., Lim, K. O., Lopez-Jaramillo, C.,

Lyall, A., Magnotta, V., Mandl, R. C., Mathalon, D. H., McCarley, R. W., McCarthy-Jones, S., McDonald, C., McEwen, S., McIntosh, A., Melicher, T., Mesholam-Gately, R. I., Michie, P. T., Mowry, B., Mueller, B. A., Newell, D. T., O'Donnell, P., Oertel-Knöchel, V., Oestreich, L., Paciga, S. A., Pantelis, C., Pasternak, O., Pearlson, G., Pellicano, G. R., Pereira, A., Pineda Zapata, J., Piras, F., Potkin, S. G., Preda, A., Rasser, P. E., Roalf, D. R., Roiz, R., Roos, A., Rotenberg, D., Satterthwaite, T. D., Savadjiev, P., Schall, U., Scott, R. J., Seal, M. L., Seidman, L. J., Shannon Weickert, C., Whelan, C. D., Shenton, M. E., Kwon, J. S., Spalletta, G., Spaniel, F., Sprooten, E., Stäblein, M., Stein, D. J., Sundram, S., Tan, Y., Tan, S., Tang, S., Temmingh, H. S., Westlye, L. T., Tønnesen, S., Tordesillas-Gutierrez, D., Doan, N. T., Vaidya, J., Van Haren, N. E., Vargas, C. D., Vecchio, D., Velakoulis, D., Voineskos, A., Voyvodic, J. Q., Wang, Z., Wan, P., Wei, D., Weickert, T. W., Whalley, H., White, T., Whitford, T. J., Wojcik, J. D., Xiang, H., Xie, Z., Yamamori, H., Yang, F., Yao, N., Zhang, G., Zhao, J., Van Erp, T. G., Turner, J., Thompson, P. M. & Donohoe, G. (2018). Widespread white matter microstructural differences in schizophrenia across 4322 individuals : Results from the ENIGMA Schizophrenia DTI Working Group. *Molecular Psychiatry*, 23(5), 1261-1269.

Kiefer, J. & Wolfowitz, J. (1952). Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, 23(3), 462 - 466.

Kim, H., Hong, S., Styner, M., Piven, J., Botteron, K. & Gerig, G. (2020). Hierarchical Geodesic Modeling on the Diffusion Orientation Distribution Function for Longitudinal DW-MRI Analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Kim, K. H., Do, W. J. & Park, S. H. (2018). Improving resolution of MR images with an adversarial network incorporating images with different contrast. *Medical Physics*, 45(7), 3120-3131.

Kingma, D. P. & Ba, J. L. (2015). Adam : A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.

Lan, H., Toga, A. & Sepehrband, F. (2020). SC-GAN : 3D self-attention conditional GAN with spectral normalization for multi-modal neuroimaging synthesis. *bioRxiv*, 1.

Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions : Examples Employing Nuclear Magnetic Resonance. *Nature*, 190-191.

LeCun, Y., Haffner, P., Bottou, L. & Bengio, Y. (1999). Object recognition with gradient-based learning. Dans *Shape, contour and grouping in computer vision* (pp. 319-345). Springer.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. & Shi, W. (2017). *Photo-realistic single image super-resolution using a generative adversarial network*.

- Lehmann, T. M., Gonner, C. & Spitzer, K. (1999). Survey : Interpolation methods in medical image processing. *IEEE transactions on medical imaging*, 1049-1075.
- Lei, Y., Harms, J., Wang, T., Liu, Y., Shu, H. K., Jani, A. B., Curran, W. J., Mao, H., Liu, T. & Yang, X. (2019). MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. *Medical Physics*, 46(8), 3565-3581.
- Li, C. & Wand, M. (2016). Precomputed real-time texture synthesis with markovian generative adversarial networks. *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*.
- Li, Y., Huang, J. B., Ahuja, N. & Yang, M. H. (2019). *Joint image filtering with deep convolutional networks*.
- Long, J., Shelhamer, E. & Darrell, T. (2015, June). Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Lu, G., Zhou, Z., Song, Y., Ren, K. & Yu, Y. (2019). Guiding the One-to-One Mapping in CycleGAN via Optimal Transport. *Proceedings of the AAAI Conference on Artificial Intelligence*, 4432-4439.
- Mansfield, P. (1984). *Real-time echo-planar imaging by NMR*.
- Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. *Journal of Physics C : Solid State Physics*, L55-L58.
- Mansfield, P. & Grannell, P. K. (1973). NMR'diffraction'in solids ? *Journal of Physics C : solid state physics*, L422.
- Mao, X., Li, Q., Xie, H., Lau, R. Y., Wang, Z. & Paul Smolley, S. (2017). Least squares generative adversarial networks. *Proceedings of the IEEE international conference on computer vision*, pp. 2794-2802.
- Melzer, T. R., Watts, R., MacAskill, M. R., Pitcher, T. L., Livingston, L., Keenan, R. J., Dalrymple-Alford, J. C. & Anderson, T. J. (2013). White matter microstructure deteriorates across cognitive stages in Parkinson disease. *Neurology*, 1841–1849.
- Mirza, M. & Osindero, S. (2014). Conditional Generative Adversarial Nets.
- Neuner, I., Kupriyanova, Y., Stöcker, T., Huang, R., Posnansky, O., Schneider, F., Tittgemeyer, M. & Shah, N. J. (2010). White-matter abnormalities in Tourette syndrome extend beyond motor pathways. *NeuroImage*, 1184-1193.

- Ni, M., Lei, J., Cong, R., Zheng, K., Peng, B. & Fan, X. (2017). Color-Guided Depth Map Super Resolution Using Convolutional Neural Network. *IEEE Access*, 26666-26672.
- Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q. & Shen, D. (2018). Medical Image Synthesis with Deep Convolutional Adversarial Networks. *IEEE Transactions on Biomedical Engineering*, 65(12), 2720-2730.
- Oouchi, H., Yamada, K., Sakai, K., Kizu, O., Kubota, T., Ito, H. & Nishimura, T. (2007). Diffusion Anisotropy Measurement of Brain White Matter Is Affected by Voxel Size : Underestimation Occurs in Areas with Crossing Fibers. *American Journal of Neuroradiology*, 1102-1106.
- Pennec, X., Sommer, S. & Fletcher, T. (2019). Riemannian Geometric Statistics in Medical Image Analysis. *Riemannian Geometric Statistics in Medical Image Analysis*, 1-636.
- Poot, D. H., Jeurissen, B., Bastiaensen, Y., Veraart, J., Van Hecke, W., Parizel, P. M. & Sijbers, J. (2013). Super-resolution for multislice diffusion tensor imaging. *Magnetic Resonance in Medicine*, 103-113.
- Purcell, E. M., Torrey, H. C. & Pound, R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.*, 37–38.
- Querbes, O., Aubry, F., Pariente, J., Lotterie, J.-A., Démonet, J.-F., Duret, V., Puel, M., Berry, I., Fort, J.-C., Celsis, P. et al. (2009). Early diagnosis of Alzheimer's disease using cortical thickness : impact of cognitive reserve. *Brain*, 2036–2047.
- Radford, A., Metz, L. & Chintala, S. (2016). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.
- Ramsey, C. R. & Oliver, A. L. (1998). *Magnetic resonance imaging based digitally reconstructed radiographs, virtual simulation, and three-dimensional treatment planning for brain neoplasms* (Rapport n°10).
- Ramsey, C. R., Arwood, D., Scaperoth, D. & Oliver, A. L. (1999). Clinical application of digitally-reconstructed radiographs generated from magnetic resonance imaging for intracranial lesions. *International Journal of Radiation Oncology Biology Physics*, 45(3), 797-802.
- Rheault, F., De Benedictis, A., Daducci, A., Maffei, C., Tax, C. M. W., Romascano, D., Caverzasi, E., Morency, F. C., Corrivetti, F., Pestilli, F., Girard, G., Theaud, G., Zemmoura, I., Hau, J., Glavin, K., Jordan, K. M., Pomiecko, K., Chamberland, M., Barakovic, M., Goyette, N., Poulin, P., Chenot, Q., Panesar, S. S., Sarubbo, S., Petit, L. & Descoteaux, M. (2020). Tractostorm : The what, why, and how of tractography dissection reproducibility. *Human Brain Mapping*, 41(7), 1859-1874. doi : <https://doi.org/10.1002/hbm.24917>.

- Riegler, G., Rüther, M. & Bischof, H. (2016). ATGV-net : Accurate depth super-resolution. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 268-284.
- Ronneberger, O., Fischer, P. & Brox, T. (2015). U-Net : Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, pp. 234-241.
- Rosenblatt, F. (1958). *The perceptron : A probabilistic model for information storage and organization in the brain* (Rapport n°6).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. & Chen, X. (2016). Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2234-2242.
- Sanaat, A., Arabi, H., Mainta, I., Garibotto, V. & Zaidi, H. (2020). Projection Space Implementation of Deep Learning-Guided Low-Dose Brain PET Imaging Improves Performance over Implementation in Image Space. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 61(9), 1388-1396.
- Sánchez, I. & Vilaplana, V. (2018). Brain MRI super-resolution using 3D generative adversarial networks. *arXiv preprint arXiv :1812.11440*, 1-8.
- Schilling, K. G., Blaber, J., Huo, Y., Newton, A., Hansen, C., Nath, V., Shafer, A. T., Williams, O., Resnick, S. M., Rogers, B., Anderson, A. W. & Landman, B. A. (2019). Synthesized b0 for diffusion distortion correction (Synb0-DisCo). *Magnetic Resonance Imaging*, 64, 62-70.
- Schilling, K. G., Rheault, F., Petit, L., Hansen, C. B., Nath, V., Yeh, F.-C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., Fisch-Gomez, E., Pizzolato, M., Ocampo-Pineda, M., Schiavi, S., Canales-Rodríguez, E. J., Daducci, A., Granziera, C., Innocenti, G., Thiran, J.-P., Mancini, L., Wastling, S., Cocozza, S., Petracca, M., Pontillo, G., Mancini, M., Vos, S. B., Vakharia, V. N., Duncan, J. S., Melero, H., Manzanedo, L., Sanz-Morales, E., Peña-Melián, Á., Calamante, F., Attyé, A., Cabeen, R. P., Korobova, L., Toga, A. W., Vijayakumari, A. A., Parker, D., Verma, R., Radwan, A., Sunaert, S., Emsell, L., De Luca, A., Leemans, A., Bajada, C. J., Haroon, H., Azadbakht, H., Chamberland, M., Genc, S., Tax, C. M. W., Yeh, P.-H., Srikanthana, R., McKnight, C., Yang, J. Y.-M., Chen, J., Kelly, C. E., Yeh, C.-H., Cochereau, J., Maller, J. J., Welton, T., Almairac, F., Seunarine, K. K., Clark, C. A., Zhang, F., Makris, N., Golby, A., Rathi, Y., O'Donnell, L. J., Xia, Y., Aydogan, D. B., Shi, Y., Fernandes, F. G., Raemaekers, M., Warrington, S., Michielse, S., Ramírez-Manzanares, A., Concha, L., Aranda, R., Meraz, M. R., Lerma-Usabiaga, G., Roitman, L., Fekonja, L. S., Calarco, N., Joseph, M., Nakua, H.,

- Voineskos, A. N., Karan, P., Grenier, G., Legarreta, J. H., Adluru, N., Nair, V. A., Prabhakaran, V., Alexander, A. L., Kamagata, K., Saito, Y., Uchida, W., Andica, C., Masahiro, A., Bayrak, R. G., Gandini Wheeler-Kingshott, C. A., D'Angelo, E., Palesi, F., Savini, G., Rolandi, N., Guevara, P., Houenou, J., López-López, N., Mangin, J.-F., Poupon, C., Román, C., Vázquez, A., Maffei, C., Arantes, M., Andrade, J. P., Silva, S. M., Raja, R., Calhoun, V. D., Caverzasi, E., Sacco, S., Lauricella, M., Pestilli, F., Bullock, D., Zhan, Y., Brignoni-Perez, E., Lebel, C., Reynolds, J. E., Nestrasil, I., Labounek, R., Lenglet, C., Paulson, A., Aulicka, S., Heilbronner, S., Heuer, K., Anderson, A. W., Landman, B. A. & Descoteaux, M. (2020). Tractography dissection variability : what happens when 42 groups dissect 14 white matter bundles on the same dataset ? *bioRxiv*, 1.
- Seehusen, D. A., Reeves, M. & Fomin, D. (2003). Cerebrospinal fluid analysis. *American family physician*, 1103–1108.
- Shi, F., Cheng, J., Wang, L., Yap, P. T. & Shen, D. (2016). Super-resolution reconstruction of diffusion-weighted images using 4D low-rank and total variation. *Mathematics and Visualization*.
- Son, S. J., yong Park, B., Byeon, K. & Park, H. (2019). Synthesizing diffusion tensor imaging from functional MRI using fully convolutional networks. *Computers in Biology and Medicine*, 103528.
- Song, X., Dai, Y. & Qin, X. (2017). *Deep depth super-resolution : Learning depth super-resolution using deep convolutional neural network*.
- Song, X., Dai, Y., Zhou, D., Liu, L., Li, W., Li, H. & Yang, R. (2020). Channel Attention based Iterative Residual Learning for Depth Map Super-Resolution. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 5630-5639.
- Sotropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., Van Essen, D. C., Ugurbil, K. & Behrens, T. E. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*, 125-143.
- Srivastava, A., Jermyn, I. & Joshi, S. (2007). Riemannian analysis of probability density functions with applications in vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Stanescu, T., Hans-Sonke, J., Stavrev, P. & Fallone, B. G. (2006). 3T MR-based treatment planning for radiotherapy of brain lesions. *Radiology and Oncology*, 1.
- Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F. & Descoteaux, M. (2020). TractoFlow : A robust, efficient and reproducible diffusion MRI pipeline leveraging Nextflow

- and Singularity. *NeuroImage*, 218, 116889.
- Thomas, C., Frank, Q. Y., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A. & Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences*, 111(46), 16574-16579.
- Tournier, J., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H. & Connelly, A. (2019). MRtrix3 : A fast, flexible and open software framework for medical image processing and visualisation. *bioRxiv*, 116137.
- Tuch, D. S. (2004a). Q-ball imaging. *Magnetic Resonance in Medicine*, 1358-1372.
- Tuch, D. S. (2004b). Q-ball imaging. *Magnetic Resonance in Medicine : An Official Journal of the International Society for Magnetic Resonance in Medicine*, 52(6), 1358-1372.
- Turner, R., Le Bihan, D., Maier, J., Vavrek, R., Hedges, L. K. & Pekar, J. (1990). Echo-planar imaging of intravoxel incoherent motion. *Radiology*, 407–414.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E. & Ugurbil, K. (2013). The WU-Minn Human Connectome Project : An overview. *NeuroImage*, 62-79.
- Wang, T., Lei, Y., Fu, Y., Wynne, J. F., Curran, W. J., Liu, T. & Yang, X. (2021). A review on medical imaging synthesis using deep learning and its clinical applications. 22(1), 11-36.
- Wang, Z., Chen, J. & Hoi, S. C. H. (2020). Deep Learning for Image Super-resolution : A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1.
- Williams, R. W. & Herrup, K. (1988). The control of neuron number. *Annual Review of Neuroscience*, 423–453.
- Yap, P. T., An, H., Chen, Y. & Shen, D. (2014). Fiber-driven resolution enhancement of diffusion-weighted images. *NeuroImage*, 939-950.
- Yeh, F.-C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J. C., Vettel, J. M. & Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage*, 178, 57-68. doi : <https://doi.org/10.1016/j.neuroimage.2018.05.027>.
- Yi, X., Walia, E. & Babyn, P. (2019). Generative adversarial network in medical imaging : A review. *Medical Image Analysis*, 101552.
- Yi, Z., Zhang, H., Tan, P. & Gong, M. (2017). DualGAN : Unsupervised Dual Learning for Image-to-Image Translation. *Proceedings of the IEEE International Conference on Computer*

Vision, 2017-October, 2868-2876.

Zhao, H., Gallo, O., Frosio, I. & Kautz, J. (2016). Loss Functions for Image Restoration With Neural Networks. *IEEE Transactions on Computational Imaging*, 47-57.

Zhong, J., Wang, Y., Li, J., Xue, X., Liu, S., Wang, M., Gao, X., Wang, Q., Yang, J. & Li, X. (2020). Inter-site harmonization based on dual generative adversarial networks for diffusion tensor imaging : Application to neonatal white matter development. *BioMedical Engineering Online*, 4.

Zhu, J. Y., Park, T., Isola, P. & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision*, 2242-2251.